A HIGH-SPEED PARALLEL PIPELINE A/D CONVERTER TECHNIQUE IN CMOS

by

Cormac S. G. Conroy

Memorandum No. UCB/ERL M94/9

15 February 1994

A HIGH-SPEED PARALLEL PIPELINE A/D CONVERTER TECHNIQUE IN CMOS

by

Cormac S. G. Conroy

Memorandum No. UCB/ERL M94/9

15 February 1994

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Abstract

A High-Speed Parallel Pipeline A/D Converter Technique in CMOS

by

Cormac Séamus Gerard Conroy

Doctor of Philosophy in

Engineering — Electrical Engineering and Computer Sciences

University of California at Berkeley

Professor Paul R. Gray, Chair

Within electronic signal processing, analog-to-digital conversion is a key function required to interface between "real-world" analog signals and digital computers. The particular focus of this research is a resolution of 7–10 bits and a sampling frequency of 50–100 megasamples per second (MS/s). Example applications are high-speed waveform acquisition, HDTV, image processing, and digital communications. Monolithic integrated circuit (IC) solutions with these sample rate and resolution specifications already exist in bipolar technology, for example, the folding and interpolation architecture. In CMOS, however, the speed has been limited to about 50 MS/s. The motivation for CMOS is the high levels of integration possible — the ultimate goal is to integrate the A/D converter (ADC) on the same IC as a digital signal processor in order to implement a complete analog/digital (A/D) interface. In summary, the key objective of this work is to achieve the maximum possible A/D converter throughput, with reasonable power and area, in CMOS.

This research proposes a *new architecture for high-speed A/D conversion in CMOS* consisting of a parallel time-interleaved array of pipelined ADC's: *a parallel pipeline array*. The work builds on and combines two existing concepts in A/D architectures: (a) pipelined multistage ADC's and (b) parallel time-interleaved converter arrays.

Integrated circuit realizations of pipeline architectures have received considerable attention over the last five years; specifically, CMOS implementations employing switched-capacitor (SC) techniques. A major portion of this dissertation is devoted to a detailed review and exposition of the operation of pipelined multistage ADC's at both the algorithm/dc transfer characteristic level and the implementation level. Error sources and nonidealities are discussed, with particular focus on SC approaches. Furthermore, a generalized approach to analysis of multistage A/D conversion is described, which is applicable to a large class of ADC architectures.

The use of parallelism and time-interleaving in A/D converter architectures is well known: the basic concept is to increase throughput by means of multiple parallel signal paths. A very high performance data acquisition system has been reported by HP in bipolar and GaAs technologies; Black and Hodges have described a parallel array of successive-approximation converters. However, this is the first time that the concepts of pipelining and parallelism have been combined in an ADC topology.

Performance limitations for parallel pipeline ADC's fall into two main categories: (i) problems associated with the use of parallelism and (ii) problems associated with the limited speed of the individual channel.

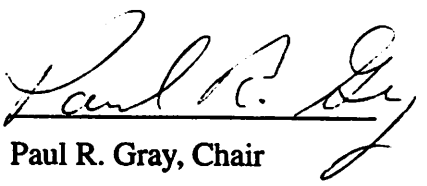
First, a major problem associated with the use of parallelism is that the hardware cost increases in direct proportion to the number of parallel paths. In order to keep power and area reasonable to facilitate implementation on a single IC, key resources such as resistor strings, bias circuitry, and clock generation circuitry are *shared* over the array. Another important issue arising from the use of parallelism is the difficulty of path matching in the analog domain. Offset, gain, and timing mismatches between the multiple channels give rise to *fixed pattern effects*, which in the frequency domain are manifested as spurious harmonics. These tones occur at multiples of the individual channel sampling frequency, in the case of offset mismatch, and as sidebands around multiples of the channel sampling frequency, in the case of gain and timing mismatches, respectively. A detailed review of

the effects of these inter-channel mismatches at the signal processing level is presented, including quantitative analysis and qualitative intuitive interpretations. At the circuit implementation level, in order to minimize the effect of mismatches, techniques are employed such as the use of a common resistor string DAC for all the channels and the use of appropriate autozeroing.

Second, within each channel, the speed of an individual SC pipeline ADC is limited by the closed-loop settling time of the sample-and-hold (S/H) and interstage residue amplifiers. To address this issue, an analytical approach to the design of any single-stage op amp for maximum speed in a given CMOS technology is described in detail. For circuit realization, the op amp topology used is a simple and intrinsically fast, mostly-NMOS, fully-differential, non-folded cascode operational transconductance amplifier.

A prototype parallel pipeline array ADC consisting of a time-interleaved combination of four SC pipelined multistage ADC's was implemented in a 1-μm CMOS technology and 8-bit resolution at a sample rate of 85 MS/s was obtained; this is the fastest 8-bit CMOS ADC reported to date. Signal-to-Noise-plus-Distortion was 41 dB for an input sinusoid of 40 MHz.

The contributions of this thesis are a survey and detailed review of multistage A/D conversion architectures; a generalized analytical approach for multistage ADC's; a systematic procedure for speed optimization of SC gain stages; a unified treatment of the effects in the signal processing domain of offset, gain, and timing mismatches in parallel channels; introduction and experimental demonstration of a new ADC topology — the parallel pipeline array; and insights into the key tradeoffs and limitations associated with the use of parallelism in A/D converter architectures.



Paul R. Gray, Chair

Acknowledgments

Being a grad student at UC Berkeley is a very special privilege — there are many people I would like to thank.

I wish to express my sincere appreciation to my research advisor Prof. Paul R. Gray. I have been very honored and privileged to have worked under his supervision. I am deeply indebted to him for his guidance and support and for providing an excellent environment and facilities for research. I have benefitted a lot from his deep intuition for circuits and systems, from his emphasis on clarity in presentations, and from being exposed to his supremely high technical and professional standards. From him I have learned the importance both of stepping back and looking at the big picture and asking "what's really happening here?" *and* of focussing on the details.

I would like to thank Prof. Robert G. Meyer for his encouragement throughout my time here, for giving many inspiring seminars, and for teaching two great graduate classes on analog IC design from which I learned a lot. I also appreciate his input and perspective on my research, in particular relating to technology choice issues in high-speed mixed-signal circuit design. I am grateful to Prof. Edward A. Lee for some very beneficial interaction on the topic of DSP analysis of mismatches in parallel channels and for encouraging me to study that topic — which became Chapter 6. Thanks also to Prof. David A. Hodges for his enthusiasm and encouragement and advice with various things. I would like to thank Profs. Gray, Lee, Meyer, and Speed for taking time out of their busy schedules to be on my Qualifying Exam Committee, and to Profs. Gray, Lee, and Speed for reviewing this dissertation.

I am very privileged to have taken classes from many excellent teachers at UCB. In particular, I would like to mention Prof. Chenming Hu and Prof. Terence P. Speed, both of whom have an extraordinary ability to communicate clearly the key underlying intuition and concepts in very complex material — Semiconductor Devices (EECS 231) and Linear Models (Statistics 230A) respectively. Both of these classes were outstanding.

Throughout my time here Greg Uehara has been a terrific colleague and a great friend. I have learned a lot from his intuition on circuits and his ways of thinking about IC's, his vision for system-focussed mixed-signal design, his attention to implementation robustness, and his emphasis on clear technical presentations — shared over countless cappuccinos and many dinners at Juan's! His generosity and encouragement over the years have been absolutely great. ALOHA!!

Much of the material in this thesis is built on some key work on CMOS pipeline A/D converters done by Steve Lewis. Our interaction continued after Steve graduated and while he was at AT&T Bell Labs and more recently as a professor at UC Davis. I appreciate his advice and encouragement and many enlightening technical exchanges over the years. For a long time Yuh-Min Lin was the A/D converter and sample-and-hold guru in 550 Cory — Yuh-Min was always extremely helpful in sharing his intuition and patiently answering my questions, and always amazed me by the extent of his knowledge and insight into issues at the device level, circuit level, and signal processing levels. I learned a lot from Yuh-Min! I have been lucky to have had Dave Cline as a colleague — Dave did the layout for the 1-μm chip described in Chapter 8, gave indispensable help with the design and testing, and I have benefitted from many discussions with him on amplifier design and pipeline ADC's. Thomas Cho has also been a terrific and inspiring colleague — I've learned a lot from many great interactions with him and I really appreciate his intuition and good humor. Often over the last few years he has graciously asked me for feedback on many of his really neat ideas — invariably my response was "wow!". Much of the material in Chapters 3, 4, and 5 concerned with issues in pipeline converters and ways of thinking about pipeline converters and the associated circuits owes a lot to interactions and discussions over the last five years with Steve, Yuh-Min, Dave, and Thomas.

I'm deeply indebted to many other students in the group over the years. Beomsup Kim gave me help and advice with lots of things — especially clock generation using tapped delay lines — and has always been very helpful and encouraging. Gani Jusuf and Weijie Yun were great colleagues and office-mate for many years. Ken Nishimura deserves special mention: as well as working on his own Ph.D., he did a superb job managing and maintaining our computer systems and software, while sharing his unending, undying enthusiasm for *FrameMaker* :-). I have had many great discussions with Ed Liu from which I always learned something. I would like to thank Joe Greco for designing the digital data acquisition system, memory board, and associated software used in the A/D testing. I have had many discussions with Robert Neff that were always enlightening. I've enjoyed working alongside and sharing, exchanging, and bouncing ideas around with Chris Rudell, Dave Helman, Keith Onodera, Tim Hu, and Todd Weigandt. More recently it's been great getting to know the new gang of students in the group — Andy Abo, George Chien, Jeff Ou, Sekhar Narayanaswami, and Srenik Mehta. Thanks also to Shoichiro Tada of Matsushita Corporation for some useful discussions.

In other research groups in EECS Department, I would like in particular to acknowledge Monte Mar for many interesting discussions, and for sharing his thoughtful and very interesting insights on many aspects of mixed-signal design, signal processing, and testing, and for the use of his data analysis programs for A/D testing. I also would like to express my appreciation to Alfred Yeung; Anantha Chandrakasan; Andy Burstein, especially for *ext2spice*; Arya Behzad; Beorn Johnson, Tom Quarles, and Wayne Christopher for providing support for various incarnations of *Nutmeg* and *Spice3*; Cindy Keys; Eric Boskin; Hong-June Park, for help with Non-Quasi Static Spice; Joey Doernberg, for advice and help with many things, including borrowing a HP pulse generator at a critical time!; Karti Mayaram, especially for his patient help with *CODECS*; Nhat Nguyen; Peter Kennedy for his perspective and insight and enthusiasm for *cool* circuits; Randy Allmon of DEC; Ranjit Gharpurey; Sam Sheng; and Tony Stratakos.

Special thanks to Andy Abo, Cindy Keys, Dave Cline, Keith Onodera, Sam Sheng, Sekhar Narayanaswami, and Thomas Cho for feedback and comments and corrections on versions of various portions of this thesis.

Thanks also to Dr. Hemant Thapar of IBM Corporation for his encouragement and support during the writing of this dissertation.

Throughout my time here I have been the beneficiary of indispensable administrative and technical support from the staff of the EECS Department and the Electronics Research Laboratory. I would like to thank Ken Lutz for all his help and advice with lab issues, acquiring parts and equipment, testing setups, surface mount issues, board layout and fabrication, and generally advice on lots of things. Ken knows a lot! and was always helpful. On other computer, equipment, and technical matters, Alex Para, Bill Baringer, Caryn Dombroski (root@diva!), the late Kirk Thege, and Sue Mellers assisted me whenever I needed their help and expertise. For help with many administrative matters, and in general for helping to make Cory Hall a pleasant place to be, I would especially like to mention Barbara Allen; Carol Block; Carol Erickson; Chere Hyek; Di Hsiang; Eleta Cook; Elise Mills — for help and advice with lots of things; Flora Oviedo; Irena Stanczyk-Ng; Jeff Wilkinson; Jean Root; Jeanene Hayes and Maryann Norris — for their efficiency and help with purchasing; Maureen Master; all the ERL payroll staff, especially Julie Aquino for making sure I got paid (!); Loretta Lutcher; Pat Hernan; Pearl Tranter; Ruth Tobey; Sheila Humphreys — for her support and encouragement; Sue DeVries; Tahani Sticpewich and Tom Boot — who always helped me whenever I needed assistance with anything from Fedexes to faxes to Frame; Tito Gatchalian; Mike Furutani and Walter in ERL Receiving; and Valyrie Brown.

There have been many friendly faces in the Graduate Student Office over the years who provided indispensable help and guidance in dealing with graduate school matters. I would like to thank Beatrice Lamotte d'Incamps — especially for some inspired words at a critical time; Carol Lynn Stewart; Geneviève Thiébaut — for her warmth and humanity — merci beaucoup!; and Heather Brown — who has been absolutely terrific and has been extremely helpful with many things.

I am lucky to have met and been in contact with many terrific people while I've been a student here. In particular, I would like to express my heartfelt gratitude and appreciation to Alberto García; Beomsup Kim and Seung Hee Choi; Bernhard Eschermann; Bob Ried — the first person I met in Cory; Carlos Kirjner; Chedseda Chinrungrueng; Chris Perleberg; Cindy Keys — for her friendship and encouragement and support, and lots of low-fat, de-caf mocha discussions about life and careers and *high* frequencies, and for keeping me out of elevators (!); Cormac Herley; Dave and Lisa Kelly; Deirdre McAuliffe Bauer; Dev Chen — for sharing his wisdom and perspective on many things; Eric Boskin; Elise Mills; Ellen Sentovich (no gaps, complete sentences!); Elizabeth Andersen — for her wonderful friendship and support; Farshid Moussavi; Finbarr Crispie — for help with many things, and Clara Porcella; Gary Mandrusiak; George Kesidis; Glen Andersen and Simone Drinkwater; Gustavo de Veciana — a great friend; Hamid Savoj; Hélène du Masle and Søren Hein; Henry Chang and Pora Park; Irena Stanczyk-Ng; Jaijeet Roychowdhury — for his bubbling optimism (!); Javier Alonso; Jim Morehead and Steve Creagh — for being great buddies of mine since I came to Berkeley and for many great times together; Joe Greco and Barbara Stieglitz; Joe Harrington — cara iontach ar fad: *"Corcaigh Abú!"*; John and Susan Ryan; Kamal Bharucha; Kamran Nemati; Karti Mayaram; Kathleen and Edward Kovach and all their family for their friendship and hospitality and many wonderful times; Kaveh Niazi; Keshab Parhi — especially for his encouragement during my first year here; Marina Pantazidou; Mary-Lynn Lidstone — for her wonderful friendship and support and great sense of humor; Maureen Master — for her terrific wit, and who made sure I never took myself too seriously; Mike Riemer; Nhat Nguyen — whose friendship and encouragement have been great; Rajeev Murgai — for inspiring me and for some interesting discussions about music and meaning *(meaning?!)*; Renu Mehra; Rob Pfile — who now really is a senior grad student; Robert Hass — for inspiring me; Roberto Suarez and Angela; Sarah Banks — for her wonderful friendship, for lots of (guilt-free!) great times, and for being my conscience (!) — "C'était sur le lac de Côme. Une après-midi d'avril. Tu peux le croire?" — and Peter Larsen; Serge Hustin et

Marguerite Deschamps; Sherry Lee; Shuvra Bhattacharyya and Milu; Sovarang "Armani" (!) Leang; Takis Konstantopoulos — for his friendship and companionship and many great discussions about Yeats, Machado, Tarrega, Cafafy, Borges, ... at the Med' and Au Coquelet — "Caminante, no hay camino, / se hace camino al andar"; Thomas Cho and Ok Kyung Suk; Tony Miranda; Tony Sarkis; Valerie Taylor; Valyrie Brown; Weijie Yun and Haifang Guo; Yuh-Min Lin and Julie Hsu; Yvonne O'Connor; and Zeina Daoud.

I would like to thank the EECS grad students from Ireland — occasionally known as *eireheads@diva* (!) — with whom I overlapped while I've been at Cal: Joe Higgins, Liam Murphy, Orla Feely, and Peter Kennedy. I benefitted a lot from their individual and collective help and advice regarding grad school and academic matters and many other things, and I really appreciate their constant support and encouragement and friendship. Thanks also to Claudia Thio and Joe Higgins for their hospitality over the years and for feeding me many times!

I would especially like to acknowledge a great group of friends, the "Russell Street gang", especially Cammie Edwards, Soo-in Hwang, Kathy van Leuwen, Bob Young, Lisa Zuckerman, Miguel Helft, Rick Apgar, Stephanie Blackburn, and Tracy Friedman — the best housemates in the world! Thanks for everything you guys — the bike rides, the Nabolom cakes, the Yogurt Park breaks, the chats and joking around in the kitchen, the wonderful weddings!, for listening and cheering me up many times, and lots, lots more! Plum Heaven lives on!!

At this time, I remember a great friend and housemate and a wonderful human being, Joe Wiest, with whom I shared many great times and who encouraged me so much.

Finally, the people who made it all possible.

First, my parents Séamus and Veronica Conroy for all their support, encouragement, and love throughout my life. But for them, I would never have had the privilege of this experience. Thank you also to all my family especially my sisters for their love.

And, some absolutely key people and very special friends. Nothing that I write here can come close to what I want to say but I'll try: Mireille Broucke, for being a wonderful friend for the last six years and for changing my life in so many positive ways; Anne Kovach, for her special friendship and deep humanity and wisdom about life and lots more; Peter Kennedy, for always being there when I needed propping up and for being the best friend that anyone could ever have; Kathy van Leuwen, for everything, and especially for always listening when life got complicated; and Tracy Friedman, for her encouragement and belief in me.

RESEARCH FUNDING ACKNOWLEDGMENTS

This research was sponsored by the National Science Foundation, the California Micro Program and Texas Instruments. Chip fabrication was by Texas Instruments. I would like to express my deep gratitude to the Circuits Research Group at SPDC, Texas Instruments, Dallas, for all their invaluable help; in particular, Dr. Richard Hester (now with the Department of Electrical Engineering and Computer Engineering, Iowa State University), Dr. Khen-Sang Tan (now with the Institute of Microelectronics, Singapore), and Dr. Michiel de Wit. I also would like to thank AT&T Bell Labs for providing the printed circuit board used in the digital data acquisition system and to Steve Lewis for all his help. Thanks also to Sean X. Lee of FIB Lab, Inc. for microsurgery on some prototype IC's.

Table of Contents

Chapter 1 Introduction

1.0 OVERVIEW

Within electronic signal processing, analog-to-digital (A/D) conversion is a key function required to interface between "real-world" analog signals and digital computers. An analog-to-digital converter (ADC) performs three conceptually distinct operations. First, it *samples* a continuous-valued, continuous-time input signal to give a discrete-time signal. Second, it *amplitude-quantizes* that signal to one of a number of (usually equally-spaced) levels giving a discrete-valued signal. Third, it *assigns a code* or label to that level — usually a binary digital word.

There are many ways to categorize ADC's. One is in terms of *key attributes* such as

- resolution — number of bits,
- speed — number of samples per second,
- latency — number of clock cycles per conversion,
- input range — allowable range (in volts) of the analog input signal.

Note that in this dissertation the terms ADC *speed, sampling rate, sampling frequency,* and *throughput* are used interchangeably.

An alternative viewpoint, which could be said to constitute a *user-* or *system-view,* is to think in terms of *performance* and *cost.* In general, *performance* relates to functional characteristics such as speed, input bandwidth, number of bits of resolution, dynamic range, etc. *Cost* usually relates to quantities such as power dissipation and chip area. Circuit fabrication technology, whether CMOS, BiCMOS, bipolar, GaAs, or hybrid, is also a key contributor to cost, and throughout this dissertation technology choice issues are addressed whenever relevant. The *user* of an A/D converter may be a board-level system designer using a standalone ADC, or a VLSI system designer using an ADC block inside a more complex chip. The main focus of the user is with the system context and application, and the primary concern is, therefore, does the A/D converter meet the system

requirements and at what cost? Other important considerations include: how hard is it to drive the ADC? how many precision external components are required? how difficult is it to test? Depending on the targeted market, other factors such as implementation complexity and its implications for design time and product time-to-market may merit attention also.

A third view is that of the A/D designer who categorizes ADC's according to circuit architecture within a taxonomy of topologies — e.g., flash, two-step flash, pipeline, successive approximation, Δ–Σ, etc. This viewpoint motivates questions such as: which architecture gives the best throughput per unit power or area for a given technology? or which topology gives the lowest power, for a given speed requirement, assuming that both CMOS and BiCMOS technologies of specific device performance characteristics are available?

Finally, a related hybrid view is that of the system-focused, mixed-signal IC designer who — as well as asking the question does the ADC meet the requirements with minimum cost — might also ask: are there modifications possible in other parts of the system, either at the chip level or signal processing level, which might lessen some stringent A/D requirements? or what is the most appropriate ADC implementation, given the nature of a particular system function? Obviously, in this dissertation only some of these topics are addressed. However, the purpose of these introductory paragraphs is to acknowledge the existence of a broad spectrum of issues and considerations that impact the design of monolithic ADC's.

Clearly, in the parameter space of speed–resolution–power–technology, many different converter algorithms, architectures, and topologies exist. The particular focus of this research is a resolution in the range 7–10 bits, speed of 50–100 megasamples per second (abbreviated henceforth to MS/s), and implementation on a single integrated circuit (IC) using CMOS technology. There is an ever-increasing need for monolithic

ADC's with these resolution and sample rate specifications for such applications as waveform acquisition, instrumentation, digital video and High-Definition TV (HDTV) [5], [139], [142], [149], radar, robotics, industrial and medical imaging, high-speed digital communications [116], [117], [119], and data detection in the magnetic recording channel [118], [19]. There is also a trend in many communications and signal processing systems for the analog/digital interface to move closer to the signal source [102], necessitating higher ADC sample rates. Integrated circuit solutions with the above speed and resolution capabilities already exist in bipolar technology — for example the folding and interpolation architecture has been shown to be an effective approach [107], [150]–[154]. However, in previous CMOS implementations, the speed has been limited to about 50 MS/s [52], [87]. A key motivation for CMOS is that higher levels of integration and lower power are possible than in bipolar implementations, and ultimately it would be desirable to integrate the analog/digital interface function on the same IC as a digital signal processor.

In summary, the motivation for this work is to maximize the ADC sample rate and achieve a resolution of about 8 bits, with reasonable power and area, in CMOS. In order to achieve these objectives, this research builds on two existing A/D converter architecture concepts: pipelining and parallelism.

First, *pipelined multistage* ADC architectures have received considerable attention over the last five years [71]–[76], [79]–[80], [110]–[111], [161]–[162]. They operate by having a number of cascaded stages, each of which performs a low-resolution A/D conversion, a low-resolution D/A conversion, and a subtraction to give an analog remainder or *residue*, which is amplified and passed to the next stage. These stages can be configured to operate in a pipelined manner, i.e., the first stage operates on the most recent input sample, and concurrently the second stage operates on the gained-up residue from the previous sample, etc.

Second, *time interleaving* is a well-known idea for obtaining high throughput A/D conversion by means of parallelism. If M ADC's are operated in a parallel time-interleaved fashion, overall system sampling rate can potentially increase by the same factor M. The concept of time-interleaving has been applied previously in very high speed applications for instrumentation [109], [114], [141], and in an IC consisting of an array of successive-approximation ADC's [8].

This research proposes a new architecture for high-speed A/D conversion in CMOS consisting of a parallel time-interleaved array of pipelined multistage ADC's — a *parallel pipeline array.* This is the first time that the concepts of pipelining and parallelism have been combined in an A/D converter topology [21], [22].

Performance limitations for parallel pipeline ADC's fall into two categories: (a) problems associated with the parallelism and (b) the throughput of the individual pipelined channel, as shown schematically in Fig. 1.0. These issues are now addressed.

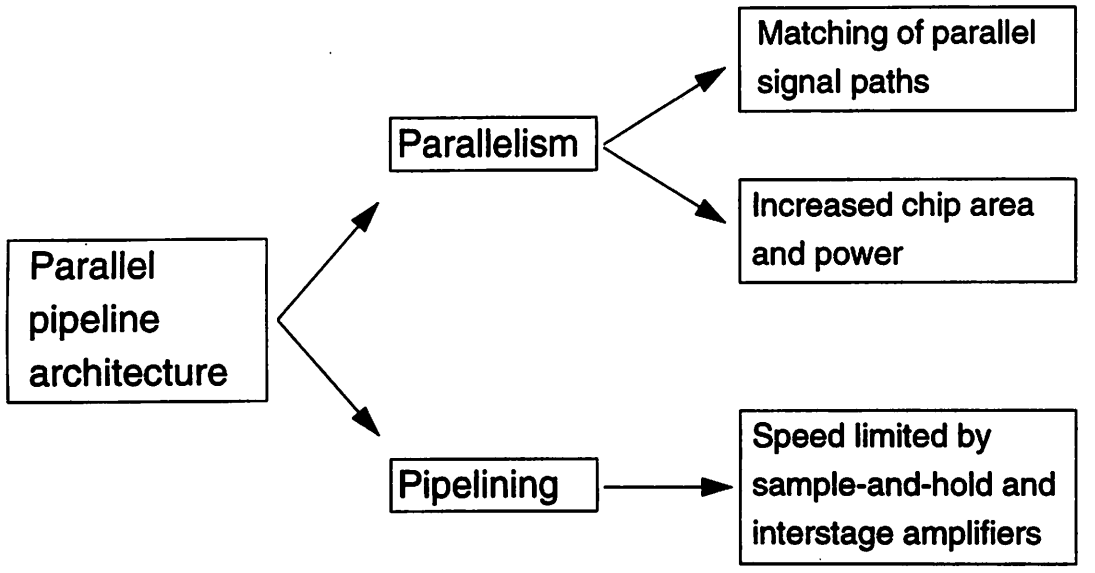


Fig. 1.0 Summary of key problems in the design of parallel pipeline ADC's.

Clearly, converter throughput is improved by the use of more parallelism in the signal path, but the use of parallel signal paths in ADC's has received only limited application for

two main reasons: (i) the difficulties of path matching and (ii) the dramatically increased area and power required for the implementation.

First, the use of parallelism in the analog domain introduces problems due to mismatches. Offset, gain, and timing mismatches between the multiple parallel signal paths give rise to *fixed pattern* effects, which in the frequency domain are manifested as spurious harmonics [8], [53], [105], [109]. Consider a time-interleaved A/D system consisting of M ADC's in parallel, with overall sampling rate F_s, and therefore each individual path or *channel* sampling at rate F_s/M. The presence of offset mismatches results in undesired frequency components in the output spectrum at *multiples of the individual channel sampling frequency* F_s/M. In the case of gain and timing mismatches, the undesired components occur as *sidebands centered around multiples of* F_s/M. In this work, these effects are minimized by appropriate offset correction and gain matching techniques. (Note that the precise effects of gain and timing mismatches are somewhat different — this is elucidated later.)

Second, a key problem associated with parallel time-interleaved A/D converters is hardware cost. To date, the use of such architectures in monolithic integrated circuit ADC's has been limited due to the greatly increased silicon area required. However, if hardware resources are *shared* across the multiple channels, then it is possible to keep area and power reasonable. In this way hardware cost does not increase in direct proportion to the number of channels.

The new parallel pipeline architecture described in this dissertation alleviates many of the above difficulties by employing a number of approaches and circuit techniques, described in more detail in the following chapters. These include sharing of circuitry such as bias circuits and resistor strings over all channels; minimization of fixed pattern noise effects due to inter-channel mismatches by appropriate autozeroing, and by the use of a common resistor string DAC for all the channels; use of digital error correction to ease

comparator accuracy requirements; and an on-chip delay-locked loop (shared over the multiple channels) for generation of low-jitter, multiphase sampling clocks.

Another major performance limitation relates to the individual channel throughput: within each channel, given that the implementation employs a CMOS, switched-capacitor approach, the operating speed of each pipeline ADC is limited by the speed of the input sample-and-hold (S/H) circuit and the interstage residue amplification circuits [71], [73]. The speed of these blocks is determined by the closed loop settling time of an operational amplifier (op amp): this therefore constitutes the key circuit design challenge for pipeline ADC's. This issue is addressed (a) by choosing a simple, and hence intrinsically fast, op amp topology — a fully-differential, mostly NMOS, non-folded cascode circuit and (b) by employing analytical techniques, supported by extensive simulation, to choose the S/H and interstage gain amplifier parameters in order to give the optimum speed possible within the available CMOS technology.

A prototype 8-bit 4-channel parallel pipeline ADC has been designed and fabricated in a 1-μm VLSI CMOS technology [21], [22]. Measured results have demonstrated that 8-bit resolution at a sampling rate of 85 MS/s has been achieved making this the fastest 8-bit CMOS ADC reported to date. Signal-to-Noise-plus-Distortion Ratio (SNDR) was 41 dB for an input sinusoid of 40 MHz.

1.1 STRUCTURE OF DISSERTATION

The outline of the thesis is as follows.

Chapter 2 contains a broad overview of the applications of high-speed ADC's and gives the research context for this work. First, some important and interesting applications are described, and then some general issues and trends in mixed-signal IC system design are considered. Finally, the performance characteristics of some recently-reported monolithic, high-speed, medium-resolution ADC's are presented and discussed.

In Chapter 3, pipelined multistage ADC architectures are reviewed. The chapter includes an exposition of converter algorithm and coding issues, dc characteristics, digital error correction techniques, and error sources. Implementations and the associated nonidealities are also described, with primary focus on switched-capacitor approaches.

Chapter 4 builds on the material presented in Chapter 3 and proposes a generalized, systematic approach for analyzing multistage A/D systems. This allows many issues such as digital correction and non-binary radix architectures to be considered within a common unified framework. There is also some discussion of calibration techniques.

Before describing the parallel pipeline architecture in detail, two important relevant topics are discussed independently in the next two chapters: (a) settling time optimization of single-stage MOS op amps and (b) general system issues relating to mismatches in parallel analog channels.

Chapter 5 presents an analysis and optimization of settling time in switched-capacitor gain stages used in pipeline ADC's, and addresses the problem of how to design such a stage in a technology-limited way. The focus is primarily a single-pole model, with some allusion to second-order system analysis.

In Chapter 6, the issue of mismatches between multiple parallel analog signal paths is considered, and some analytical approaches are presented to examine the effects of gain, offset, and timing mismatches, both qualitatively and quantitatively.

Chapter 7 reviews previous uses of parallelism in A/D converters, and then describes the key circuit and system implementation aspects of the parallel pipeline architecture, in particular the combination of pipelining and time-interleaving, and its implications. Performance limitations and the proposed solutions are discussed. Some issues concerning generation of multiphase sampling clocks are addressed also.

In Chapter 8, experimental results from a prototype chip are presented, and finally Chapter 9 contains a summary of the work and conclusions.

Chapter 2 Monolithic, High-Speed, Medium-Resolution ADC's — Review of Applications and Recent Trends

2.0 INTRODUCTION

This chapter presents a general applications background and context for research on monolithic, high-speed, medium-resolution A/D converters. First, some applications of high-speed ADC's are described in Section 2.1. This section is not intended to be a comprehensive review: rather, it constitutes a sampling of important applications, and provides some background. Section 2.2 discusses some recent trends in mixed-signal design which impact A/D converters. Finally, in Section 2.3, recent developments in monolithic ADC's are described.

2.1 APPLICATIONS

2.1.0 Overview

As stated in Chapter 1, A/D conversion is a ubiquitous function required in a myriad of applications. In this section, the focus is on describing applications of medium-resolution, high-speed A/D converters. The word "medium" is interpreted here as a resolution of 6–10 bits, and the term "high-speed ADC" is defined (somewhat arbitrarily) to mean operating at a sample rate of 1 MS/s or higher. The applications for ADC's with these characteristics are classified into three main groups: (i) high-speed data acquisition, (ii) digital video, TV, imaging, and (iii) in a broad sense, "digital communications".

2.1.1 High-Speed Data Acquisition

High-speed data acquisition is the traditional application for fast A/D converters. Instrumentation applications such as waveform digitizers and digital sampling

oscilloscopes require arbitrarily high sample rates to observe fast single-shot events. This is the domain of relatively expensive technologies as exemplified by GaAs [109], high-speed bipolar [63], [64], [90], [109], [114], [141], and hybrid implementations [109], [114], [141]. Cost is usually not the dominant consideration. Because of the requirement for highest possible speed, this area has traditionally been where parallel, time-interleaved A/D converter architectures have been employed. The current state-of-the-art is an 8-bit, 8-GS/s data acquisition system [93], which is composed of four interleaved 2-GS/s hybrid microcircuits. Each hybrid contains a sampler chip, passive filters, and two dual 500-MS/s 8-bit bipolar ADC chips (i.e., two ADC's on each of the two chips). Thus, the entire 8-GS/s system comprises sixteen interleaved A/D converters. Some further examples and discussion of parallel time-interleaved ADC topologies are given in Section 7.1.1.

2.1.2 Digital Video, TV, Imaging

This section first describes applications specific to video and TV. Some broader issues are then considered, and finally other imaging applications are briefly summarized.

Digital Video and TV: There are three ways to represent color information [100]; these are now reviewed in the context of conventional TV.

1. The first method entails using a full-bandwidth signal for each of the three *primary color components* — Red, Green, and Blue (RGB). An example of this usage is a color TV camera, which contains a sensor that generates three analog voltages related to the RGB intensities. In PC's and workstations, and computer graphics environments, the primary colors are represented digitally in the screen memory, and are converted to analog levels to drive the color CRT screen using video DAC's, RAMDAC's, etc. In the digital domain, the RGB signals are usually represented by 8-bit digital words, hence the term "24-bit color".

2. Often video processing takes advantage of some properties of the human visual system in order to avoid transmitting full-bandwidth RGB signals. Specifically, the eye is more sensitive to luminance (brightness) information than to chrominance (color) information,

and so linear combinations of the RGB components are used, chosen in such a way that the full bandwidth is needed only on the luminance signal, and less bandwidth is needed for the chrominance signals. There are some historical reasons for this also; namely, to maintain compatibility with monochrome systems. One such mode of color representation employs three signals known as *color difference components*, which are linear combinations of the RGB components, and are referred to as the luminance (Y) and the color difference signals (R–Y) and (B–Y) — usually denoted by Pr and Pb, respectively. For digital video processing, the widely-used CCIR 601 standard specifies that the luminance is sampled at 13.5 MS/s and the two color difference signals at 6.75 MS/s, all at 8-bit resolution. These ratios of sampling frequencies are referred to as "4 : 2 : 2". The sample rates correspond to a 525-line, 60-field/s system with 858 samples (720 active) per total line for luminance. This format is used in studio processing and non-broadcast digital video applications. Note that the sample rates of 13.5 MS/s and 6.75 MS/s were chosen to be compatible with *both* NTSC and PAL composite formats discussed next.

3. For transmission and broadcasting, a *composite* color format such as NTSC or PAL is used. Composite color uses a single channel with the luminance component occupying the full bandwidth and two lower-bandwidth color difference signals being transmitted in the same frequency spectrum using Quadrature Amplitude Modulation (QAM) about a color subcarrier frequency F_{sc}. The standard NTSC video signal has a bandwidth of about 4.2 MHz [11] and F_{sc} is 3.57954525 MHz. It is usually sampled at a rate equal to four times the color subcarrier frequency — i.e., 14.3 MS/s. For the PAL standard used in Europe, the luminance bandwidth is about 5.5 MHz, F_{sc} is 4.43361875 MHz, and $4F_{sc}$ gives a sample rate of 17.7 MS/s. Note that there is a significant system advantage to sampling at $4F_{sc}$ since it greatly simplifies the demodulation process in discrete time by only requiring the values $\sin(n\pi/2)$, i.e., $-1, 0, +1, 0, -1, 0, \ldots$, etc. [100]. This is a general property of QAM schemes, and arises also in Section 2.1.3 in the discussion of digital radio receivers.

High Definition TV: HDTV has a greatly increased spatial resolution as compared to conventional TV, and requires a correspondingly higher electrical signal bandwidth. Specifically, it employs (a) approximately twice the horizontal resolution, (b) twice the vertical resolution, and (c) a wider aspect ratio — 16:9 as compared to 4:3 — resembling a movie theater format. Charge-Coupled Device (CCD) sensors for use in HDTV cameras have been reported with 70 dB dynamic range [1], [2]. HDTV quality is approximately equivalent to 35-mm motion picture film as projected in a typical theater [115], and it is likely that HDTV as a medium will be used for film production and also for high-end applications such as publishing. The luminance bandwidth is approximately 30 MHz, and the proposed standards suggest sampling at 74.25 MS/s. A professional HDTV digital video tape recorder (VTR) for studio use has been described in which a luminance signal is sampled at 74.25 MS/s, and two color-difference signals are sampled at 37.125 MS/s, all at 8-bit resolution, giving a total data rate of 1.2 Gb/s [149]. At present, the most likely scenario is that an all-digital HDTV standard will be adopted in the U.S. [160]. Further details on system specifications and signal processing issues may be found in [5], [139]. Some issues related to the use of 10 bits rather than 8 to take better advantage of the dynamic range of HDTV CCD imagers are discussed in [142].

Some Broader Issues Related to Digital Video and TV Signal Processing: As a medium, video has a number of characteristics that have a direct impact on ADC architecture choice. First, in almost all applications, there are no particular latency constraints on any signal processing operation. Second, the signal bandwidth of video, which depends on the horizontal and vertical resolution, the frame rate, etc., is defined by standards such as NTSC. Accordingly, the dynamic performance of ADC's at input frequencies higher than the well-defined peak bandwidth is usually not important. Furthermore, in applications such as camcorders, in which the signal to be digitized is the output of a CCD, which is a sampled-data analog signal and already in the discrete-time domain, the input bandwidth requirements on the ADC are greatly reduced [65], [84].

Higher Integration Video Signal Processing IC's: In the consumer desktop video arena, there is a constantly growing demand, driven by cost requirements, for IC's with greater levels of functionality and integration. Single-chip multimedia audio codecs (Coder/Decoders) are now very common, and undoubtedly video codecs will become increasingly widely used in multimedia systems. Furthermore, following the emergence in digital audio of ADC's and DAC's with greatly enhanced on-chip intelligence and DSP functionality, such as on-chip compression/decompression, it seems reasonable that a similar trend will occur in video applications such as video conferencing and digital video on personal computers, perhaps over the next 2–3 years, and initially in low-end, cost-driven or low-power applications. Some examples of this trend — albeit with relatively modest DSP capability — are a video signal processing IC incorporating three 27-MS/s 8-bit ADC's [40] and a digital NTSC decoder [103].

Other Digital Imaging Applications: There is a wide variety of imaging applications with relatively little standardization or regulation, ranging from digitization of 35-mm slides — e.g., as used in the Kodak Photo CD system — to high-quality professional image processing applications using HDTV [6]. In real-time applications in robotics and manufacturing that require functions such as machine-vision or pattern-recognition, images are first captured, and then undergo various DSP-intensive operations — for example, 2-D filtering — that finally result in object characterization or recognition [113]. Typical requirements are 8-bit resolution and speeds of 10–100 MS/s. Medical imaging encompasses various types of diagnostic applications with a range of speed and dynamic range requirements [12], [27]. The signal produced in ultrasound is an acoustic waveform at a typical frequency of around 10 MHz, and requires a sample rate of 20–30 MS/s with a resolution of 10–12 bits [102]. Other modalities such as Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) require much higher resolutions such as 14–16 bits; another important application is digitization of plain film (X-rays).

2.1.3 Digital Communication

The third application area of A/D converters could broadly be described as *digital communication.* A simplified but reasonably general block diagram view of the physical layer of a digital communications or telecommunications system, appropriate for both baseband and passband modulation schemes, is shown in Fig. 2.0.

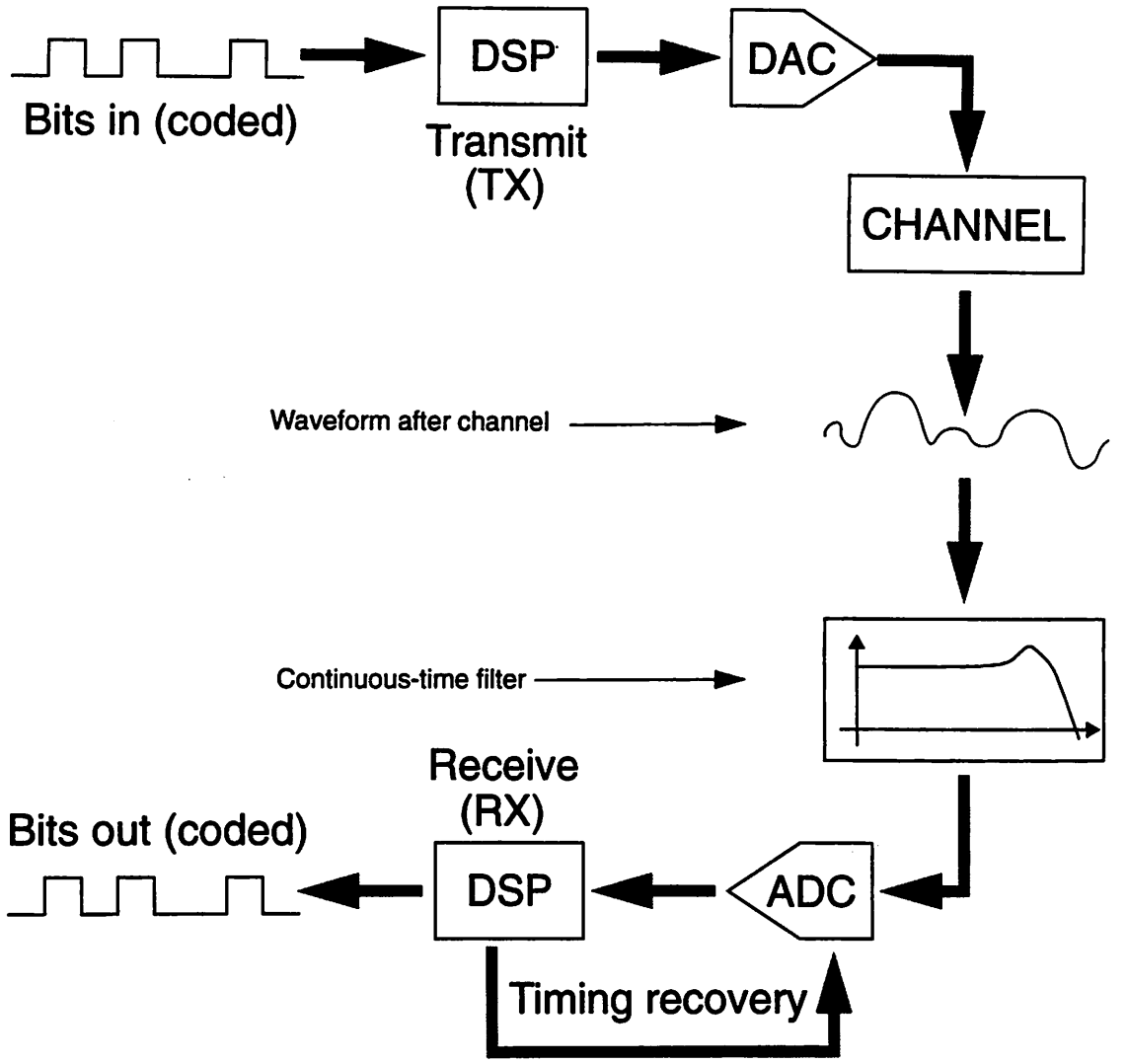


Fig. 2.0 Simplified system-level view of signal processing required in a typical digital data communications or telecommunications application.

The figure indicates the bits entering at the transmit (TX) end, and passing though some digital processing and a D/A converter, before being sent through the "channel". A

representation of the signal waveform corrupted by noise, distortion, intersymbol interference (ISI), etc., is shown at the receive (RX) end. Typically, this received signal goes through a continuous-time analog amplification/filtering/equalization block before being digitized by the ADC. A timing-recovery loop is also indicated. This general configuration is pertinent to a wide variety of applications. In particular, some examples of (i) wireless and (ii) non-wireless communications systems are now described briefly.

Digital Radio and Wireless Personal Communications: Recently there has been explosive growth in the areas of cordless and cellular phones, wireless LAN's, and personal communication services all based on digital communication techniques [32]. There are also a number of more traditional applications in digital radio such as long distance digital microwave radio systems [96]. Another application is Global Positioning System (GPS): examples of receiver architectures are described in [42], [98].

Passband Pulse Amplitude Modulation (PAM) schemes such as Quadrature Amplitude Modulation with M possible complex-valued symbols (M-QAM) are widely used to maximize spectral efficiency on the radio channel [68]. In the conventional, widely-used heterodyne system configuration, the receiver performs analog demodulation (in one or two steps) of the Inphase (I) and Quadrature (Q) components of an Intermediate Frequency (IF) signal before digitization. The IF center frequency F_{IF} is typically in the 10's or 100's of MHz. However, *all-digital* architectures have been proposed [116], [117] in which the IF signal is directly sampled at the rate $F_s = 4F_{IF}$; this simplifies the complex demodulation for the exact same reasons mentioned earlier for sampling at $4F_{sc}$ in TV signal processing. Of course this requires an A/D converter capable of operating at that sample frequency. In terms of resolution requirements, studies have indicated that satisfactory SNR performance is possible for 256-QAM with 8-bit quantization [116].

Another illustration of the trends in wireless communications systems is a novel transceiver architecture incorporating the approach of passband sampling, or *sampling*

demodulation [120]. This has been proposed for a receiver front-end in a spread spectrum based system for a portable multimedia terminal.

Commonly-used, standard IF frequencies such as 70 MHz, 85 MHz, 140 MHz, etc., which were previously the exclusive domain of Surface Acoustic Wave (SAW) filters are now coming into reach of CMOS and BiCMOS analog front-end IC's — note for example the operating frequency of some of the disk drive read channel chips — and this seems likely to lead to higher levels of integration in mixed-signal front-end IC's [119]. Simultaneously, the upper limit of "baseband" will tend to increase, driven by the demand for higher data-rate wireless communication and the trend towards higher bit-rate digital radio modems. Clearly, the combination of (a) increasing data-rate requirements, (b) limited spectral resources, and (c) greatly increased usage of wireless communications services will continue to drive research and development in high data-rate, high spectral-efficiency, low-power, high-integration transceivers using mixed-signal chips.

Finally, note that performing more signal processing in the digital domain allows algorithms of far greater complexity to be implemented than with purely analog signal processing, and leads to the capability of accommodating different data rates, modulation formats such as M-QAM and trellis-coded modulation, different filter specifications, etc. This is particularly important for communication channels where spectral efficiency and capacity is important. In the cellular arena, code division multiple access (CDMA) has received a lot of attention: issues regarding synchronization are addressed in [132].

High-Speed Data Communications: There is a plethora of high-speed *wired*, or *wireline*, (as opposed to *wireless*) data communications and networking applications using two principal media: (i) copper and (ii) coaxial cable.

The category "copper" includes various grades of shielded and unshielded twisted pair. Current 10-Mb/s Ethernet will soon be superseded by significantly higher data-rate LAN's driven by ever-increasing bandwidth requirements in, for example, networked

multimedia systems and networked workstation environments. Schemes such as 100 Mb/s Ethernet and CDDI (Copper Distributed Data Interface) are emerging, some of which will use complex mixed-signal front-ends. Asymmetric Digital Subscriber Line (ADSL) has been proposed to utilize the existing telephone subscriber loop infrastructure to provide a unidirectional "video-to-the-home" service [15]. At the receive end, this will require an ADC sampling on the order of a few MS/s with a resolution of 12–13 bits.

There is a vast existing cable TV infrastructure employing coaxial cable, which has a bandwidth on the order of 500 MHz. The next-generation cable TV systems will feature greatly enhanced multimedia-type services, possibly including interactive digital TV and "video-on-demand", and will have many applications in medicine, business, education, and entertainment, all of which may be seen as part of the evolution towards a national "information super-highway" [36]. Such systems will require an intelligent "digital set-top box" [159] with vastly increased functionality over the present-generation equivalent. This unit will receive an analog signal from the cable and perform some type of high-speed demodulation function, such as 64-QAM; this will involve sampling at a resolution of 8–10 bits and at a rate of at least 10 MS/s, depending on the compression scheme employed, the kind of data, etc. The resultant digital data will then be decompressed to give NTSC. With the likely introduction of an all-digital HDTV standard incorporating compression and a *digital* broadcast and transmission format such as M-QAM, future systems will undoubtedly be based on HDTV, and demand a correspondingly higher sample rate.

Although many of the above-mentioned systems are not yet widely installed, it seems inevitable that the area of high-speed digital communication over coaxial cable will provide fertile ground for innovation in high-integration, mixed-signal front-ends incorporating some combination of a continuous-time filter, an A/D converter, and a digital signal processing block, as depicted generically in Fig. 2.0.

Disk Drive Read Channel: Recently, the application of partial response signaling with maximum likelihood detection (PRML) to the magnetic recording channel has received considerable attention [19], [118]. This DSP-based approach requires an A/D converter typically with 6-bit resolution sampling at the channel data rate. Currently, the fastest reported commercial implementation [108] is a 65-MHz IC implemented in a 6-GHz npn, 1-GHz pnp, BiCMOS process [123], and incorporating a 6-bit flash ADC using 64 comparators. It is clear that with magnetic storage densities continuing to increase [51], there will be a trend towards higher data rates [16], [147] necessitating faster embedded ADC's running at speeds significantly in excess of 100 MS/s. Keeping power consumption reasonable will be a key issue.

2.2 SOME ISSUES AND TRENDS IN MIXED-SIGNAL IC DESIGN

This section discusses some general trends in mixed-signal IC's, all of which have an impact on A/D converter design. In general there is a common thread linking these issues: each is associated with *system partitioning or system integration and the associated cost/performance tradeoffs.*

Location of the Analog/Digital Interface in Signal Processing Systems: There is a continuing trend in complex signal processing systems for the "A/D interface to move closer to the signal source" implying that instead of going through an entirely continuous-time analog front-end, consisting of linear and nonlinear analog processing that finally produces at its output a digital 1 or 0, the received signal undergoes some analog processing and is then digitized and passed to a DSP. Some instances of this trend are oversampling analog/digital front ends, direct (digital) demodulation, in which the ADC is directly digitizing IF signals with demodulation being performed in the digital domain, and baud-rate sampling for digital communications applications [102]. An important issue is the flexibility, programmability, and ease of adaptability afforded by performing signal processing in the digital domain as compared to in the analog domain.

Ever-Increasing Levels of Integration: There is a constant thrust towards higher levels of integration, driven by the desire for the lower system cost that results from greater functionality on a single chip. Fewer chips leads to smaller boards and physically smaller end products; in particular, this is true for high-volume applications. This trend is resulting in highly-complex mixed-signal chips with ADC's *embedded* as one block. CMOS and BiCMOS have a clear advantage in this context because of their higher levels of integration, and will probably be the technologies of choice for large mixed analog-digital chips, for example, in future wireless communications systems.

"Intelligent" A/D Converters: Another trend, which could be viewed as a specific case of the more general trend discussed in the previous paragraph, is the emergence over the last five years of "intelligent" monolithic ADC's — sometimes referred to as *intelligent interface components*. These chips have a significant amount of on-board DSP capability, control functionality, and programmability, incorporate on-chip calibration of analog components, and are typically implemented in CMOS. Hitherto, they have been targeted for use in data acquisition applications at around 12-bit resolution and with sample rates of a few hundred kS/s [25], [97], [101].

Design Complexity and Cost: Another important consideration is that although a mixed-signal IC solution may be more complex and require a longer design time than a "brute-force" digital one consisting of a standalone A/D converter and a DSP chip, the resulting (analog) implementation may be cheaper — less silicon area and/or less power — when manufactured in quantities of millions. This is especially true for low- or medium-resolution applications (6–8 bits). In these cases, matching of passive analog components is often good enough. Of course, the same market forces in the PC and consumer arena that give rise to high volumes also tend to require rapid time-to-market! A key system-level issue that must be resolved is where to do the system partitioning between analog and digital, taking into account the associated costs.

Noise/Coupling Between Analog and Digital: It is clear that the problems associated with interference and coupling between the analog and digital portions of a complex mixed-signal IC, which arise from (a) capacitive coupling and (b) magnetic coupling due to bonding wire and package inductance, merit careful attention at all levels of the design. For example, at the device and process technology level [131], one issue is the widespread use of epitaxial-based CMOS technologies using a low-resistivity p+ substrate, which tends to exacerbate the noise coupling problem, as compared to the higher-resistivity substrates used previously. Careful modeling of the entire electrical environment including the package has been described in the context of a mixed-signal IC for a PRML disk drive read channel [118]. Strategies at the system timing level for noise management in Δ–Σ A/D converters are described in [137]. Another trend is the increasing use of on-chip decoupling capacitors. However, coupling considerations may limit the level of integration in RF applications, and may motivate more widespread adoption of other interconnect technologies such as Multi-Chip Modules (MCM's).

System-Based Latency Constraints: For both embedded applications inside a large mixed-signal chip and lower integration solutions, latency of the A/D converter becomes an issue whenever a timing-recovery loop is present. For this reason, long pipelines with many stages may not be acceptable. In the PRML disk drive read channel implementations reported by IBM [118], [108], the ADC topology employed is a 6-bit flash using 64 comparators — a minimum-latency approach. In other applications such as video processing, latency is not a constraint. More broadly, this distinction could be used as a means to differentiate the various applications classes.

Low Voltage and Low Power: For portable systems, personal digital assistants, etc., low power consumption is a critical issue, and will impact the analog circuit design. Initially, dual-supply systems may be employed, but the trend towards operation at reduced supplies will dictate that issues such as (i) design of low-voltage op amps and (ii) the problem of turning on MOS switches in low-voltage electrical environments be addressed.

Necessity for a Holistic System View: From the above, it is clear that there is a large set of applications where detailed attention to the interrelationship between system issues and circuit issues is required [147]. An example of this approach involves examination of the signal processing function being performed and the characteristics of the A/D interface in a particular application, and determination of the appropriate ADC topology *from these system considerations* [147]. The key point is that the resulting solutions implemented in silicon are better if the "big picture" is considered from the outset. In summary, this *holistic* system view takes account of the symbiotic relationship between the circuit-level issues and the signal processing and algorithm issues.

2.3 RECENT TRENDS IN HIGH-SPEED MONOLITHIC ADC'S

Having reviewed some applications and system issues in the previous section, the focus of this section is on issues and trends in monolithic A/D converters, including architecture considerations and circuit design details. However, as mentioned above, it should be borne in mind throughout this discussion that even though an ADC may be viewed as a black box with particular specifications, in complex analog/digital interfaces the suitability of a particular A/D topology and the relative importance of certain performance characteristics *depends on the system context.* This is true for standalone ADC's, and is especially relevant for an ADC embedded within a large mixed-signal chip.

In order to provide some feeling for recent trends in high-speed, monolithic ADC's, some tables and graphs are given in this section comparing the performance of IC's reported in the literature approximately over the last five years, on the basis of speed, power, architecture, etc. Initially the data is categorized by technology; this is followed by a comparison based on sample rate normalized to technology f_T.

2.3.0 CMOS ADC's

In Fig. 2.1, resolution is plotted versus sample rate for some high-speed CMOS ADC's.

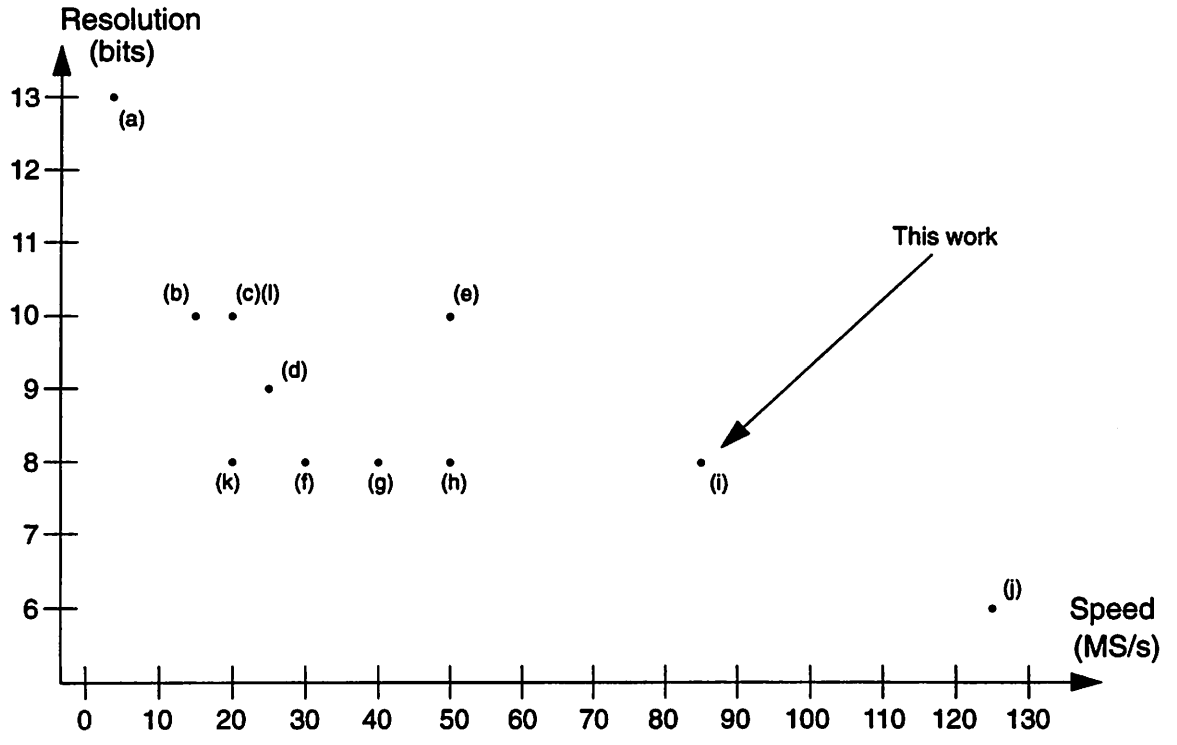


Fig. 2.1 Resolution and speed of some recently-reported high-speed monolithic CMOS ADC's.

Additional details regarding these implementations are tabulated below.

	Bits	MS/s	P(mW)	L(μm)	C_{in}(pF)	Architecture	Source	Year	References
(a)	13	2.5	100	3.0	4.0	3 bits/stage pipeline	UCB	'90	[79], [80]
(b)	10	15	95	0.8	—	4 bits/stage pipeline	Hitachi	'92	[88]
(c)	10	20	240	0.9	0.5	1 bit/stage pipeline	AT&T	'91	[75]
(d)	9	25	100	1.3	—	Two-step subranging	Hitachi	'91	[58]
(e)	10	50	900	0.8	5.0	4 bits/stage pipeline	NEC	'92	[161], [162]
(f)	8	30	180	1.0	—	Full flash	TI	'91	[122]
(g)	8	40	105	1.4	11.0	Two-step subranging	Sony	'89	[31]
(h)	8	50	225	0.8	—	Two-step subranging	Hitachi	'90	[87]
(i)	8	85	1100	1.0	2.0	Parallel pipeline	UCB	'92	[21], [22]
(j)	6	125	200	1.5	—	Full flash	Mic. Netw.	'92	[91]
(k)	8	20	200	2.0	16.0	Two-step subranging	Hitachi	'88	[86]
(l)	10	20	30	0.8	12.0	Two-step interpolation	Matsushita	'93	[65]

Note that in the column for input capacitance (C_{in}), "—" indicates that the value of C_{in} was not reported in the cited reference. The following are some general observations based on the data for CMOS ADC's given in the above table.

For 8-bit resolution, the CMOS solution at sampling frequencies up to about 50 MS/s, but in practice used more commonly in video applications at around 20 MS/s, is the two-step, subranging architecture; for example, data points (g) and (h) in Fig. 2.1. (A speed capability of 20 MS/s is appropriate for both of the standard digital video sample rates mentioned in Section 2.1.2, namely, 13.5 MS/s and 14.3 MS/s.) Since about 1988, many papers have appeared in the literature on variations and refinements of the two-step, subranging topology; in particular, from companies in Japan, including Sony [30], [31], Hitachi [58], [86], [87], Toshiba [145], [146], and Mitsubishi [45], [46]. These ADC's have been used extensively in high-volume video and consumer electronics applications. Some of the two-step, subranging architectures are considered further in Section 7.1.2, and certain features that enable the high sample rates to be achieved are discussed.

One of the problems with the two-step architecture is that the input capacitance, while low in comparison to that of a full flash ADC, is still relatively high, often 10–16 pF. The significantly lower input capacitance of the pipeline architecture is one of its key advantages and leads to ease of driveability and overall ease of use. Note that for all the CMOS ADC's discussed in this section, the input capacitance is a *switched capacitance*, and so may also be viewed as being resistive, with resistance given by $1/(F_s C_{in})$.

A related problem associated with the two-step subranging implementations is the integrity of the sampling instant. First, there is a degradation due to sampling time mismatch across the bank of comparators causing the effective sampling instant to be poorly defined. Second, at the transistor level, the actual sampling technique itself is often sensitive to input-signal-dependent charge injection errors and not as robust as in typical pipeline implementations. These factors tend to result in relatively poor dynamic behavior.

More detailed information on dynamic performance characteristics — specifically, the Signal-to-Noise-plus-Distortion Ratio (SNDR) as a function of analog input signal frequency — is now tabulated for some of the CMOS ADC's listed in the previous table, and results from additional published implementations are included also. If SNDR measurements were not reported, this is indicated by "—". Before considering this data, it is worthwhile to recall that the SNR (in dB) of an ideal, distortionless, N-bit ADC when digitizing a full-scale sinewave is (6N + 1.76).

Bits	MS/s	P(mW)	L(μm)	C_{in}(pF)	Architecture	Signal/(Noise+Disto.)	Source	Ref.
10	15	95	0.8	—	4 bits/stage pipeline	—	Hitachi	[88]
10	20	240	0.9	0.5	1 bit/stage pipeline with S/H	5.0 MHz 60 dB 15.0 MHz 57 dB 19.9 MHz 55 dB	AT&T	[75]
10	40	900	0.8	5.0	4 bits/stage pipeline	1.0 MHz 58.0 dB 5.0 MHz 56.6 dB 20.0 MHz 48.0 dB	NEC	[162]
10	15	250	1.0	3.2	Two-step recycling	0.1 MHz 52 dB 7.5 MHz 45 dB	Univ. Illinois	[129]
10	15	30	0.8	12.0	Two-step interpolation	1.0 MHz 55 dB 3.5 MHz 50 dB 5.5 MHz 40 dB	Matsut.	[65]
9	25	100	1.3	—	Two-step subranging	—	Hitachi	[58]
8	30	180	1.0	—	Full flash with robust sampling technique	1.2 MHz 45 dB 15.0 MHz 45 dB	TI	[122]
8	40	105	1.4	11.0	Two-step subranging	—	Sony	[31]
8	20	50	1.0	10.0	Recycling two-step subranging	2.0 MHz 45 dB 9.8 MHz 41 dB	Mitsub.	[46]
8	20	50	0.8	—	Two-step subranging	—	Toshiba	[146]
8	50	600	1.0	1.5	Subtractor-type two-step subranging with S/H	10 MHz 49.3 dB 30 MHz 45.8 dB 50 MHz 41.7 dB	NTT	[52]
8	50	225	0.8	—	Two-step subranging	—	Hitachi	[87]
8	85	1100	1.0	2.0	Parallel pipeline	4 MHz 43.5 dB 10 MHz 43.1 dB 20 MHz 42.5 dB 40 MHz 40.5 dB	UCB	[22]
8	20	200	2.0	16.0	Two-step subranging	—	Hitachi	[86]
6	125	125	1.5	—	Full flash, asynch. AZ	—	M. Net.	[91]

It is clear that the ideal SNDR is never achieved — the best that is attained is within about 1–2 dB, and in most cases the deviation from ideal is much greater. Furthermore, in many of the two-step subranging architectures, the SNDR falls off significantly (often by 6–10 dB) as the input signal approaches the Nyquist frequency $F_s/2$. Thus, the *effective number of bits* (ENOB) at high input frequencies is typically 1–2 *bits* less than the nominal ADC resolution. Of course, it should be remembered that in certain applications, this may not matter. In general, the pipeline architectures (including the implementation described later in this dissertation) display very good SNDR behavior with increasing input slew rates. Note that the full flash ADC chip reported by TI [122] achieves good high-frequency dynamic performance by adding — in front of *each* comparator — a fully differential sampling network and employing a more robust sampling scheme that eliminates signal-dependent charge injection errors.

In order to accommodate high input signal frequencies, and assuming that the source impedance of the driving source is not the limit, it is clear that the higher the capacitance presented by input acquisition/sampling circuitry, the lower must be the effective switch on-resistance. However, realizing low MOSFET on-resistance entails using large devices, which requires silicon area and consumes power in the associated clock driver circuitry. Accordingly, better A/D converter dynamic performance tends to be associated with lower input capacitance — which results either from (a) using a front-end S/H or (b) using a multistage topology with a small number of bits per stage, which automatically leads to a low input capacitance.

As implemented in the Sony and Hitachi papers, the two-step, subranging architecture does not require any op amps. This enables relatively low power dissipation to be achieved, and furthermore is an important consideration given the trend towards lower power supply voltages, since op amp design becomes very difficult with reduced supplies. Some two-step flash architectures do contain a front-end S/H, but these implementations tend to have relatively higher power: an example is the 8-bit, 50-MS/s, two-stage,

subtractor-type, subranging ADC described in [52], which has a power consumption of 600 mW, but better SNDR.

To summarize the preceding discussion: a key point to note from the data presented above is that speed and resolution alone do not tell the whole story, and are not always sufficient in order to characterize ADC performance, especially if the application involves input signals significantly above about $F_s/20$. Many papers do show the reconstructed output on an oscilloscope from a beat frequency test, which is useful qualitative information, but does not constitute the quantitative performance characterization provided by SNDR as a function of input signal frequency.

The two-step subranging architecture has been extended above 8 bits; for example, a 9-bit implementation is described in [58], and two-step *recycling* architectures have been reported at the 10-bit level [128], [129], but in general as the resolution increases from 8 to 10 bits pipelined multistage architectures begin to appear more. The three 10-bit data points indicated on Fig. 2.1 all use a pipeline approach. At higher resolutions, a 12-bit two-step flash ADC using a recycling topology and employing self-calibration [61] has been reported, and is in production. At this resolution, however, pipelined multistage architectures have been demonstrated to have significant advantages in terms of reduced hardware cost [79], [80].

It is noteworthy that many of the two-step subranging ADC implementations at 8-bit resolution — especially for consumer electronics applications — do not use differential signal paths. Clearly, there is a substantial hardware cost associated with a fully differential implementation. Depending on the electrical environment and the application, the common-mode noise rejection benefit from a fully differential signal path may not be necessary. Standalone ADC's are usually required to accept a single-ended input. In embedded ADC's, where the signal path is already differential for dynamic range and power supply rejection reasons, it is often necessary to maintain a differential nature.

2.3.1 Bipolar and BiCMOS ADC's

The performance characteristics of some monolithic A/D converters implemented in bipolar technologies are tabulated below.

Bits	MS/s	P(W)	Supplies	f_T(GHz)	Architecture	Source	Year	References
6	400	2.70	±5.0 V	5.0	Analog encoding (folding)	HP	'84	[23]
6	30	0.012	+3.0 V	8.4	Full flash	Hitachi	'87	[47], [48]
6	2000	2.00	–5.2 V	25.7	Full flash	NTT	'88	[155]
6	1000	2.30	–5.2 V	13.0	2 interleaved flash ADC's	Matsushita	'91	[90]
8	50	0.30	+5.0 V	7.5	Folding and interpolation	Philips	'87	[150]
8	100	0.80	–5.2 V	9.0	Folding and interpolation	Philips	'88	[151]
8	650	0.85	–4.5 V	13.0	Folding and interpolation	Philips	'92	[152], [153]
10	300	4.00	–5.2 V	11.0	Interpolated-parallel	Matsushita	'92	[63], [64]
10	50	0.75	+5.0 V	3.0	2-stage (5+6) with folding	Philips	'92	[154]
10	75	0.80	±5.0 V	4.0	2-stage (4+7) with folding	UCLA	'93	[20]

In general, for ADC sample rates of 50 MS/s or above, bipolar implementations tend to predominate. Particularly noteworthy is the *folding and interpolation* architecture developed and in production (at 8-bit resolution) by Philips [107], which has yielded extremely high sample rates at moderate power levels. As may be seen, other architectures employing the concepts of folding and interpolation have also been reported. The parallel time-interleaved data acquisition systems reported by HP [109], [114], [93], have used bipolar ADC's employing a folding, or *analog encoding*, approach in the individual signal paths.

Finally, some examples of BiCMOS implementations are examined. Performance data is listed below.

Bits	MS/s	P(mW)	Supplies	f_T (GHz)	L(μm)	Architecture	Source	Year	Ref.
10	100	950	–5.0 V	11.0	0.8	Pipelined subranging	NEC	'93	[125]
10	30	750	±5.0 V	7.5	1.2	Two-step interpolation	Matsushita	'90	[89]
10	20	1000	±5.0 V	2.0	2.0	4 bits/stage pipeline	AD	'90	[111]
10	40	700	+5.0 V	9.0	0.8	Two-step subranging	Toshiba	'89	[144]

Relatively few high-speed standalone BiCMOS ADC's have been reported. As may be seen from the above tabulated data, the performance has not been particularly impressive, and the power levels are relatively high. However, some complex mixed-signal chips implemented in BiCMOS technologies have been reported that incorporate embedded BiCMOS ADC's [118], [108].

2.3.2 Comparison of ADC Performance in Different Technologies

A comparison across technologies usually entails some compromises and approximations, but often yields useful insights. One approach is to compare ADC implementations of a certain resolution, on the basis of sample rate as a fraction of technology f_T. This metric, F_s/f_T, does not take account of power or hardware cost: it simply gives a measure of how well the intrinsic speed of the technology is being utilized. However, it is appropriate for this dissertation, since the objective of this work is to

achieve the highest possible throughput for a given CMOS technology. In Fig. 2.2 below, resolution is plotted against F_s/f_T for all the previously listed examples. For CMOS, an estimate of the NMOS f_T is used based on the channel length, which gives an f_T of 6 GHz for a 1-μm device [120]. It is assumed that the f_T scales as $1/L^2$. For the BiCMOS ADC's, the f_T used is that of the bipolar npn.

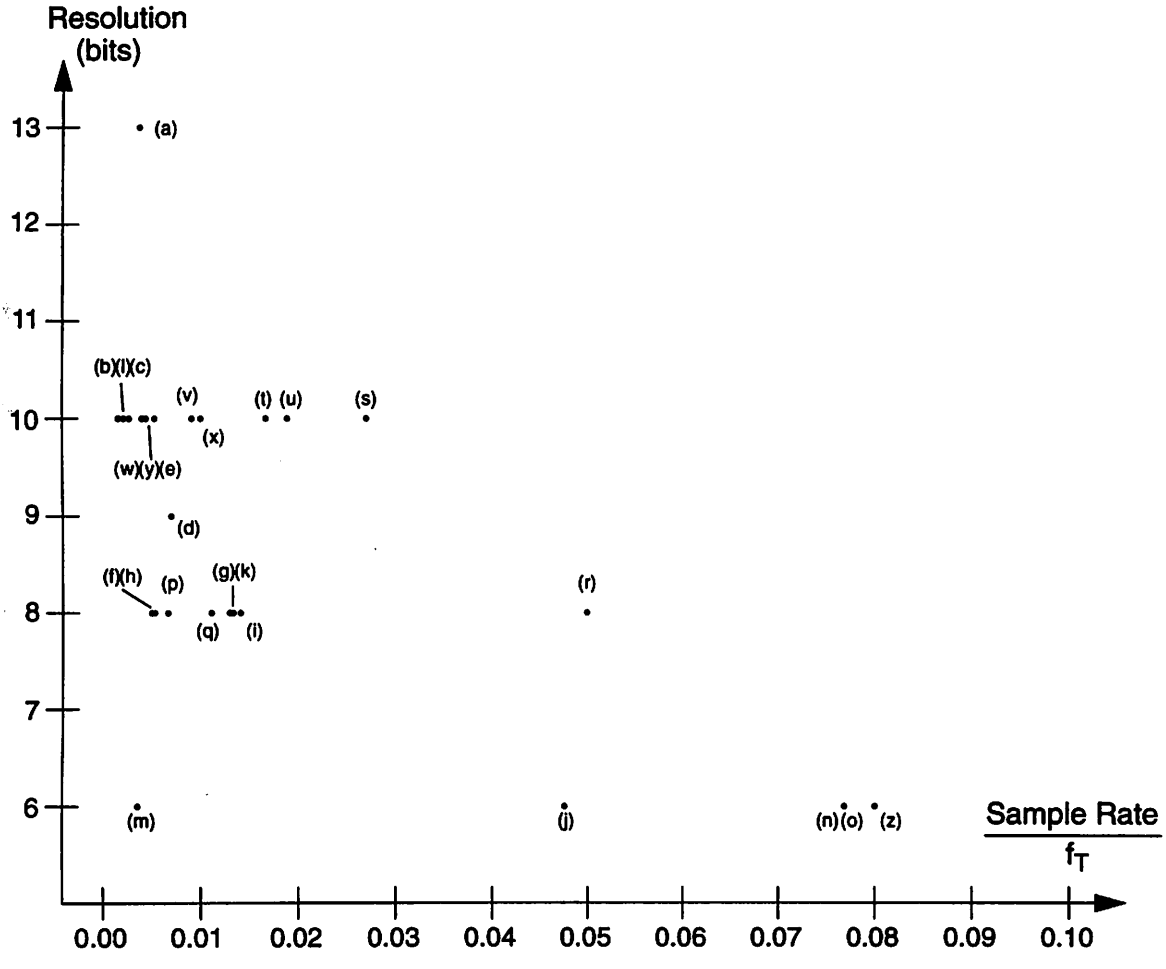


Fig. 2.2 Resolution plotted against speed normalized to technology f_T for high-speed ADC's.

Some details regarding these implementations are tabulated below. The data is grouped by technology: CMOS, bipolar, and BiCMOS.

	Bits	MS/s	P(mW)	L or f_T	Architecture	f_T/Speed	Source	Year	Ref.
CMOS									
(a)	13	2.5	100	3.0 µm	3 bits/stage pipeline	267	UCB	'90	[80]
(b)	10	15	95	0.8 µm	4 bits/stage pipeline	625	Hitachi	'92	[88]
(c)	10	20	240	0.9 µm	1 bit/stage pipeline	370	AT&T	'91	[75]
(d)	9	25	100	1.3 µm	Two-step subranging	142	Hitachi	'91	[58]
(e)	10	50	900	0.8 µm	4 bits/stage pipeline	188	NEC	'92	[162]
(f)	8	30	180	1.0 µm	Full flash	200	TI	'91	[122]
(g)	8	40	105	1.4 µm	Two-step subranging	77	Sony	'89	[31]
(h)	8	50	225	0.8 µm	Two-step subranging	188	Hitachi	'90	[87]
(i)	8	85	1100	1.0 µm	Parallel pipeline	70	UCB	'92	[22]
(j)	6	125	200	1.5 µm	Full flash	21	Mic. Netw.	'92	[91]
(k)	8	20	200	2.0 µm	Two-step subranging	75	Hitachi	'88	[86]
(l)	10	20	30	0.8 µm	Two-step interpolation	469	Matsushita	'93	[65]
Bipolar									
(z)	6	400	2700	5.0 GHz	Analog encoding (folding)	13	HP	'84	[23]
(m)	6	30	12	8.4 GHz	Full flash	280	Hitachi	'87	[48]
(n)	6	2000	2000	25.7 GHz	Full flash	13	NTT	'88	[155]
(o)	6	1000	2300	13.0 GHz	2 interleaved flashes	13	Matsushita	'91	[90]
(p)	8	50	300	7.5 GHz	Folding/interpolation	150	Philips	'87	[150]
(q)	8	100	800	9.0 GHz	Folding/interpolation	90	Philips	'88	[151]
(r)	8	650	850	13.0 GHz	Folding/interpolation	20	Philips	'92	[153]
(s)	10	300	4000	11.0 GHz	Interpolated-parallel	37	Matsushita	'92	[64]
(t)	10	50	750	3.0 GHz	2-stage (5+6) with folding	60	Philips	'92	[154]
(u)	10	75	800	4.0 GHz	2-stage (4+7) with folding	53	UCLA	'93	[20]
BiCMOS									
(v)	10	100	950	11.0 GHz	Pipelined subranging	110	NEC	'93	[125]
(w)	10	30	750	7.5 GHz	Two-step interpolation	250	Matsushita	'90	[89]
(x)	10	20	1000	2.0 GHz	4 bits/stage pipeline	100	AD	'90	[111]
(y)	10	40	700	9.0 GHz	Two-step subranging	225	Toshiba	'89	[144]

The following are some observations on the above data.

The CMOS, 8-bit, two-step subranging architectures — such as (g), (h), and (k) — achieve sample rates roughly in the range $f_T/200$–$f_T/80$, and have relatively low power dissipation. At resolutions of 10 bits or higher, it is apparent that pipelined multistage architectures in CMOS and BiCMOS are becoming prevalent, as exemplified by implementations (a), (b), (c), (e), (v), and (x).

Generally, at resolutions of 8–10 bits, sample rates of $f_T/100$ are achievable; $f_T/50$ appears to represent a practical technology limit: realizing A/D converters with sample rates above that is extremely difficult. The work reported in this dissertation achieves a speed of $f_T/70$: this is the highest technology-normalized sample rate ever reported in CMOS for an ADC with resolution greater than 6 bits, and it compares favorably with ADC implementations in other technologies. At 6-bit resolution, some bipolar IC's with sample rates of approximately $f_T/13$ have been reported — for example, (n), (o), and (z) above.

Architectures employing folding work well in bipolar technology and take advantage of the high device transconductance; the implementations labelled (s), (r), and (z) above are notable examples. This thesis does not cover these topics in any detail, but it is appropriate to mention that the two ideas of folding and interpolation are quite distinct. Various forms of interpolation techniques have appeared in CMOS such as the reference feedforward technique [136], the two-residue architecture [83], and the two-step capacitive interpolation topology [65]. In general, it appears that further embodiments of interpolation have significant potential especially for applications where DNL is more important than INL and for applications where low power is the primary objective.

Chapter 3 Multistage A/D Converter Architectures and Implementations

3.0 INTRODUCTION

This chapter presents a review and examination of pipelined multistage A/D converter (ADC) architectures. Broadly, it discusses (i) converter system- and algorithm-level issues such as dc transfer characteristics, coding considerations, and error correction strategies, and also (ii) circuit implementation issues, error sources, and nonidealities. Throughout the chapter, an attempt is made to present and discuss multistage ADC's from a number of different viewpoints, including both qualitative, intuitive explanations, and also systematic, analytical approaches. First, in Section 3.1, the basic principles are introduced, and in Section 3.2, the concepts of overrange detection, coding redundancy, and digital error correction are reviewed. Some qualitative intuitive explanations are also given to provide insight into how multistage ADC's operate. Section 3.3 briefly discusses the accuracy requirements on the DAC and gain blocks used within multistage ADC's. Next, Section 3.4 presents various implementation options for pipelined multistage ADC's with emphasis on CMOS switched-capacitor techniques. Finally, in Section 3.5, error sources and nonidealities are discussed such as offset errors, gain errors (for example, due to finite op amp open-loop gain), and thermal noise.

3.1 PIPELINED MULTISTAGE ADC ARCHITECTURES

The functional block diagram for a pipelined multistage (or multistep) A/D converter

is shown in Fig. 3.0 below.

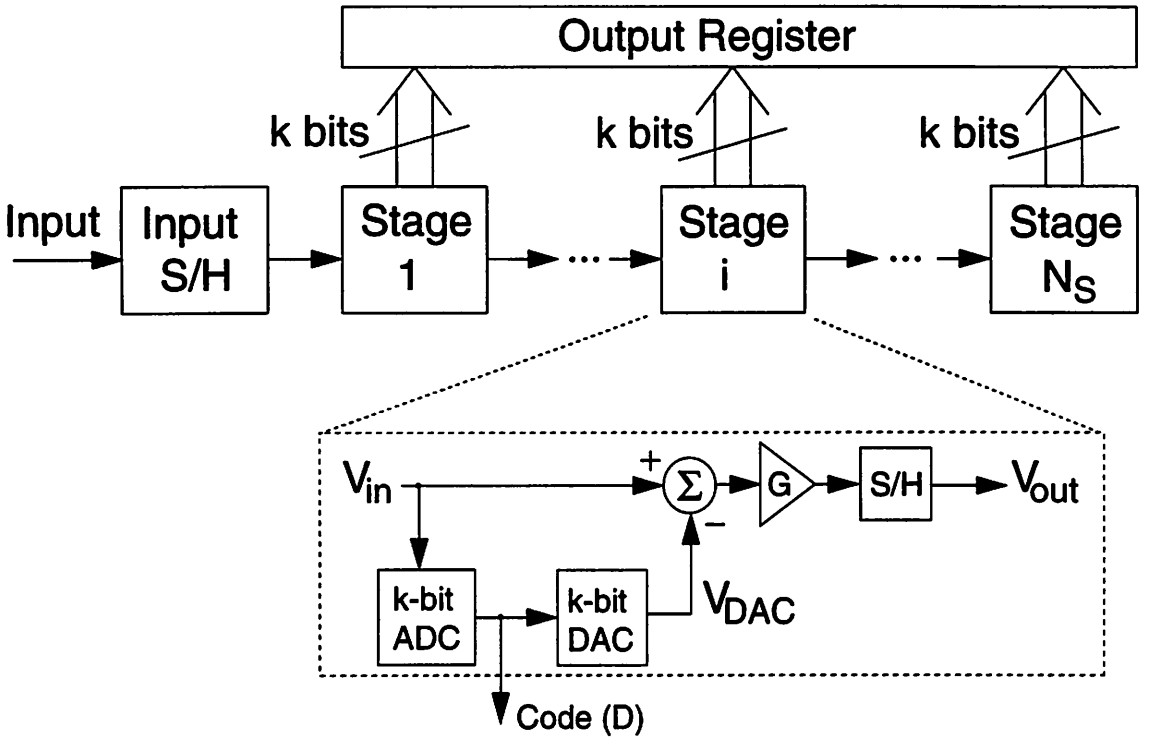


Fig. 3.0 Basic functional block diagram of a pipelined multistage ADC.

The system consists of a cascade of N_S blocks or stages. Assume at this point that all stages are identical. As indicated on the figure, the input V_{in} to each stage is connected to an ADC, typically consisting of a bank of comparators, which performs a k-bit A/D conversion. Digital logic (not shown) (a) performs the encode function to generate the k-bit ADC digital output code D and (b) drives the k-bit DAC to produce V_{DAC} — a quantized analog estimate of the input. This DAC output is then subtracted from the input to give an analog remainder or *residue,* $(V_{in} - V_{DAC})$, which is gained up by amount G, and transferred to the next stage. Typically k is a relatively small number of bits, usually in the range 2–5. The *interstage gain* G is usually proportional to 2^k. (At this point, for simplicity, G may be assumed to be equal to 2^k.) However, the case of $G = 1$, which corresponds to a subtractor only, i.e., no "gain", is relevant for analysis of subranging and two-step flash architectures, and will also be discussed. The k-bit ADC and k-bit DAC

within one stage of a multistage ADC are sometimes referred to as an A/D subconverter or sub-ADC, and a D/A subconverter or sub-DAC, respectively [73]. The last stage is different from the previous stages: it consists only of the sub-ADC block, and does not contain the DAC or gain function.

There are a number of ways to specify unambiguously the dc characteristics of a single stage of a multistage ADC. First, the parameters of the individual components within the stage can be specified, i.e., the ADC threshold levels or decision levels, the digital codes, the mapping of codes to DAC levels, and finally the interstage gain. As an example, a simple 2-bit stage that operates on a unipolar signal is shown in Fig. 3.1.

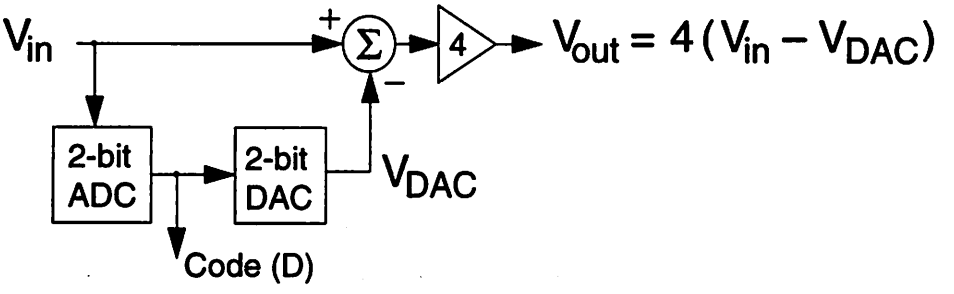


Fig. 3.1 Functional schematic of a 2-bit stage within a multistage A/D converter.

The parameters that specify completely this block are listed below.

Input range:	[0, 1] (normalized to V_{ref})
3 ADC threshold levels:	{1/4, 1/2, 3/4}
4 digital codes:	{0, 1, 2, 3} or {00, 01, 10, 11} in binary representation
4 corresponding DAC levels:	{0, 1/4, 1/2, 3/4}
Interstage gain:	G = 4
Number of bits:	k = 2

For convenience, the ADC threshold levels and DAC reference levels are normalized to V_{ref}. This convention is followed throughout this chapter when listing stage parameters. In the above example, the input signal lies between 0 and V_{ref} — i.e., one polarity only, or *unipolar*. Whenever the signal can assume both positive and negative values, such as in a differential implementation, it is said to be *bipolar*. The comparator

levels are spaced $V_{ref}/4$ apart and divide the range into $2^2 = 4$ intervals. Therefore, one Least Significant Bit (LSB) *with respect to the stage* has amplitude $V_{ref}/4$. The symbol D will be used to denote a digital code; D may be represented either as a decimal integer or as an unsigned binary number having some number of bits. Depending on the context, one or other or both of these representations may be most appropriate.

The above table/listing implies the relationship between threshold levels, digital codes, and DAC levels. For example, (referring to the table) when $0 < V_{in} < 1/4\,V_{ref}$, the digital code generated by the stage is $D = 00\,(\text{binary}) = 0\,(\text{decimal})$, which maps to the DAC level 0; when $1/4\,V_{ref} < V_{in} < 1/2\,V_{ref}$, the digital code is $D = 01$, which maps to the DAC level $V_{ref}/4$, etc. Obviously, for this case, the number of bits (k) is 2 and the interstage gain (G) is $2^2 = 4$.

Graphical representations of ADC dc characteristics are also extremely useful. The static transfer curve of an A/D converter is commonly presented as a graph of the ADC digital output code plotted against the analog input voltage V_{in}, and has the familiar staircase appearance. However, for the purposes of this chapter, this transfer curve is defined more carefully. Rather than considering the output *code* as such, the focus will be on the *underlying discrete-valued analog voltage* $\tilde{V}_{in}$ *represented by that code.* The notation $\tilde{V}_{in}$ is intended to emphasize that, fundamentally, A/D conversion involves amplitude-quantization: an ADC takes a continuous-valued input signal V_{in}, and at the output yields a digital code D representing a discrete-valued signal $\tilde{V}_{in}$, which is an *approximation* or *estimate* of the original input V_{in}. Usually, $\tilde{V}_{in}$ is directly proportional to the numerical value of the integer code, sometimes with an offset of $\pm 1/2$ LSB. Always, the output code *represents* or *maps to* a quantized analog approximation of the input, and a plot of $\tilde{V}_{in}$ versus V_{in} exhibits a staircase appearance. The mapping from V_{in} to $\tilde{V}_{in}$ may be considered a staircase or piecewise-constant approximation. Note that the threshold level spacing is uniform in the above example — this is not the case in some more general cases to be considered later.

In the same way that the transfer curve of a complete ADC may be examined, it is similarly possible to consider the corresponding plot for a single stage — this will also have a staircase shape. For a single stage, $\tilde{V}_{in}$ — the discrete-valued approximation of the input signal — is simply the DAC voltage generated by that stage. For the 2-bit stage whose parameters are given above, the input-output characteristic from V_{in} to $\tilde{V}_{in}$ is shown in Fig. 3.2 with the ADC digital outputs corresponding to each interval indicated.

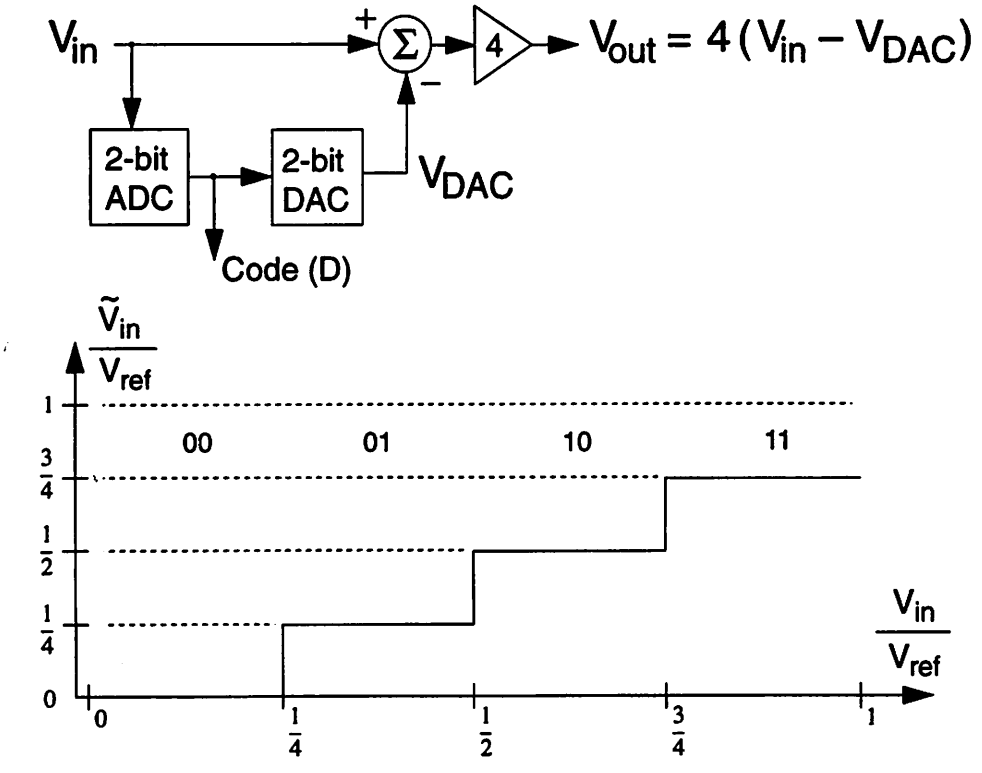


Fig. 3.2 Static transfer characteristic of a single 2-bit pipeline stage from stage input (V_{in}) to DAC output voltage ($\tilde{V}_{in} = V_{DAC}$).

In the analysis and design of multistage ADC's, it is also very important to examine the dc static transfer characteristic of a single stage *from the input* V_{in} *to the stage output* V_{out}, *i.e., the continuous-valued, amplified residue.* Note that (a) V_{out} is the voltage which forms the input to the next stage, and (b) V_{out} is not the DAC voltage — it is not an approximation to V_{in}. Shown in Fig. 3.3 is the static input-output transfer characteristic from V_{in} to V_{out} of the 2-bit pipeline stage, with the ADC digital outputs corresponding

to each interval indicated on the graph.

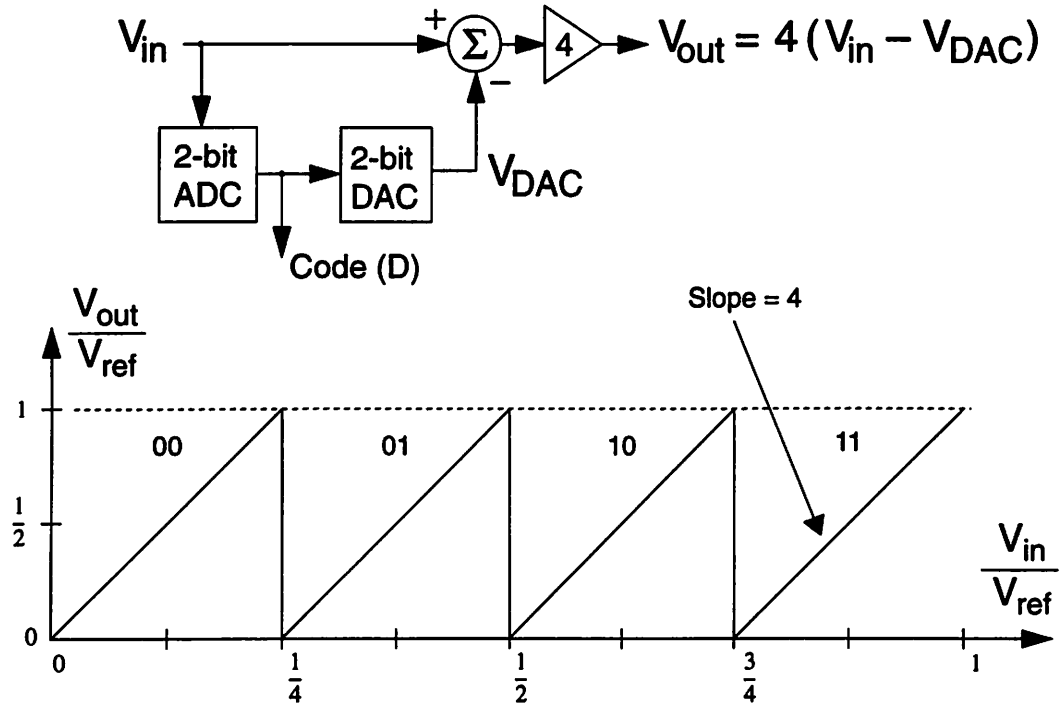


Fig. 3.3 Static transfer characteristic of a single 2-bit pipeline stage from stage input (V_{in}) to gained-up residue (V_{out}). The interstage gain of 4 leads to a slope of 4, as shown.

This sawtooth-like shape for the input-output relationship is very characteristic of multistage architectures. Qualitatively, each time V_{in} exceeds a particular threshold or decision level, the ADC digital code increases by one, which causes the DAC analog output V_{DAC} to increase by one LSB with respect to the stage. This in turn causes the output of the subtractor block to *decrease* by one LSB and so V_{out} — the final output of the gain block — must decrease by an amount equal to the gain G multiplied by the DAC LSB step size. In the example shown in Fig. 3.3, this is equal to V_{ref}, which is Full Scale for this unipolar case. Therefore, each time V_{in} exceeds a comparator decision level, V_{out} jumps discontinuously by an amount equal to minus Full Scale. Clearly, the slope of the positive-slope, straight-line regions of the input-output characteristic of the block is equal to the stage gain G.

In summary, a table listing of the stage parameters and both of the above-mentioned dc transfer characteristics — (i) from input V_{in} to amplitude-quantized output $\tilde{V}_{in}$ and (ii) from input V_{in} to V_{out}, the gained-up residue of the stage — are useful and informative in the analysis and design of multistage ADC's.

By placing a sample-and-hold (S/H) between each pair of stages, as shown in Fig. 3.0, and employing a two-phase clocking scheme, the stages in the multistage A/D converter can be isolated from each other and system operation may be pipelined. In a *pipelined multistage ADC* the first stage operates on the most recent input sample while the second stage operates on the gained-up residue from the previous sample, etc. Therefore, the analog signal associated with a particular input sample flows down the cascade of stages, and the S/H within each stage functions as an analog pipeline latch. The details of the pipelined nature of the system and issues such as clocking and implementation of the S/H and interstage gain blocks are discussed later in this chapter. In this section and Sections 3.2–3.3, the focus is on the static behavior; the S/H blocks shown in Fig. 3.0 are ignored; and the fact that the subsequent implementation uses a CMOS switched-capacitor approach is irrelevant. In fact, multistage, *non-pipelined, ripple-through* implementations are possible in bipolar technology [9], [54].

Before proceeding, a note regarding terminology is warranted. In the literature, the term *"pipelined"* or *"pipeline"* A/D converter has usually been applied to multistage ADC's *with* interstage gain — i.e., nondegenerate gain ($G > 1$). Examples of such architectures are described in [57], [71]–[76], [79]–[80], [110]–[111], [126]–[127], [135]–[136], [161]–[162]. These ADC topologies contain explicit hardware — such as a SC gain block with closed-loop gain G — to implement the residue amplification; in general, implementation of this gain function requires an op amp. Signal range is usually constant along all stages in the pipeline, and is generally determined by the op amp output voltage swing. (Note that in some unusual implementations, the reference voltages and signal swing decrease *and* an interstage gain greater than one is employed [161]–[162]).

The particular case of a multistage architecture with an interstage gain of unity is sometimes referred to as a *subranging* topology; the "gain" block is a subtractor (G = 1) [52], [125], [147]. More usually, however, the term "subranging" implies *no interstage gain hardware* — i.e., the subtraction is implemented *implicitly* by selecting an appropriate subrange of a set of reference levels, usually generated by a resistor string [28], [31], [58], [87], [145]–[146]. The terms "two-step flash" and "half-flash" [86] are also applied to this topology. No op amps are required; sampling is performed directly on the comparator capacitors. In either case, whether or not explicit subtractor hardware is present, in subranging ADC's the analog signal range decreases progressing from the Most Significant Bit stages to the Least Significant Bit stages.

Henceforth in this dissertation the following interpretations are generally adopted. Unless stated to the contrary, *"subranging"* ADC's are assumed to contain no explicit interstage gain hardware and thus no op amps. *"Pipelined multistage A/D converters"*, or simply *"pipelines"* (the main focus of the dissertation) are assumed to incorporate interstage gains with G > 1, and therefore these ADC's contain explicit interstage gain hardware including op amps. It will be made clear whenever the "subtractor" architecture is being referred to.

Finally, it should be emphasized that the reason to mention subranging converters at all is that many issues relating to dc input-output characteristics, threshold accuracy, and digital error correction are common to (i) pipelined multistage, (ii) subranging, and (iii) subtractor architectures, and it is instructive to consider them all in a unified way.

In the example of Fig. 3.2 and Fig. 3.3, the ADC threshold levels are also DAC levels. However, this is not necessarily always true. Shown in Fig. 3.4 is the static input-output transfer characteristic of a 3-bit pipeline stage having bipolar input and output signal ranges. As before, the ADC digital output codes corresponding to each interval are indicated. Note that in this case, the sub-ADC decision levels are offset from sub-DAC levels.

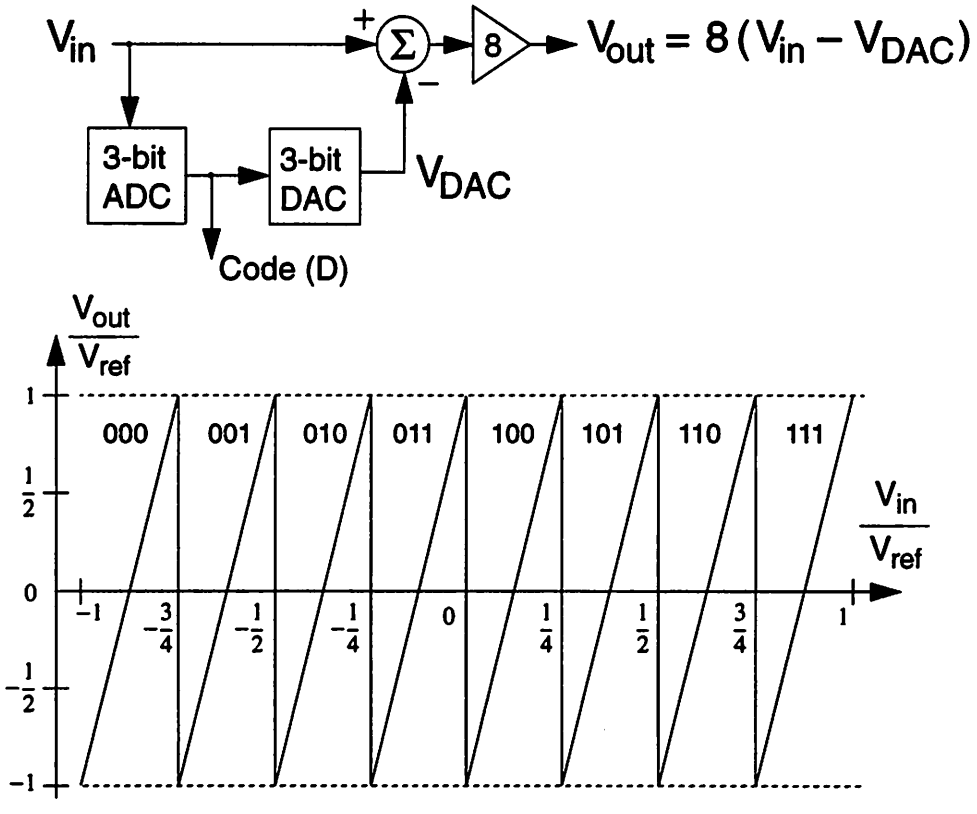


Fig. 3.4 Static transfer characteristic of a single 3-bit pipeline stage from stage input (V_{in}) to gained-up residue (V_{out}).

This stage has the following parameters.

Input range:	[−1, +1] (normalized to V_{ref})
7 ADC threshold levels:	{−3/4, −1/2, −1/4, 0, 1/4, 1/2, 3/4}
8 digital codes:	{0, 1, 2, 3, 4, 5, 6, 7} or
	{000, 001, 010, 011, 100, 101, 110, 111}
8 corresponding DAC levels:	{−7/8, −5/8, −3/8, −1/8, 1/8, 3/8, 5/8, 7/8}
Interstage gain:	G = 8
Number of bits:	k = 3

Note that in general, the discontinuous jumps in the stage input-output characteristic occur at the sub-ADC threshold levels, and the zero-crossings with positive slope occur at the DAC levels. (This latter statement will be modified, of course, if a *desired* systematic offset is present in the stage input-output relationship — i.e., if the stage input-output relationship has the form $V_{out} = G\,(V_{in} - V_{DAC}) + V_{os_{Desired}}$.) The transfer characteristic from V_{in} to $\tilde{V}_{in} = V_{DAC}$ for the ADC stage depicted in Fig. 3.4 is shown below.

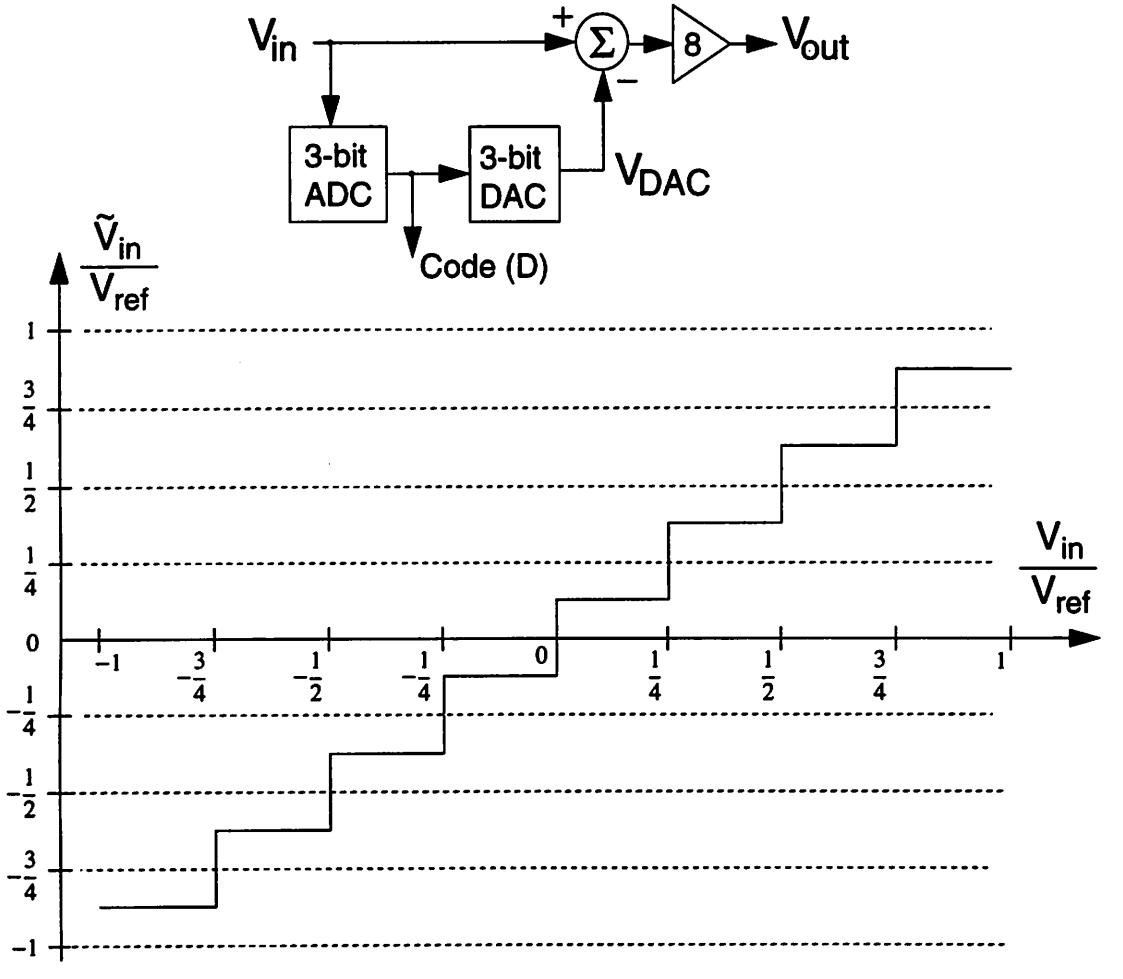


Fig. 3.5 Transfer curve from analog input to DAC output for a 3-bit stage with bipolar input. Note that in this case, the DAC levels are offset from the ADC levels by 1/2 LSB of the stage.

It is apparent that the DAC levels are offset from the ADC levels by $V_{ref}/8$, which is 1/2 LSB at 3 bits (the stage resolution).

So far, it has been assumed that each stage in a multistage ADC corresponds to a physically different circuit block. However, a class of A/D converters known as *recirculating* or *recycling* architectures exists, in which, instead of employing N_S distinct stages in a cascaded configuration as in Fig. 3.0, the *same* stage is used N_S times — i.e., the residue output is fed back to the *same* stage $N_S - 1$ times, as shown in Fig. 3.6 below.

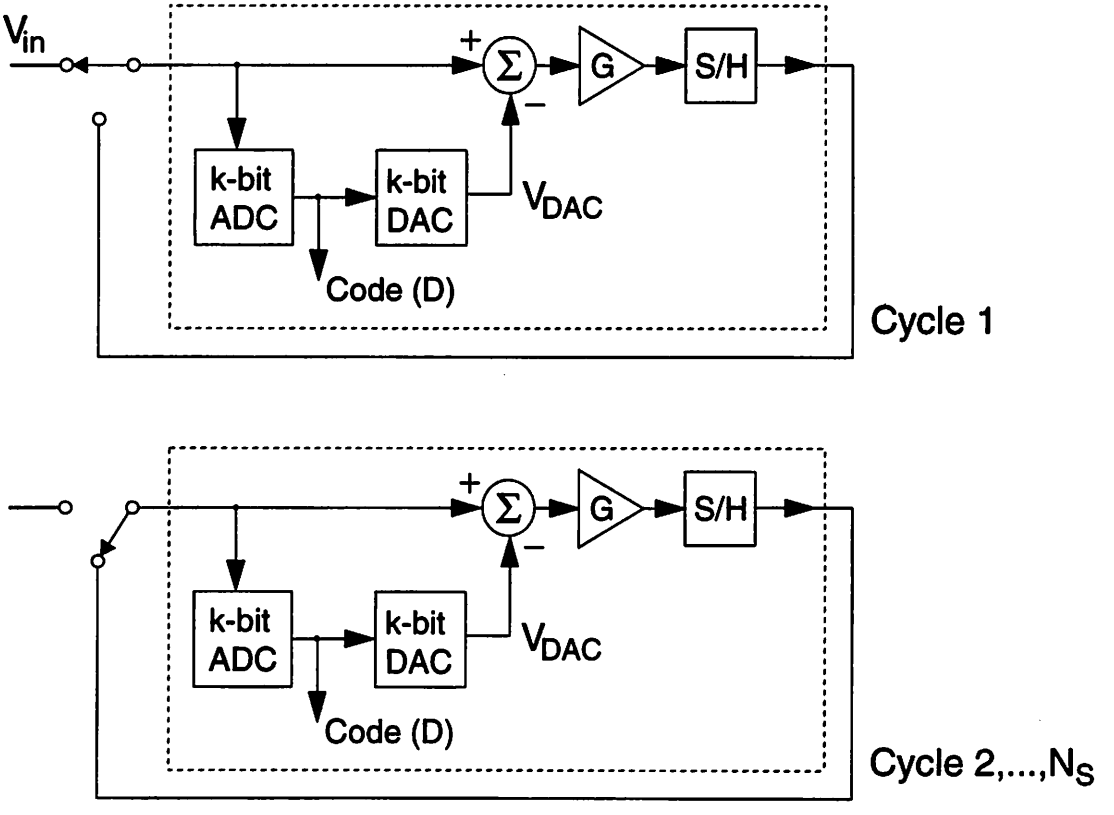


Fig. 3.6 Block diagram of a recirculating (or recycling) ADC requiring N_S cycles per conversion.

Although recycling architectures are not the main topic of this dissertation, they are alluded to occasionally, and it is instructive to consider the similarities and differences between recirculating and multistage topologies, as follows.

In general, pipelined multistage architectures achieve high throughput by using concurrency in time: at any particular instant, each stage is operating on the residue associated with some input sample, and so a number of samples are being processed

concurrently. In contrast, recycling architectures employ concurrency in space, reusing the same hardware multiple times, and therefore can process only a single sample at a time. For pipelines, the *latency* is proportional to the number of stages; similarly, for recycling architectures, the latency is proportional to the number of times the single stage is used. However, because of the concurrency, the *throughput* of pipeline ADC's is essentially independent of the latency, whereas for recirculating topologies, throughput is inversely proportional to the number of stages and hence inversely proportional to the latency.

However, despite these fundamental differences with respect to hardware cost, throughput, and latency, the basic dc input-output relationships and many considerations relating to static nonidealities are the same for both recirculating and non-recirculating topologies. Furthermore, many issues discussed later in this chapter and also in Chapter 4 regarding code assignment and digital correction apply in an identical manner to both classes of A/D architectures. For dc analysis purposes, a recirculating architecture may be viewed as a multistage A/D converter with all stages having exactly identical parameters. Conversely, in some sense, a pipelined multistage ADC may be thought of as an *unraveled* recirculating architecture.

Note that it is possible to shift the gain element in the recirculating A/D converter so that it comes *before* the subtractor block, as shown below in diagram (b) of Fig. 3.7.

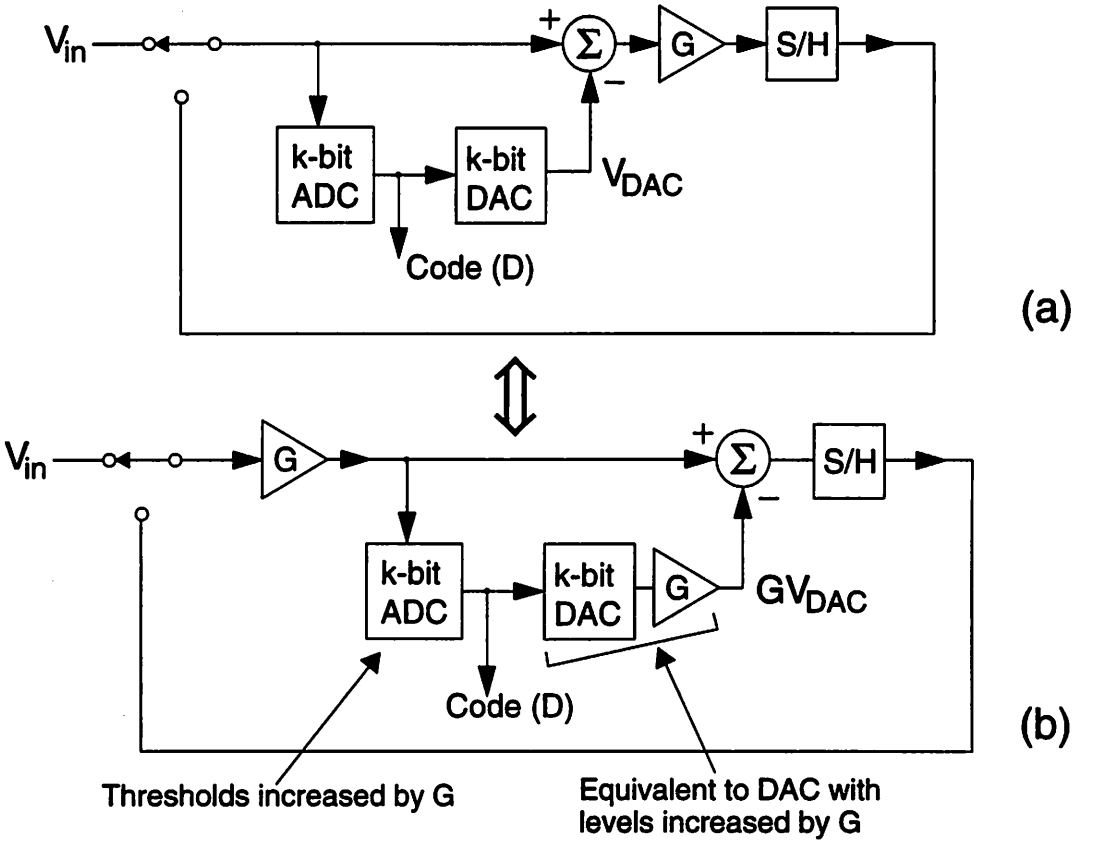


Fig. 3.7 Two configurations for a recycling A/D converter architecture with identical stage dc transfer characteristics. In the bottom figure (b) the gain block G has been moved back before the subtractor, requiring the ADC thresholds to be increased by the same factor G and a gain of G to be incorporated into the DAC.

Both configurations shown in Fig. 3.7 have identical static transfer characteristics for an individual stage. Comparing V_{in} to a particular threshold, in configuration (a), is equivalent to comparing GV_{in} to G times that same threshold, in configuration (b). When topology (b) is used with gain $G = 2$, the resulting 1-bit/stage recycling architecture is often referred to as a "cyclic" or "algorithmic" A/D converter [92], [77], [101], [95], [56]. Algorithmic architectures are discussed further in Section 4.2. An example of a 10-bit recycling two-step A/D converter employing a 5-bit stage and an interstage gain of 32 is described in [128], [129]. Subranging recycling architectures are reported in [45], [46].

3.2 THRESHOLD OFFSETS AND OVERRANGE DETECTION

A problem with the configurations shown hitherto in Fig. 3.3 and in Fig. 3.5 is the sensitivity to errors/offsets in the threshold levels — i.e., it is assumed (a) that precise threshold levels are available and (b) that the subtraction is performed perfectly with zero offset. In an actual implementation, generating accurate thresholds requires comparators with low offset that can respond to small input signals. In CMOS, for resolution requirements beyond 3–4 bits, this invariably necessitates relatively high power, high complexity, often an additional clock phase per conversion to perform autozeroing, and a penalty in comparator speed. As in many instances in analog design, *accuracy costs.* Therefore, there is a strong motivation to reduce the dependence on accurate threshold levels.

3.2.0 The Problem: ADC Errors Due to Decision Level Offsets

The behavior of the stage shown in Fig. 3.4 with an offset in one of the thresholds is illustrated in Fig. 3.8. Note that in the following discussion and examples the focus is on nonideal threshold levels — it is assumed unless otherwise stated that the DAC and gain block in the stage are ideal. DAC and gain errors will be discussed from an ADC system viewpoint in Section 3.3 and later from an implementation viewpoint in Section 3.5.

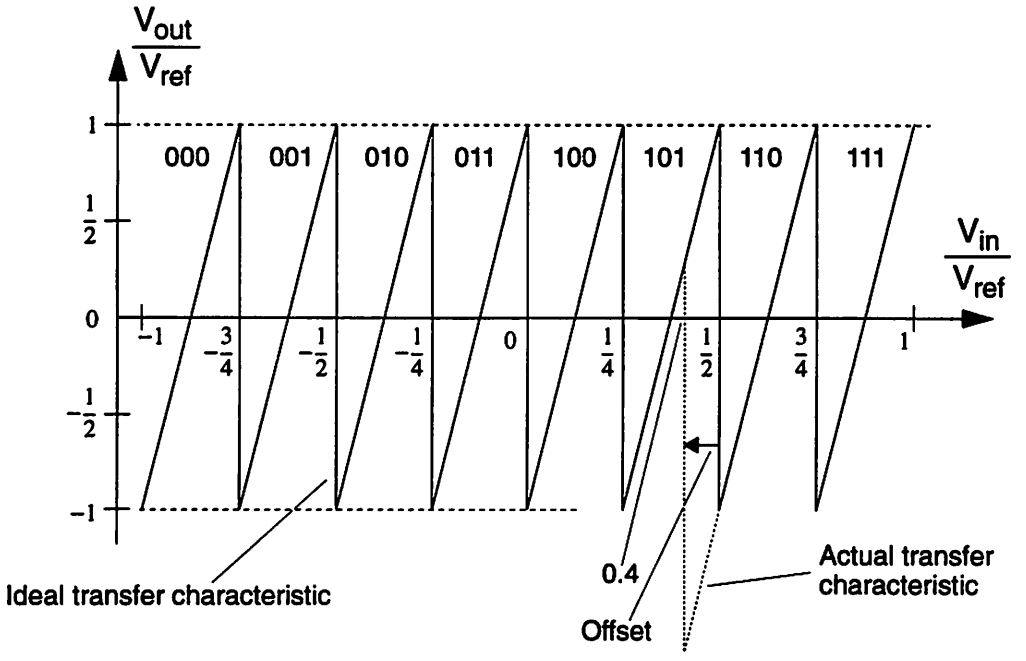


Fig. 3.8 Static transfer characteristic from stage input V_{in} to gained-up residue V_{out} resulting from the decision level ideally at 0.5 having an offset (indicated by the arrow) causing it to be shifted to 0.4. The ideal input-output relationship is shown with continuous lines, and the nonideal actual characteristic, which includes the effect of the threshold offset, has dotted lines wherever it deviates from ideality.

If the decision level has the offset indicated, then — assuming (i) that the DAC levels and gain are ideal and (ii) that the gain block is capable of handling a signal of magnitude greater than V_{ref} at its output — the actual stage output V_{out} is as shown in the dotted line in Fig. 3.8. A key point is that near the nonideal threshold, V_{out} is *more negative than* $-V_{ref}$ *and so outside the range of the next stage.* This may be confirmed by a numerical example. With reference to Fig. 3.8, suppose the input is $V_{in} = 0.41V_{ref}$. Since the threshold nominally at $0.5V_{ref}$ is now at $0.4V_{ref}$, the digital output code is D = 6 = 110 and the corresponding DAC value produced is $0.625V_{ref}$. Therefore, the amplified residue output of the stage is

$$V_{out} = G(V_{in} - V_{DAC}) = 8(0.41V_{ref} - 0.625V_{ref}) = -1.72V_{ref} < -V_{ref}$$

The terms *underrange* and *overrange* are used to describe the situations where the gained-up residue is more negative than $-V_{ref}$ or more positive than $+V_{ref}$ respectively.

More generally, because of the finite linear range present in any practical analog circuit implementation, there is an implicit hard-limiter, or saturation nonlinearity, associated with the interstage gain block. Alternatively, each stage in the multistage ADC can be regarded as having a clipping/limiting function in front of it. If the gained-up residue becomes so large — because of a threshold offset — that it is affected by this limiting/clipping, then errors in the overall ADC transfer characteristic will result. In particular, some digital output codes will be missing. Note that in many CMOS pipeline ADC implementations, the nominal range at the input and output of each stage is the same, and set by the voltage swing obtainable from an op amp within the interstage gain block. A final observation from Fig. 3.8: it is apparent that if V_{out} is less than $-V_{ref}$, the stage digital output D is too large and should be made more negative, implying that a subtraction operation is required. Conversely, V_{out} being greater than $+V_{ref}$ implies that the output code is too low and should be made more positive, thus requiring an addition.

It is of interest to examine the effect of offset in a decision level within one stage on the overall transfer characteristic of a complete A/D — which of course is ultimately of most interest to a user. It is not immediately obvious what the transfer characteristic of a *cascade* of stages looks like when there is an error. A two-stage ADC consisting of two 2-bit stages with an error in the first stage is shown schematically in Fig. 3.9.

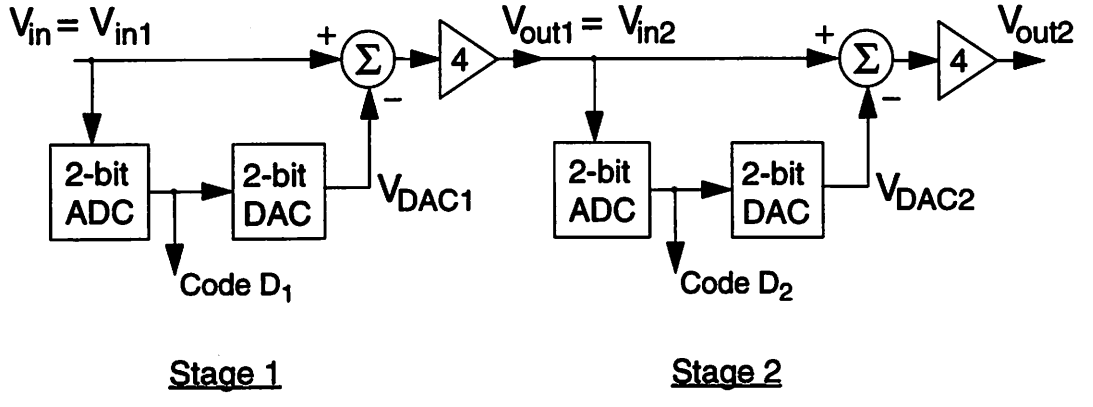


Fig. 3.9 Multistage ADC consisting of two 2-bit stages.

Since the focus of this example is on only these two stages, the gain function of the second stage is not used and may be omitted, as depicted in Fig. 3.10.

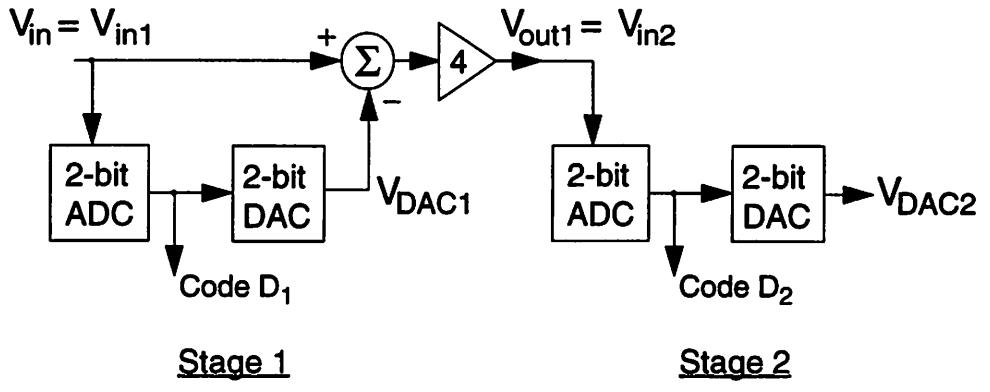


Fig. 3.10 Circuit of interest: ADC consisting of two 2-bit stages.

Some input-output relations with various threshold and subtractor offsets are shown in the following set of figures. Note that in order to minimize the clutter on the diagrams, stage input and output voltages *normalized to* V_{ref} are used. The normalized quantities are denoted by lowercase letters — i.e., $v_{in} = V_{in}/V_{ref}$, $v_{DAC} = V_{DAC}/V_{ref}$, etc. In each figure the following quantities are shown:

v_{out1} — the amplified residue from the first stage;

v_{DAC1} — the DAC output voltage produced by the first stage;

$v_{DAC2}/4$ — the DAC output voltage produced by the second stage divided by 4 to indicate its relative contribution or "weight" towards $\tilde{v}_{in}$;

$\tilde{v}_{in}$ — the overall final amplitude-quantized output of the two-stage ADC; the corresponding digital code is indicated also.

First, for reference, the ideal transfer characteristics are shown in Fig. 3.11. Next, in Fig. 3.12, the threshold level at 0.75 in the first stage has an offset causing it to be shifted to 0.6. In Fig. 3.13, the threshold at 0.75 in stage 1 is shifted in the other direction to 0.83. In Fig. 3.14 the subtractor at the output of the first stage has an offset of +0.25. Finally, in Fig. 3.15, the subtractor in stage 1 has an offset of –0.125.

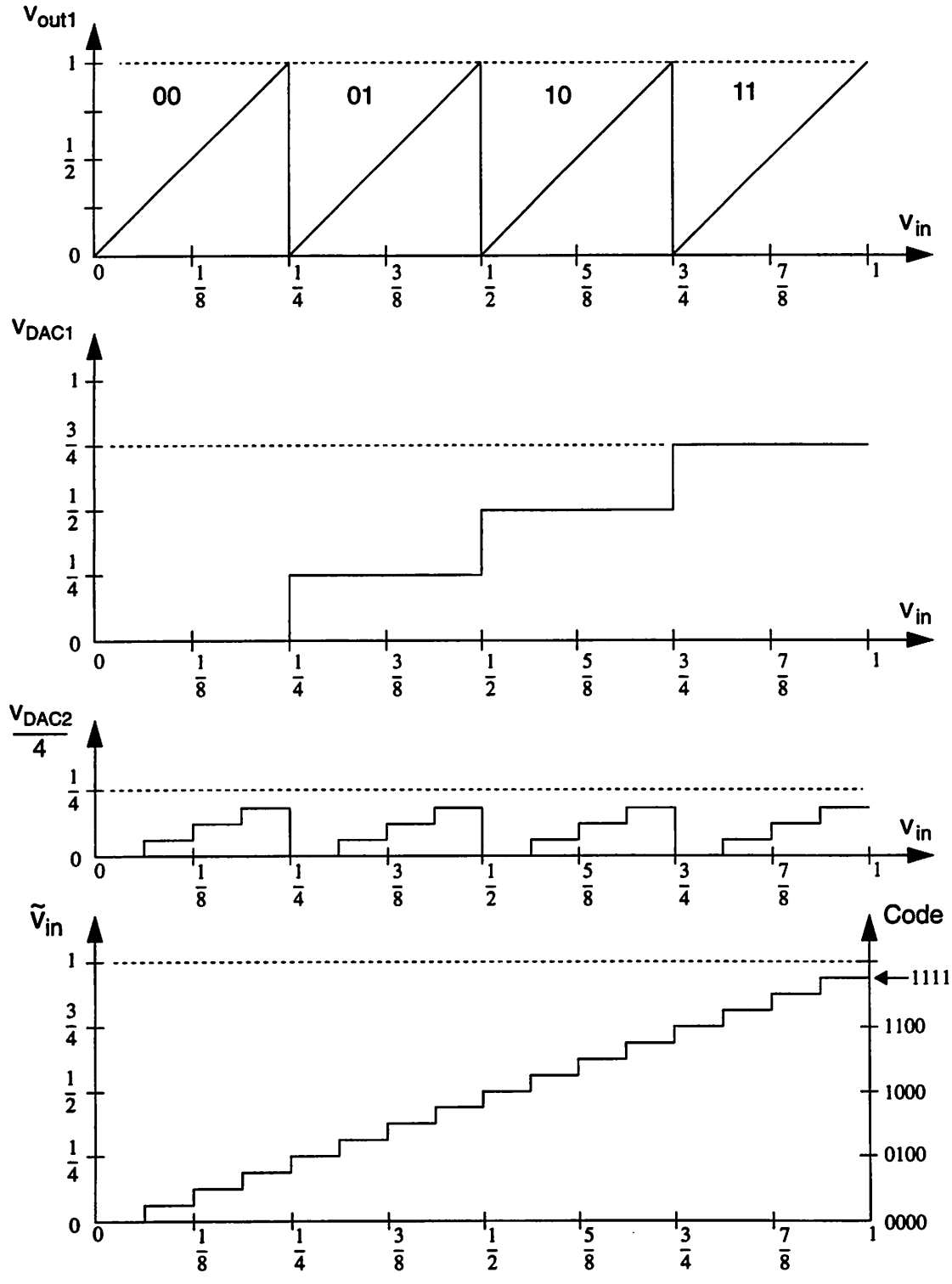


Fig. 3.11 Ideal transfer curves of two-stage ADC.

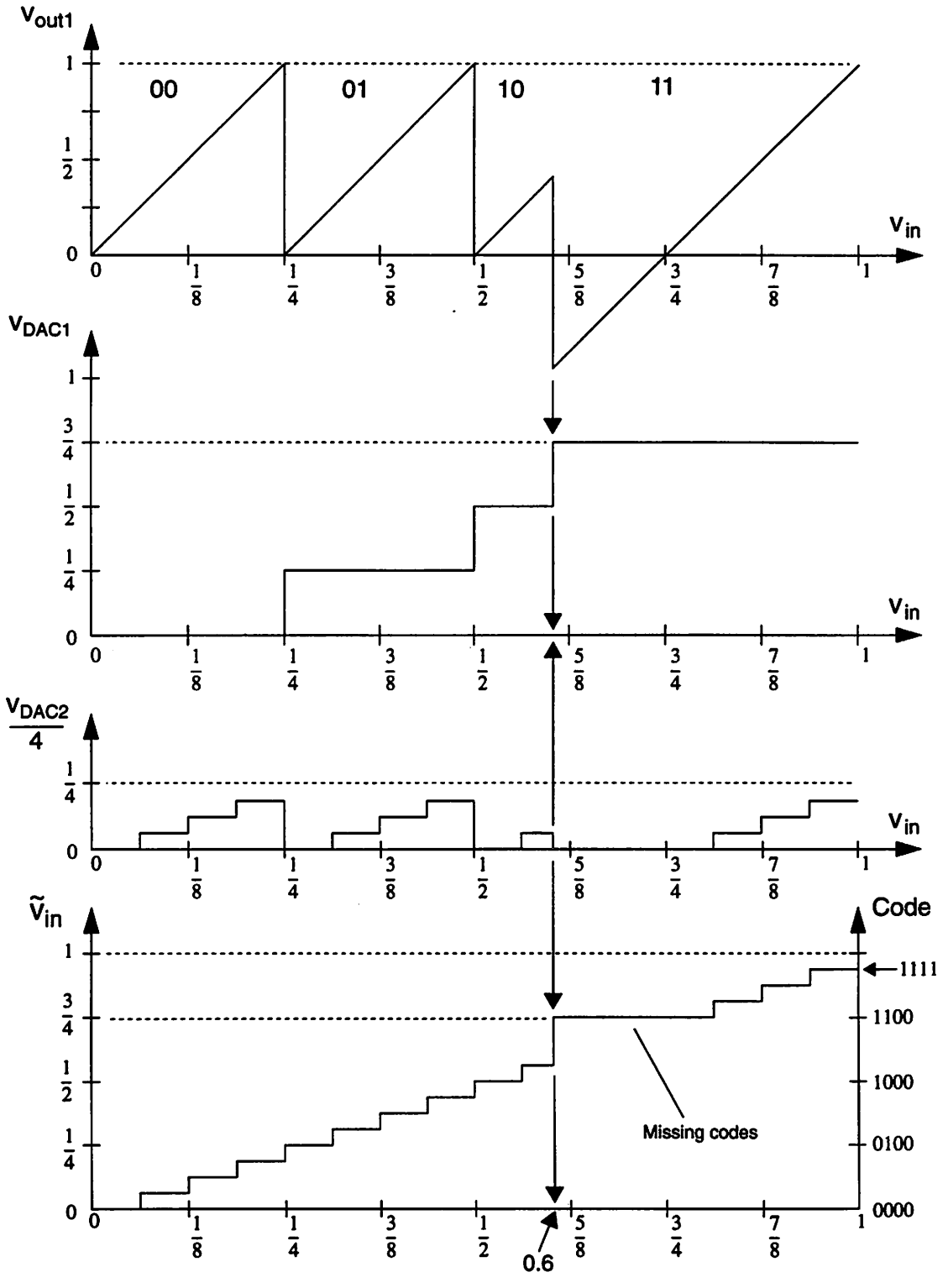


Fig. 3.12 Transfer curves of two-stage ADC with threshold level at 0.75 in stage 1 shifted to 0.6.

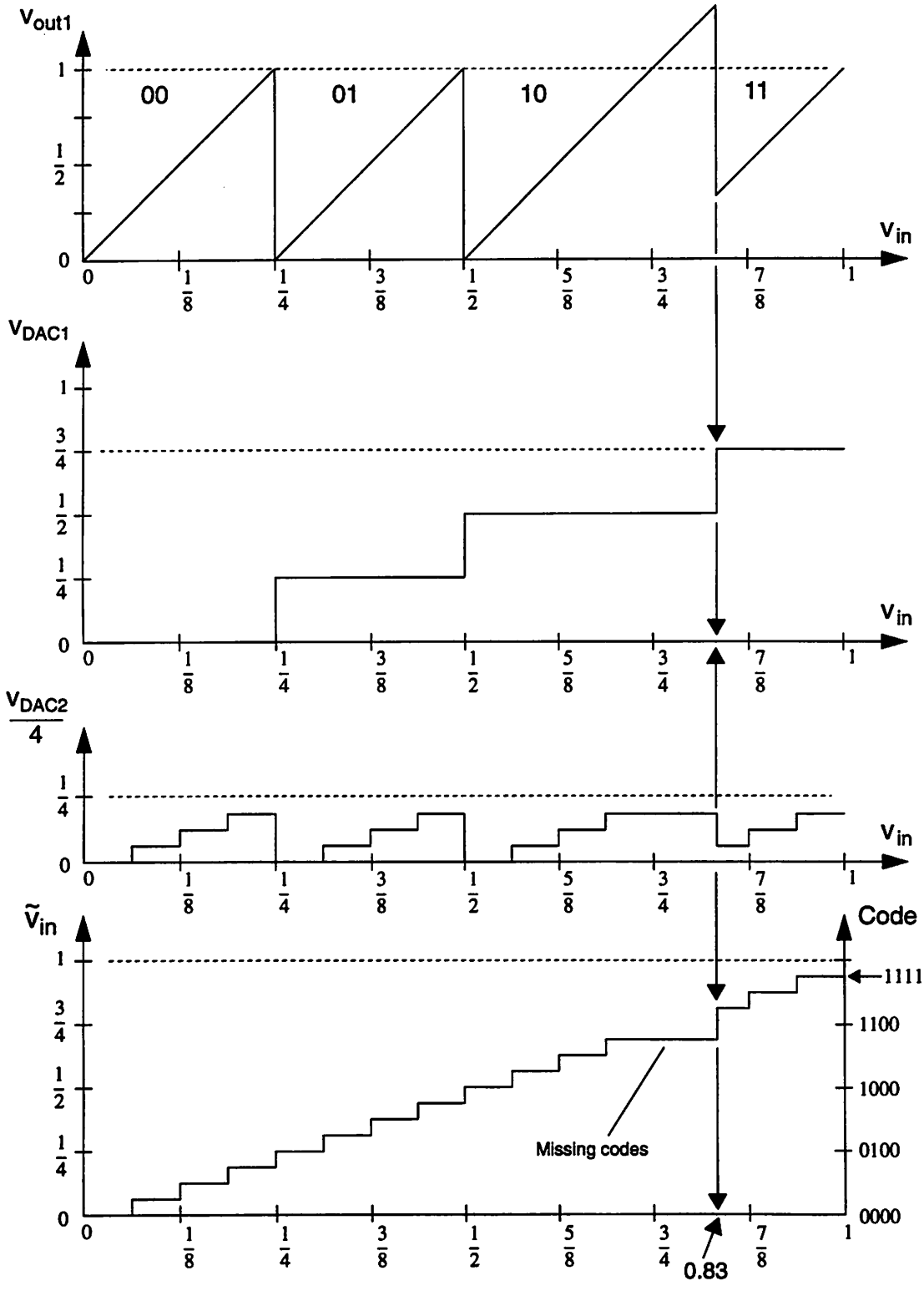


Fig. 3.13 Transfer curves of two-stage ADC with threshold level at 0.75 in stage 1 shifted to 0.83.

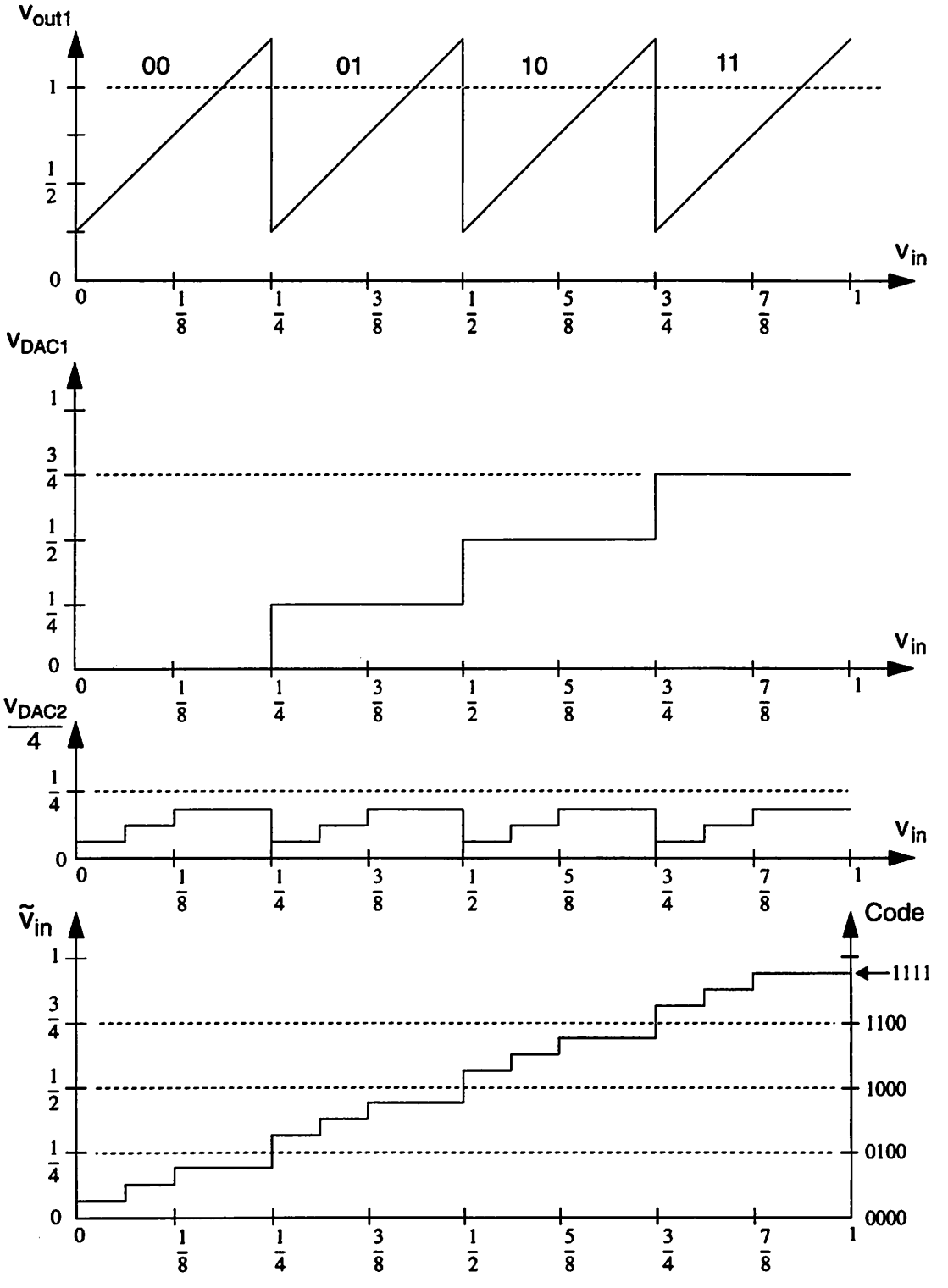


Fig. 3.14 Transfer curves of two-stage ADC with offset of +0.25 at output of stage 1.

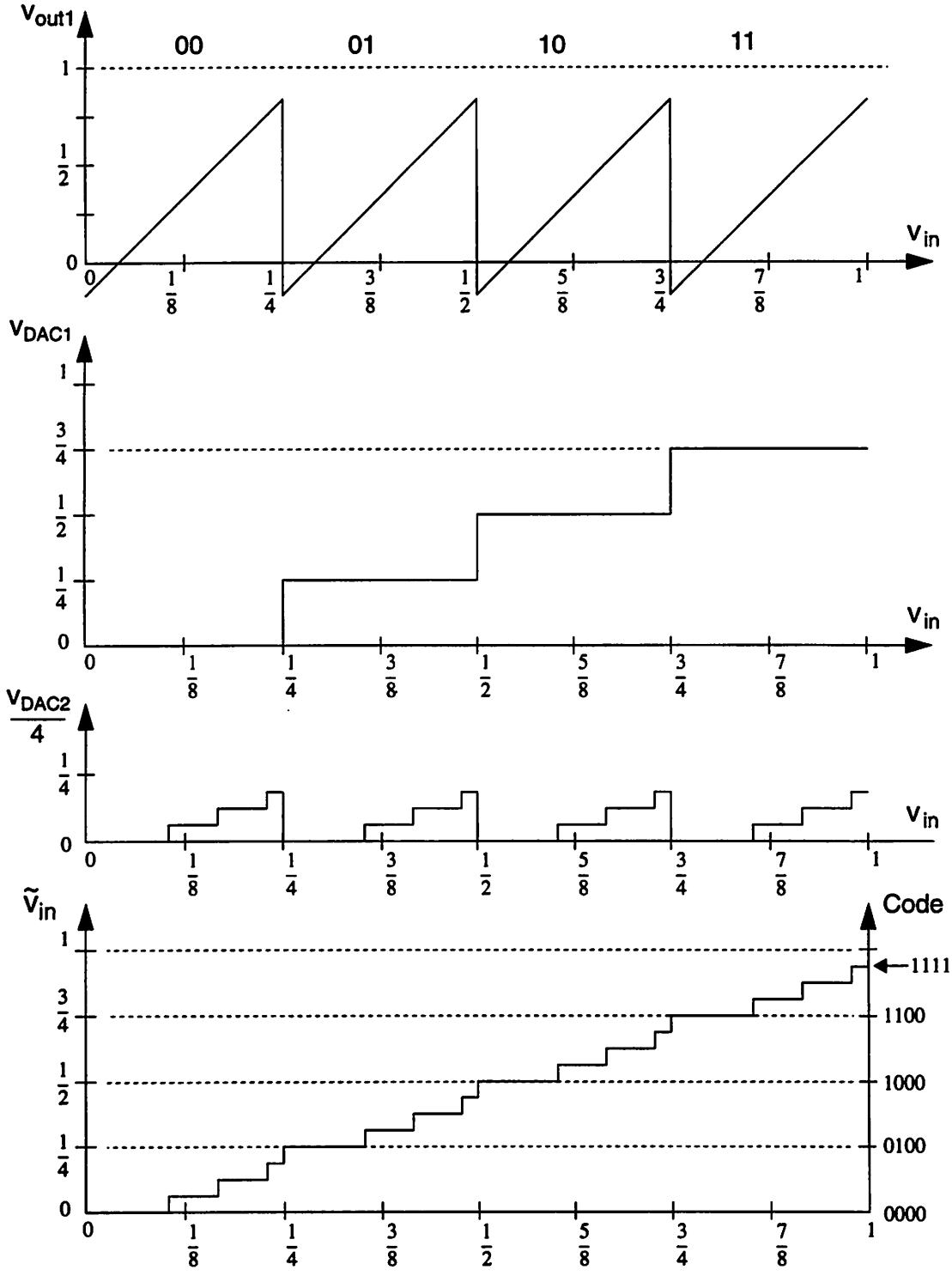


Fig. 3.15 Transfer curves of two-stage ADC with offset of –0.125 at output of stage 1.

Note that an offset in the subtractor block results in a repeated pattern, whereas an offset at a particular threshold causes errors only in the neighborhood of that threshold. Clearly, from the preceding discussion and examples, either sufficiently accurate decision levels must be generated, or else some method to detect and correct overranges and underranges is required. By *sufficiently accurate* it is apparent that the accuracy of the decision levels is required to be *compatible with the overall resolution of the portion of the ADC remaining.*

3.2.1 Solution Strategy #1: Extended Range

Conceptually, there are two ways to alleviate the problem of generating accurate decision threshold levels: (a) expand the range of the following stage in order to accommodate some amount of overrange due to threshold offsets, or (b) decrease the interstage gain, and so use less of the range under normal (ideal) conditions, but be able to guarantee that nonideal conditions (comparator offsets present) do not result in overrange. From the point of view of the ADC algorithm and coding these two schemes are often equivalent — as discussed later. Both approaches — expanding the range and decreasing the interstage gain — are referred to in the literature by the terms *digital error correction, overrange detection,* and *coding redundancy,* or simply *redundancy.* The appropriateness of the word *redundancy* will become clearer later. Essentially it entails the use of extra codes, which are introduced at each stage of the multistage A/D conversion process, and which are in addition to the minimum needed if all the decision levels are ideal and error-free. This section describes the use of extended range; Sections 3.2.2 and 3.2.3 discuss some approaches using reduced interstage gain.

Consider the example of a two-stage, 6-bit ADC with interstage gain of 8 shown in Fig. 3.16 below. In stage 1, the threshold level corresponding to the code transition 101→110, which is ideally at 0.5, actually occurs at 0.4, as shown. The decision levels and corresponding codes for stage 2 are shown on the right-hand side of figure. In the conventional configuration using no overrange detection, the input range of the second stage is from –1 to +1, and its output codes go from 000 to 111. However, if the range of the second stage is expanded by a factor of two, as shown, then it can accommodate amplified residues from the first stage that vary from –2 to +2, and can generate *additional codes* (i) above 111 and (ii) below 000, that is, negative codes, as indicated. The detailed behavior is now examined in three particular cases: (a) with no errors in the first stage, (b) with the indicated threshold offset in stage 1, but *without* overrange detection, and (c) with the first stage offset and *with* overrange detection.

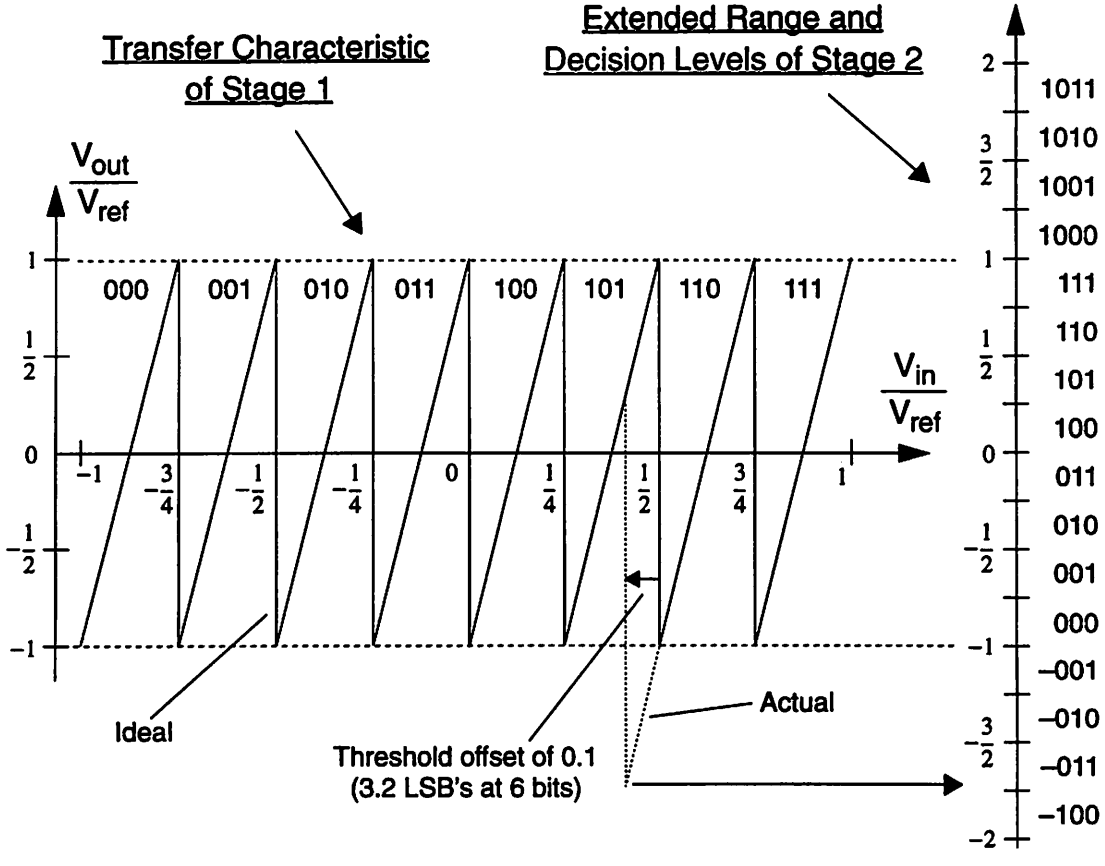


Fig. 3.16 Extended range and additional codes in stage 2 to deal with threshold offsets in stage 1.

Referring to Fig. 3.16, the behavior in these three cases is detailed below with input voltage $V_{in} = 0.41V_{ref}$.

Case (a): ideal thresholds.

Output code from stage 1: $D_1 = 5 = 101$

DAC output from stage 1: $V_{DAC,5} = 0.375V_{ref}$

Gained-up residue passed to stage 2: $V_{out,1} = V_{in,2} = 8\,(0.41 - 0.375)\,V_{ref} = 0.28V_{ref}$

Output code from stage 2: $D_2 = 5 = 101$

Overall output from the two-stage ADC is obtained from the addition:

$$\begin{array}{r} 101 \\ \quad 101 \\ \hline 101101 \end{array}$$

i.e., the final output code is 101101.

Case (b): first stage threshold at 0.5 shifted to 0.4.

Output code from stage 1: $D_1 = 6 = 110$

DAC output from stage 1: $V_{DAC,5} = 0.625V_{ref}$

Gained-up residue passed to stage 2: $V_{out,1} = V_{in,2} = 8\,(0.41 - 0.625)\,V_{ref} = -1.72V_{ref}$

Output code from stage 2: $D_2 = 0 = 000$

Overall output from the two-stage ADC is obtained from the addition:

$$\begin{array}{r} 110 \\ \quad 000 \\ \hline 110000 \end{array}$$

i.e., the final output code is 110000.

Case (c): first stage threshold ideally at 0.5 shifted to 0.4, extended range in second stage.

Output code from stage 1: $D_1 = 6 = 110$

DAC output from stage 1: $V_{DAC,5} = 0.625V_{ref}$

Gained-up residue passed to stage 2: $V_{out,1} = V_{in,2} = 8\,(0.41 - 0.625)\,V_{ref} = -1.72V_{ref}$

Output code from stage 2: $D_2 = 0 = -011$

Overall output from the two-stage ADC is obtained from the addition:

$$\begin{array}{r} 110 \\ \quad -011 \\ \hline 101101 \end{array}$$

i.e., the final output code is 101101.

It may be seen that Case (c), which has a large threshold offset but uses an extended range in the second stage, yields the same correct result (101101) as does Case (a), in which ideal decision levels are assumed. Not surprisingly, the result in Case (b) is significantly in error: 110000 instead of 101101.

An alternative second stage coding scheme and digital correction algorithm [129] is shown below in Fig. 3.17, on the right. For comparison, the scheme of Fig. 3.16, which was employed in the above numerical examples, is given on the left-hand side. It is apparent that the two methods are equivalent; the only difference is that the one on the right below uses a twos complement representation for levels below the nominal range.

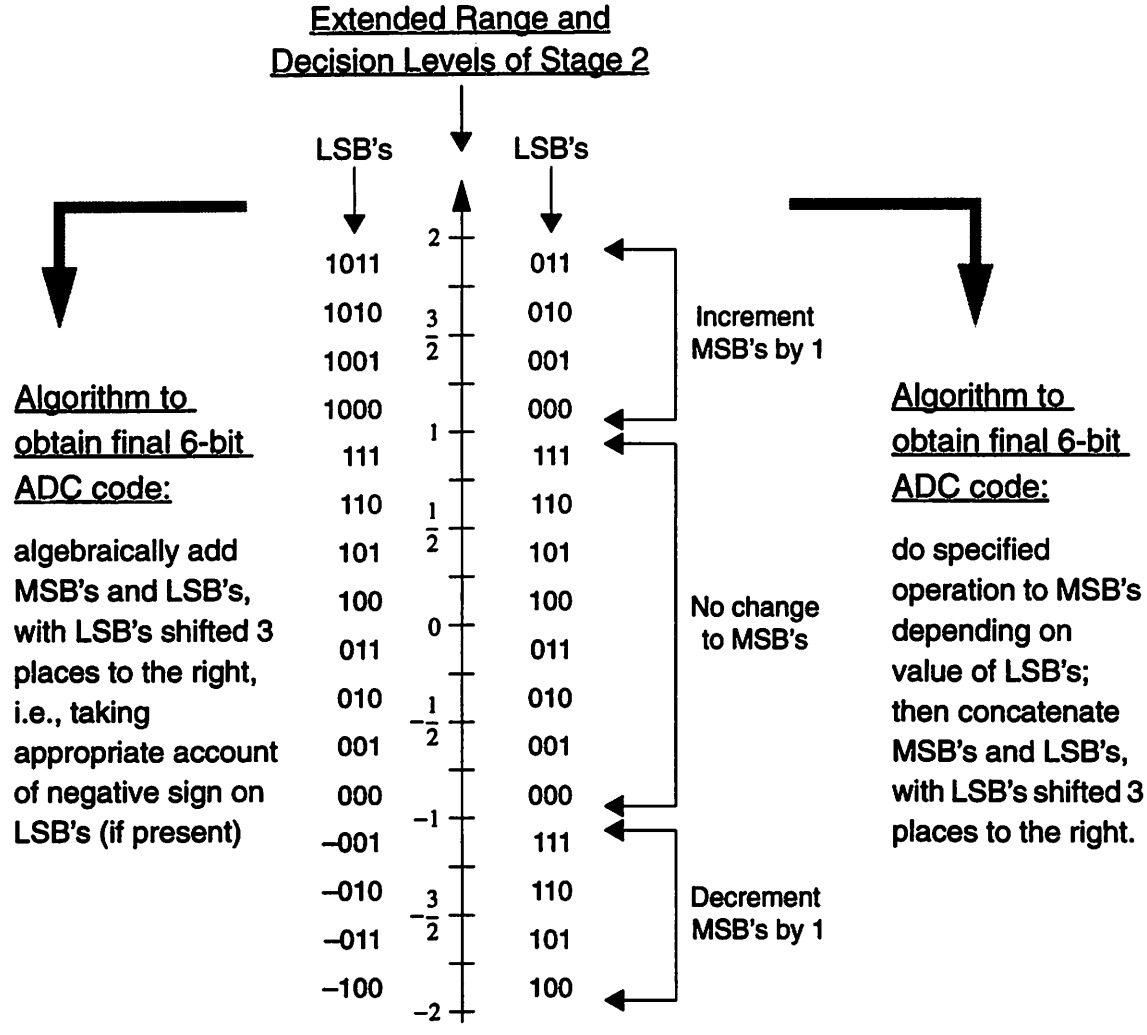


Fig. 3.17 Alternative ways to implement overrange detection algorithms.

The following points are pertinent.

Note 1: Extending the range of the second stage relaxes the accuracy requirement on the decision levels: the decision levels are no longer required to have a precision compatible with the resolution of the entire multistage ADC.

Note 2: From Fig. 3.16, it may be seen that overrange detection involves an addition if the gained-up residue is too big, and a subtraction if it is too small; that is, the codes used in the portion of the extended range *above* the nominal range are more *positive* than the nominal codes, and the codes are more *negative* in the region *below* the nominal range respectively. For example, Case (c) above involved a negative code and consequently a subtraction operation. The alternative scheme shown on the right-hand side of Fig. 3.17 constitutes a somewhat different way to implement the subtraction, but implicitly does also use negative codes. Another overrange detection technique that leads to a simpler implementation that *avoids the necessity for subtractions* is described in Section 3.2.3.

Note 3: An extended-range scheme similar to that shown on the right-hand side of Fig. 3.17 has been applied in a 10-bit recycling two-step flash architecture with an interstage gain of 32 [129]. However, for implementation reasons discussed in more detail later, in multistage ADC's *using an interstage gain greater than 1,* it is more common to keep the signal range constant and decrease the gain. This approach obviates the requirement that the analog circuits operate over an increased range, which is usually difficult to achieve at a constant power supply.

Note 4: The approach of expanding the range of the subsequent stage is very commonly used, however, in two-step flash and subranging architectures *with no interstage gain.* With respect to dc input-output relationships, subranging ADC architectures may be viewed as multistage topologies employing an interstage gain of 1, and so the nominal signal swing is already greatly reduced after the first stage. Consequently, expanding the range in stages after the first stage is quite feasible.

An example of the approach of expanding the range of the subsequent stage in a two-step *subranging* architecture is illustrated in Fig. 3.18 below.

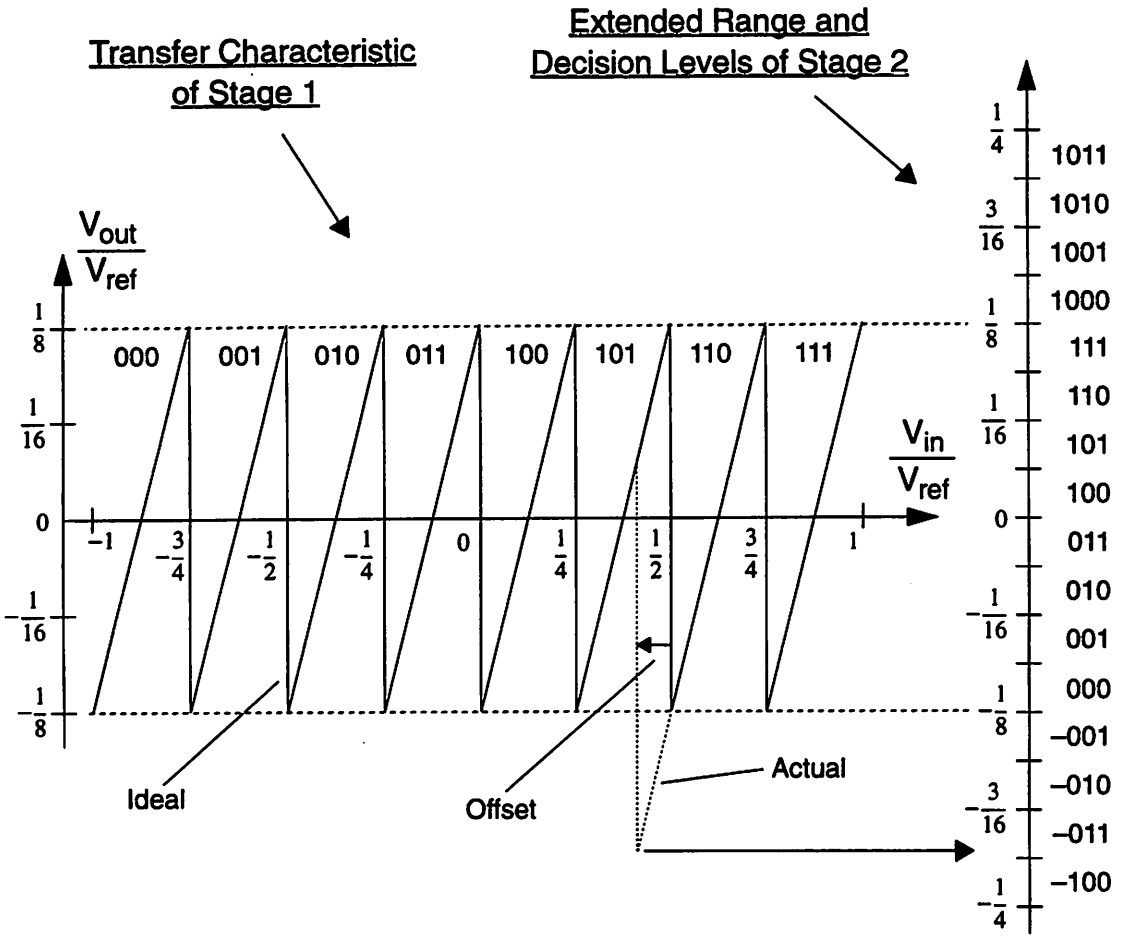


Fig. 3.18 Overrange detection by using a doubling of the range of the next stage — here as applied to a multistage ADC with 3 bits resolved per stage and an interstage gain of 1. As before, the ideal input-output relationship is shown with continuous lines, and the nonideal characteristic, which includes the effect of the threshold offset, has dotted lines wherever it deviates from ideality. Schemes such as this are typically used in two-step subranging architectures.

Here, the following stage range is expanded by a factor of two; therefore, a full extra bit or a factor of two of redundancy is used. The nominal dc parameters of the stage are identical to those listed for the example of Fig. 3.4, with the exception that the interstage gain is $G = 1$.

In other cases, it is known — for example, from process characterization data for offset variation, or from test results from previously-fabricated chips — that a few extra comparators on either side of the nominal range are sufficient to take account of the worst case threshold offset. Shown in Fig. 3.19 is a scheme employing two (rather than four) additional decision levels above and below the nominal range in the second stage. Note that in the implementation of two-step subranging ADC's it is usually straightforward and cheap to add extra threshold levels by simply tapping off from the main resistor string.

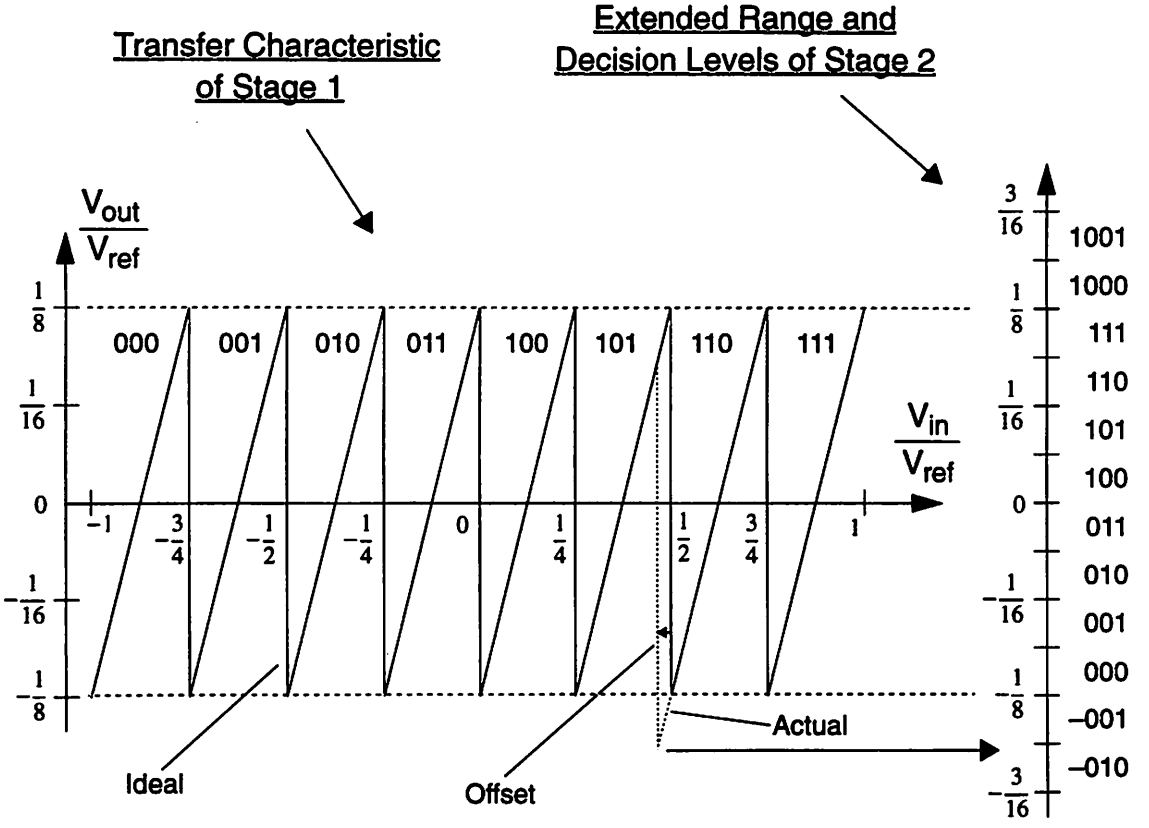


Fig. 3.19 Use of two extra comparators above and below the nominal range for overrange and underrange detection. Approaches similar to this are widely used in two-step flash/subranging architectures.

Schemes similar to that shown in Fig. 3.19 have been employed in 8-bit two-step subranging architectures reported by Sony [30], [31] and Hitachi [86], [87]; typically, 1–3 additional comparators are used on each side of the range. Some of these architecture are discussed from an A/D converter topology perspective in Section 7.1.2.

3.2.2 Solution Strategy #2: Reduced Interstage Gain, "Naive" Coding

As mentioned earlier, conceptually it is feasible to deal with the overrange/underrange problem by decreasing the nominal interstage gain. However, the coding becomes complicated and non-obvious. As an example, consider the stage dc transfer characteristic shown in Fig. 3.20 below. Here, the interstage gain has been reduced from 4 to 2. As may be seen, under nominal conditions the amplified residue output V_{out} lies in the range $[-1/2V_{ref}, +1/2V_{ref}]$, and even with threshold and subtractor offsets of magnitude up to $V_{ref}/4$, V_{out} lies within $[-V_{ref}, V_{ref}]$, which is the range of the next stage. However, it is not immediately obvious how to assign the digital codes. Since the interstage gain has been reduced by two, the *significance* or *weight* of the next stage digital outputs has been increased by two, corresponding to a left shift by one bit. Therefore, it seems plausible that the codes from adjacent stages should be combined additively *but with one bit overlapped.*

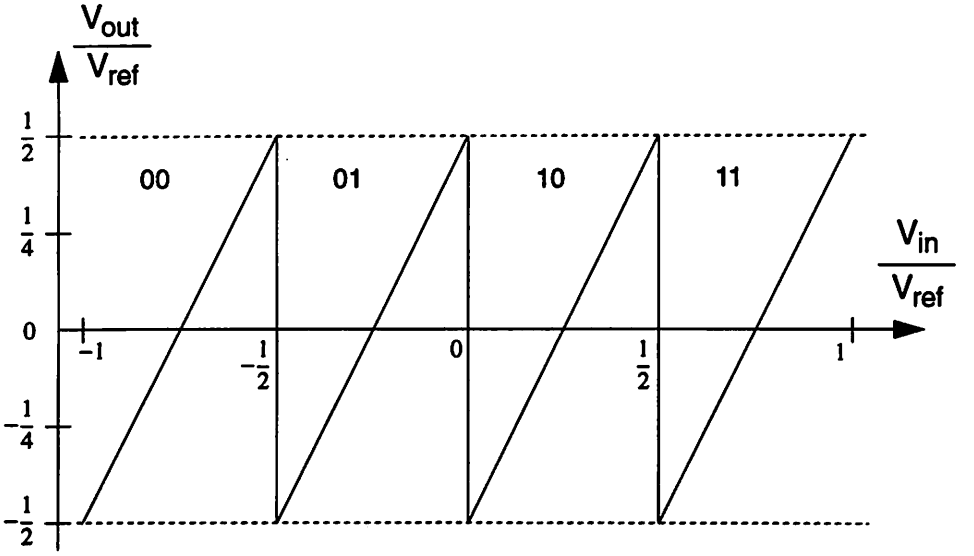


Fig. 3.20 Transfer curve of a 2-bit stage with reduced interstage gain and first attempt at coding.

One possible coding assignment is shown in Fig. 3.20. This does not give the usual or "expected" digital output code, however, as may be seen from the following example. Consider a multistage ADC composed of a cascade of 3 of these stages. Since one bit of redundancy is used in stages 1 and 2, the overall resolution is 4 bits. Let the input voltage,

output code, and DAC voltage from each stage i be denoted by V_i, D_i, and $V_{DAC,i}$ respectively. Suppose the input is $V_{in} = V_1 = +\varepsilon$, where ε is a positive voltage of magnitude much smaller than 1 LSB at the overall ADC resolution. (For convenience, in the following computations the V_{ref} scaling constant is omitted.) The outputs from the three stages are as follows:

$$D_1 = 2 = 10 \Rightarrow V_{DAC,1} = 0.25 \Rightarrow V_2 = 2(\varepsilon - 0.25) = -0.5 + 2\varepsilon,$$

$$D_2 = 1 = 01 \Rightarrow V_{DAC,2} = -0.25 \Rightarrow V_3 = 2[(-0.5 + 2\varepsilon) - (-0.25)] = -0.5 + 4\varepsilon,$$

$$D_3 = 1 = 01.$$

Since the overall analog input range is from $-V_{ref}$ to $+V_{ref}$, the ADC digital output is expected to be within 1 LSB of the mid-scale code, i.e., 1000. Now, using the scheme suggested above, the overall output from the 3-stage ADC is obtained from the addition

$$\begin{array}{r} 10 \\ 01 \\ 01 \\ \hline 1011 \end{array}$$

which is clearly not what would be expected.

A second possible coding assignment is shown in Fig. 3.21 below.

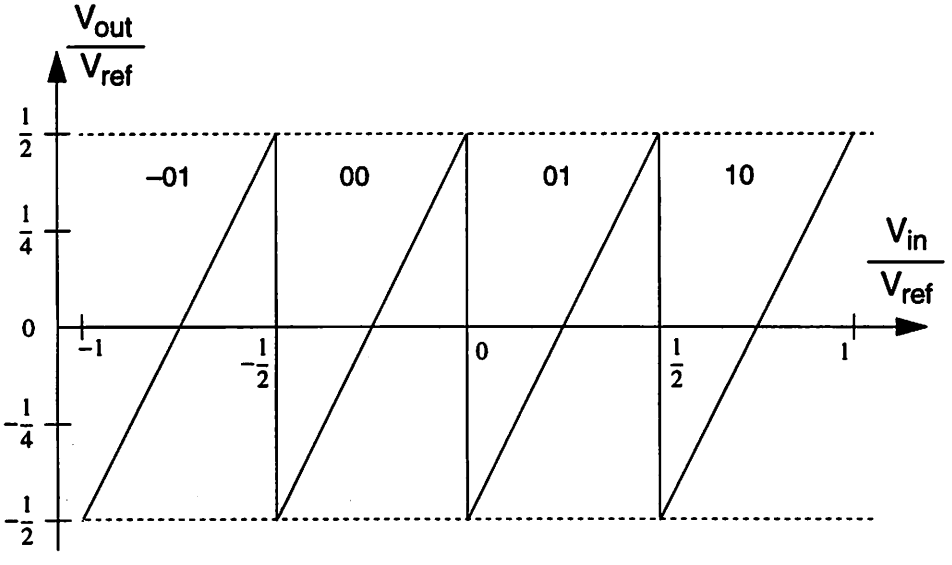


Fig. 3.21 Transfer curve of a 2-bit stage with reduced interstage gain and another possible code assignment scheme.

However, this coding does not give the expected answer either. Consider a multistage ADC composed of a cascade of 3 of these stages. (Again, for convenience, the V_{ref} scaling constant is omitted from the calculations.) As in the previous case, let V_i, D_i, and $V_{DAC,i}$ be the input voltage, output code, and DAC output voltage from each stage i respectively. Again, suppose the input is $V_{in} = V_1 = +\varepsilon$. The resulting stage outputs are

$$D_1 = 2 = 01 \Rightarrow V_{DAC,1} = 0.25 \Rightarrow V_2 = 2(\varepsilon - 0.25) = -0.5 + 2\varepsilon,$$

$$D_2 = 1 = 00 \Rightarrow V_{DAC,2} = -0.25 \Rightarrow V_3 = 2[(-0.5 + 2\varepsilon) - (-0.25)] = -0.5 + 4\varepsilon,$$

$$D_3 = 1 = 00.$$

The overall output from the 3-stage ADC in this case is obtained from the addition

$$\begin{array}{r} 01 \\ 00 \\ \quad 00 \\ \hline 0100 \end{array}$$

Once again, the result is not what would be expected from a conventional ADC.

It is apparent that both of the above combinations of threshold levels, codes, and DAC levels result in a large offset in the output digital codes. However, it somehow seems intuitive that since this scheme has allowed for threshold errors by reducing the interstage gain by a factor of 2, there should be some way to make it work. *Well, stay tuned!* In the next section a solution is presented, and later — in Section 4.2 — it is shown that the above scheme *can* be made to work and how.

3.2.3 Solution Strategy #3: Reduced Interstage Gain, Modified Coding

The fact that the above "naive" scheme results in an offset in the ADC output code seems to suggest the idea of introducing a deliberate offset in the transfer characteristic. An elegant solution to this issue, which resolves the complications mentioned above, is to shift the sub-ADC and sub-DAC levels up by 1/2 LSB of the stage. This approach — referred to henceforth as *reduced interstage gain with modified coding* — has been

described in [43] and [140], and is discussed in detail in the specific context of CMOS pipelined multistage ADC's in [73]. For a 2-bit case, shifting the ADC and DAC levels by 1/2 LSB relative to their values in Fig. 3.20 results in the transfer characteristic shown below in Fig. 3.22.

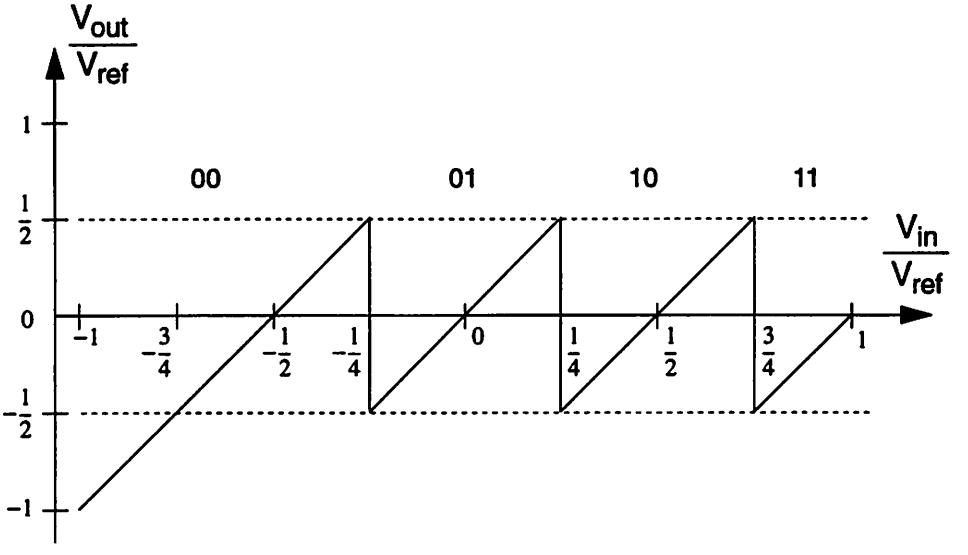


Fig. 3.22 Static transfer characteristic from stage input to amplified residue output of 2-bit stage with interstage gain reduced from 4 to 2, and *shifted* threshold and DAC levels.

The complete dc parameters of the stage are listed below.

Input range:	[–1, +1] (normalized to V_{ref})
3 ADC threshold levels:	{–1/4, 1/4, 3/4}
4 digital codes:	{0, 1, 2, 3} or
	{00, 01, 10, 11}
4 corresponding DAC levels:	{–1/2, 0, 1/2, 1}
Interstage gain:	G = 2
Number of bits:	k = 2 (net, after digital correction: k = 1)

Note that the term "number of bits" must be used carefully when discussing stages incorporating redundancy: it should be clearly stated whether the number of bits being referred to is the "raw" uncorrected number, or the *net* stage resolution, i.e., after coding redundancy has been removed. Henceforth in this chapter, wherever appropriate — in particular, when listing the dc parameters of stages — both values will be explicitly given.

The input-output characteristic with reduced gain and modified coding for a 3-bit stage is shown below.

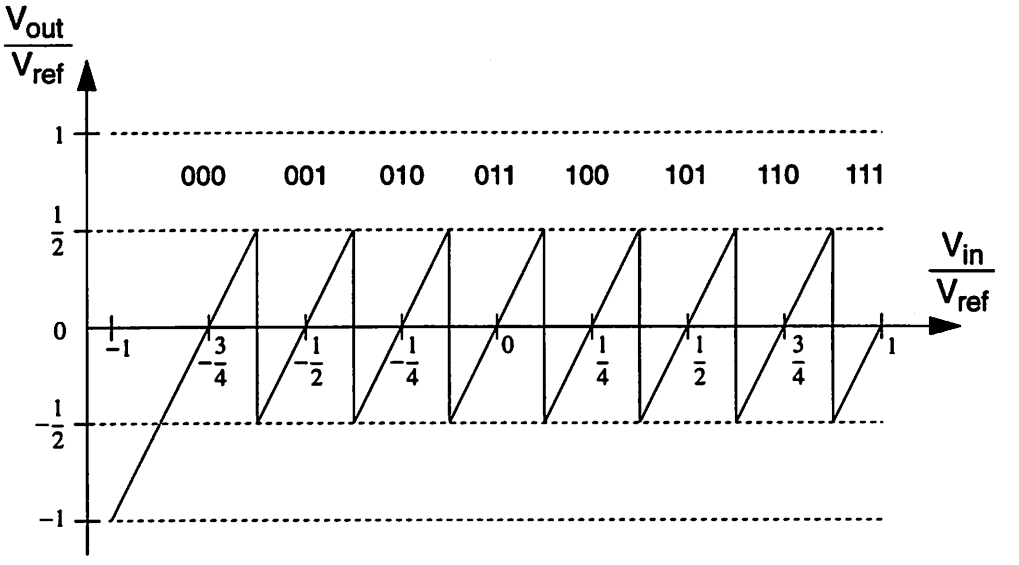


Fig. 3.23 Transfer characteristic from stage input to gained-up residue of a 3-bit pipeline stage with shifted ADC threshold levels and DAC levels and with an interstage gain of 4 [73].

The dc parameters of the stage are:

Input range:	[–1, +1] (normalized to V_{ref})
7 ADC threshold levels:	{–5/8, –3/8, –1/8, 1/8, 3/8, 5/8, 7/8}
8 digital codes:	{0, 1, 2, 3, 4, 5, 6, 7} or
	{000, 001, 010, 011, 100, 101, 110, 111}
8 corresponding DAC levels:	{–3/4, –1/2, –1/4, 0, 1/4, 1/2, 3/4, 1}
Interstage gain:	G = 4
Number of bits:	k = 3 (net, after digital correction: k = 2)

Using the dc input-output relation and coding scheme indicated on Fig. 3.23, the digital error correction logic becomes very simple: the 3-bit digital outputs from each stage are simply added together with one bit overlapped between adjacent stages. For example — assuming an ADC consisting of 4 stages — if the digital outputs for stages 1–4 are 001, 100, 101, and 111 respectively, then the overall output is computed by means of the addition shown below.

```
  001
  100
  101
  111
---------
010011011
```

Note that this coding scheme also avoids subtraction — i.e., only positive numbers are added to form the final ADC output code.

This technique for digital error correction / overrange detection has been employed in some practical implementations of CMOS pipelined multistage ADC's [71], [73], [80]. In these architectures, each pipeline stage except the last has one bit of redundancy (since the last stage cannot be corrected.) One further advantage of this technique, as compared with the *naive coding* approach discussed earlier, is that it automatically yields "nice" output codes without any unexpected digital offset; specifically, when the input is near ±Full Scale, the output code is near all zeros or all ones, as appropriate, and when the input is at midscale, the output code is 011111..., etc.

At this point in the discussion, the following question arises: why does this "modified coding" scheme work? Why is it that the configurations of ADC threshold levels, code assignment, and DAC levels shown in Fig. 3.22 and Fig. 3.23 require a simple overlapped addition to compute the overall ADC output code, as described above, whereas the earlier "naive coding" schemes — e.g., as depicted in Fig. 3.20 and Fig. 3.21 — do not work as "nicely"? This question is answered later in Section 4.2 in the context of a generalized analysis of multistage A/D conversion schemes.

Finally, for the purposes of comparison, it is instructive to view side by side the three approaches: (a) the standard/classical stage input-output transfer characteristic with no redundancy, (b) the scheme with gain reduced by two and "naive" coding, and (c) the scheme with gain reduced by two and using the modified scheme mentioned above. The transfer curves for each of the three cases for a 3-bit stage are given in Fig. 3.24.

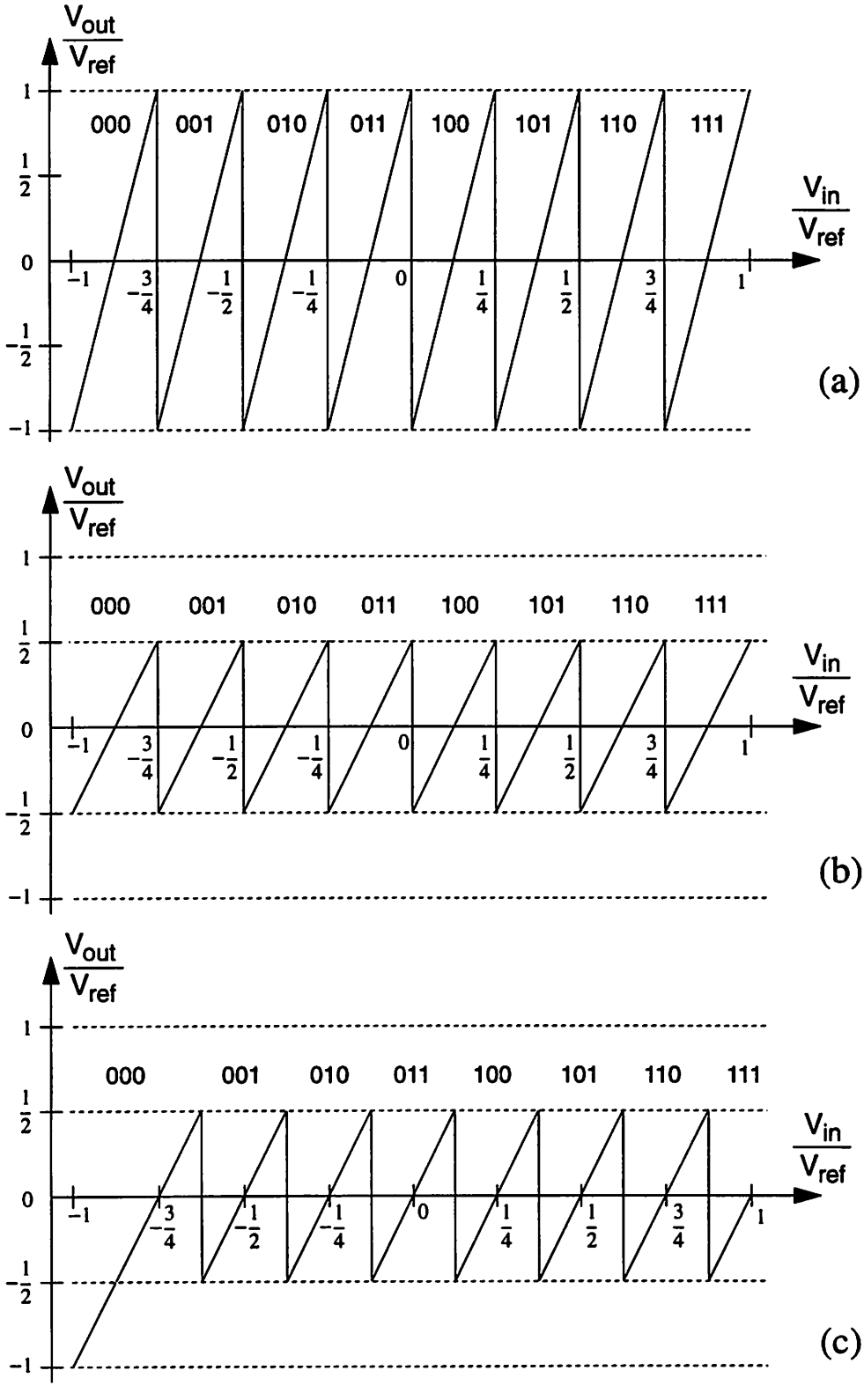


Fig. 3.24 Transfer characteristics from stage input to amplified residue of three 3-bit stages.

Summary of Digital Correction Using Reduced Interstage Gain with Modified Coding: At this point, it is appropriate to reiterate clearly the motivation for this entire discussion: *in multistage A/D converters, reduced interstage gain permits large decision level offsets and so leads to a more robust implementation, and avoids the cost of generating precise threshold levels.* The 3-bit stage of Fig. 3.23 is repeated in Fig. 3.25 below and also shown on the plot are two examples of the maximum allowable threshold offset. Note that these maximum permissible offsets are defined by the requirement that the stage output must be within the range of the next stage. *The accuracy of the decision levels is therefore independent of the overall accuracy of the A/D converter.*

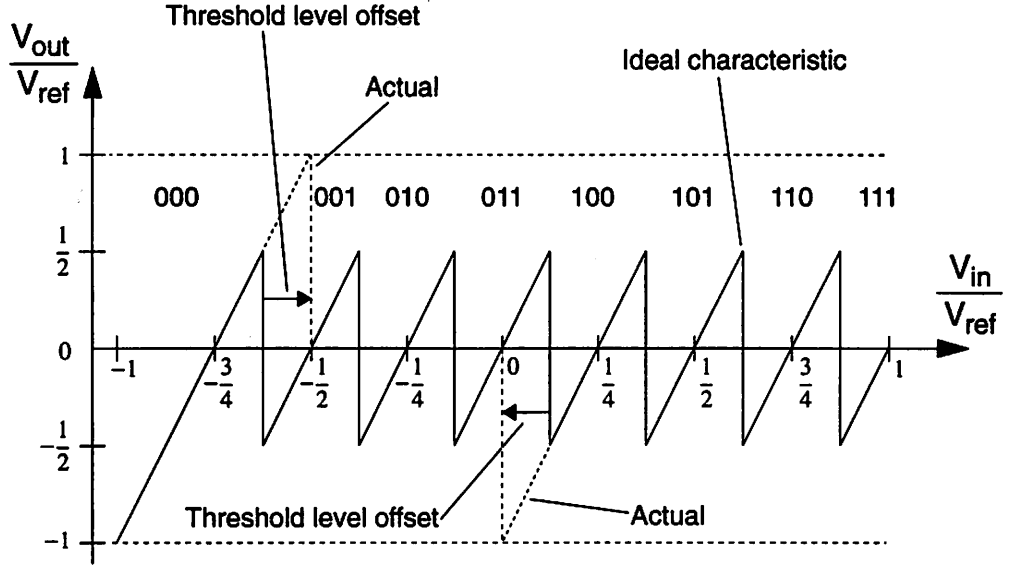


Fig. 3.25 Static transfer characteristic from stage input to gained-up residue output of a 3-bit pipeline stage with shifted ADC threshold levels and DAC levels and with an interstage gain of 4 [73]. Also indicated on the graph, in dotted lines, is the actual input-output relationship in the presence of the *largest* allowable threshold level offset, i.e., $V_{ref}/8$.

It is apparent from Fig. 3.25 that even if the comparators have offsets of up to $\pm V_{ref}/8$ the stage output is still within $\pm V_{ref}$ — i.e., within the range of the next stage — and so the final ADC output code is still correct (assuming ideal DAC levels and ideal interstage gains). Note that the stage decision levels closest to zero are $\pm V_{ref}/8$, which implies that *offsets are tolerated up to the smallest magnitude thresholds of an individual stage.*

In general, given a stage input-output transfer characteristic, in order to ascertain how much threshold and subtractor offset can be tolerated, the following question should be asked: how much can each decision level shift and still maintain the amplified residue within the range of the next stage? Usually this is immediately apparent from a plot of the input-output relation by determining how much any threshold may shift to the left or to the right in order to cause the amplified residue to equal the minimum or maximum values respectively of the next stage range.

Finally, it is appropriate to mention the *cost* of redundancy and digital error correction. The most obvious one is that since the effective or net resolution per stage is reduced, more stages are required for a given total A/D converter resolution, which results in the following two problems.

1. In applications where latency is important the ADC may be required to satisfy some specified latency constraint; this may limit the number of pipelined stages, since in a pipelined implementation, latency is directly proportional to the number of stages. Alternatively, of course, this issue can be addressed by modifying some aspect of the overall system architecture or application in order to make the latency less stringent.
2. To first order, more stages requires more hardware and leads to increased silicon area and power dissipation. However, this issue can be addressed by means of scaling down the sizes of the later stages, and is discussed further in Chapter 8.

3.2.4 Multistage A/D Conversion: Some Interpretations

What is really happening when digital error correction is used in a multistage ADC? The following are two interpretations.

Multistage A/D Conversion as a Multistep Search Procedure: In general, A/D conversion can be envisioned as a search operation, and multistage A/D conversion can be thought of as performing that search in multiple steps — i.e., a sequence of *zooms*. The interstage gain corresponds loosely to a "zoom factor". When redundancy is employed, the amount zoomed is less, and the final decision is postponed until the analog input signal can be "seen" better with respect to the ADC threshold levels. An attempt is made to represent this schematically in Fig. 3.26.

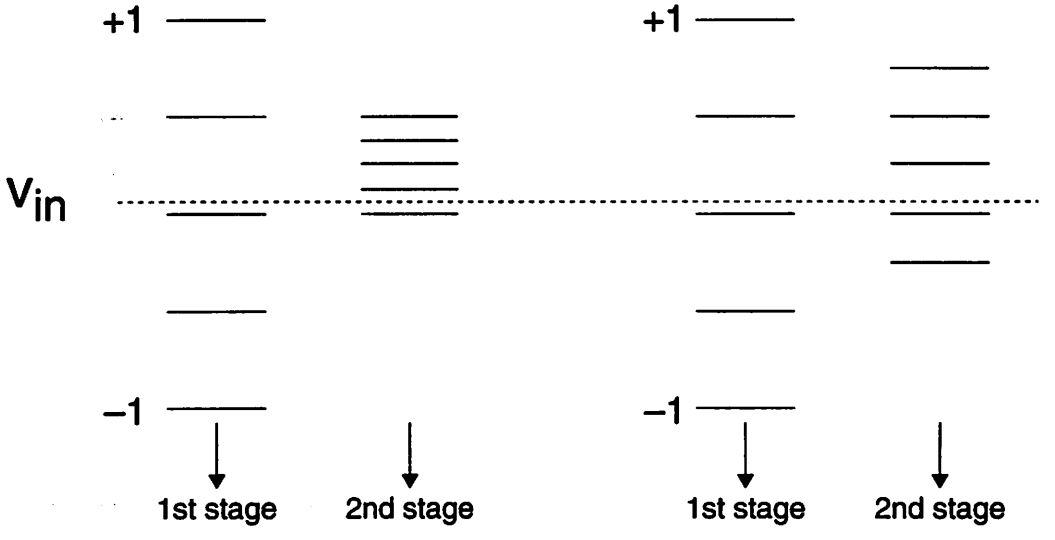


Fig. 3.26 Two views of multistage A/D conversion: on the left, the standard way, and on the right, with a reduced interstage gain.

On the left is shown the standard way: the first stage makes a 2-bit decision; this is followed by a "zoom-in" by a factor of $2^2 = 4$. The range of the second stage exactly matches one LSB of the first stage with no margin for error. An offset in the first stage has the effect of moving the second stage decision levels either above or below V_{in}: this corresponds to an out-of-range condition. In the diagram on the right, on the other hand, the "zoom-in" factor is reduced from 4 to 2, and the range of the second stage covers not

only one LSB of the first stage, but extends above it and below it also; that is, there is some margin or tolerance for an upward or downward shift in the first stage.

Multistage A/D Conversion as a Decomposition into Weights: Another way to think about multistage A/D conversion is as a decomposition of an input signal v_{in} into a linear combination of N_S *weights,* where N_S is the number of stages. Mathematically, this may be expressed as

$$v_{in} = \sum_{i=1}^{N_S} w_i + \varepsilon \tag{3.0}$$

where N_S is the number of stages, w_i is the "weight" generated by stage i, and ε is the error — the quantization error. Omitting the error term gives $\tilde{v}_{in}$, the amplitude-quantized approximation of the input, i.e.,

$$\tilde{v}_{in} = \sum_{i=1}^{N_S} w_i \tag{3.1}$$

Each weight w_i is chosen from a set of possible values for stage i. The weights from the *successive* stages always satisfy

$$w_1 > w_2 > w_3 \ldots > w_{N_S} \tag{3.2}$$

which corresponds to the intuition that the first stage is the most significant and the later stages are less significant. Two particular ADC algorithms are now formulated in terms of (3.1).

First, classical multistage A/D conversion with straightforward coding is considered. In this case, the weights from different stages are orthogonal with respect to their binary representations — i.e., there is no redundancy. For example, in the case of a two-stage ADC with 3 decision levels and 4 weights per stage, the first stage weight has 4 possible values given by

$$w_1 \in \{0000, 0100, 1000, 1100\} \tag{3.3}$$

and the second stage weight has 4 possible values given by

$$w_2 \in \{0000, 0001, 0010, 0011\} \tag{3.4}$$

It may be seen that all possible pairs (w_1, w_2) are orthogonal in a bitwise sense — i.e., there are no "1" overlaps. The orthogonality can be thought of as being with respect to an inner product computed by first taking the logical AND of corresponding bits, and then summing the results.

Second, "digital error correction" with a factor of two of "redundancy" is considered. In this case, the weights $\{w_i\}$ from different stages are *not* orthogonal and may overlap — but *by at most one bit.* For example, in the case of a 5-bit two-stage ADC with 7 decision levels and 8 weights per stage, the first stage weight has 8 possible values given by

$$w_1 \in \{00000, 00100, 01000, 01100, 10000, 10100, 11000, 11100\} \tag{3.5}$$

and the second stage weight has 8 possible values given by

$$w_2 \in \{00000, 00001, 00010, 00011, 00100, 00101, 00110, 00111\} \tag{3.6}$$

In this case, note that, for example, the pair of weights $(w_1, w_2) = (10100, 00110)$ are not orthogonal since they have a common "1" overlap — in bit position 3.

Finally, it is apparent that a *generalized* decomposition should be possible with no particular restrictions on the $\{w_i\}$; however, in that case it would be required to measure and calibrate the weights [70], [57]. The preceding ideas of "decomposition" into "weights" and A/D conversion as an approximation problem motivate a more general approach to multistage A/D conversion that is pursued in Chapter 4.

3.3 EFFECTS OF GAIN AND DAC ERRORS

This section briefly reviews the effects on the dc transfer characteristics of a multistage ADC of errors in (i) the interstage gain block and (ii) the DAC levels within a stage. This topic is separated from the discussion of threshold offsets and the associated digital error correction techniques, since DAC errors and interstage gain errors result in effects fundamentally different — both qualitatively and quantitatively — from threshold offsets, and the digital correction techniques discussed in Section 3.2 do not apply. Exact analysis of the Integral Nonlinearity (INL) of a multistage ADC due to an individual DAC or gain error is complicated. Here, the approach is to compute the error in the amplified analog residue from a particular stage, and compare this error with the LSB size at that point in the multistage A/D system. Differential Nonlinearity (DNL) is somewhat more tractable, and exact conditions to guarantee no missing codes are derived.

3.3.0 Analysis of Interstage Gain Errors

The effects of nonideal interstage gain are considered (i) in simple multistage A/D conversion schemes and (ii) in schemes using reduced interstage gain and digital correction.

Stages Using Simple Coding With No Redundancy

The transfer characteristic from analog input V_{in} to amplified residue V_{out} of a simple 2-bit stage is shown in Fig. 3.27.

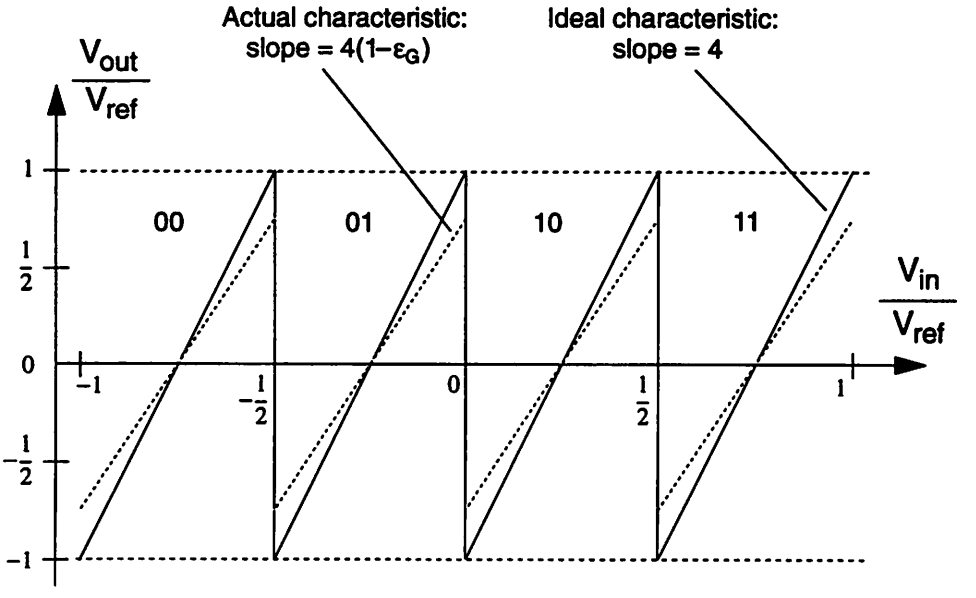


Fig. 3.27 Static input-output relationship of a 2-bit stage with nominal interstage gain of 4. The ideal transfer characteristic is shown with continuous lines; the nonideal curve resulting from a gain error has dotted lines.

The nominal (ideal) gain is 4; however, due to a gain error, the actual gain is given by $G = 4(1 - \varepsilon_G)$, and hence the actual input-output characteristic has slope $4(1 - \varepsilon_G)$, as indicated. The maximum value of the gained-up residue output is therefore reduced from V_{ref} to $V_{ref}(1 - \varepsilon_G)$.

In order to examine the effect of this gain error, consider what happens when V_{in} is negative but arbitrarily close to zero — i.e., $V_{in} = -\delta V$, where δV is a positive voltage of magnitude much smaller than 1 LSB at the overall ADC resolution. Referring to Fig. 3.27, it may be seen that the resulting digital code is $D = 01$, and the DAC value generated is $V_{DAC} = -(1/4)\,V_{ref}$.

The stage output is then given by

$$V_{out} = G(V_{in} - V_{DAC}) \tag{3.7}$$

$$= G(-\delta V - (-\tfrac{1}{4}V_{ref})) \tag{3.8}$$

i.e.,

$$V_{out} = G(V_{ref}/4) \tag{3.9}$$

since by assumption, δV is negligible in comparison to $V_{ref}/4$.

Next, substituting the expression for the actual gain G into (3.9) gives

$$V_{out} = 4(1-\varepsilon_G)(V_{ref}/4) \tag{3.10}$$

$$= (1-\varepsilon_G)V_{ref} \tag{3.11}$$

and thus there is an amplitude error equal to $\varepsilon_G V_{ref}$. The amplified residue V_{out} is now passed down the pipeline. Let the resolution *remaining* after this stage be n_2 bits. Then, if the error due to the nonideal gain in the first stage is required to be *less than x LSB's of the remaining resolution,* this implies that

$$\varepsilon_G V_{ref} < \frac{x}{2^{n_2}} \cdot (2V_{ref}) \tag{3.12}$$

which yields

$$\varepsilon_G < \frac{x}{2^{n_2-1}} \tag{3.13}$$

In order to cause a missing code, the amplitude error in the gained-up residue must be equal to 1 LSB — i.e., corresponding to $x = 1$. Therefore, to avoid missing codes, the following condition must hold:

$$\varepsilon_G < \frac{1}{2^{n_2-1}} \tag{3.14}$$

Note that, if a gain error is large enough to cause a missing code, then a missing code will occur every time that particular gain error is encountered in the transfer characteristic. For example, in the case of Fig. 3.27, the resulting error pattern will repeat at intervals of $(2V_{ref})/4 = V_{ref}/2$. This is one way to identify a gain error from DNL and INL plots.

Note that in a practical design, to ensure robustness and to allow for settling time errors and matching errors, the error contribution due to nonideal gain is typically required to be less than 1/4 LSB of the remaining resolution.

Stages Using Reduced Interstage Gain and Modified Coding

An example of the effect of nonideal gain in a stage employing redundancy is illustrated in Fig. 3.28.

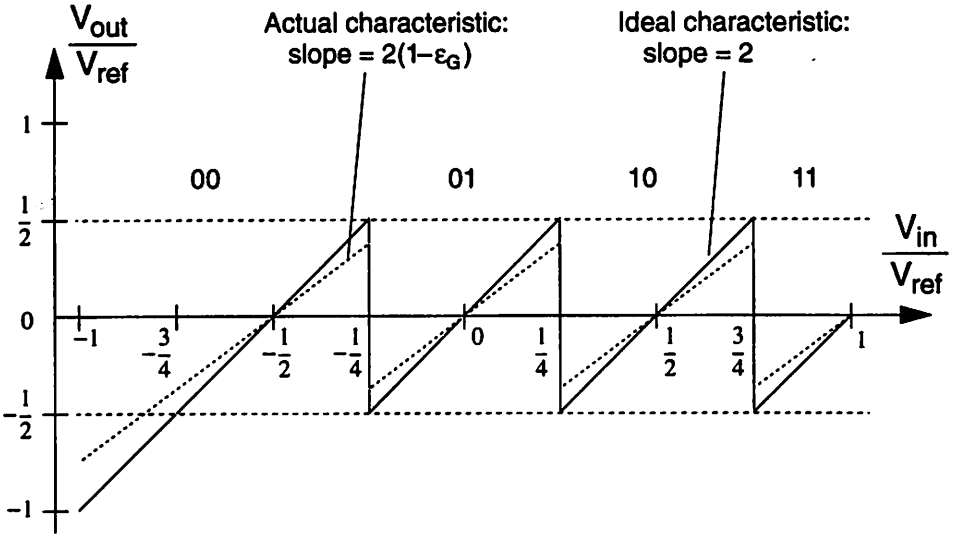


Fig. 3.28 Static input-output relationship of a 2-bit stage with nominal interstage gain of 2. The ideal transfer characteristic is shown with continuous lines; the nonideal curve resulting from a gain error has dotted lines.

Proceeding as before, in this case, the nominal (ideal) gain is 2, and the actual gain is given by $G = 2(1 - \varepsilon_G)$. In order to examine the effect of this gain error, consider what happens when the input is $V_{in} = -\delta V + (1/4)\, V_{ref}$ — i.e., less than but arbitrarily close to $1/4 V_{ref}$. From Fig. 3.28 it may be seen that the resulting digital code is $D = 01$, and the DAC value generated is $V_{DAC} = 0$. Therefore, the stage output is

$$V_{out} = G(V_{in} - V_{DAC}) \tag{3.15}$$

$$= G(\frac{1}{4}V_{ref} - \delta V) \tag{3.16}$$

i.e.,

$$V_{out} = G(V_{ref}/4) \tag{3.17}$$

Since $G = 2(1 - \varepsilon_G)$, this gives

$$V_{out} = 2(1 - \varepsilon_G)(V_{ref}/4) \tag{3.18}$$

$$= (1/2)(1 - \varepsilon_G)V_{ref} \tag{3.19}$$

i.e., an amplitude error equal to $(\varepsilon_G V_{ref})/2$.

Note, however, that this analysis has been with respect to Fig. 3.28, which assumed *ideal* thresholds. If a worst-case threshold offset occurs, i.e., $\pm V_{ref}/4$, then an output amplitude error of magnitude $\varepsilon_G V_{ref}$ results. Again, assume that the resolution remaining after this stage is n_2. Note that this is the net resolution in the remaining stages treated as one composite stage, i.e., after any digital error correction or redundancy has been accounted for. Therefore, if the gain error from the first stage is required to contribute *less than x LSB's of the remaining resolution,* this implies that

$$\varepsilon_G V_{ref} < x \cdot \frac{1}{2^{n_2}} \cdot (2V_{ref}) \tag{3.20}$$

which yields

$$\varepsilon_G < \frac{x}{2^{n_2 - 1}} \tag{3.21}$$

as before.

Here, *because the decision levels are shifted,* in order to cause a missing code, it is enough that the amplitude error in the gained-up residue be equal to 1/2 LSB of the total

remaining resolution — i.e., corresponding to $x = 1/2$. Therefore, to avoid missing codes, the following condition must hold:

$$\varepsilon_G < \frac{1}{2^{n_2}} \tag{3.22}$$

(Note that letting $x = 1/2$ in (3.14) and (3.21) gives conditions that agree with the results derived in [76].)

The gain accuracy condition derived above implies that for a given overall ADC resolution, if a high number of bits is resolved in the first stage, then n_2 is low, and thus (3.21) implies that a large relative gain error ε_G is allowed. Furthermore, because of the reduced interstage gain when redundancy is used, a given relative gain error, when referred to the input, is more important. However, as is explained in greater detail later, in a practical implementation of the gain block using, for example, switched-capacitor techniques, as the number of bits resolved in the first stage increases, the closed-loop gain also increases, and thus the feedback factor decreases, causing the overall *loop gain* to decrease. Therefore, the op amp open-loop dc gain requirement, which is the key circuit challenge, tends not to change very much.

To see this, recall that in any closed-loop feedback system, the gain error ε_G is given by the reciprocal of the loop gain, that is,

$$\varepsilon_G = \frac{1}{Af} \tag{3.23}$$

where A is the open-loop gain of the forward path, i.e., the op amp, and f is the feedback factor. Substituting this expression into (3.21) gives

$$\frac{1}{Af} < \frac{x}{2^{n_2-1}} \tag{3.24}$$

Let n_1 and 2^{n_1} be the net number of bits resolved in the first stage and the gain of the first stage respectively. A basic approximation (although quite optimistic for most SC implementations) is that the feedback factor f of the first stage gain block is the reciprocal of its closed loop gain, i.e.,

$$f = \frac{1}{2^{n_1}} \tag{3.25}$$

Therefore, (3.24) implies that the following condition is necessary in order for the gain error from the first stage to contribute less than x LSB's of the remaining resolution:

$$\frac{1}{A} < \frac{x}{2^{n_1 + n_2 - 1}} \tag{3.26}$$

But $n_1 + n_2$ is equal to the total ADC resolution n_{tot}. Therefore, the open-loop gain requirement for the op amp in the first stage is set by the *overall resolution,* and given by

$$\frac{1}{A} < \frac{x}{2^{n_{tot} - 1}} \tag{3.27}$$

Choosing $x = 1/2$ gives $A > 2^{n_{tot}}$, which is the minimum op amp gain required.

3.3.1 Analysis of DAC Errors

The case of a DAC error in an individual stage of a multistage ADC is now examined, as depicted below in Fig. 3.29.

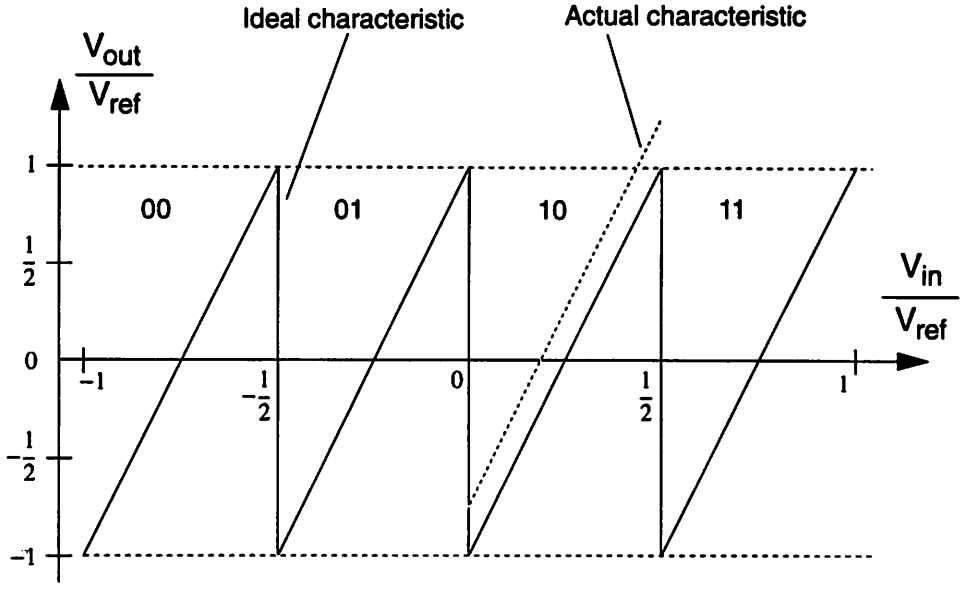


Fig. 3.29 Static input-output relationship of a 2-bit stage with nominal interstage gain of 4. The ideal transfer characteristic is shown with continuous lines; the nonideal curve resulting from a DAC error has dotted lines.

Clearly an error ΔV_{DAC} in the DAC value will result in an error in the final residue equal to $(G\Delta V_{DAC})$. In order for this to be less than x LSB's of the remaining stages requires

$$G(\Delta V_{DAC}) < \frac{x}{2^{n_2}} \cdot (2V_{ref}) \tag{3.28}$$

which yields

$$\frac{\Delta V_{DAC}}{2V_{ref}} < \frac{x}{2^{n_2}} \cdot \frac{1}{G} \tag{3.29}$$

Therefore, the DAC *relative accuracy* $[\Delta V_{DAC}/(2V_{ref})]$ should be consistent with the composite resolution of the current stage *and* the remaining stages.

This is also apparent from the basic input-output relationship of a single stage:

$$V_{out} = G(V_{in} - V_{DAC}) \tag{3.30}$$

Since V_{DAC} is subtracted directly from V_{in} it is clear that any error in V_{DAC} will directly degrade the accuracy of the analog signal.

3.3.2 Summary and Discussion

- In multistage ADC's the accuracy of the DAC levels and the residue gains within each stage cannot be compromised: the greater the overall system resolution, the more stringent are the accuracy requirements for the DAC levels and the interstage gains.
- In general, the relative accuracy of the interstage gain in one stage of a multistage ADC is required to be consistent with the resolution remaining after that stage.
- The open-loop dc gain requirement on the first stage op amp is set by the overall resolution of the ADC.
- In general, the relative accuracy of the DAC levels in one stage of a multistage ADC is required to be consistent with the resolution of that stage and all remaining stages.
- However, since the effective resolution remaining *decreases* as the analog signal progresses successively from stage to stage in a multistage ADC, the gain accuracy and DAC level accuracy requirements are relaxed accordingly in later stages.

The result of these accuracy constraints is that multistage ADC's above about 10 bits need some form of calibration for the interstage gain and DAC levels. Self-calibration approaches using the technique of trim capacitor arrays have been reported for a 13-bit algorithmic topology [101], for a 12-bit two-step recycling converter [61], and for a 13-bit, 6-stage pipeline architecture [79], [80]. Purely digital calibration approaches for multistage ADC's are described in [69], [70], [57]. These will be discussed briefly in Chapter 4 in the context of a generalized analysis of multistage A/D conversion.

3.4 IMPLEMENTATIONS OF PIPELINE ADC'S

Some possible implementations of multistage A/D converters are now presented, with primary focus on switched-capacitor techniques. Obviously, this could not possibly include *all* topologies ever published — rather, the aim is to give a representative survey elucidating the similarities and differences and the relative advantages and disadvantages of various approaches.

In Section 3.4.0 a general switched-capacitor pipeline ADC stage is described and its input-output relation derived. Section 3.4.1 discusses a number of particular examples of CMOS switched-capacitor pipeline ADC's from the literature: these are seen to be special cases of the general stage presented in Section 3.4.0. Finally, Section 3.4.2 discusses various issues that involve considerations both at the ADC algorithm level and at the circuit implementation level.

Note that although most of the examples throughout this section show single-ended circuits, the actual IC implementations are usually in a differential manner. (Hence the emphasis on bipolar signals in earlier sections.)

3.4.0 General Switched-Capacitor Pipeline ADC Stage

A general model for an individual stage of a switched-capacitor (SC) pipeline ADC is shown in Fig. 3.30. It is assumed that the op amp has infinite gain and zero offset. (The effect of finite gain, nonzero offset, and other errors and nonidealities are considered later in Section 3.5.) A pipeline ADC consists of a cascade of these stages, with adjacent stages operating on opposite phases. Digital pipeline latches are used to align, i.e., *synchronize,* the stage output codes that correspond to a particular input sample.

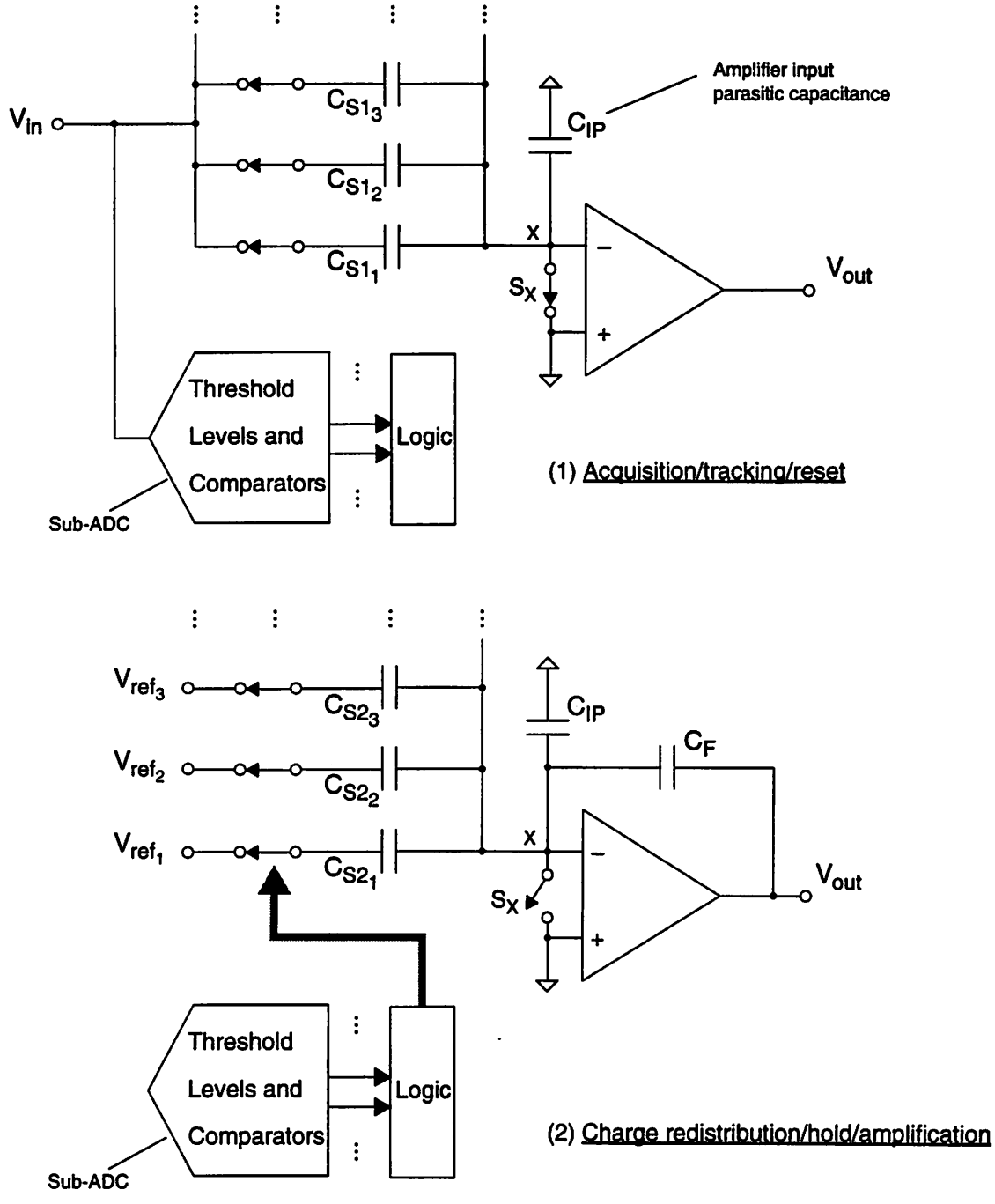


Fig. 3.30 Schematic of a general SC pipeline ADC stage showing the two phases of operation. In phase 1, the analog input is acquired on the capacitors, and sampling switch S_X at the summing node X is closed. Also during phase 1, the input is connected to the sub-ADC. In phase 2, the sampling switch S_X is opened, the capacitors are switched to appropriate reference voltages, depending on the sub-ADC decision, and the amplifier is in a closed-loop configuration with feedback capacitor C_F. (Note that the arrow in the lower figure indicates that the logic controls *all* the switches — i.e., not just C_{S2_1}.)

There are two phases of operation, as shown in Fig. 3.30. During phase 1, the acquisition/tracking/reset phase, which is shown in the upper diagram on Fig. 3.30, the input voltage V_{in} is acquired on capacitors $\{C_{S1_i}\}$. The op amp block is an operational transconductance amplifier (OTA) with a high output impedance. (In fact, in this section, and throughout most of this dissertation, whenever the term *op amp* is used, what is generally implied is an OTA.) The term *reset* refers to the fact that the op amp input terminals are reset to ground. The op amp output terminal is *not* left floating during the input acquisition phase, despite what appears in Fig. 3.30. As mentioned above, in many cases the implementation employs fully-differential circuitry, and during this reset phase, the common-mode feedback circuitry is reset, and the op amp outputs are reset to their common-mode value. However, what is occurring at the output node during the acquisition/tracking phase is not fundamental to the A/D conversion function, rather it is what is happening at the summing node — i.e., the negative input of the op amp — since at that node the input charge is sampled, redistributed, and ultimately gives rise to the output voltage. Therefore, in summary, the diagrams in this section are intended to emphasize the essence of what is happening with respect to the multistage A/D conversion function, and so, whenever possible, peripheral details such as output reset are omitted.

Also during the reset/acquisition phase (phase 1), the signal is being acquired by the comparators in the sub-ADC. (Note that the details of the comparator timing are not covered in this section — it is assumed that the comparators are implemented using a two-phase clocked approach.)

At the end of this acquisition/tracking phase the sampling switch is opened, which defines the sampling instant, and the stored charge at node X is given by

$$Q_X = C_{S1_1}(0 - V_{in}) + C_{S1_2}(0 - V_{in}) + C_{S1_3}(0 - V_{in}) + \ldots \qquad (3.31)$$

i.e.,

$$Q_X = -(\sum_i C_{S1_i}) V_{in} \tag{3.32}$$

During phase 2, the charge redistribution/hold/amplification phase, shown in the lower schematic on Fig. 3.30, the comparators latch, and their outputs pass through some logic that determines which capacitors $\{C_{S2_j}\}$ are connected to which voltage references $\{V_{ref_j}\}$ in order to perform the DAC function. The notation is defined as follows: V_{ref_j} denotes the *value* of the voltage reference that is connected to capacitor C_{S2_j}. In general, V_{ref_j} depends on the digital result from the comparators, since depending on the sub-ADC decision, C_{S2_j} is connected to a different physical reference voltage source.

Also in this phase (phase 2), the amplifier goes into closed-loop configuration with feedback capacitance C_F, and the conserved sampled charge Q_X is redistributed over the capacitors $\{C_{S2_j}\}$ and C_F, giving rise to voltage V_{out} at the amplifier output. In a typical pipelined multistage ADC implementation, as V_{out} is settling to its final value, it is being acquired by the next stage in the pipeline, which is operating 180° out of phase.

Note 1: There is a parasitic capacitance C_{IP} due to the op amp input capacitance shown connected to the summing node in Fig. 3.30. However, it stores no charge during the reset phase (phase 1), and has no effect either in phase 2, since in this analysis the op amp is assumed to be ideal. However, C_{IP} is explicitly included in the diagrams since it *is* considered later (in Section 3.5.0) in the context of examining the effects of finite op amp open-loop gain and non-zero offset.

Note 2: Capacitor C_F is not shown explicitly in the circuit diagram for the signal acquisition phase (phase 1). This is because there are two possible ways in which the feedback capacitor can be used during this phase: either (i) it may be reset to ground and so store no charge — in which case it is irrelevant anyway, or (ii) it may be connected to V_{in} and thus used for signal acquisition, and so actually be included within the $\{C_{S1_i}\}$.

Note 3: The notation C_{S1_i} and C_{S2_j} is intended to indicate the function of the capacitances during the two phases: the $\{C_{S1_i}\}$ are the source capacitances associated with the acquisition/tracking phase (phase 1), and the $\{C_{S2_j}\}$ are the source capacitances associated with the charge redistribution phase (phase 2), respectively. Of course, the $\{C_{S1_i}\}$ and $\{C_{S2_j}\}$ often *are* physically the same capacitors, but this is not necessarily the case. In fact, all the capacitors $\{C_{S1_i}\}$ employed in the acquisition phase are present in the charge redistribution phase, that is, are contained within the combination of $\{C_{S2_j}\}$ and C_F. However, *additional* capacitors may be involved in phase 2 — specifically, C_F, the feedback capacitance. In general, therefore, considering the notation and diagrams of Fig. 3.30, the following inequality may be written:

$$(\sum_i C_{S1_i}) \leq (\sum_j C_{S2_j}) + C_F \tag{3.33}$$

with equality holding when the feedback capacitor is shared during both acquisition and charge redistribution, and no additional capacitors are switched in or connected to the summing node X during the closed-loop phase.

Note 4: It is convenient to assume that the two phases are nominally of equal duration, implying that the system is driven by two-phase clocks having nominal 50% duty cycle. However, this assumption is not necessary — especially in the context of parallel time-interleaved signal paths. Some of the issues involved will be addressed in Chapter 7.

At the end of the charge redistribution/amplification phase, the charge Q_X at node X can be written as

$$Q_X = C_{S2_1}(0 - V_{ref_1}) + C_{S2_2}(0 - V_{ref_2}) + \ldots + C_F(0 - V_{out}) \tag{3.34}$$

i.e.,

$$Q_X = -\sum_j (C_{S2_j} V_{ref_j}) - C_F V_{out} \tag{3.35}$$

Equating the two expressions (3.32) and (3.35) for the conserved charge at the summing node yields

$$-(\sum_i C_{S1_i})\,V_{in} = -\sum_j (C_{S2_j} V_{ref_j}) - C_F V_{out} \tag{3.36}$$

which can be solved to give V_{out}, the final output of the stage:

$$V_{out} = \frac{(\sum_i C_{S1_i})\,V_{in} - \sum_j (C_{S2_j} V_{ref_j})}{C_F} \tag{3.37}$$

The notation C_{S1T} is now introduced to denote the total capacitance used for acquisition of the analog input voltage during phase 1:

$$C_{S1T} = \sum_i C_{S1_i} \tag{3.38}$$

i.e., C_{S1T} is the *total input sampling capacitance*. Using this definition, (3.37) can now be rewritten more succinctly in the form

$$V_{out} = \frac{C_{S1T}}{C_F}\left(V_{in} - \frac{\sum_j (C_{S2_j} V_{ref_j})}{C_{S1T}}\right) \tag{3.39}$$

Equation (3.39) is the input-output relationship for the general SC pipeline ADC stage shown in Fig. 3.30. The following points are pertinent to this equation.

Note 1: It is apparent that the term C_{S1T}/C_F is the interstage gain, and the term $\sum_j (C_{S2_j} V_{ref_j})/C_{S1T}$ is the DAC voltage of the pipeline stage.

Note 2: The gain C_{S1T}/C_F is the ratio of the total source capacitance used in the acquisition/tracking phase to the feedback capacitance used in the amplification phase.

Note 3: The DAC voltage term $\sum_j (C_{S2_j} V_{ref_j}) / C_{S1T}$ may be regarded as the sum of the charges associated with all the (C_{S2_j}, V_{ref_j}) combinations appropriate for the particular sub-ADC decision, divided by C_{S1T}, the total source capacitance used in phase 1. Thus, in some sense, it can be considered the output of a digitally-programmable array of capacitors and reference voltages.

Note 4: The objective of this analysis is to show that *different implementations of SC pipeline stages are simply different ways to implement (3.39).* Some specific cases are (a) using a single voltage reference and an array of capacitors, (b) using a single source capacitance and an array of voltage references from a resistor string, and (c) using some combination of both, i.e., multiple capacitors and multiple reference voltages, possibly with some careful choice of nominal values to give non-binary interstage gain. These are discussed further in the following section.

Note 5: If Q_{in} and Q_{DAC} are defined to be the charges associated with the input voltage acquisition and the DAC function, respectively, i.e.,

$$Q_{in} = (\sum_i C_{S1_i}) V_{in} \tag{3.40}$$

and

$$Q_{DAC} = (\sum_j C_{S2_j} V_{ref_j}) \tag{3.41}$$

then (3.37) — another version of the input-output equation — may be written in the form

$$V_{out} = \frac{Q_{in} - Q_{DAC}}{C_F} \tag{3.42}$$

and so the circuit may also be viewed as a *charge processing* block: it acquires and samples charge Q_{in} associated with the input signal, generates DAC charge Q_{DAC}, and finally subtracts them and dumps the difference onto feedback capacitor C_F to produce output voltage V_{out}.

3.4.1 Implementation Examples

The key aspects of five SC pipeline ADC implementations are now described. All of these examples may be viewed as special cases of the general SC pipeline stage depicted in Fig. 3.30.

Implementation Example 1: Pipeline stage with k bits and resistor-string DAC

One implementation of a pipeline stage in CMOS using SC circuits is shown in a simplified form in Fig. 3.31 below [71], [73].

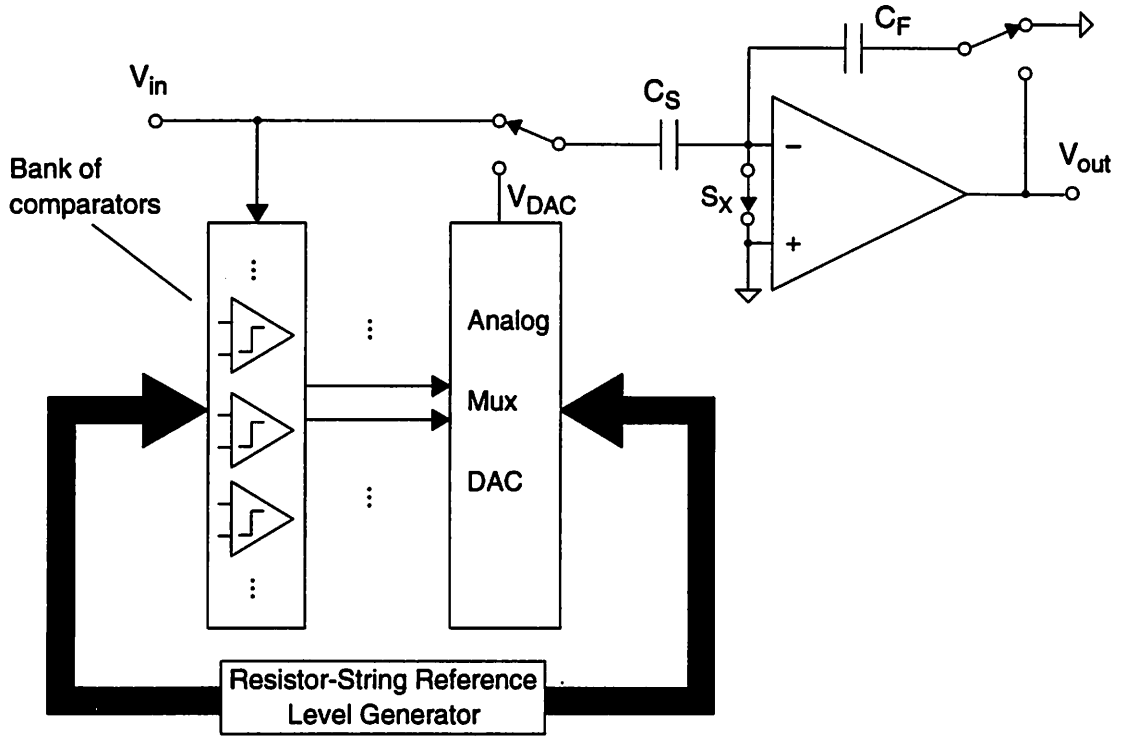


Fig. 3.31 Schematic of switched-capacitor implementation of a pipeline ADC stage. The stage sub-ADC implementation is a flash topology using a bank of comparators, as shown. The thresholds are generated by a resistor string, which also provides the levels used by the analog multiplexer sub-DAC. The interstage gain block is a standard SC gain stage with gain given by C_S/C_F.

The stage resolution is typically in the range 2–4 bits; however, the discussion here is kept general and assumes k bits. The parallel pipeline architecture prototype described in Chapter 8 employs a stage similar to this with 3-bit resolution. The switch configurations

shown in Fig. 3.31 correspond to one phase — the signal acquisition/tracking/reset phase. During the other phase — the charge redistribution/hold/amplification phase — the switches are in the opposite state; in particular, switch S_X is open. The input V_{in} is connected to a bank of comparators, which perform the k-bit A/D conversion. Digital logic (not shown) (a) performs the encode function to generate the k sub-ADC outputs, and (b) drives the k-bit DAC —here implemented using a resistor string — to produce V_{DAC}. The input sampling, subtraction, and interstage gain functions are performed by the SC gain stage shown, which consists of a single sampling capacitor C_S and a feedback capacitor C_F; closed-loop gain is $G = C_S/C_F$. The overall input-output relation of the stage is given by

$$V_{out} = \frac{C_S}{C_F}(V_{in} - V_{DAC}) \tag{3.43}$$

Note that the reference level generator for *both* the ADC and DAC is a resistor string, which can be shared over multiple parallel channels in a time-interleaved ADC architecture. Other shareable hardware resources are bias circuitry for the op amps and comparators and any clock generation circuitry. In a differential implementation, the resistor string is implemented differentially [73].

Implementation Example 2: Pipeline stage with k bits and capacitive DAC

A pipeline stage using a purely capacitive DAC is illustrated in Fig. 3.32; the figure shows a single-ended implementation suitable for a unipolar analog input in the range [0, V_{ref}].

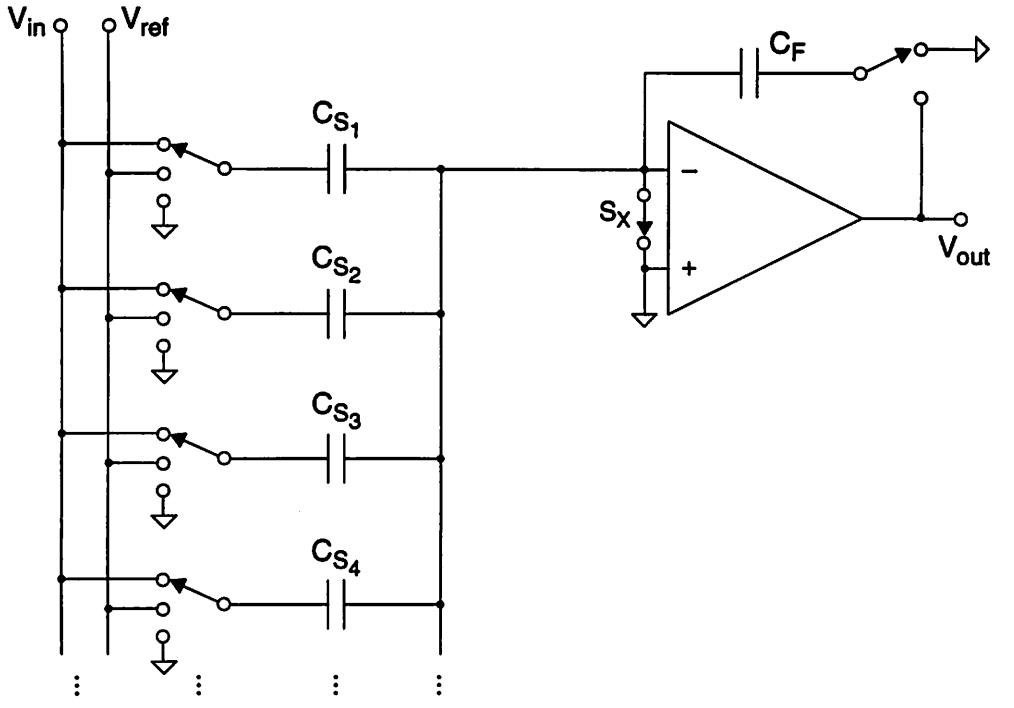


Fig. 3.32 Schematic of a SC implementation of a pipeline stage employing a capacitive DAC consisting of an array of unit capacitors $\{C_{S_i}\}$ and a voltage reference V_{ref}. The diagram indicates the acquisition/tracking/reset phase.

Here, the capacitors $\{C_{S_i}\}$ are nominally equal and are laid out as unit capacitors; this is sometimes referred to as a *linear array.* Depending on the decisions in the sub-ADC, some capacitors are connected to V_{ref} and some are connected to ground. Let M be the total number of unit capacitors, and let D unit capacitors $C_{S_1}, C_{S_2}, \ldots, C_{S_D}$ be connected to V_{ref}, where D can have the values 0, 1, ..., M. The stage transfer characteristic is then given by

$$V_{out} = \frac{\left(\sum_{i=1}^{M} C_{S_i}\right) V_{in} - \left(\sum_{i=1}^{D} C_{S_i}\right) V_{ref}}{C_F} \tag{3.44}$$

As before, let $C_{ST} = \sum_{i=1}^{M} C_{S_i}$ be the total capacitance used for input acquisition. Thus,

$$V_{out} = \frac{C_{ST}V_{in} - \left(\sum_{i=1}^{D} C_{S_i}\right)V_{ref}}{C_F} \tag{3.45}$$

$$= \frac{C_{ST}}{C_F}\left(V_{in} - \frac{\left(\sum_{i=1}^{D} C_{S_i}\right)}{C_{ST}}V_{ref}\right) \tag{3.46}$$

Note that the above two equations are special cases of (3.37) and (3.39) respectively.

From the expression $\left(\sum_{i=1}^{D} C_{S_i}\right)/C_{ST}$ on the right-hand side of (3.46), it is apparent that the DAC function is implemented by using the capacitor array to generate a programmable fraction of V_{ref}.

For a k-bit stage with interstage gain 2^k, $M = 2^k$ unit capacitors are used in the array, and the feedback capacitance is nominally equal to the array unit capacitance, that is, $C_{S_i} = C = C_F$. Therefore, from (3.46), the nominal input-output relationship is given by

$$V_{out} = M\left(V_{in} - \frac{D}{M}V_{ref}\right) \tag{3.47}$$

and thus, the nominal interstage gain is M, and the nominal DAC values are $(D/M)\,V_{ref}$, where $D = 0, 1, \ldots, M$. If a different nominal gain is required, C_F can be chosen differently; for example, it can be made equal to 2C.

A binary-weighted capacitor array is also possible: a 10-bit CMOS recycling two-step architecture using a 5-bit binary-weighted capacitor array to implement the DAC function is described in [128], [129].

Implementation Example 3: Differential implementation with capacitive DAC

An implementation of a 3-bit pipeline stage with a gain of 4 and employing a capacitive DAC [80] is shown in Fig. 3.33 below. The schematic indicates some of the details involved in a differential implementation.

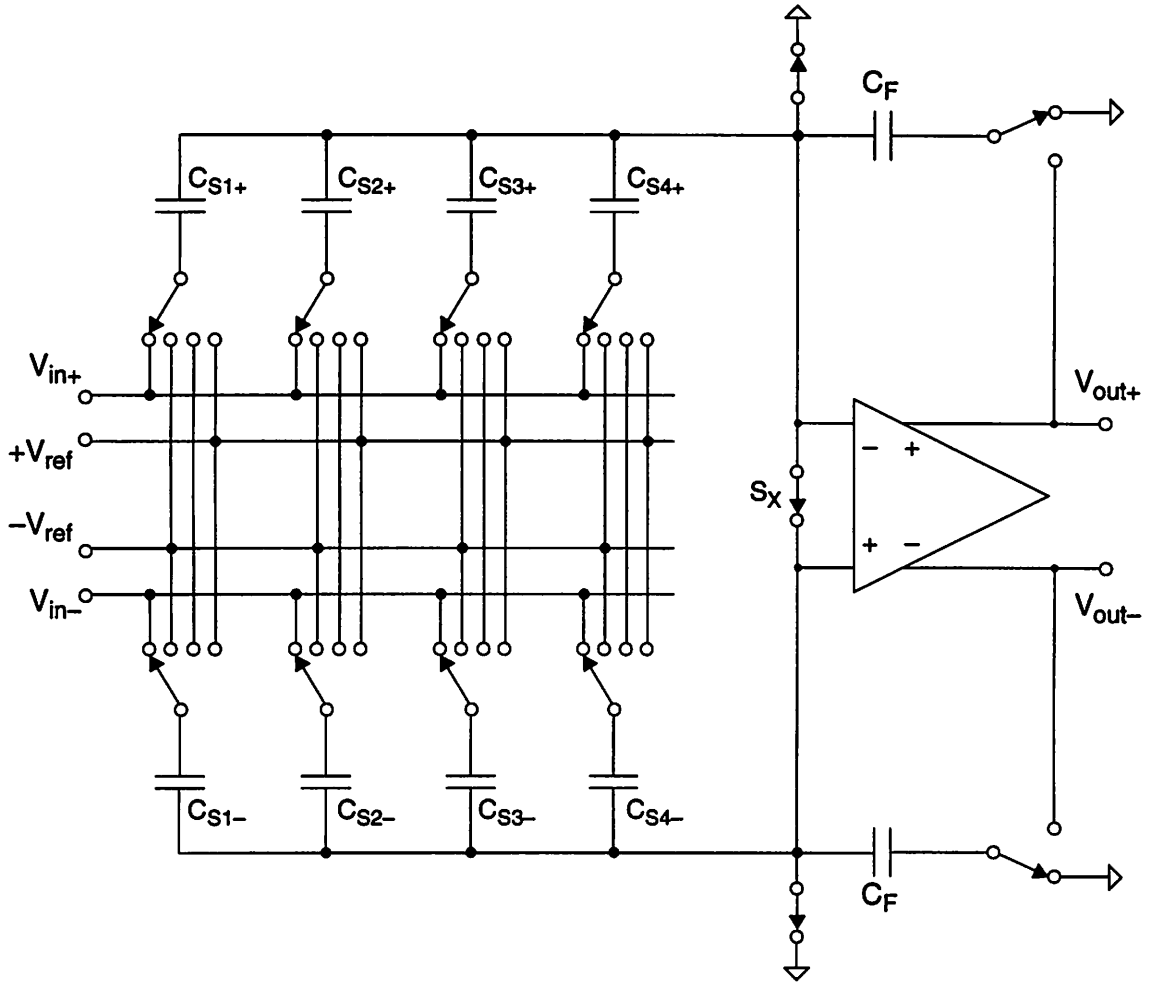


Fig. 3.33 Differential implementation of a 3-bit pipeline stage using a capacitive DAC [80]. All the capacitors are nominally equal to a unit capacitance C.

Note that this is a 3-bit stage and can generate 9 (= $2^3 + 1$) DAC levels; however, it requires only 4 unit capacitors on each side of the differential path. This is possible since each capacitor can be switched to 3 possible voltage references: $-V_{ref}$, 0, and $+V_{ref}$. In the differential implementation shown in Fig. 3.33, the "differential zero" is produced by shorting corresponding capacitors in the positive and negative sides of the array together.

One possible switching scheme to implement the DAC function in the circuit of Fig. 3.33 is now given. The table below specifies the connections between the capacitors on the positive side of the array and the reference voltages, and the resulting DAC voltage for each switch configuration. The symbols "+", "0", and "–" imply that the particular capacitor is connected to $+V_{ref}$, 0, and $-V_{ref}$ respectively. For the capacitors on the negative side of the array, all "+" and "–" symbols are swapped. Note, however, that the operation of the entire differential circuit is apparent by considering the single-ended behavior — i.e., of one side.

C_{S1+}	C_{S2+}	C_{S3+}	C_{S4+}	V_{DAC}/V_{ref}
–	–	–	–	–1
–	–	–	0	–3/4
–	–	0	0	–1/2
–	0	0	0	–1/4
0	0	0	0	0
+	0	0	0	+1/4
+	+	0	0	+1/2
+	+	+	0	+3/4
+	+	+	+	+1

The DAC voltage values given above follow directly from application of (3.39) and from the associated discussion in Section 3.4.0; that is, from the expression for the DAC voltage of a general SC pipeline stage: $\sum_j (C_{S2_j} V_{ref_j}) / C_{S1T}$. The nominal gain of the stage is $C_{ST1}/C_F = 4C/C = 4$, as mentioned above. Further discussion of the transfer characteristic of this stage may be found in Algorithm Example 11 in Appendix 4.A.

Implementation Example 4: "One bit per stage" with shared feedback capacitor

A variation on the schemes mentioned above is to use the same physical capacitor for both acquisition and residue amplification — i.e., to *share* the feedback capacitor C_F. In many pipeline ADC applications, the technique of sharing the feedback capacitor is a key idea and has certain extremely important advantages. An example of a configuration with a shared feedback capacitor is shown in Fig. 3.34; this circuit has been employed in [75].

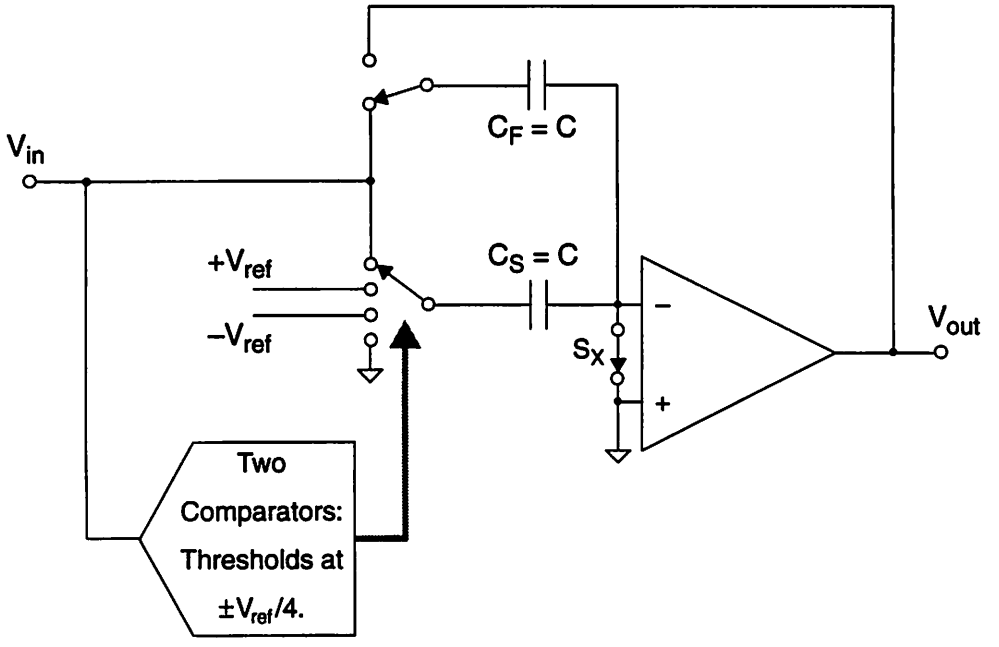


Fig. 3.34 Single-ended schematic of a gain-of-2 stage with shared feedback capacitor. The comparator decision levels are at $\pm V_{ref}/4$.

The operation is as follows. During the acquisition/tracking phase, the switch configuration is as shown in Fig. 3.34, and the input charges up the parallel combination of capacitors C_S and C_F; both C_S and C_F have the same nominal value C. Simultaneously, the comparators make their decisions. During the closed-loop amplification phase, switch S_X is open, C_F is connected to the op amp output, and C_S is connected to one of the three possible voltages $\{-V_{ref}, 0, +V_{ref}\}$ depending on the outcome of the comparator decisions. Let D be the digital code from the stage, where $D \in \{0, 1, 2\}$, and let $V_{ref,D}$ be the corresponding reference voltage selected.

The final stage output voltage is

$$V_{out} = \frac{(C_S + C_F)\, V_{in} - C_S V_{ref,D}}{C_F} \tag{3.48}$$

$$= \left(\frac{C_S + C_F}{C_F}\right)\left[V_{in} - \left(\frac{C_S}{C_S + C_F}\right) V_{ref,D}\right] \tag{3.49}$$

Since nominally $C_S = C_F = C$, the above becomes

$$V_{out} = 2\left(V_{in} - \frac{1}{2} V_{ref,D}\right) \tag{3.50}$$

and therefore, by comparing (3.50) to the general ADC stage input-output relationship given by

$$V_{out} = G\left(V_{in} - V_{DAC,D}\right) \tag{3.51}$$

it is apparent that in this case the interstage gain is 2, and the effective DAC voltage is $1/2 V_{ref,D}$, which has three possible values: $\{-V_{ref}/2, 0, +V_{ref}/2\}$.

This transfer characteristic is discussed further in Section 4.2 in Algorithm Example 3. It is often referred to as a "one bit per stage" scheme, since the net resolution per stage after digital correction and redundancy removal is one bit. What is particularly elegant about the pipeline ADC stage shown in Fig. 3.34 is that the *circuit configuration* using a gain-of-two stage with shared feedback capacitor and the *algorithm* using two threshold levels and a 3-level DAC match each other perfectly. A further key advantage of this configuration, discussed in more detail later, is the improved feedback factor obtained. Finally, note that an example of a SC 1-bit/stage pipeline ADC architecture using an interstage gain of two and sharing the feedback capacitor but employing no redundancy — i.e., using only one decision level in each stage — has been described in [126], [127].

Implementation Example 5: One bit per stage, non-binary radix

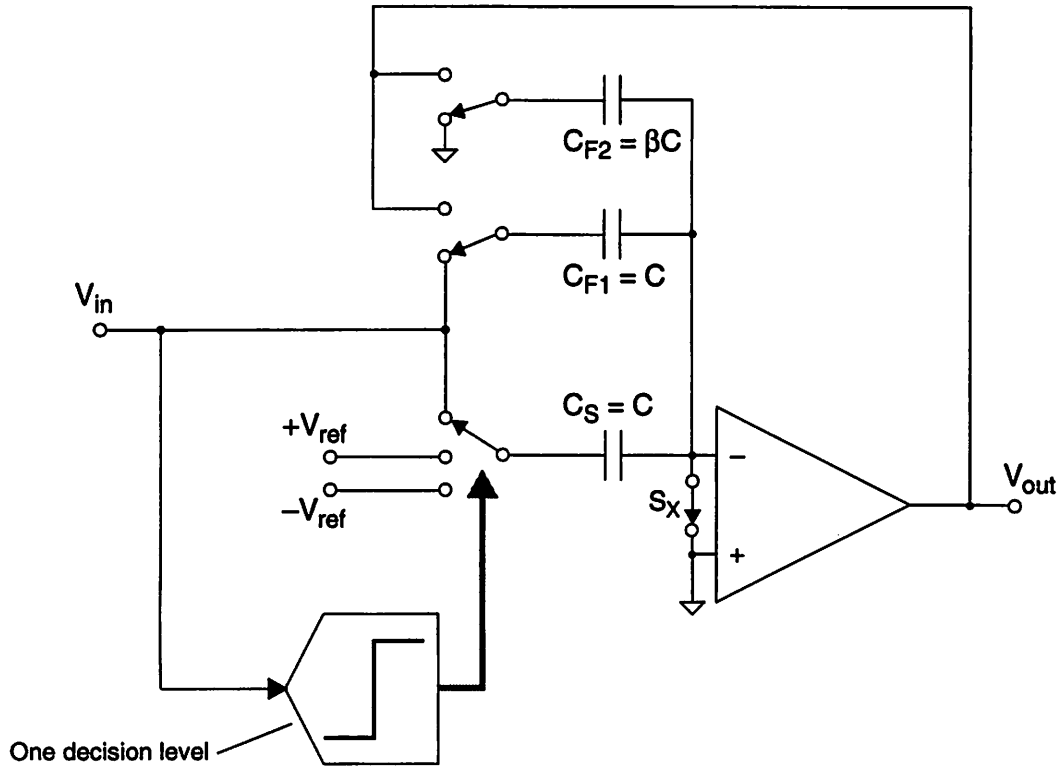


Fig. 3.35 Single-ended schematic of a pipeline stage similar to the "radix-1.93" stage used in [57].

A circuit similar to that used in [57] is shown in Fig. 3.35. Note that only one comparator is used: its decision level is at zero. During the acquisition/tracking phase, the switch configuration is as shown in Fig. 3.35, and the input charges up the parallel combination of capacitors C_S and C_{F1}, both of which have the same nominal value C. Simultaneously, the comparator is making its decision. In the next phase — the closed-loop amplification phase — the usual operations occur: switch S_X opens, the comparator output is latched, C_S is connected to one of the two possible reference voltages, $-V_{ref}$ or $+V_{ref}$, depending on the comparator decision, and C_{F1} is connected to the op amp output. A key distinguishing feature of this circuit, however, is that an additional capacitor C_{F2} is connected in parallel with C_{F1} during the charge redistribution phase.

The final stage output voltage is

$$V_{out} = \frac{(C_S + C_F)\, V_{in} - C_S V_{ref,D}}{C_F} \tag{3.48}$$

$$= \left(\frac{C_S + C_F}{C_F}\right)\left[V_{in} - \left(\frac{C_S}{C_S + C_F}\right) V_{ref,D}\right] \tag{3.49}$$

Since nominally $C_S = C_F = C$, the above becomes

$$V_{out} = 2\,(V_{in} - \frac{1}{2} V_{ref,D}) \tag{3.50}$$

and therefore, by comparing (3.50) to the general ADC stage input-output relationship given by

$$V_{out} = G\,(V_{in} - V_{DAC,D}) \tag{3.51}$$

it is apparent that in this case the interstage gain is 2, and the effective DAC voltage is $1/2 V_{ref,D}$, which has three possible values: $\{-V_{ref}/2,\ 0,\ +V_{ref}/2\}$.

This transfer characteristic is discussed further in Section 4.2 in Algorithm Example 3. It is often referred to as a "one bit per stage" scheme, since the net resolution per stage after digital correction and redundancy removal is one bit. What is particularly elegant about the pipeline ADC stage shown in Fig. 3.34 is that the *circuit configuration* using a gain-of-two stage with shared feedback capacitor and the *algorithm* using two threshold levels and a 3-level DAC match each other perfectly. A further key advantage of this configuration, discussed in more detail later, is the improved feedback factor obtained. Finally, note that an example of a SC 1-bit/stage pipeline ADC architecture using an interstage gain of two and sharing the feedback capacitor but employing no redundancy — i.e., using only one decision level in each stage — has been described in [126], [127].

Implementation Example 5: One bit per stage, non-binary radix

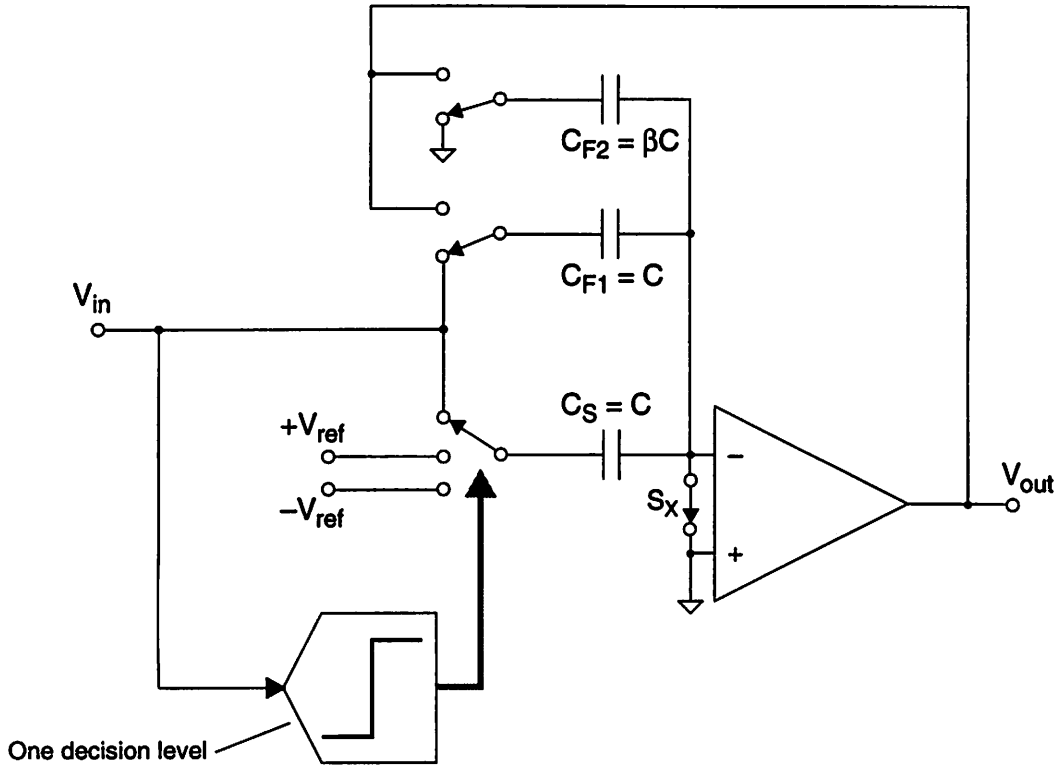


Fig. 3.35 Single-ended schematic of a pipeline stage similar to the "radix-1.93" stage used in [57].

A circuit similar to that used in [57] is shown in Fig. 3.35. Note that only one comparator is used: its decision level is at zero. During the acquisition/tracking phase, the switch configuration is as shown in Fig. 3.35, and the input charges up the parallel combination of capacitors C_S and C_{F1}, both of which have the same nominal value C. Simultaneously, the comparator is making its decision. In the next phase — the closed-loop amplification phase — the usual operations occur: switch S_X opens, the comparator output is latched, C_S is connected to one of the two possible reference voltages, $-V_{ref}$ or $+V_{ref}$, depending on the comparator decision, and C_{F1} is connected to the op amp output. A key distinguishing feature of this circuit, however, is that an additional capacitor C_{F2} is connected in parallel with C_{F1} during the charge redistribution phase.

As in the previous example, let $V_{ref,D}$ denote the particular reference voltage selected to perform the DAC function, depending on the comparator decision D, as follows: $V_{ref,0} = -V_{ref}$ and $V_{ref,1} = +V_{ref}$. The final output voltage of the stage is then given by

$$V_{out} = \frac{(C_S + C_{F1})\,V_{in} - C_S V_{ref,D}}{C_{F1} + C_{F2}} \tag{3.52}$$

$$= \left(\frac{C_S + C_{F1}}{C_{F1} + C_{F2}}\right)\left[V_{in} - \left(\frac{C_S}{C_S + C_{F1}}\right) V_{ref,D}\right] \tag{3.53}$$

Since nominally $C_S = C_{F1} = C$ and $C_{F2} = \beta C$, the nominal input-output relationship is therefore given by

$$V_{out} = \frac{2}{1+\beta}\left(V_{in} - \frac{1}{2}V_{ref,D}\right) \tag{3.54}$$

which may be rewritten as

$$V_{out} = G\,(V_{in} - V_{DAC,D}) \tag{3.55}$$

where the interstage gain is $G = \frac{2}{1+\beta}$; $V_{DAC,D} = -\frac{1}{2}V_{ref}$, when $D = 0$; and $V_{DAC,D} = +\frac{1}{2}V_{ref}$, when $D = 1$.

The value of β must be chosen to guarantee that even in the presence of the worst-case threshold offset, S/H offset, and capacitor mismatch, the output of each pipeline stage will be inside the range of the next stage. In the implementation described in [57], $\beta = 0.035$. This gives a nominal closed-loop gain $G = 2/1.035 = 1.93$, and thus motivates the description *"radix-1.93"*.

Implementation Example 6: One bit per stage — bipolar implementation

The final example of this section is a non-pipelined *ripple-through* A/D converter implemented in bipolar technology. At any instant, a single sample is being processed by the converter. This scheme uses an interstage gain of 2 and a 3-level DAC. The circuit configurations corresponding to the 3 DAC levels are shown in Fig. 3.36 below.

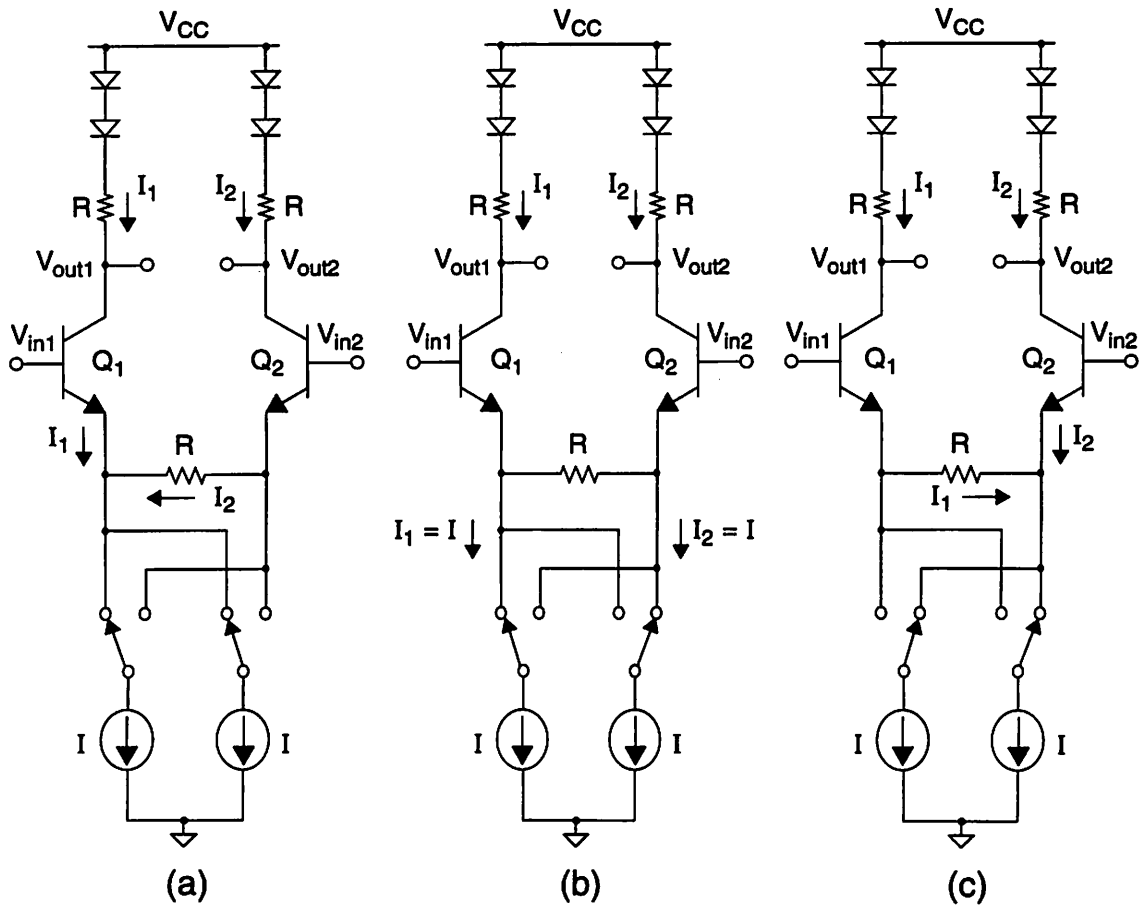


Fig. 3.36 Simplified schematic of circuit used in a bipolar ripple-through architecture [54].

In order to derive the transfer characteristic, the switch configuration labelled (c) in Fig. 3.36 is considered. The diodes in the collector branches are base-emitter diodes of BJT's and so have a voltage drop of V_{BE}. Writing Kirchoff's Voltage Law for the collector branches gives

$$V_{CC} = V_{out1} + 2V_{BE1} + I_1R = V_{out2} + 2V_{BE2} + I_2R \tag{3.56}$$

Next, examining the voltage drops in the vicinity of the input transistors yields

$$V_{in1} - V_{BE1} - I_1R = V_{in2} - V_{BE2} \tag{3.57}$$

$$\Rightarrow V_{BE2} - V_{BE1} = V_{in2} - V_{in1} + I_1R \tag{3.58}$$

Rearranging (3.55) gives

$$V_{out2} - V_{out1} = 2V_{BE1} - 2V_{BE2} + I_2R - I_1R \tag{3.59}$$

and substituting for $(V_{BE1} - V_{BE2})$ from (3.58) yields

$$V_{out2} - V_{out1} = 2(V_{in1} - V_{in2} + I_1R) + I_2R - I_1R \tag{3.60}$$

$$= 2(V_{in1} - V_{in2}) + 2I_1R + I_2R - I_1R \tag{3.61}$$

$$= 2(V_{in1} - V_{in2}) + (I_1 + I_2)R \tag{3.62}$$

$$= 2(V_{in1} - V_{in2}) + 2IR \tag{3.63}$$

$$= 2(V_{in1} - V_{in2} + IR) \tag{3.64}$$

Finally, using the differential quantities $V_{in} = V_{in1} - V_{in2}$ and $V_{out} = V_{out2} - V_{out1}$ gives

$$V_{out} = 2(V_{in} + IR) \tag{3.65}$$

Similarly, analysis of configurations (a) and (b) gives respectively

$$V_{out} = 2(V_{in} - IR) \tag{3.66}$$

and

$$V_{out} = 2(V_{in}) \tag{3.67}$$

From inspection of (3.65)–(3.67) it is apparent that these three equations constitute the input-output relationship for a gain-of-two stage in a multistage ADC, with 3 possible

DAC levels $\{-IR, 0, +IR\}$. Note that the static transfer characteristic of this stage is essentially identical to that of Implementation Example 4: both schemes employ 2 decision levels, 3 DAC levels, and an interstage gain of 2. This stage input-output relationship is discussed further in Algorithm Example 3 in Chapter 4.

Some other examples of bipolar or BiCMOS implementations of multistage A/D converters are an 8-bit, bipolar, ripple-through architecture [9]; a 10-bit, BiCMOS, 3-stage pipeline [111], [110]; a 10-bit, BiCMOS implementation that uses pipelining and interpolation [89]; and a 10-bit, BiCMOS, pipelined subranging architecture [125].

3.4.2 Some Issues Relating to Algorithms and Implementations

This section discusses various issues which involve considerations at both (a) the ADC algorithm / transfer characteristic level *and* (b) the circuit implementation level.

Issue 1: Having a threshold at zero in a pipeline stage dc transfer characteristic

Shown in Fig. 3.37 are two pipeline stage dc transfer characteristics: (a) a classical scheme and (b) a scheme with reduced interstage gain and modified coding; these input-output relationships have been presented previously in Fig. 3.4 and Fig. 3.23 respectively, and issues relating to digital error correction and redundancy have been discussed. The aim here is to focus on another aspect of these two transfer characteristics: specifically, note that scheme (b) below uses *shifted* sub-ADC levels and *shifted* sub-DAC levels, as compared to scheme (a), and, in particular, scheme (b) does not have a threshold at zero.

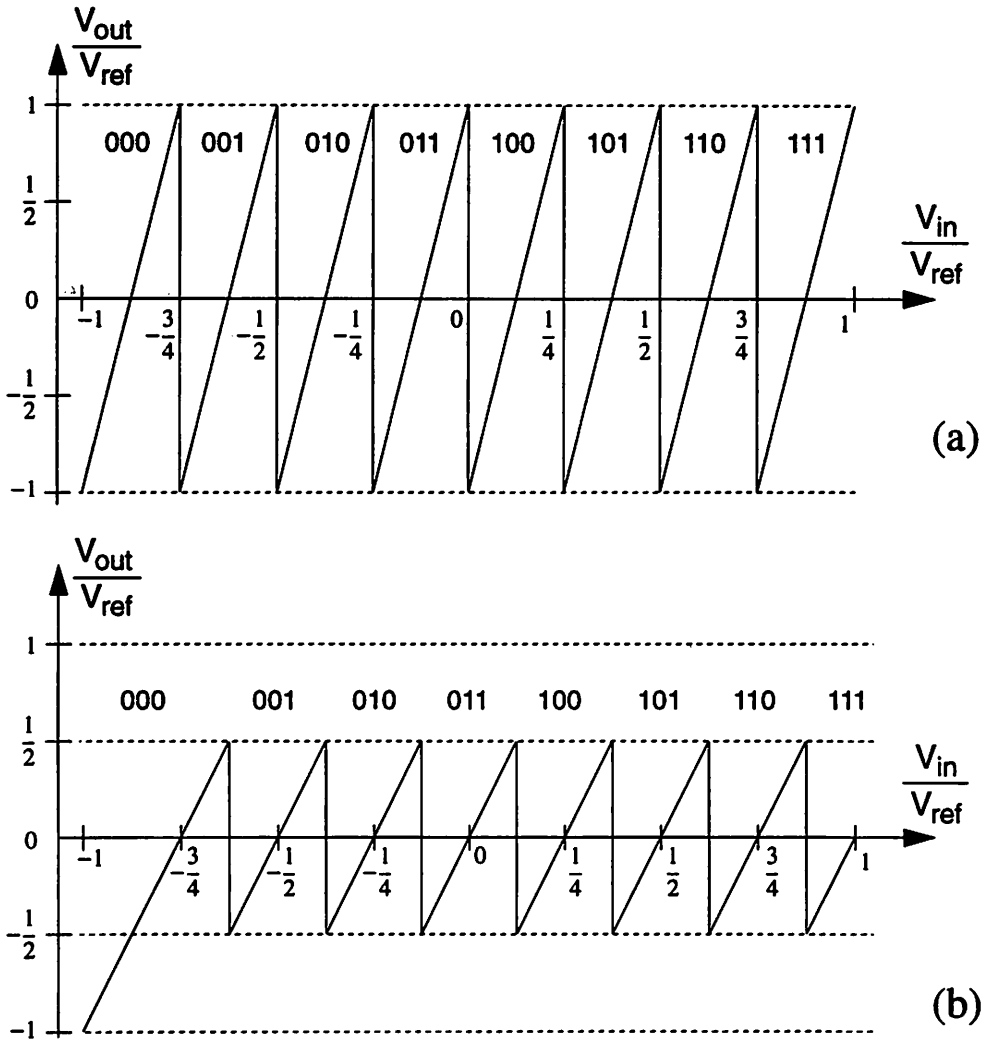


Fig. 3.37 Two 3-bit pipeline stage dc input-output relationships; (b) uses shifted DAC levels.

The following two points are pertinent.

First, observe that in scheme (b) above, the sub-DAC levels are {–3/4, –1/2, –1/4, 0, 1/4, 1/2, 3/4, 1}, which are at the LSB points with respect to the 3-bit stage input-output characteristic, whereas the sub-DAC levels in (a) are {–7/8, –5/8, –3/8, –1/8, 1/8, 3/8, 5/8, 7/8}, which are at the 1/2-LSB points with respect to the 3-bit stage. The relevance of this point is the following: *often*, the implementation of the DAC is simpler if the DAC levels are at LSB points rather than at 1/2-LSB points.

In particular, if a capacitive DAC is used, then in scheme (b) only *unit capacitors* are needed, whereas in the case of (a), *half-unit capacitors* are required. Implementation Example 3 and Implementation Example 4 in Section 3.4.1 are examples of SC pipeline implementations with the DAC levels at the stage LSB points and using unit capacitors. In some sense, the DAC levels are "more important", since their accuracy is more critical to the overall A/D converter function, and so it seems reasonable to tailor the transfer curve to make the DAC implementation easier. Since invariably the accuracy requirements on the sub-ADC decision levels are relaxed by mean of an overrange detection or digital correction scheme, the necessity for the sub-ADC thresholds to be at 1/2-LSB points does not usually have a significant impact on the implementation complexity. Note that on this basis, as may be seen from Fig. 3.24, the "naive coding" approach mentioned earlier is less desirable than the "modified coding" scheme.

Second, a dc transfer curve having a transition at zero, as compared to having the smallest magnitude thresholds at ±1/2 LSB, often causes noise — specifically, toggling of the ADC output code — when the input is near zero, or when there is no input; this phenomenon may be considered *idle-channel noise*. However, if the overall input-referred offset of the converter is comparable to 1 LSB, then the offset effectively shifts the overall ADC static transfer characteristic in an unpredictable way, and thus, both configurations are equally susceptible to unwanted toggling of the overall output code. Furthermore, note that use of a scheme that does not have a threshold nominally at zero has no effect on the noise associated with the *actual* thresholds — it simply means that noise *when the input is near zero* is reduced. However, if a scheme with shifted thresholds is employed in a pipelined multistage architecture with a small number of bits per stage, then the *most significant stages* in the pipeline are essentially immune from any spurious switching, and so the overall noise of the circuit due to switching of internal digital circuitry may be less. In the literature, the two types of transfer curves, namely, those with a threshold at zero (such as scheme (a) in Fig. 3.37) and those with shifted thresholds (such as scheme (b) in Fig. 3.37) are referred to as *mid-rise* and *mid-tread* respectively.

Issue 2: Interstage gain or reduced signal swing?

Another question which arises is: why use an interstage gain at all, that is, why not simply scale the references? There are two extremes here: (i) subranging architectures, which have no interstage gain (i.e., a gain of 1, which may be viewed as a degenerate case of interstage gain), and (ii) pipelined multistage architectures, which usually have an interstage gain sufficient to maintain the same nominal signal range or swing in all stages. There are some examples of hybrid implementations that incorporate some combination of both. The key problem associated with reduced signal swing is that the offset requirement on the threshold levels becomes very stringent.

Issue 3: Location of interstage gain block in recirculating architectures

A related issue is the effect of having the gain block located at the front of the stage. This is particularly relevant in algorithmic and recycling ADC's; specifically, compare the two implementations shown in Fig. 3.7. For a given achievable range for the internal analog circuits, having the gain in front means that the allowable input signal range has been decreased by an amount equal to that gain — i.e., the dynamic range has been decreased.

Issue 4: S/H in front?

Throughout this chapter, since the focus has been on static behavior, dynamic effects associated with sampling of a high-frequency analog input signal have not been discussed. However, the sampling function is fundamental to high-speed A/D conversion. In some implementations [80], [73], [75], in order to make the input sampling process as robust as possible, a switched-capacitor S/H stage is placed in front of the entire pipeline as shown in Fig. 3.0. This also obviates the necessity for the input to drive the comparators directly. In terms of the layout, the "pipeline stage" is then regarded as having the S/H in front rather than at the end, as described in Section 7.2 for the parallel pipeline architecture.

Issue 5: Single-ended or differential?

A single-ended implementation requires approximately half the hardware (power, area, etc.) of a fully-differential approach, and is usually significantly less complex; for example, no common-mode feedback circuitry is required. Most of the 8-bit two-step subranging implementations use single-ended circuits; for example, those described in [30], [31], [86], [87]. The necessity for differential implementations seems clearer at resolutions greater than 8 bits. In many of the CMOS SC pipeline ADC's reported that use a differential signal path, the range is directly determined by the op amp swing, and the reference levels for the DAC's, etc., are chosen accordingly, that is, matched to the swing. Obviously, a differential implementation doubles the signal swing — assuming the same circuits are used. However, swing may be less of an issue in ADC's that do not need op amps. Other issues such as how much noise is present on the power supplies — or, more generally, the entire electrical environment — are also relevant. In circuits where dynamic range is critical, and high power-supply rejection ratio (PSRR) is required, a differential implementation is essential.

3.5 ERROR SOURCES AND NONIDEALITIES IN SC PIPELINE ADC IMPLEMENTATIONS

In this section, various error source and nonidealities present in switched-capacitor (SC) implementations of pipelined multistage A/D converters are discussed. Specifically, this section analyzes some key nonidealities in the switched-capacitor gain block circuit. Gain and offset errors are considered in Section 3.5.0 by examining a general SC stage. Section 3.5.1 discusses some capacitor matching issues. Sampled thermal noise, i.e., kT/C noise, is analyzed in Section 3.5.2. Some pertinent implementation issues include the effect of finite op amp gain, the effect of the feedback factor, the use of a shared feedback capacitance, etc. This section does not explicitly address any issues related to speed of the S/H and interstage gain blocks — the topic of settling time in SC gain stages is considered in Chapter 5.

3.5.0 Gain and Offset

Although the offset requirements on the threshold levels and gain blocks can be relaxed by means of digital error correction, nevertheless, it is of interest to examine the offset performance of a SC gain stage, especially in the context of an A/D converter using parallel analog signal paths, since mismatches between the input-referred offsets of the various channels give rise to fixed pattern noise — a very important issue (see Chapter 6). A single-ended version of a general SC gain stage is shown in Fig. 3.38.

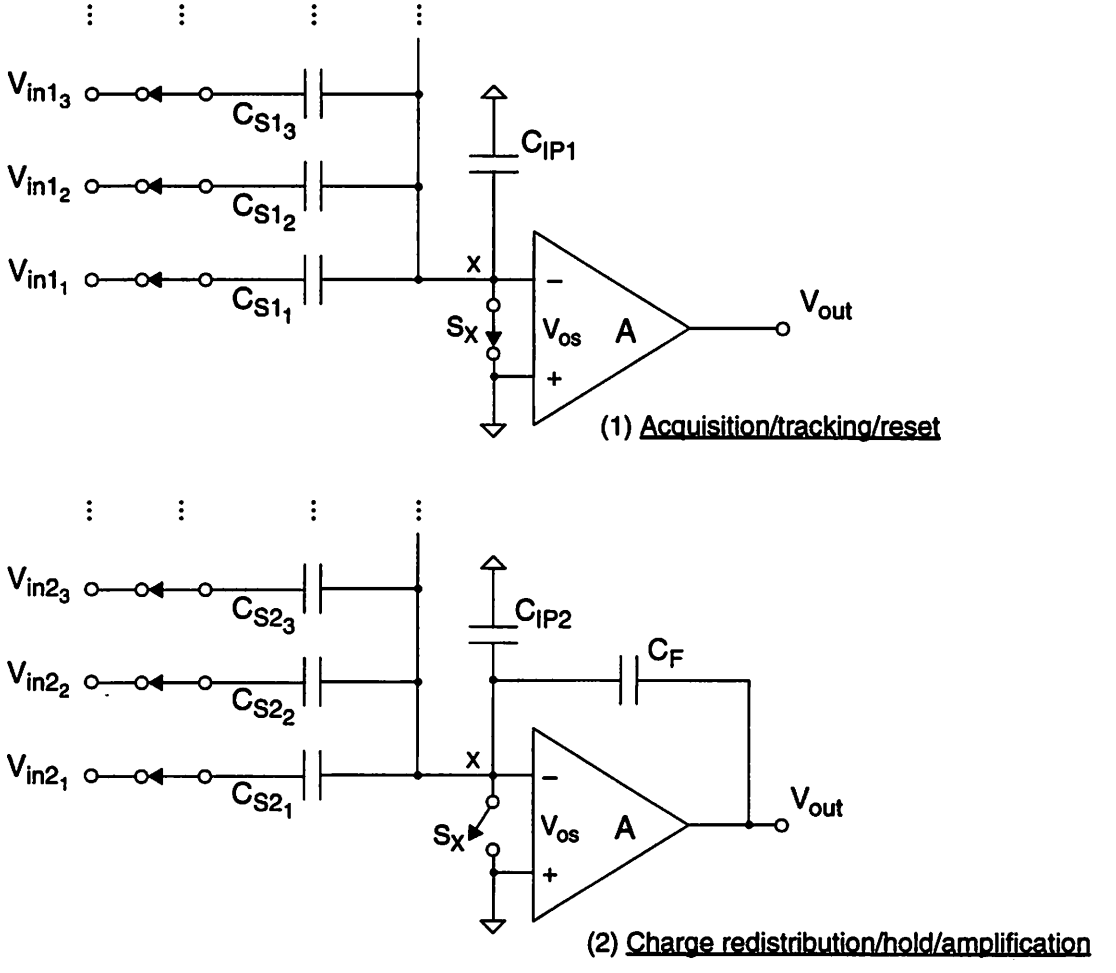


Fig. 3.38 Two phases of operation of a generalized multi-input switched-capacitor gain block. During phase 1, one set of analog inputs $\{V_{in1_i}\}$ are acquired on the capacitors, and sampling switch S_X at the summing node X is closed. In phase 2, the sampling switch S_X is opened, the capacitors are switched to another set of voltages sources $\{V_{in2_j}\}$, which may depend on a sub-ADC decision, and the amplifier is in a closed-loop configuration with feedback capacitor C_F. This model includes op amp gain A and offset V_{OS}.

This key building block is widely used in SC signal processing — not only in ADC's. The input-output relationship of the stage is now derived, including the effects of op amp offset V_{os} and finite op amp open-loop gain A, as shown on Fig. 3.38.

First, consider the top diagram in Fig. 3.38, which shows the reset/acquisition/track mode. The notation is similar to that used in Section 3.4.0 where a general pipeline stage was considered: all the capacitors $\{C_{S1_i}\}$ used in the acquisition phase are present in the charge redistribution phase, that is, are contained within the combination of $\{C_{S2_j}\}$ and C_F. However, *additional* capacitors may be involved in phase 2, in particular, C_F.

Clearly, the voltages $\{V_{in1_i}\}$ connected during phase 1 can represent different physical voltage sources. Alternatively, if the sampling instants are delayed or staggered using a multiphase clocking scheme, the $\{V_{in1_i}\}$ can be samples of one signal source at different times, thereby giving rise to an implementation of an FIR or transversal filter [147]. Filtering is not the subject of this dissertation and is not covered here. However, it should be noted that the following analysis does apply to such SC transversal filter structures.

During phase 1, sampling switch S_X is closed, and so the op amp inputs are shorted out, and the charge at the summing node (labelled X) is given by

$$Q_X = C_{S1_1}(0 - V_{in1_1}) + C_{S1_2}(0 - V_{in1_2}) + C_{S1_3}(0 - V_{in1_3}) + \ldots \tag{3.68}$$

i.e.,

$$Q_X = -\sum_i (C_{S1_i} V_{in1_i}) \tag{3.69}$$

There is some parasitic capacitance C_{IP1} at the summing node but it stores no charge.

Next, the switch configuration during the amplification or charge redistribution mode is shown in the lower diagram on Fig. 3.38. In pipeline ADC's, the voltages $\{V_{in2_j}\}$ are sub-DAC reference voltages, but this analysis is not specific to a particular application: it is simply assumed here that there is a set of voltages/signals $\{V_{in2_j}\}$ connected to the capacitors during phase 2.

The charge Q_X is now redistributed on the combination of capacitors $\{C_{S2_j}\}$, C_F, and C_{IP2}, and the voltage at node X becomes V_X. Q_X may thus be expressed as

$$Q_X = C_{S2_1}(V_X - V_{in2_1}) + C_{S2_2}(V_X - V_{in2_2}) + \ldots + C_{IP2}V_X + C_F(V_X - V_{out}) \quad (3.70)$$

i.e.,

$$Q_X = -\sum_j (C_{S2_j}V_{in2_j}) + (\sum_j C_{S2_j})V_X + C_{IP2}V_X + C_FV_X - C_FV_{out} \quad (3.71)$$

The notation C_{S2T} is used to denote the total capacitance switched to input voltage sources during phase 2, i.e.,

$$C_{S2T} = \sum_j C_{S2_j} \quad (3.72)$$

The input-output relation of the op amp including finite gain and non-zero offset is

$$V_{out} = A(0 - V_X - V_{os}) \quad (3.73)$$

which gives

$$V_X = -(\frac{V_{out}}{A} + V_{os}) \quad (3.74)$$

Charge is conserved at the summing node X; equating the expressions (3.69) and (3.71) for the sampled and conserved charge Q_X respectively yields

$$-\sum_i (C_{S1_i}V_{in1_i}) = -\sum_j (C_{S2_j}V_{in2_j}) + C_{S2T}V_X + C_FV_X + C_{IP2}V_X - C_FV_{out} \quad (3.75)$$

which gives

$$C_FV_{out} = (C_{S2T} + C_{IP2} + C_F)V_X + \sum_i (C_{S1_i}V_{in1_i}) - \sum_j (C_{S2_j}V_{in2_j}) \quad (3.76)$$

Substituting the expression for V_X from (3.74) into (3.76) yields

$$C_F V_{out} = -(C_{S2T} + C_{IP2} + C_F)\left(\frac{V_{out}}{A} + V_{OS}\right) + \sum_i (C_{S1_i} V_{in1_i}) - \sum_j (C_{S2_j} V_{in2_j}) \quad (3.77)$$

and solving this for V_{out} gives

$$V_{out} = \frac{\sum_i (C_{S1_i} V_{in1_i}) - \sum_j (C_{S2_j} V_{in2_j}) - (C_{S2T} + C_{IP2} + C_F) V_{os}}{C_F\left(1 + \frac{C_{S2T} + C_{IP2} + C_F}{AC_F}\right)} \quad (3.78)$$

With reference to Fig. 3.38, the feedback factor f of the closed-loop system of phase 2 is defined to be the transfer function "looking back" from the output node to the op amp inverting input, as shown explicitly in Fig. 3.39.

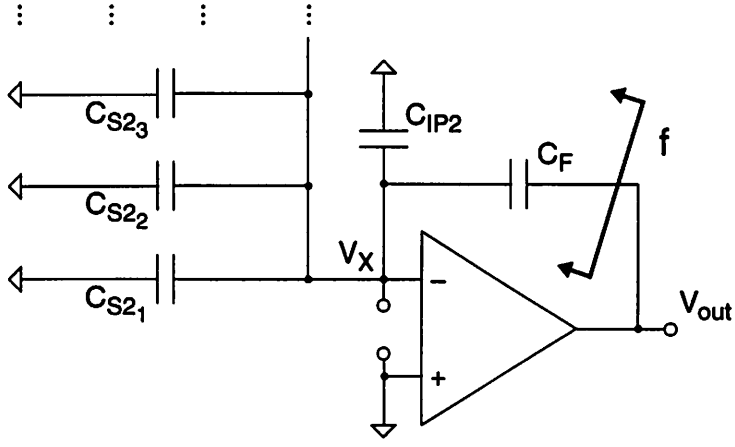


Fig. 3.39 Relevant circuit for calculation of the feedback factor f for the closed-loop amplification phase of operation of the SC gain block of Fig. 3.38.

Clearly, the feedback factor is due to a capacitive divider effect, and is given by

$$f = \left.\frac{V_X}{V_{out}}\right|_{\text{Op amp removed}} = \frac{C_F}{C_{S2T} + C_{IP2} + C_F} \quad (3.79)$$

The use of the subscript "2" emphasizes that the feedback factor refers to the configuration during phase 2 — the charge redistribution or closed-loop amplification phase. Note that f is the ratio of the feedback capacitance C_F to the *total* capacitance at the summing node in

the closed-loop phase, $C_{S2T} + C_{IP2} + C_F$, where C_{S2T} is the total source capacitance switched to input voltage sources during phase 2, and C_{IP2} is the input capacitance of the amplifier. Thus, if C_{S2T}, C_F, and C_{IP2} are comparable in value, minimizing C_{S2T} by *sharing the feedback capacitor,* that is, using it also for acquisition, is beneficial.

Replacing the feedback factor expression in equation (3.78) by the symbol f yields

$$V_{out} = \frac{\sum_i (C_{S1_i} V_{in1_i}) - \sum_j (C_{S2_j} V_{in2_j})}{C_F (1 + \frac{1}{Af})} - \frac{V_{OS}}{f(1 + \frac{1}{Af})} \tag{3.80}$$

which can also be written as

$$V_{out} = \frac{\sum_i (C_{S1_i} V_{in1_i}) - \sum_j (C_{S2_j} V_{in2_j})}{C_F} \cdot \frac{1}{(1 + \frac{1}{Af})} - \frac{V_{OS}}{f} \cdot \frac{1}{(1 + \frac{1}{Af})} \tag{3.81}$$

The term 1/Af may be viewed as a gain error ε_G. The above equation then becomes

$$V_{out} = \frac{\sum_i (C_{S1_i} V_{in1_i}) - \sum_j (C_{S2_j} V_{in2_j})}{C_F} \cdot \frac{1}{(1 + \varepsilon_G)} - \frac{V_{OS}}{f} \cdot \frac{1}{(1 + \varepsilon_G)} \tag{3.82}$$

Note that, as before, it is possible to regard this circuit as a charge difference amplifier. That is, letting $Q_{S1} = \sum_i (C_{S1_i} V_{in1_i})$ and $Q_{S2} = \sum_j (C_{S2_j} V_{in2_j})$ in (3.82) gives

$$V_{out} = (\frac{Q_{S1} - Q_{S1}}{C_F}) \cdot \frac{1}{(1 + \varepsilon_G)} - \frac{V_{OS}}{f} \cdot \frac{1}{(1 + \varepsilon_G)}) \tag{3.83}$$

It is apparent that the gain accuracy of the stage is set by the *dc loop gain* (T = Af) of the closed-loop system. Therefore, when considering gain error sources, attention must be given to *both* the op amp open-loop gain *and* the feedback factor. For relatively low interstage gain — roughly ≤4 — there can be significant benefit obtained from sharing the feedback capacitor; at higher interstage gains the improvement obtained is negligible.

Summary:

The main result of this section (3.82) is repeated below.

$$V_{out} = \frac{\sum_i (C_{S1_i} V_{in1_i}) - \sum_j (C_{S2_j} V_{in2_j})}{C_F} \cdot \frac{1}{(1+\varepsilon_G)} - \frac{V_{OS}}{f} \cdot \frac{1}{(1+\varepsilon_G)} \qquad (3.84)$$

This equation gives the input-output relation for a general SC gain stage, in the presence of finite op amp gain A and non-zero op amp offset V_{OS}, as shown in Fig. 3.38. Stage operation consists of a reset/acquisition phase, followed by a closed-loop charge redistribution and amplification phase, as follows.

- During phase 1, voltages $\{V_{in1_i}\}$ are acquired on capacitors $\{C_{S1_i}\}$.
- Sampling switch S_X is then turned off thus conserving charge at the summing node.
- During phase 2, voltages $\{V_{in2_j}\}$ are connected to capacitors $\{C_{S2_j}\}$; the amplifier is in a closed-loop configuration with feedback capacitance C_F.
- The feedback factor f of the closed-loop configuration is given by the ratio of the feedback capacitance to the total capacitance at the summing node:

$$f = \frac{C_F}{C_{S2T} + C_{IP2} + C_F} \qquad (3.85)$$

- The final output voltage V_{out} is given by (3.84).

Expressions for the *effects* of finite op amp gain and non-zero offset are as follows.

- The gain error of the general SC stage is

$$\varepsilon_G = \frac{1}{Af} = \frac{C_{S2T} + C_{IP2} + C_F}{AC_F} \qquad (3.86)$$

- The output offset V_{OSout} of the general SC stage is

$$V_{OSout} = \frac{V_{OS}}{f} \cdot \frac{1}{1+\varepsilon_G} \approx \frac{V_{OS}}{f} \qquad (3.87)$$

3.5.1 Capacitor Matching

In this section, the effect of capacitor mismatch on gain accuracy of a SC gain stage is examined for two circuit configurations: (i) using a feedback capacitor that is not shared (i.e., not used during the acquisition phase) and (ii) using a shared feedback capacitor. The discussion demonstrates another benefit obtained from sharing the feedback capacitor: a reduction in gain variance.

First, consider a switched-capacitor gain stage using M nominally identical capacitors $C_1, C_2, \ldots, C_M$, and a separate, physically distinct feedback capacitor C_F. Assume that C_F and all the $\{C_i\}$ are nominally identical, with capacitance value C. The closed-loop gain G of the stage is given by

$$G = \frac{C_1 + C_2 + C_3 + \ldots + C_M}{C_F} \tag{3.88}$$

and has nominal value equal to M. Assume the capacitances $\{C_i\}$ and C_F are independent and identically-distributed random variables. The variance of the gain is

$$\mathrm{Var}(G) = \sum_{i=1}^{M} \mathrm{Var}\left(\frac{C_i}{C_F}\right) = M\sigma^2_{\Delta C/C} \tag{3.89}$$

where $\sigma_{\Delta C/C}$ is the standard deviation of the capacitor relative mismatch — and has a typical value 0.001, or 0.1%, for a 1 pF capacitor implemented in a CMOS technology using a poly-poly or some other special-purpose precision capacitor structure.

Next, if one of the sampling capacitors — for example, C_1 — is *shared*, i.e., as well as being used for acquisition, is *also* used in the amplification phase as the feedback capacitor, then the above expression (3.88) for the closed-loop gain becomes

$$G = \frac{C_1 + C_2 + C_3 + \ldots + C_M}{C_F} = \frac{C_1 + C_2 + C_3 + \ldots + C_M}{C_1} \tag{3.90}$$

Hence

$$\mathrm{Var}(G) = \mathrm{Var}\left(1 + \sum_{i=2}^{M} \left(\frac{C_i}{C_F}\right)\right) = 0 + (M-1)\,\sigma^2_{\Delta C/C} \tag{3.91}$$

Thus, the variance of the gain has decreased from $M\sigma^2_{\Delta C/C}$ to $(M-1)\,\sigma^2_{\Delta C/C}$, and so the standard deviation of the gain has decreased by a factor of $\sqrt{(M-1)/M}$. It may be seen that for small values of M, a significant improvement is obtained. In particular, for $M = 2$ the standard deviation is reduced by 0.71; for $M = 4$ there is a reduction of 0.87.

3.5.2 kT/C Noise

Thermal noise sampled onto the sampling capacitors, commonly known as kT/C noise, is an important consideration particularly at resolutions higher than 10 bits. Shown in Fig. 3.40 below are the single-ended equivalent circuits of a SC gain stage during its two phases of operation with the relevant thermal noise sources due to the switches indicated. (Thermal noise due to the op amp transistors is not considered here [14], [18].)

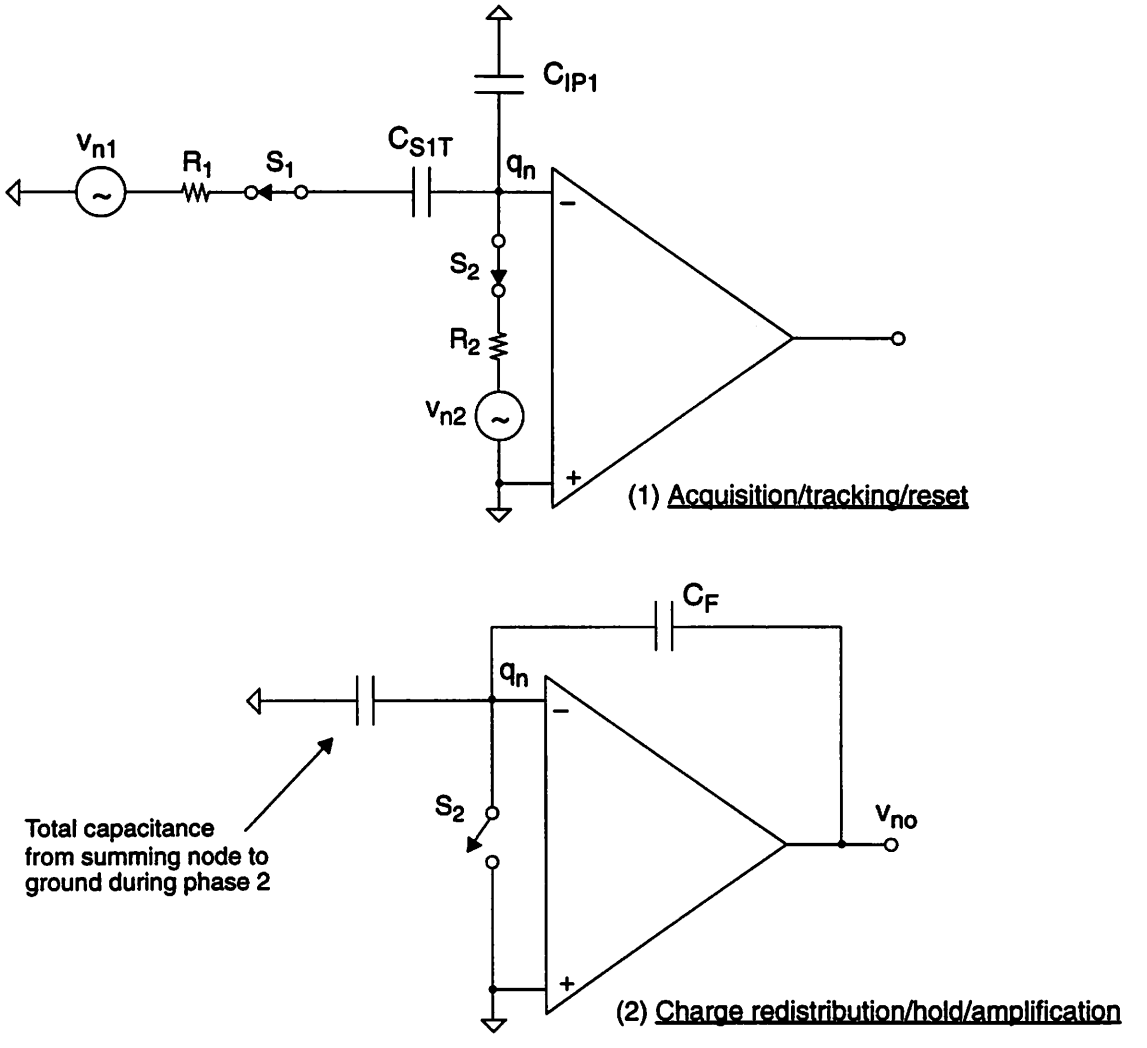


Fig. 3.40 Equivalent circuits for analysis of sampled thermal noise in a SC gain stage.

The notation and conventions are similar to Fig. 3.38; $C_{S1T} = \sum_i C_{S1_i}$ is the total capacitance used for acquisition of the input voltage (or input voltages) during phase 1.

Referring to Fig. 3.40, during the sampling acquisition phase, the noise sources v_{n1} and v_{n2}, associated with the on-resistances R_1 and R_2 of switches S_1 and S_2 respectively, feed into the summing node. At the sampling instant (when S_2 opens), noise charge q_n is stored on the summing node. During the charge redistribution phase, q_n is transferred to the feedback capacitor C_F and the output noise voltage v_{no} is given by

$$v_{no} = \frac{q_n}{C_F} \tag{3.92}$$

In order for (3.92) to be useful, it is necessary to know the variance of q_n. For a simple one-pole R-C circuit, it is known that the total thermal noise energy when integrated over all frequencies is kT/C (in volts squared), and due to aliasing, all this energy appears in the baseband [39]. However, since the network of Fig. 3.40 is a 2nd-order network, it is not obvious that the same expression holds. As shown in Appendix 3.A, the total noise charge power (i.e., variance) for the network of Fig. 3.40 is given by

$$\text{Var}(q_n) = \overline{q_n^2} = kT(C_{S1T} + C_{IP1}) \tag{3.93}$$

$$= kTC_{T1} \tag{3.94}$$

where

$$C_{T1} = C_{S1T} + C_{IP1} \tag{3.95}$$

that is, C_{T1} is the total capacitance at the summing node during the acquisition phase, i.e., the phase during which the noise is sampled. Note that this is equivalent to the total noise obtained in a first-order R-C system with capacitance C_{T1}.

Using (3.92), and assuming a differential implementation, σ_{od}^2, the differential voltage noise variance at the output of the stage due to sampled thermal noise, is therefore given by

$$\sigma_{od}^2 = \text{Var}(v_{nod}) = \frac{2\text{Var}(q_n)}{C_F^2} \tag{3.96}$$

and substituting from (3.94) yields

$$\sigma_{od}^2 = \frac{2kTC_{T1}}{C_F^2} \tag{3.97}$$

(Note that, as shown in Appendix 3.A, *both* differential S/H circuit configurations — with a center switch or without a center switch — result in this doubling of the noise power as compared to a single-ended configuration.)

For a stage with a known closed-loop gain G, the differential output noise σ_{od}^2 may be referred to the input to give the stage input-referred differential noise variance σ_{id}^2. Thus,

$$\sigma_{id}^2 = \mathrm{Var}(v_{nid}) = \frac{\mathrm{Var}(v_{nod})}{G^2} \tag{3.98}$$

$$= \frac{2kTC_{T1}}{C_F^2 G^2} \tag{3.99}$$

Since the ratio C_F/C_{T1} is simply the feedback factor f of the closed-loop system, equations (3.97) and (3.99) above become respectively

$$\sigma_{od} = \sqrt{\frac{2kT}{C_{T1}}} \cdot \frac{1}{f} \tag{3.100}$$

and

$$\sigma_{id} = \frac{\sigma_{od}}{G} = \frac{1}{G} \cdot \sqrt{\frac{2kT}{C_{T1}}} \cdot \frac{1}{f} \tag{3.101}$$

By comparing (3.87) and (3.100), it may be seen that noise and offset are affected similarly by the feedback network: both undergo an enhancement or amplification by the amount 1/f. This behavior is a general characteristic of closed-loop feedback systems, and the quantity equal to the reciprocal of the feedback factor is often referred to as the *noise gain*. In summary, the key quantities for kT/C noise in the SC gain stage are C_{T1}, the total capacitance at the summing node during the acquisition phase, and f, the feedback factor.

3.6 SUMMARY

The objective throughout this chapter has been to give a systematic presentation of multistage A/D converter structures and their operation, including basic dc functionality, code assignment and digital error correction issues, key static nonidealities, and also a broad survey of implementations with emphasis on switched-capacitor approaches. The primary focus throughout the chapter has been on the individual stage within a multistage ADC: the static input-output behavior was discussed in Sections 3.1–3.4; implementations and some associated issues and nonidealities were dealt with in Sections 3.5–3.6. Wherever possible, connections between the algorithms and the implementations have been pointed out. As well as concentrating on the architectures commonly known as "pipelines", recirculating and subranging topologies have been included in the discussion.

Pipelined Multistage A/D Converter Architectures

Some key properties of pipelined multistage ADC's are now listed.

- The primary reason to use any multistage architecture as compared to a full flash topology is that hardware cost — i.e., power, area, and associated characteristics such as input capacitance — does not increase exponentially with the resolution.
- Any A/D conversion process involves decisions: using a class of techniques known as *digital error correction* or *redundancy* or *overrange detection,* the accuracy of the decision threshold levels may be greatly relaxed.
- The cost of this redundancy is increased number of stages in the pipeline; this always entails increased latency.
- The accuracy of the interstage gain and DAC levels within each stage cannot be compromised: the greater the overall system resolution, the more stringent are the accuracy requirements for the DAC levels and the interstage gains.

Switched-Capacitor Gain Stage

A key component in switched-capacitor (SC) implementations of pipeline ADC's is the SC gain stage. In the latter half of this chapter, the SC gain block was examined in detail; the versatility of the SC gain block was demonstrated justifying its widespread usage in multistage ADC's.

The feedback factor f for a SC gain stage in its closed-loop configuration is given by

$$f = \frac{C_F}{C_{S2T} + C_{IP2} + C_F}$$

i.e., the ratio of the feedback capacitance C_F to the *total* capacitance at the summing node in the closed-loop phase. Note that the reciprocal of feedback factor ($1/f$) is not equal to the closed-loop gain of the stage, although in some circumstances, they may be close. Expressions for the *effects* of finite op amp gain and non-zero offset were obtained, as follows.

- Gain error:

$$\varepsilon_G = \frac{1}{Af} = \frac{C_{S2T} + C_{IP2} + C_F}{AC_F} \tag{3.102}$$

  where A is the op amp open-loop gain.

- Output offset:

$$V_{OSout} = \frac{V_{OS}}{f} \cdot \frac{1}{1 + \varepsilon_G} \approx \frac{V_{OS}}{f} \tag{3.103}$$

  where V_{OS} is the op amp input-referred offset.

- kT/C noise at output:

$$\sigma_o = \sqrt{\frac{2kT}{C_{T1}}} \cdot \frac{1}{f} \tag{3.104}$$

  for a differential implementation.

The feedback factor can be improved by reducing the total capacitance at the summing node during the closed-loop phase. One way to achieve this is to share the feedback capacitor — i.e., to use it both for acquisition and for amplification. An example of this technique was described: a one bit per stage SC pipeline ADC implementation using a gain of 2 and a 3-level DAC [75]. More generally, sharing the feedback capacitor is a manifestation of a deeper idea: *feedback is good.* More feedback, that is, a larger feedback factor, helps: it improves dc loop gain accuracy, and minimizes noise and offset.

As will be discussed further in Chapter 5, maximizing f also tends to improve speed, and for the op amps used in these applications, which are usually one-stage operational transconductance amplifiers, stability is usually not a serious constraint.

Other Issues

The issue of optimization of pipelined multistage architectures has not been addressed in this chapter. Some previous work has concentrated on minimizing die area and proposed a resolution of 3–4 bits per stage [72], [135]. Another recent approach has included some consideration of speed [76], and concluded that for maximum conversion rate, the minimum possible stage resolution should be selected. The strategy of using the smallest possible number of bits per stage in order to achieve maximum A/D converter throughput was also recommended in [59] in the context of hybrid and module implementations. Further systematic approaches for low power [14] and high dynamic range performance objectives [18] are also possible; these approaches take account of speed and accuracy requirements, and also include the scaling of later stages in the pipeline.

APPENDIX

3.A kT/C NOISE ANALYSIS OF MOS S/H STAGE

3.A.0 Introduction and Motivation

This appendix relates to the discussion in Section 3.5.2 concerning thermal noise in switched-capacitor (SC) gain stages. Two widely-used differential sampling networks are shown below in Fig. 3.41 and Fig. 3.42 [73], [80]. The purpose of this analysis is to compute the sampled thermal noise of the switches.

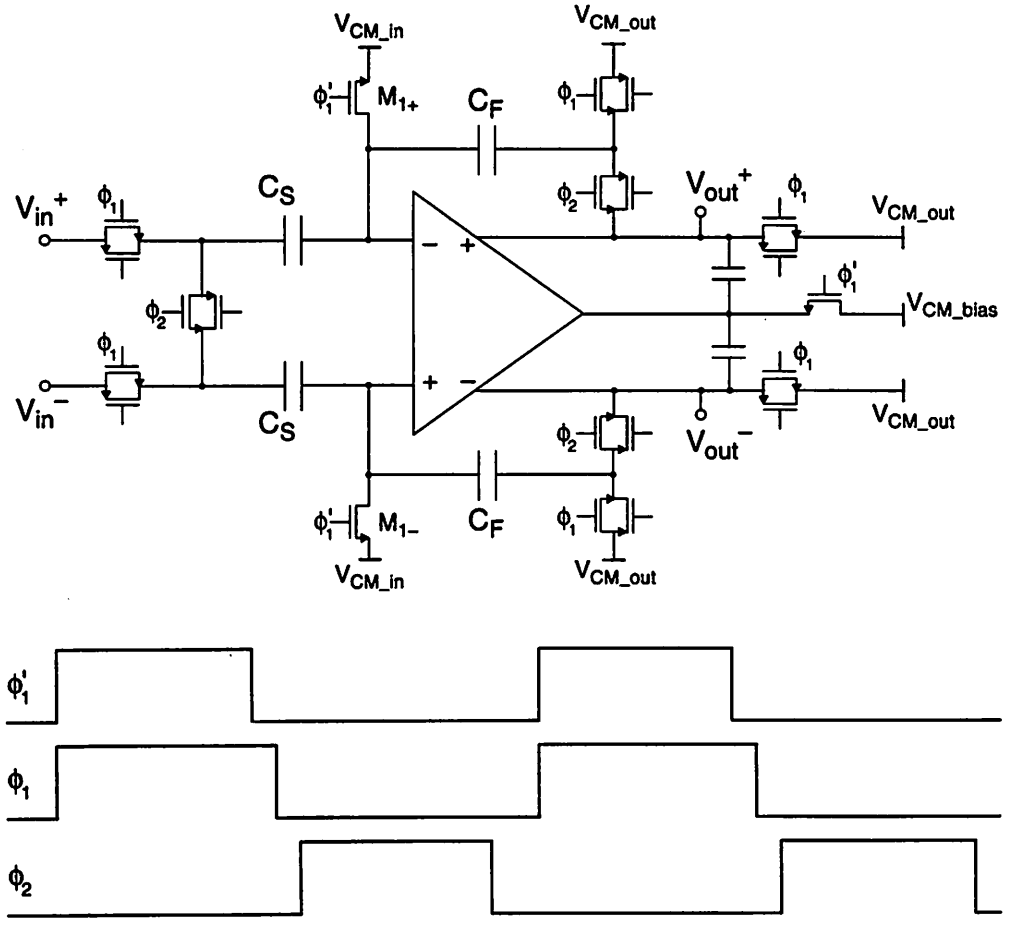


Fig. 3.41 Differential SC gain stage with clock phases indicated. V_{CM_in}, V_{CM_out}, and V_{CM_bias} are dc bias levels generated elsewhere. Sampling occurs at the instant that transistors M_{1+} and M_{1-} are turned off — i.e., on the high-to-low transition of ϕ_1'.

The configuration of Fig. 3.41 does not use a center switch, whereasthe configuration shown below in Fig. 3.42 does.

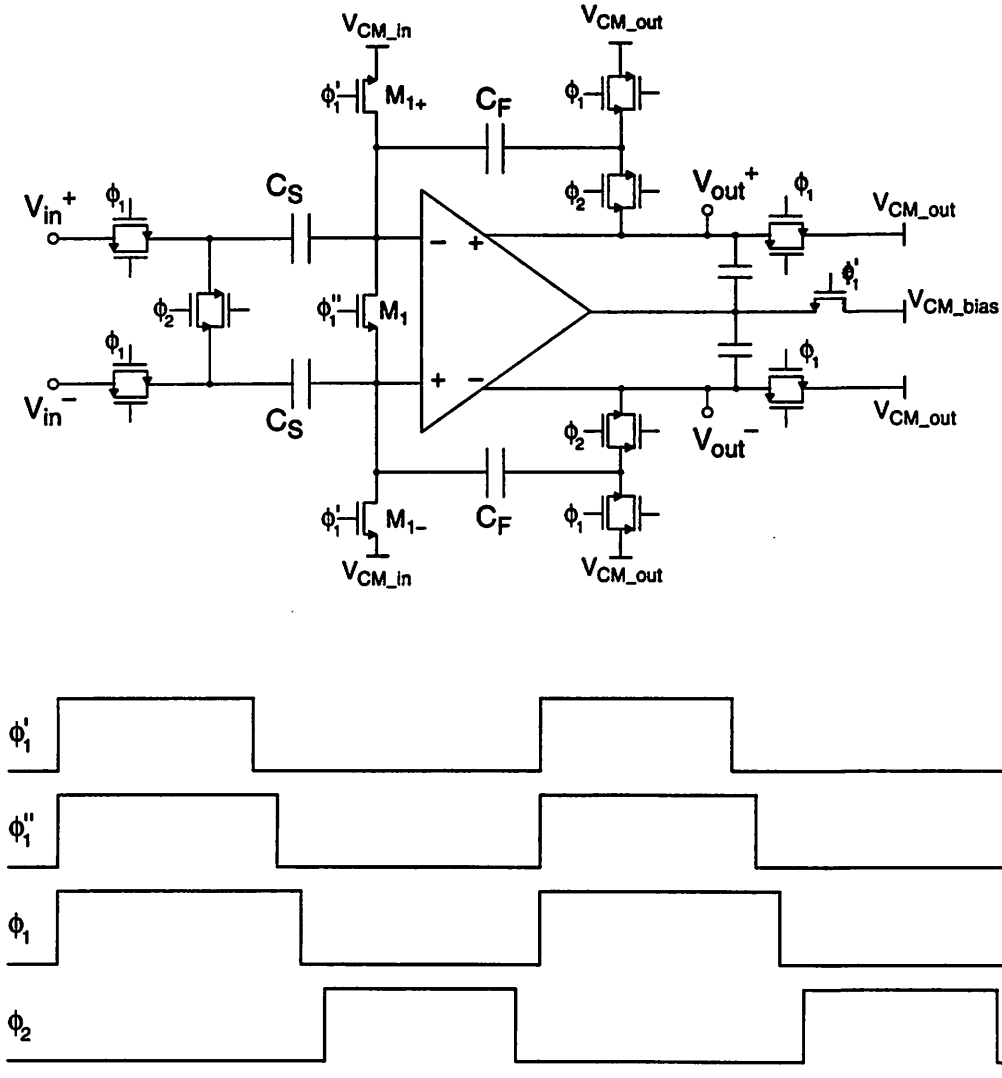


Fig. 3.42 Differential SC gain stage with clock phases indicated. V_{CM_in}, V_{CM_out}, and V_{CM_bias} are dc bias levels generated elsewhere. The input common-mode level is reset when transistors M_{1+} and M_{1-} are turned off — i.e., on the high-to-low transition of ϕ_1'. *Sampling* occurs at the instant that transistor M_1 is turned off — i.e., on the falling edge of ϕ_1''.

A single-ended equivalent circuit for the differential sampling network *without* center switch of Fig. 3.41 is shown in Fig. 3.43 below. Thermal noise sources are indicated.

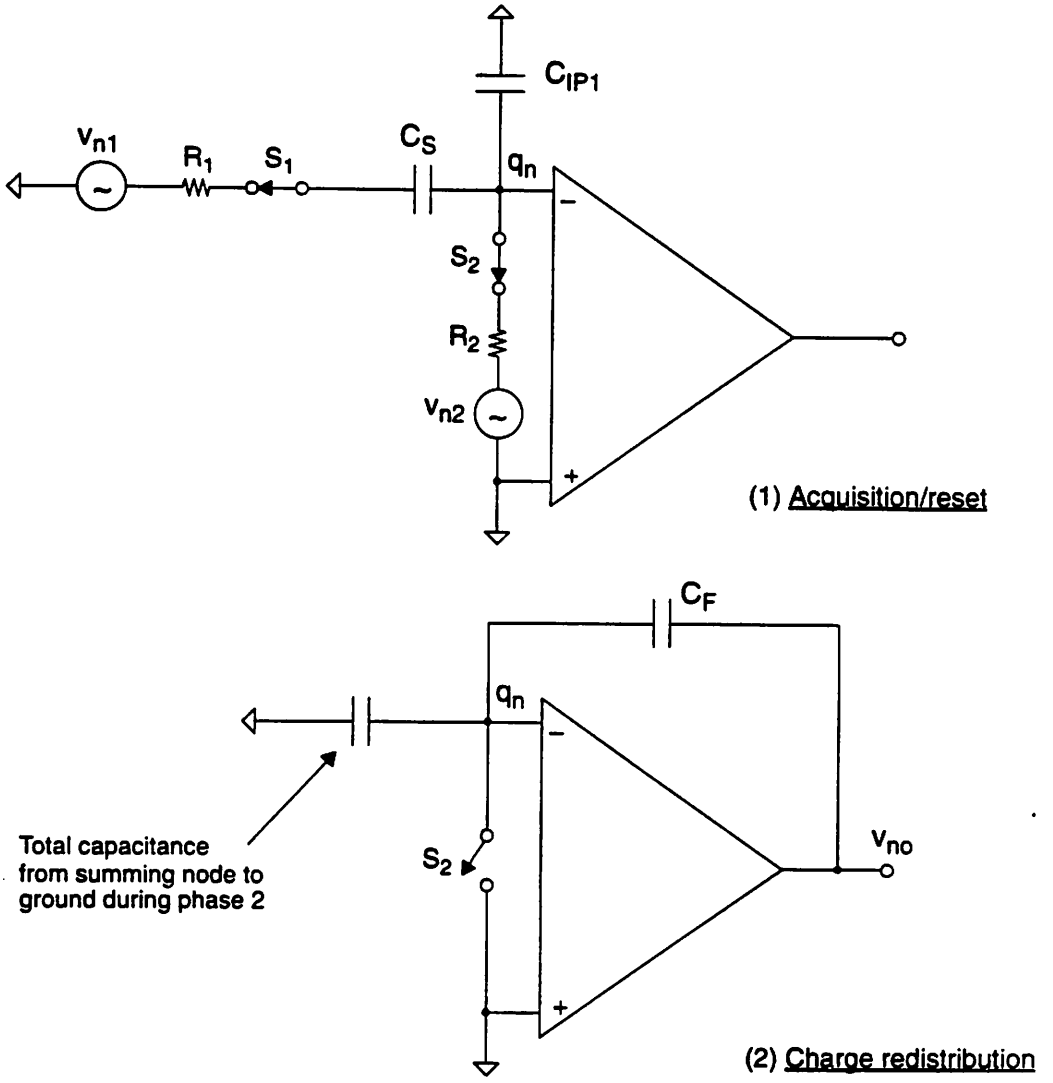


Fig. 3.43 Equivalent differential-mode half-circuits (i.e., single-ended circuits) for analysis of kT/C noise in the switched-capacitor gain stage *without* a center switch shown in Fig. 3.41.

Switch S_2 and resistance R_2 in Fig. 3.43 correspond to transistors M_{1+} or M_{1-} in Fig. 3.41, and switch S_1 and resistance R_1 in Fig. 3.43 correspond to the CMOS switches on the input side of the sampling capacitors in Fig. 3.41. Capacitor C_{IP1} is the input parasitic capacitance of the amplifier — usually it consists primarily of the gate capacitance of the input devices. The subscript 1 in C_{IP1} is intended to emphasize that this is the parasitic capacitance at the summing node *during phase 1*, the acquisition phase.

The key equivalent circuit representing the acquisition phase is shown in Fig. 3.44.

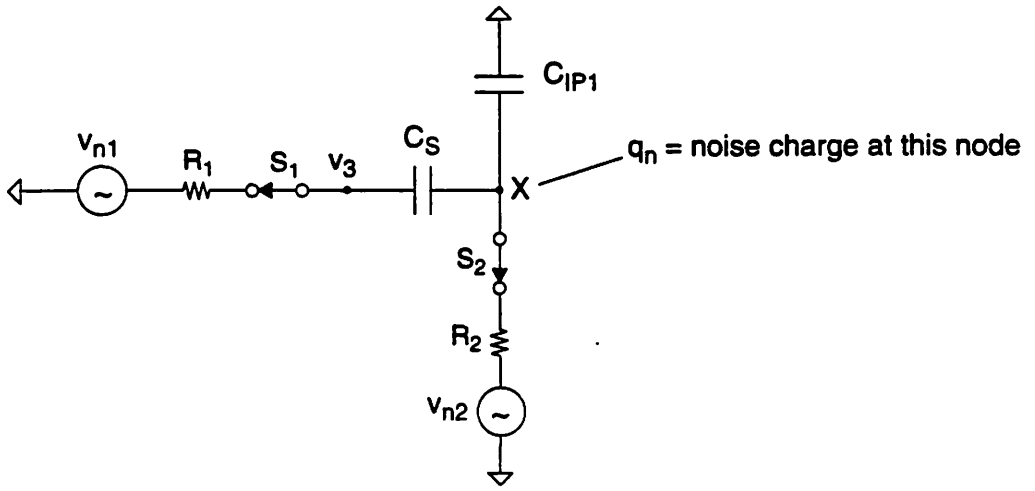


Fig. 3.44 Equivalent circuit for noise analysis in switched-capacitor gain stage showing noise acquisition and filtering.

Referring to the above diagram, during the sampling/acquisition phase, the noise sources v_{n1} and v_{n2}, which correspond to the on-resistances R_1 and R_2 of switches S_1 and S_2 respectively, feed into the summing node, and at the sampling instant (when S_2 opens) noise charge q_n is stored on the summing node. During the charge redistribution phase, q_n is transferred to the feedback capacitor C_F and the output noise voltage v_{no} is given by

$$v_{no} = \frac{q_n}{C_F} \tag{3.105}$$

Since the network of Fig. 3.44 is not a simple first-order R-C structure, it is not immediately obvious what the total noise variance is. This observation motivates the following analysis of kT/C noise in networks more complicated than simple R-C structures. The objective is to find the variance of q_n. It is assumed that noise aliasing results in the total noise energy appearing in the baseband [39].

3.A.1 Analysis of Network Without Center Switch

To ease the analysis complexity, the notation is simplified to that shown in Fig. 3.45.

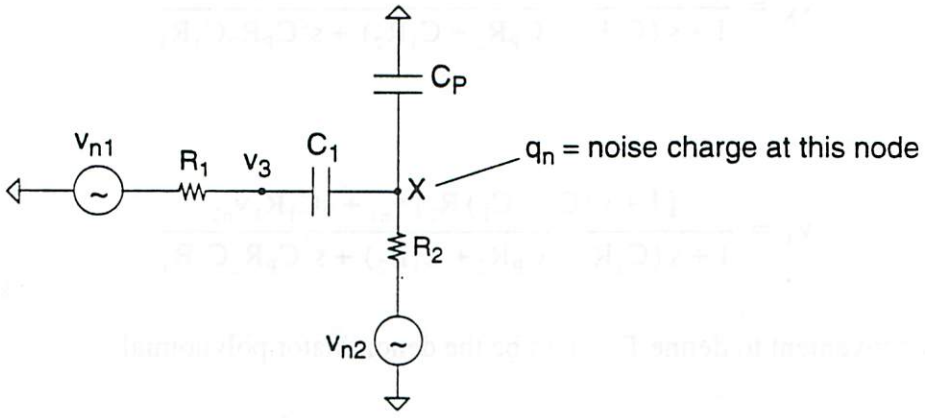


Fig. 3.45 Equivalent circuit for analysis of kT/C noise in a SC gain stage — simplified notation.

This model shown in Fig. 3.45 is applicable for analysis of either (a) a single-ended S/H circuit or (b) a differential circuit *without* a center switch, since in the case without a center switch the two halves of the differential circuit are clearly independent for noise purposes and may be analyzed individually. The final noise is just the sum of the individual noise powers. When a center switch is present, there is some correlation between the noise samples at the positive and negative summing nodes: that complicates the analysis.

The method of analysis is as follows. When switch S_2 goes off, filtered noise charge from the two thermal noise sources v_{n1} and v_{n2} is sampled at node X. The voltage at node X is denoted by v_X. A useful quantity is voltage v_3, which is defined as the voltage at the node between R_1 and C_1, as shown in Fig. 3.45. The noise charge q_n sampled at the summing node X is the sum of the charges on capacitors C_P and C_1, and is therefore given by

$$q_n = C_p (v_X - 0) + C_1 (v_X - v_3) \tag{3.106}$$

$$= (C_p + C_1) v_X - C_1 v_3 \tag{3.107}$$

A simple nodal analysis of the circuit of Fig. 3.45 gives

$$v_X = \frac{(sC_1R_2)\,v_{n1} + (1+sC_1R_1)\,v_{n2}}{1+s\,(C_1R_1+C_PR_2+C_1R_2)+s^2C_PR_2C_1R_1} \tag{3.108}$$

and

$$v_3 = \frac{[1+s\,(C_1+C_P)\,R_2]\,v_{n1} + sC_1R_1v_{n2}}{1+s\,(C_1R_1+C_PR_2+C_1R_2)+s^2C_PR_2C_1R_1} \tag{3.109}$$

It is convenient to define D (s) to be the denominator polynomial:

$$D(s) = 1+s\,(C_1R_1+C_PR_2+C_1R_2)+s^2C_PR_2C_1R_1 \tag{3.110}$$

The above equations (3.108) and (3.109) then simplify to

$$v_X = \frac{(sC_1R_2)\,v_{n1} + (1+sC_1R_1)\,v_{n2}}{D(s)} \tag{3.111}$$

and

$$v_3 = \frac{[1+s\,(C_1+C_P)\,R_2]\,v_{n1} + sC_1R_1v_{n2}}{D(s)} \tag{3.112}$$

respectively.

Rewriting (3.107): the noise charge is given by

$$q_n = (C_p+C_1)\,v_X - C_1v_3 \tag{3.113}$$

Substituting the expressions for v_X and v_3 in (3.111) and (3.112) respectively into (3.113), and simplifying yields

$$q_n = \frac{-C_1v_{n1} + (C_p+C_1+sC_1C_PR_1)\,v_{n2}}{D(s)} \tag{3.114}$$

Next, the following well-known relation is used relating the input and output power spectral densities in the s-domain of a linear network having transfer function H (s) :

$$S_o(s) = H(s)H(-s)S_i(s) \tag{3.115}$$

where $S_i(s)$ and $S_o(s)$ are the bilateral Laplace Transforms of the autocorrelation functions at the input and output respectively. Since each of the two thermal noise sources v_{n1} and v_{n2} contribute independently to the noise charge q_n, they may be analyzed separately and the resulting power spectral densities added.

To clarify this point, (3.114) may be rewritten in the form

$$q_n = H_1(s)v_{n1} + H_2(s)v_{n2} \tag{3.116}$$

where $H_1(s)$ and $H_2(s)$ are the transfer functions from the voltage noise sources v_{n1} and v_{n2} respectively to the noise charge q_n, that is,

$$H_1(s) = \frac{-C_1}{D(s)} \tag{3.117}$$

and

$$H_2(s) = \frac{(C_p + C_1 + sC_1C_pR_1)}{D(s)} \tag{3.118}$$

Note again the emphasis on the noise charge q_n at the summing node: q_n is what gets transferred to the feedback capacitor; therefore, q_n is the key quantity whose variance is required. Based on (3.116), $S_o(s)$, the power spectral density of q_n in the s-domain, is given by

$$S_o(s) = H_1(s)H_1(-s)S_1(s) + H_2(s)H_2(-s)S_2(s) \tag{3.119}$$

where $S_1(s)$ and $S_2(s)$ are the 2-sided power spectral densities of the white input noise sources v_{n1} and v_{n2} respectively, and have values given by the familiar thermal noise expressions, i.e.,

$$S_1(s) = 2kTR_1 \tag{3.120}$$

$$S_2(s) = 2kTR_2 \tag{3.121}$$

Substituting the above two relations into (3.119) gives

$$S_o(s) = H_1(s)H_1(-s)(2kTR_1) + H_2(s)H_2(-s)(2kTR_2) \tag{3.122}$$

and including the expressions for $H_1(s)$ and $H_2(s)$ from (3.117) and (3.118) respectively yields

$$S_o(s) = \frac{C_1^2(2kTR_1) + (C_p + C_1 + sC_1C_PR_1)(C_p + C_1 - sC_1C_PR_1)(2kTR_2)}{D(s)D(-s)} \tag{3.123}$$

$$= \frac{C_1^2(2kTR_1) + [(C_p + C_1)^2 - s^2(C_1C_PR_1)^2](2kTR_2)}{D(s)D(-s)} \tag{3.124}$$

which may be rearranged to give

$$S_o(s) = 2kT\left[\frac{C_1^2R_1 + (C_p + C_1)^2R_2 - s^2(C_1C_PR_1)^2R_2}{D(s)D(-s)}\right] \tag{3.125}$$

Following the approach described in [104], the right-hand side of (3.125) may be decomposed into partial fractions as

$$S_o(s) = \frac{(A + Bs)}{D(s)} + \frac{(A - Bs)}{D(-s)} \tag{3.126}$$

$$= S^+(s) + S^-(s) \tag{3.127}$$

where $S^+(s)$ and $S^-(s)$ are the transforms of the causal and anti-causal components respectively of the autocorrelation function of q_n. At this point, it is useful to recall that

$$D(s) = 1 + s(C_1R_1 + C_PR_2 + C_1R_2) + s^2C_PR_2C_1R_1 \tag{3.128}$$

Equating coefficients of s^0 and s^2 in (3.125) and (3.126) gives

$$2A = 2kT[C_1^2R_1 + (C_p + C_1)^2R_2] \tag{3.129}$$

and

$$2A(C_PR_2C_1R_1) - 2B(C_1R_1 + C_PR_2 + C_1R_2) = -2kT[(C_1C_PR_1)^2R_2] \tag{3.130}$$

which are easily solved to yield

$$A = kT\,[C_1^2R_1 + (C_p + C_1)^2R_2] \tag{3.131}$$

$$B = kT\,(C_p + C_1)\,(C_pR_2C_1R_1) \tag{3.132}$$

Returning to the power spectral density in the s-domain, it is useful to rearrange the expression for $S^+(s)$ in (3.126) as follows.

$$S^+(s) = \frac{A + Bs}{D(s)} \tag{3.133}$$

$$= \frac{B\,(s + A/B)}{D(s)} \tag{3.134}$$

$$= \frac{B}{C_pR_2C_1R_1} \cdot \frac{(s + A/B)}{D'(s)} \tag{3.135}$$

where $D'(s) = D(s) / (C_pR_2C_1R_1)$. Note that $D'(s)$ is a polynomial with coefficient of 1 on the s^2 term.

Now, substituting the expression for B from (3.132) into (3.135) gives

$$S^+(s) = kT\,(C_p + C_1) \cdot \frac{(s + A/B)}{D'(s)} \tag{3.136}$$

Taking the inverse Laplace transform of this gives $R_o(\tau)$, the autocorrelation function. Next, note that $D'(s)$ is a 2nd-order polynomial in s with the s^2 term having a coefficient of 1, and so using a well-known result from Laplace transform theory, the inverse transform of

$$\frac{(s + A/B)}{D'(s)}$$

has the form

$$e^{-\alpha\tau}(\cos\beta\tau + \gamma\sin\beta\tau)$$

which has value 1 at $\tau = 0$. Furthermore, recall that for the purposes of this analysis, only

the variance is needed, i.e., the autocorrelation evaluated at $\tau = 0$. Thus, the variance is simply obtained as the constant term multiplying the right-hand side of (3.136). That is,

$$R_o(0) = \mathrm{Var}(q_n) = kT(C_p + C_1) \tag{3.137}$$

which is the result required. (This result can also be obtained more directly by applying the Laplace Transform initial-value theorem to the expression for $S^+(s)$ in (3.126) [121].)

Note that this is equivalent to a simple R-C network with capacitance $(C_p + C_1)$. Specifically, this result is the same as would be obtained by assuming that a noise voltage of variance $kT/(C_p + C_1)$ is sampled onto capacitance $(C_p + C_1)$ resulting in a noise charge of variance

$$\frac{kT}{(C_p + C_1)} \cdot (C_p + C_1)^2 = kT(C_p + C_1).$$

Note that $(C_p + C_1)$ is simply the total capacitance at the summing node.

The output noise variance may now be obtained from (3.105).

$$R_o(0) = \mathrm{Var}(v_{no}) = \mathrm{Var}\left(\frac{q_n}{C_F}\right) = \frac{kT(C_p + C_1)}{C_F^2} \tag{3.138}$$

Finally, multiplying by 2 to account for the two independent halves of the differential network gives the differential voltage noise variance $\mathrm{Var}(v_{nod})$ at the output of the stage due to thermal noise in the switches of the sampling network:

$$\mathrm{Var}(v_{nod}) = \frac{2kT(C_p + C_1)}{C_F^2} \tag{3.139}$$

which may also be written as

$$\mathrm{Var}(v_{nod}) = \frac{2kTC_{T1}}{C_F^2} \tag{3.140}$$

where $C_{T1} = C_p + C_1$ is the total capacitance at the summing node (i.e., on one side of the differential input sampling network) during the acquisition phase.

3.A.2 Extension to Network With Center Switch

For a differential S/H network with a center switch the analysis is somewhat more involved. The key equivalent circuit is shown in Fig. 3.46 below.

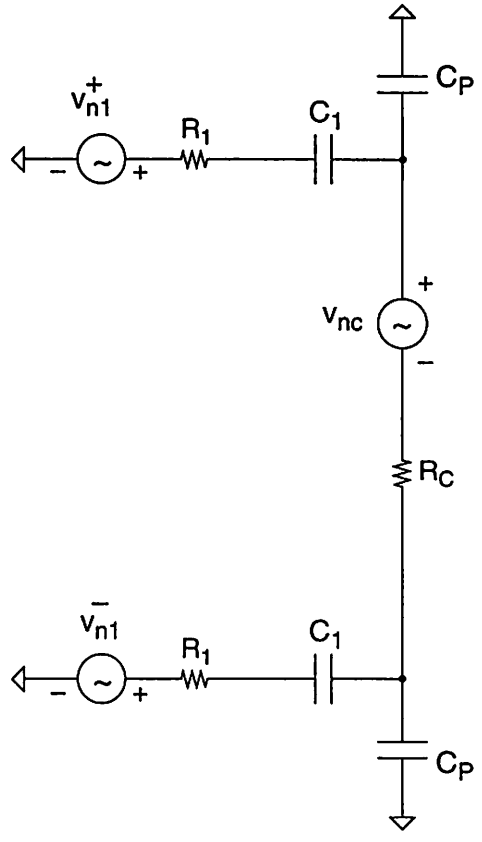


Fig. 3.46 Equivalent circuits for analysis of kT/C noise in differential switched-capacitor gain stage with center switch. R_C is the on-resistance of the center switch.

In Fig. 3.46, R_C is the on-resistance of the center switch, and R_1 corresponds to the on-resistance of the input switches. The transistors M_{1+} and M_{1-} in Fig. 3.42 are off when sampling occurs and so do not contribute any noise, and consequently are not represented in Fig. 3.46. The following analysis is based on the approach followed in [78].

In this case, since there is some correlation between the positive and negative sides of the network due to the shared center switch, the differential charge q_{nd} is considered directly. The network may be split into two halves that are correlated — as shown in Fig. 3.47.

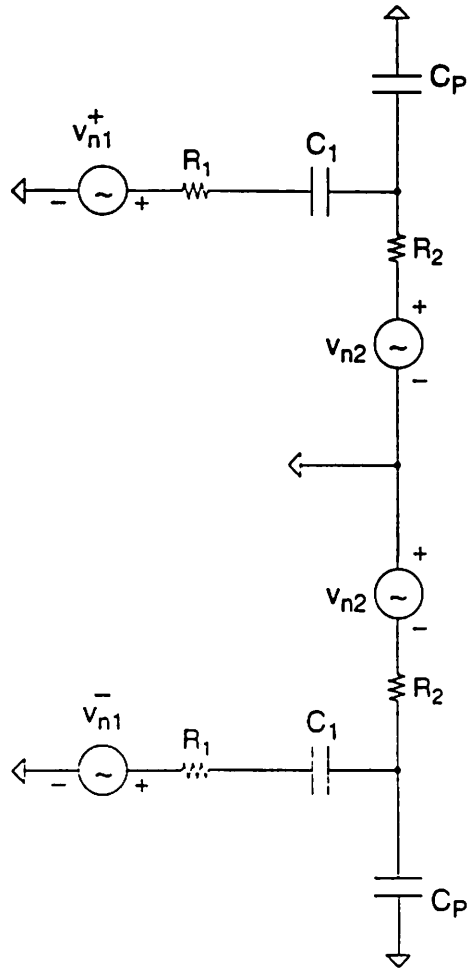


Fig. 3.47 Equivalent circuit for analysis of noise in a differential SC gain stage with center switch showing correlation between noise sources due to the thermal noise of the center switch.

Here, the *same* noise sample v_{n2} is generated on the positive and negative sides of the network, since this noise is due to one resistor — the center switch on-resistance.

The value of $S_1(s)$ is as before:

$$S_1(s) = 2kTR_1 \tag{3.141}$$

However, $S_2(s)$ is more complicated. If the value of the center switch resistance is $R_C = 2R_2$, then the 2-sided power spectral density due to R_C is

$$S_C(s) = 2kTR_C \tag{3.142}$$

If this noise power is partitioned between the positive and negative sides of the differential circuit [78], and letting v_{nc} be the noise voltage due to R_C, then the following relation must hold:

$$v_{nc} = v_{n2} + v_{n2} = 2v_{n2} \tag{3.143}$$

which gives (using the relation $R_C = 2R_2$)

$$S_2(s) = \frac{S_C(s)}{4} = \frac{2kTR_C}{4} = \frac{2kTR_2}{2} \tag{3.144}$$

This set of relationships is shown schematically in Fig. 3.48 below.

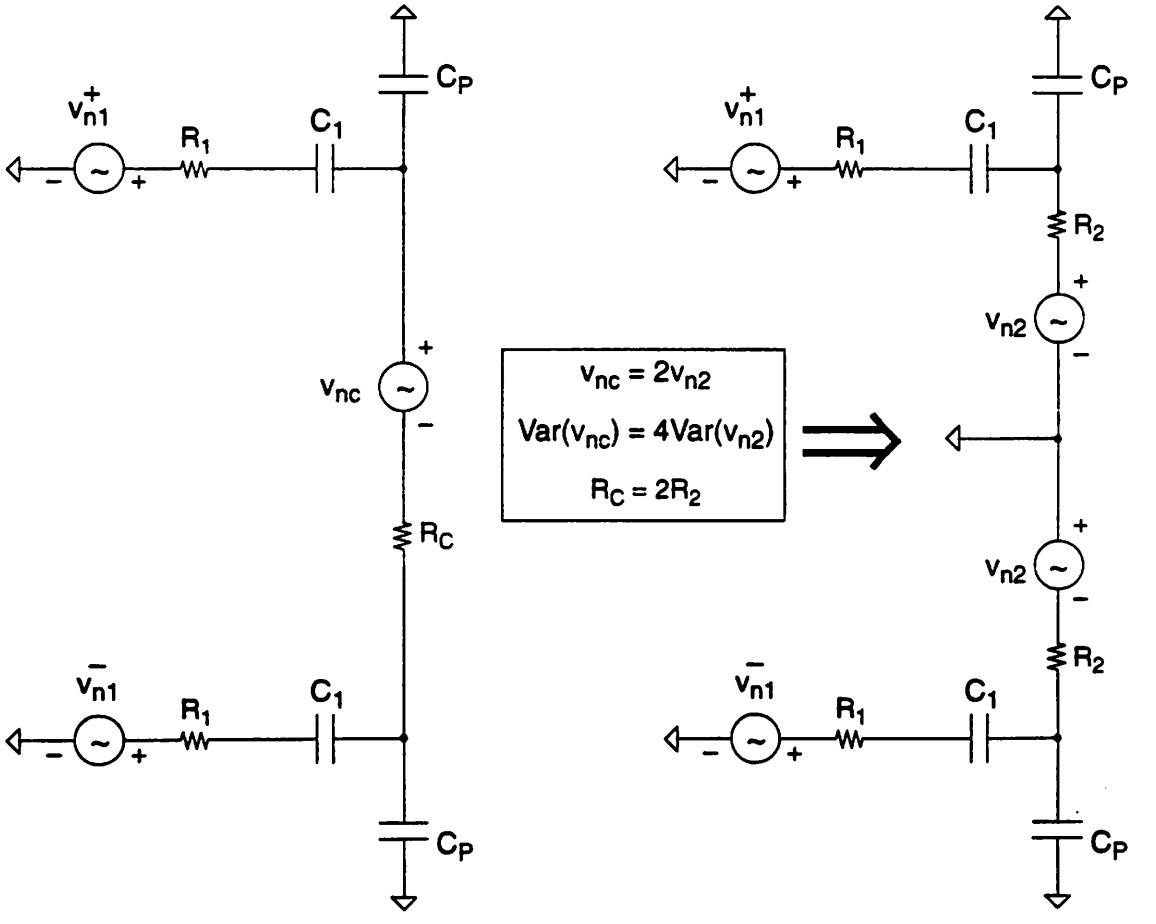


Fig. 3.48 Equivalent circuits for analysis of kT/C noise in the input network of a differential SC gain stage with center switch. On the left is the differential equivalent circuit; on the right are two correlated half-circuits that together produce the same differential noise. Note that for clarity, the polarities of the noise voltage sources are shown explicitly.

Referring to the right-hand diagram in Fig. 3.48, and adopting an approach similar to the approach employed in the case without a center switch yields the following expression for the differential noise charge:

$$q_{nd} = [H_1(s)\,v_{n1}^{+} + H_2(s)\,v_{n2}] - [H_1(s)\,v_{n1}^{-} + H_2(s)\,(-v_{n2})] \tag{3.145}$$

$$= H_1(s)\,v_{n1}^{+} - H_1(s)\,v_{n1}^{-} + 2H_2(s)\,v_{n2} \tag{3.146}$$

where $H_1(s)$ and $H_2(s)$ are the transfer functions from the noise sources v_{n1} and v_{n2} respectively to the noise charge q_n. (It is assumed the two halves of the differential network are identical and so the transfer functions from the positive and negative sides are the same.) The power spectral density of the differential noise charge is therefore given by

$$S_o(s) = H_1(s)\,H_1(-s)\,S_1(s) + H_1(s)\,H_1(-s)\,S_1(s) + 4H_2(s)\,H_2(-s)\,S_2(s) \tag{3.147}$$

$$= 2H_1(s)\,H_1(-s)\,S_1(s) + 4H_2(s)\,H_2(-s)\,S_2(s) \tag{3.148}$$

where $S_1(s)$ and $S_2(s)$ are the 2-sided power spectral densities (i.e., the white noise variance) of the input noise sources v_{n1} and v_{n2} respectively.

Finally, substituting the expressions for $S_1(s)$ and $S_2(s)$ given in (3.141) and (3.144) respectively back into (3.148) yields

$$S_o(s) = 2H_1(s)\,H_1(-s)\,(2kTR_1) + 4H_2(s)\,H_2(-s)\,\frac{(2kTR_2)}{2} \tag{3.149}$$

$$= 2\,[H_1(s)\,H_1(-s)\,(2kTR_1) + H_2(s)\,H_2(-s)\,(2kTR_2)] \tag{3.150}$$

which is the same as the single-ended expression (3.122), but multiplied by a factor of 2. Therefore, the final result for the differential output noise variance is identical to the case without the center switch, i.e.,

$$\mathrm{Var}(v_{nod}) = \frac{2kT(C_P + C_1)}{C_F^2} = \frac{2kTC_{T1}}{C_F^2} \tag{3.151}$$

Chapter 4 Generalized Theory of Multistage A/D Conversion

4.0 INTRODUCTION

The approach followed in Sections 3.1–3.2 was to begin with the classical multistage A/D conversion system using idealized components, and then determine how to modify that scheme in order to accommodate offsets in the threshold levels. It was assumed throughout that the DAC levels in each stage (i.e., the *sub-DAC levels*) and the interstage gains all equaled their nominal ideal values, to within some requisite precision. The accuracy requirements on the DAC levels and the interstage gains were analyzed in Section 3.3, and Section 3.5 discussed some pertinent circuit-level nonidealities in switched-capacitor implementations. To summarize: overall, the focus in Chapter 3 was (a) on *ideal* multistage ADC's and (b) on how to deal with certain deviations from ideal behavior present in *real* multistage ADC's. In this chapter the approach is different: here, the nonidealities of the sub-ADC, sub-DAC, and gain block are included explicitly at the outset, and the A/D conversion function is constructed without making any assumptions of ideality.

The chapter is divided into two main parts. Section 4.1 presents a generalized analysis of multistage A/D conversion. Then, in Section 4.2, various examples are given to illustrate the approach.

4.1 ANALYSIS

In order to develop this generalized view of multistage A/D conversion, the key building block is considered, namely, one individual stage of a multistage ADC. It is assumed, in general, that the input signal is bipolar. The analysis may easily be presented for a unipolar signal; furthermore, it is always possible to consider a unipolar signal as bipolar centered about some (mid-scale) reference level.

Let V_{in} be the bipolar input signal to the stage; in any real system the analog signal lies within some range, and so without any loss of generality V_{in} may be normalized to a system reference level V_{ref}. In the following analysis, it is convenient to work with a normalized input defined as

$$v_{in} = \frac{V_{in}}{V_{ref}} \tag{4.0}$$

where

$$-1 \le v_{in} \le 1 \tag{4.1}$$

i.e., lower-case letters denote normalized quantities.

Associated with this block are $M - 1$ decision elements (comparators) and the corresponding $M - 1$ decision levels or threshold levels $\{v_{C_1}, v_{C_2}, v_{C_3}, \ldots, v_{C_{M-1}}\}$, where

$$-1 < v_{C_1} < v_{C_2} < v_{C_3} < \ldots < v_{C_{M-1}} < 1 \tag{4.2}$$

These levels partition the input range into M non-overlapping intervals given by

$$[-1, v_{C_1}],\ [v_{C_1}, v_{C_2}],\ [v_{C_2}, v_{C_3}],\ \ldots,\ [v_{C_{M-2}}, v_{C_{M-1}}],\ [v_{C_{M-1}}, 1]\ .$$

Each interval is assigned a unique label, referred to as a *digital code D*. Usually it is convenient to have $D \in \{0, 1, 2, 3, \ldots, M - 1\}$, where $D = 0$ corresponds to interval $[-1, v_{C_1}]$, $D = 1$ corresponds to interval $[v_{C_1}, v_{C_2}]$, etc. To summarize so far: a set of threshold levels partitions the input range into a number of intervals, and thus any input v_{in} lies within only one of these intervals.

Each interval having code/label D also has associated with it a known, real, analog signal called a "DAC" level, $v_{DAC,D}$, which lies in the range $[-1, 1]$. Note that this model of the ADC stage explicitly separates the two functions of quantization and coding.

The input-output relation of the block — in terms of normalized voltages — is defined to be

$$v_{out} = G\,(v_{in} - v_{DAC,\,D(v_{in})}) - v_{os} \tag{4.3}$$

where $v_{DAC,\,D(v_{in})}$ is the DAC level corresponding to the unique interval within which v_{in} lies, v_{os} is an (output) offset, and G is a real number greater than or equal to one, referred to as the *stage gain*. The above is now shown schematically in Fig. 4.0.

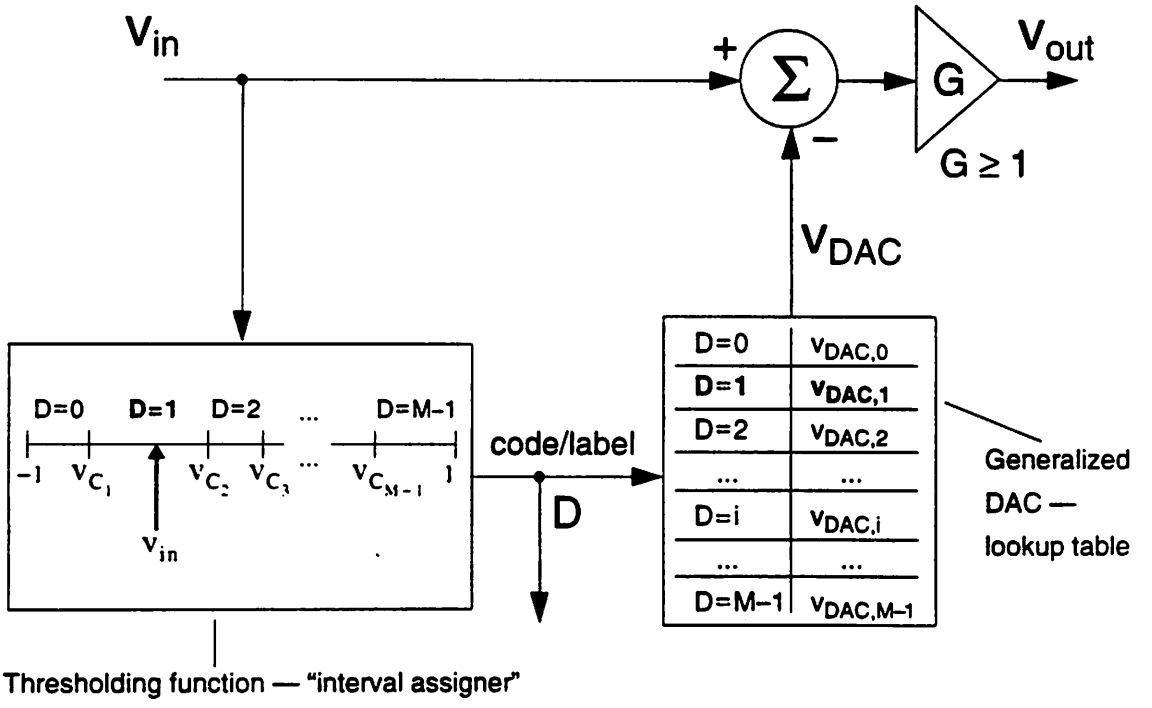


Fig. 4.0 Schematic representation of a generalized stage of a multistage ADC. In the case shown, the input v_{in} lies in the interval corresponding to label D = 1, which in turn maps to DAC value $v_{DAC,1}$.

Basic Definition:

A generalized multistage ADC is formed by a cascade of the building blocks that are defined by (4.3) and shown in Fig. 4.0.

A key assumption is the following: it is always assumed that for an individual stage, if

$$-1 \le v_{in} \le 1 \tag{4.4}$$

then

$$-1 \le v_{out} \le 1 \tag{4.5}$$

i.e., it is assumed that, if V_{in}, the input to a stage, is within the nominal range, then the stage output V_{out} is also within the range. Since in a multistage ADC this output forms the input to the next stage, this implies *that the input to the next stage is within the range of that next stage.* An alternative way to state this is as follows. The basic input-output relation for the stage, expressed in terms of quantities normalized to V_{ref}, i.e.,

$$v_{out} = G\,(v_{in} - v_{DAC,\,D(v_{in})}) - v_{os} \tag{4.6}$$

assumes that $-1 \le v_{in} \le 1$. Clearly, if v_{in} is the output of a previous stage, this therefore imposes the constraint that the output v_{out} of that stage must be within the nominal range $[-1, 1]$; otherwise, the relation defined in (4.6) is not guaranteed to hold. For emphasis, this key assumption is repeated below — this time using non-normalized quantities.

If V_{in} *is the input to a stage, and* V_{out} *is the output of the stage — i.e., the amplified residue — then it is assumed that the condition*

$$-V_{ref} \le V_{in} \le V_{ref} \tag{4.7}$$

implies

$$-V_{ref} \le V_{out} \le V_{ref} \tag{4.8}$$

In any implementation, appropriate strategies and design decisions must be taken to ensure that this holds. Essentially, this is a linearity property: v_{in} and v_{out} are required to lie within the range in which the linear relation (4.6) holds; otherwise, information is lost.

Another viewpoint is to consider each stage of the multistage A/D as having a hard limiter in front of it that limits the input signal to the specified range. As long as the signal is within the range the nonlinearity has no effect. As mentioned previously, in Section 3.2.0, because of the finite, limited range present in any practical implementation of an analog circuit block, there is an implicit hard-limiter, or clipping function associated with the interstage gain block. Therefore, each stage in a multistage ADC may be regarded as having a clipping/limiting function either in front of it or, directly after it. Since the phenomenon of clipping/limiting is *nonlinear*, its effect is to cause the A/D conversion process, which is ultimately a *linear* decomposition into a sum of weights, to break down.

Note that normalizing with respect to V_{ref} and conditions (4.4) and (4.5) above implicitly assume that the nominal range of all stages is the same, which is usually true in pipelined multistage ADC's; clearly, it is not true in subranging architectures. However, the following analysis can be performed without normalization simply by imposing the condition that at each stage i, the amplified residue is within the range of the next stage.

Denote the stages by indices 1, 2, 3, ..., N_S. Let the input to stage i be v_i; thus, v_{i+1} is the output from stage i, and is also the input to stage i+1, as indicated in Fig. 4.1 below.

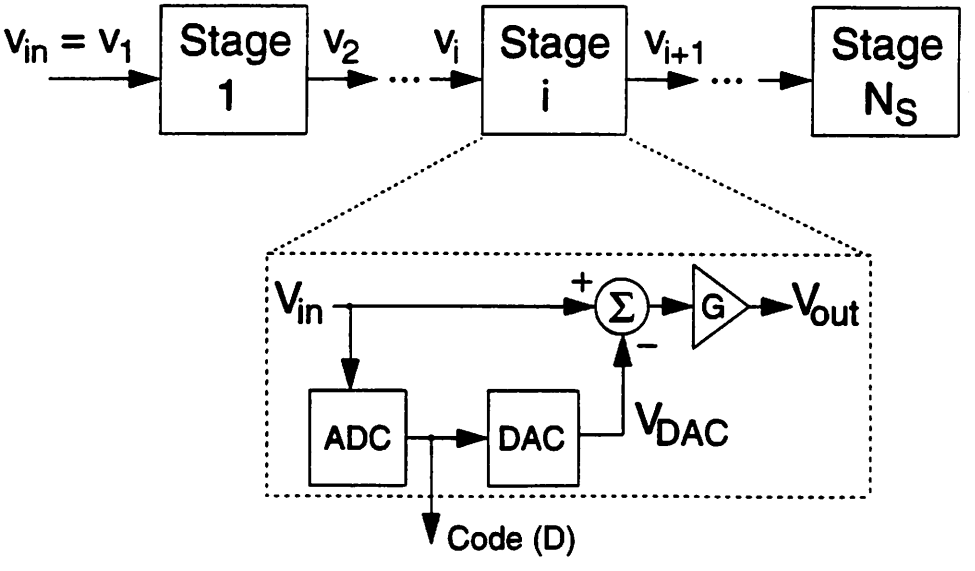


Fig. 4.1 Overall structure of a generalized multistage A/D converter.

Using this notation, the static transfer characteristic for stage i becomes

$$v_{i+1} = G_i (v_i - v_{DAC_i, D(v_i)}) - v_{os_i} \tag{4.9}$$

for $i = 1, \ldots, N_S$, where G_i, v_{os_i} are the gain and offset of stage i, respectively, and $v_{DAC_i, D(v_i)}$ is the DAC output of stage i resulting from input v_i and corresponding digital code or label $D(v_i)$. This notational complexity is intended to allow for the cases where the gains, offsets, and DAC levels are different in different stages. To simplify things somewhat, the notation $v_{DAC_i, D(v_i)}$ is abbreviated to v_{DAC, D_i}, which is defined to be the DAC level corresponding to digital code D_i that results from input v_i to stage i. The above equation (4.9) therefore becomes

$$v_{i+1} = G_i (v_i - v_{DAC, D_i}) - v_{os_i} \tag{4.10}$$

So far, therefore, a general expression (4.10) has been presented for the dc input-output relationship of each stage in a cascade of N_S stages comprising a multistage ADC. Note that this model also forms a useful and quite general building block for *simulation* of multistage converters. As mentioned previously in Section 3.1, for dc analysis purposes, a *recirculating* architecture may be viewed as a multistage ADC *with all stages exactly identical*; specifically, the gains and DAC levels are the same.

Next, (4.10) may be inverted to give the *input* v_i *as a function of the output* v_{i+1}, i.e.,

$$v_i = \frac{v_{i+1} + v_{os_i}}{G_i} + v_{DAC, D_i} \tag{4.11}$$

Thus, starting at the first stage, the input v_{in} $(= v_1)$ to the multistage ADC is given by

$$v_{in} = v_1 = \frac{v_2 + v_{os_1}}{G_1} + v_{DAC, D_1} \tag{4.12}$$

Similarly, v_2 may be expressed as

$$v_2 = \frac{v_3 + v_{os_2}}{G_2} + v_{DAC,D_2} \tag{4.13}$$

and can be substituted back into (4.12) yielding

$$v_1 = \frac{\left(\frac{v_3 + v_{os_2}}{G_2} + v_{DAC,D_2}\right) + v_{os_1}}{G_1} + v_{DAC,D_1} \tag{4.14}$$

$$= v_{DAC,D_1} + \frac{v_{os_1}}{G_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{os_2}}{G_1G_2} + \frac{v_3}{G_1G_2} \tag{4.15}$$

$$= \left(v_{DAC,D_1} + \frac{v_{os_1}}{G_1}\right) + \frac{1}{G_1}\left(v_{DAC,D_2} + \frac{v_{os_2}}{G_2}\right) + \frac{v_3}{G_1G_2} \tag{4.16}$$

Continuing in this vein, expanding the expression for v_3, gives

$$v_{in} = v_1 = \left(v_{DAC,D_1} + \frac{v_{os_1}}{G_1}\right) + \frac{1}{G_1}\left(v_{DAC,D_2} + \frac{v_{os_2}}{G_2}\right) + \frac{1}{G_1G_2}\left(v_{DAC,D_3} + \frac{v_{os_3}}{G_3}\right) + \frac{v_4}{G_1G_2G_3} \tag{4.17}$$

It may be seen that v_{in} is being expressed as a linear combination of the DAC voltages weighted by the interstage gains.

Next, recall that, the *last* stage in a multistage ADC differs from all the others because it does not have an interstage gain. Thus, v_{N_S}, the input to the last stage, can be written as

$$v_{N_S} = v_{DAC,D_{N_S}} + \varepsilon_{N_S} \tag{4.18}$$

where ε_{N_S} is the approximation error in the last stage — i.e., ultimately the quantization error of the entire A/D converter.

Therefore, for a general N_S-stage system, the input v_{in} may be written as

$$v_{in} = \left(v_{DAC,D_1} + \frac{v_{os_1}}{G_1}\right) + \frac{1}{G_1}\left(v_{DAC,D_2} + \frac{v_{os_2}}{G_2}\right) + \cdots$$

$$\cdots + \frac{1}{G_1 G_2 \cdots G_{N_S-2}}\left(v_{DAC,D_{N_S-1}} + \frac{v_{os_{N_S-1}}}{G_{N_S-1}}\right)$$

$$+ \frac{1}{G_1 G_2 \cdots G_{N_S-1}}\left(v_{DAC,D_{N_S}} + \varepsilon_{N_S}\right) \tag{4.19}$$

and so $\tilde{v}_{in}$, the quantized analog approximation of the input (as defined in Section 3.1), is simply the above expression without the last stage error term:

$$\tilde{v}_{in} = \left(v_{DAC,D_1} + \frac{v_{os_1}}{G_1}\right) + \frac{1}{G_1}\left(v_{DAC,D_2} + \frac{v_{os_2}}{G_2}\right) + \frac{1}{G_1 G_2}\left(v_{DAC,D_3} + \frac{v_{os_3}}{G_3}\right) + \cdots$$

$$\cdots + \frac{1}{G_1 G_2 \cdots G_{N_S-2}}\left(v_{DAC,D_{N_S-1}} + \frac{v_{os_{N_S-1}}}{G_{N_S-1}}\right)$$

$$+ \frac{1}{G_1 G_2 \cdots G_{N_S-1}} v_{DAC,D_{N_S}} \tag{4.20}$$

Equation (4.20) is very general model for multistage A/D conversion and may be considered a *canonical* model.

Before proceeding further, the key steps in the preceding analysis are summarized.

1. The input-output relationship between the analog input v_{in} and the gained-up residue v_{out} for a single stage in a multistage ADC is written down.
2. Notation is introduced defining the first (most significant) stage to be stage 1, and the last stage to be stage N_S; the input to stage i is denoted by v_i.
3. The input-output relation for the first stage is rearranged to give an expression for the input to the first stage (v_1) as a function of its amplified residue output (v_2).
4. The same procedure is followed to express v_2, which is equal to the output of the first stage and equal to the input of the second stage, as a function of v_3, the output of the second stage, and the procedure continues to the end of all the stages.

The only assumption is that the basic *linear* input-output relationship holds for each stage. In order for this to be true, *there must be no overranging at any point in the multistage ADC system.* Some implications of the above analysis are now discussed.

The key equation (4.20) is rewritten below.

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{os_1}}{G_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{os_2}}{G_1G_2} + \frac{v_{DAC,D_3}}{G_1G_2} + \frac{v_{os_3}}{G_1G_2G_3} + \cdots$$

$$\cdots + \frac{v_{DAC,D_{N_S-1}}}{G_1G_2\cdots G_{N_S-2}} + \frac{v_{os_{N_S-1}}}{G_1G_2\cdots G_{N_S-1}} + \frac{v_{DAC,D_{N_S}}}{G_1G_2\cdots G_{N_S-1}} \qquad (4.21)$$

The following points are pertinent to the above discussion. Notes 1–4 use (4.21) to derive accuracy requirements for the DAC levels and the gains in multistage ADC's. Notes 5–8 consider A/D conversion as a decomposition into weights and discuss some issues relating to non-binary radix approaches and digital calibration.

Note 1: DAC levels

By examining (4.21), the sensitivity of the ADC transfer characteristic to errors in any of the DAC levels or interstage gain values is immediately apparent. Clearly an error in the first stage DAC voltage v_{DAC,D_1} contributes directly to an error in the ADC output. Progressing to subsequent stages, it may be seen that, as expected, the required accuracy decreases, and *errors in a DAC level at stage i are reduced by the product of the gains up to but not including that stage.* Equivalently stated, the accuracy of a DAC level within stage i must be commensurate with the resolution of this *and* all remaining stages (as mentioned also in Section 3.3.1). Note that in two-step or subranging architectures with no interstage gain, the accuracy requirements are unchanged for all stages.

Note 2: Interstage gain values

For simplicity, assume the offset terms are zero; this assumption does not affect the analysis in any essential way. Equation (4.21) then becomes

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{DAC,D_3}}{G_1G_2} + \ldots + \frac{v_{DAC,D_{N_S}}}{G_1G_2\ldots G_{N_S-1}} \tag{4.22}$$

Let ε_{G_i} be a gain error causing the value of an interstage gain to be $G_i(1-\varepsilon_{G_i})$ instead of G_i: that is, ε_{G_i} is the *relative* error. If ε_{G_i} is small, then the approximation

$$(1-\varepsilon_{G_i}) \approx \frac{1}{1+\varepsilon_{G_i}} \tag{4.23}$$

holds. Consider the effect of gain error ε_{G_1} in the first stage. Using (4.23) and substituting back into (4.22) gives

$$\tilde{v}_{in} = \left[v_{DAC,D_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{DAC,D_3}}{G_1G_2} + \ldots + \frac{v_{DAC,D_{N_S}}}{G_1G_2\ldots G_{N_S-1}} \right]$$

$$+ \left[\frac{\varepsilon_{G_1} v_{DAC,D_2}}{G_1} + \frac{\varepsilon_{G_1} v_{DAC,D_3}}{G_1G_2} + \ldots + \frac{\varepsilon_{G_1} v_{DAC,D_{N_S}}}{G_1G_2\ldots G_{N_S-1}} \right] \tag{4.24}$$

i.e., the ideal term plus the error term due to the gain error ε_{G_1}. Since the normalized DAC voltages $\{ v_{DAC,D_i} \}$ are bounded in magnitude by 1 the error term is at worst

$$\frac{\varepsilon_{G_1}}{G_1} + \frac{\varepsilon_{G_1}}{G_1 G_2} + \ldots + \frac{\varepsilon_{G_1}}{G_1 G_2 \ldots G_{N_S-1}} = \varepsilon_{G_1} \left(\frac{1}{G_1} + \frac{1}{G_1 G_2} + \ldots + \frac{1}{G_1 G_2 \ldots G_{N_S-1}} \right) \quad (4.25)$$

It is apparent from (4.25) that *an error in interstage gain* G_i *affects all stages after stage i, but the effect at any given stage is reduced by the product of the gains up to and including that stage.* Equivalently, the relative accuracy of gain G_i must be commensurate with the resolution of all the remaining stages *after* this stage i (as mentioned also in Section 3.3.0). This may also be seen by performing a first-order Taylor expansion on (4.21). An alternative way to perform this analysis is discussed in [76].

Note 3: Offsets

Referring again to (4.21), it is possible to group the offsets together, which gives

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{DAC,D_3}}{G_1 G_2} + \ldots + \frac{v_{DAC,D_{N_S-1}}}{G_1 G_2 \ldots G_{N_S-2}} + \frac{v_{DAC,D_{N_S}}}{G_1 G_2 \ldots G_{N_S-1}}$$

$$+ \left(\frac{v_{os_1}}{G_1} + \frac{v_{os_2}}{G_1 G_2} + \ldots + \frac{v_{os_{N_S-1}}}{G_1 G_2 \ldots G_{N_S-1}} \right) \quad (4.26)$$

i.e., the overall input-referred offset is given by

$$v_{os_{tot}} = \frac{v_{os_1}}{G_1} + \frac{v_{os_2}}{G_1 G_2} + \ldots + \frac{v_{os_{N_S-1}}}{G_1 G_2 \ldots G_{N_S-1}} \quad (4.27)$$

The fact that the above equations (4.26) and (4.27) can be written in the form indicated demonstrates that offsets in the residue amplifiers in each stage eventually contribute to an overall composite input-referred offset, but *do not affect the linearity of the ADC.* This greatly relaxes the offset requirements on the interstage gain blocks, since in many real applications, an input-referred offset is either not important or can be accounted for in

some subsequent processing. Again, the key assumption in the derivation of (4.20)–(4.26) is restated: the output of each stage in the multistage system must be within the range of the next stage.

Note 4: Error budget

It is clear from (4.26) that in order to achieve a particular overall ADC accuracy of say, n bits, *each error component must contribute significantly less than 1/2 LSB of error.* Furthermore, op amp settling time errors in the gain stages can be approximated as linear gain errors, and need to be included in this error budget.

Note 5: A/D conversion as decomposition into weights that can be calibrated

Viewing A/D conversion as a decomposition into a linear summation of weights makes sense from (4.26). This leads naturally to a non-binary radix approach, as follows. It may be seen that as long as the values of the DAC levels and the interstage gains are known, the input approximation $\tilde{v}_{in}$ can be obtained. Therefore, why not *measure* the DAC weights during a calibration cycle [70], [70], [57]? To avoid overranges, the DAC values cannot be measured directly; however, their *differences* can be measured, i.e.,

$$v_{DAC,1} - v_{DAC,0}, v_{DAC,2} - v_{DAC,1}, \cdots, v_{DAC,M} - v_{DAC,M-1}$$

which may be referred to as the DAC *segments* [17]. Note that this is fundamentally different from the conventional calibration approaches using trimming techniques such as capacitive trim arrays or poly fuses: in terms of the above model of A/D conversion, these traditional approaches modify the DAC values and interstage gains to cause them to be equal to (or at least sufficiently close to) their ideal values. This usually entails some analog circuit complexity. In the approach described here, the DAC and gain values are measured and are accounted for in subsequent digital processing. Thus, the complexity is in the digital domain.

Note 6: Digital calibration by digitizing weights directly

In order to make clearer the notion of stage "weights" in a multistage ADC, it is useful to multiply across in (4.26) by $G_1G_2\ldots G_{N_S-1}$ to give (ignoring offsets for simplicity)

$$(G_1G_2\ldots G_{N_S-1})\,\tilde{v}_{in} = (G_1G_2\ldots G_{N_S-1})\,v_{DAC,D_1} + (G_2G_2\ldots G_{N_S-1})\,v_{DAC,D_2} + \ldots$$
$$+ (G_{N_S-2}G_{N_S-1})\,v_{DAC,D_{N_S-2}} + G_{N_S-1}v_{DAC,D_{N_S-1}} + v_{DAC,D_{N_S}} \quad (4.28)$$

This forms the basis for an approach to digitally self-calibrated multistage ADC's. The key sequence of observations is as follows. First, once the DAC voltages and the gains are known the A/D conversion algorithm can be performed assuming care has been taken to ensure no overranges. Second, it is possible to measure the combination of DAC voltages and gain products directly by starting at the end of the multistage ADC and digitizing these weights directly. As the measurement sequence progresses away from the LSB end, more bits are added to the measurement and the effect of the "weights" is taken into account automatically. Therefore, no multiplications are needed in the calibration!

Note 7: "Resolution" in digitally self-calibrated non-binary ADC systems

The concept of "resolution" needs to be clarified in the context of non-binary multistage ADC's. The following seems a reasonable approach. Define the effective or equivalent resolution n_{eq} of a generalized multistage A/D converter as follows:

$$\frac{\text{Max}\{\tilde{V}_{in}\} - \text{Min}\{\tilde{V}_{in}\}}{\text{Min}\{\Delta\tilde{V}_{in}\}} = 2^{n_{eq}} - 1 \quad (4.29)$$

In some sense, this is the ratio of the full-scale-range to the minimum resolvable signal (i.e., the granularity or minimum quantization level in the amplitude domain). The meaning of $\text{Max}\{\tilde{V}_{in}\}$ is obvious; the strict interpretation is as follows: "take all possible values of $\tilde{V}_{in}$ that the generalized ADC can produce, that is, all possible weighted combinations of the values contained in look-up tables of all the stages (i.e., as illustrated in Fig. 4.0), and then take the maximum." Note that (4.29) does reduce to the expected

result for standard binary-radix converters. For example, in the case of Fig. 3.4, $\text{Max}\{\tilde{V}_{in}\} = +7/8$, $\text{Min}\{\tilde{V}_{in}\} = -7/8$, and $\text{Min}\{\Delta\tilde{V}_{in}\} = 1/4$; thus, (4.29) yields

$$\frac{7/8 - (-7/8)}{1/4} = 2^{n_{eq}} - 1 \Rightarrow n_{eq} = 3$$

as expected.

Note that this effective resolution is not the same as the "Equivalent Number of Bits" (ENOB), which is an ac measurement related to Signal-to-Noise-plus-Distortion Ratio (SNDR) of an A/D converter and is usually less than the "resolution" — as mentioned in Chapter 2.

Issue 8: How much redundancy?

Another question worth considering is: how much "redundancy" is actually needed? For a given worst-case threshold offset, and interstage gain block offset, and also given worst case DAC and gain errors, by using approaches similar to those employed in Sections 3.2 and 3.3, it is possible to compute the minimum amount of redundancy required, and hence the minimum number of stages required.

4.2 APPLICATION TO SPECIFIC A/D CONVERTER EXAMPLES

In this section, various examples of multistage A/D converters are examined, and in each case the ADC operation is presented in terms of the generalized approach developed in Section 4.1. Whenever the parameters and transfer characteristics of an individual ADC stage are given, it is assumed that the stage is a particular case of the general model of Fig. 4.0, and that the corresponding multistage ADC system consists of N_S of these stages cascaded. The examples that follow contain detailed description of static input-output characteristics and digital code assignments; wherever relevant, threshold accuracy and gain accuracy requirements are also discussed. The treatment in this section is at the ADC system/algorithm level: underlying circuit implementations are not described. The objectives here are (i) to review and compare various multistage A/D converter architectures with respect to *parameters* such as number of bits per stage and interstage gain, and *algorithmic features* such as digital correction/redundancy and code assignment scheme, and (ii) to show that a large class of ADC algorithms and topologies may be viewed as special cases of the general model presented in Section 4.1.

In the examples in this section, the emphasis is on multistage A/D converter systems *with* interstage gain — i.e., nondegenerate gain (>1) — that have the same range in all stages. As mentioned previously, the analysis of Section 4.1 can be applied to subranging converters by not normalizing with respect to V_{ref}. However, the main focus of this dissertation is on the class of converters known as "pipeline ADC's", and, in the literature, the term "pipeline ADC" has usually been applied to multistage A/D converters *with* interstage gain.

Before proceeding, it is useful to present some review of binary representation of real numbers. A real number x lying between 0 and 1 (i.e., a *unipolar* number) with binary representation $B = b_1 b_2 b_3 \ldots b_N$, where the bits $\{b_i\}$ have value either 0 or 1, can be written as

$$x = \frac{b_1}{2^1} + \frac{b_2}{2^2} + \frac{b_3}{2^3} + \ldots + \frac{b_N}{2^N} \tag{4.30}$$

$$= \sum_{i=1}^{N} \frac{b_i}{2^i} \tag{4.31}$$

In some sense, the terms in the above summations can be thought of as *binary fractions*. For the purposes of this work, the following representation of a bipolar number x lying between –1 and +1 is used:

$$x = -1 + 2\sum_{i=1}^{N} \frac{b_i}{2^i} \tag{4.32}$$

This is similar to the unipolar number representation but referenced to –1, and with a total range of value 2.

Algorithm Example 1: Unipolar, 1 bit per stage

The first algorithm to be considered resolves one bit per stage, has an interstage gain of two, and operates on unipolar signals. Shown in Fig. 4.2 below is the static transfer characteristic of one individual stage from input V_{in} to amplified residue output V_{out}.

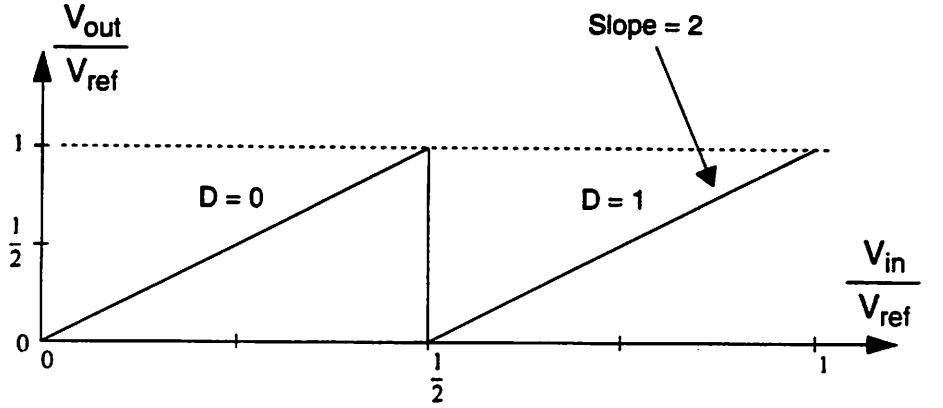


Fig. 4.2 Static transfer characteristic from stage input to gained-up residue output of a pipeline stage — conventional 1 bit/stage architecture.

The stage dc parameters are explicitly listed below.

Input range:	[0, +1] (normalized to V_{ref})
1 ADC threshold level:	{1/2}
2 digital codes:	{0, 1}
2 corresponding DAC levels:	{0, 1/2}
Interstage gain:	G = 2
Number of bits:	k = 1

The individual stage input-output relation can be written as

$$v_{out} = 2\,(v_{in} - v_{DAC,D}) \tag{4.33}$$

The starting point for the analysis is the general expression derived in Section 4.1 for the overall ADC amplitude-quantized analog output $\tilde{v}_{in}$ as a function of the stage DAC voltages and the interstage gains. That equation is repeated below. (The offset is zero.)

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{DAC,D_3}}{G_1 G_2} + \ldots + \frac{v_{DAC,D_{N_S}}}{G_1 G_2 \ldots G_{N_S-1}} \tag{4.34}$$

Since $G_1 = G_2 = \ldots = G_{N_S-1} = 2$ for this example, (4.34) becomes

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{2} + \frac{v_{DAC,D_3}}{2^2} + \frac{v_{DAC,D_4}}{2^3} + \ldots + \frac{v_{DAC,D_{N_S}}}{2^{N_S-1}} \tag{4.35}$$

$$= \sum_{i=1}^{N_S} \frac{v_{DAC,D_i}}{2^{i-1}} \tag{4.36}$$

Next, the functional relationship or mapping between the digital codes $\{D_i\}$ and the corresponding DAC levels is needed. By examining the codes $\{D_i\}$ and the corresponding DAC levels $\{v_{DAC,i}\}$ it is clear that the relation is

$$v_{DAC,D_i} = \frac{D_i}{2} \tag{4.37}$$

where $D_i = 0, 1$.

Substituting for v_{DAC,D_i} in (4.36) gives

$$\tilde{v}_{in} = \sum_{i=1}^{N_S} \frac{D_i \cdot \frac{1}{2}}{2^{i-1}} \tag{4.38}$$

$$= \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.39}$$

which is the familiar expression for the binary representation of a unipolar number.

As mentioned previously in Section 3.1 of Chapter 3, a multistage architecture with one bit per stage is closely related to the recycling "algorithmic" architecture [92], [77]. In fact, both have identical dc transfer characteristics for an individual stage, and issues related to coding and static errors apply equally to both. The block diagram for an algorithmic architecture using a unipolar signal range is shown in Fig. 4.3.

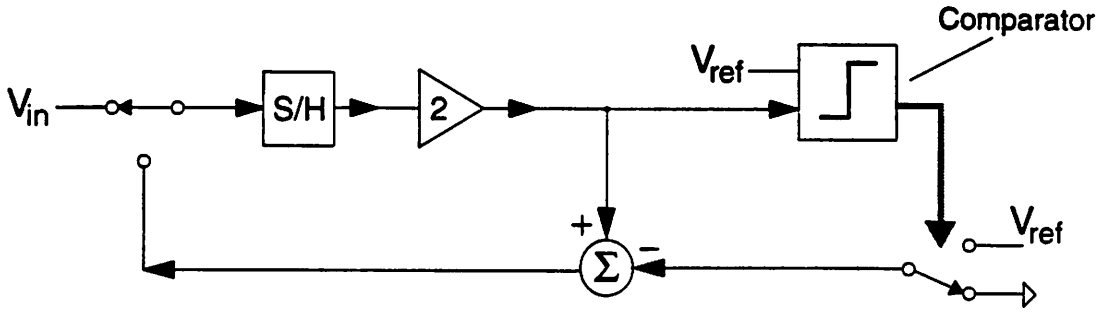


Fig. 4.3 Block diagram of an algorithmic A/D converter with unipolar range. The comparator threshold level is at V_{ref}.

The transfer characteristic for one cycle of this algorithmic topology is given by

$$V_{out} = (2V_{in} - V_{DAC}) = 2(V_{in} - V_{DAC}/2) \tag{4.40}$$

where $V_{DAC} \in \{0, V_{ref}\}$ — i.e., the "DAC" values here are twice those given in the parameter list for Fig. 4.2, and so the stage input-output relation is the same.

Algorithm Example 2: Bipolar, 1 bit per stage

In this case, as in the previous example, the interstage gain is 2 and there is one threshold level. Here, however, the input and output signal range is bipolar, namely, from $-V_{ref}$ to $+V_{ref}$. The stage input-output relationship is shown in Fig. 4.4.

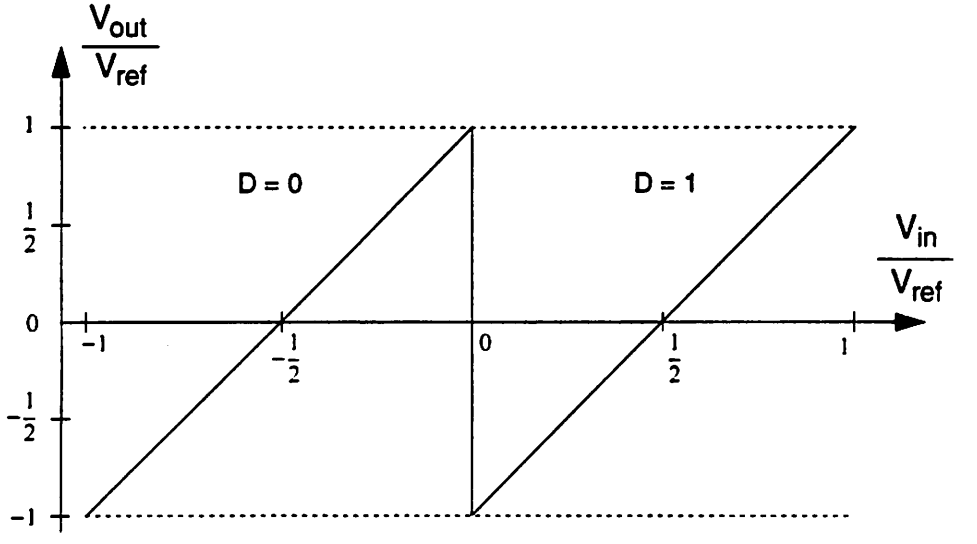


Fig. 4.4 Static transfer characteristic from stage input to gained-up residue of a bipolar 1 bit per stage architecture.

The static parameters are as follows.

Input range:	[−1, +1] (normalized to V_{ref})
1 ADC threshold level:	{0}
2 digital codes:	{0, 1}
2 corresponding DAC levels:	{−1/2, 1/2}
Interstage gain:	G = 2
Number of bits:	k = 1

The general expression for the quantized analog output of the multistage ADC is

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{DAC,D_3}}{G_1 G_2} + \ldots + \frac{v_{DAC,D_{N_S}}}{G_1 G_2 \ldots G_{N_S-1}} \tag{4.41}$$

Here, as in Algorithm Example 1, the individual stage input-output relation is

$$v_{out} = 2\,(v_{in} - v_{DAC,D}) \tag{4.42}$$

Therefore, as in the previous case, the expression for the final amplitude-quantized output of the N_S-stage ADC is

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{2} + \frac{v_{DAC,D_3}}{2^2} + \frac{v_{DAC,D_4}}{2^3} + \ldots + \frac{v_{DAC,D_{N_S}}}{2^{N_S-1}} \tag{4.43}$$

$$= \sum_{i=1}^{N_S} \frac{v_{DAC,D_i}}{2^{i-1}} \tag{4.44}$$

Again, a relation between the digital codes $\{D_i\}$ and the corresponding DAC levels is required. By inspection of the codes $\{D_i\}$ and the corresponding DAC levels $\{v_{DAC,i}\}$ it is apparent that the relation is given by

$$v_{DAC,D_i} = -\frac{1}{2} + D_i \tag{4.45}$$

where $D_i = 0, 1$.

Substituting into (4.44) yields

$$\tilde{v}_{in} = \sum_{i=1}^{N_S} \frac{-\frac{1}{2} + D_i}{2^{i-1}} \tag{4.46}$$

$$= \sum_{i=1}^{N_S} \frac{-\frac{1}{2}}{2^{i-1}} + \sum_{i=1}^{N_S} \frac{D_i}{2^{i-1}} \tag{4.47}$$

$$= -\sum_{i=1}^{N_S} \frac{1}{2^i} + 2\sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.48}$$

$$= -\left(1 - \frac{1}{2^{N_S}}\right) + 2\sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.49}$$

i.e.,

$$\tilde{v}_{in} = -1 + \frac{1}{2^{N_S}} + 2\sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.50}$$

which is the usual expression for the binary representation of a *bipolar* number between −1 and +1. The $(1/2^{N_S})$ term is simply a horizontal shift of the transfer characteristic by 1/2 LSB. (Note that for a bipolar signal range, 1 LSB at n-bit resolution has value $2/2^n$.)

An algorithmic version of this topology is shown in Fig. 4.5 below [77], [101], [95].

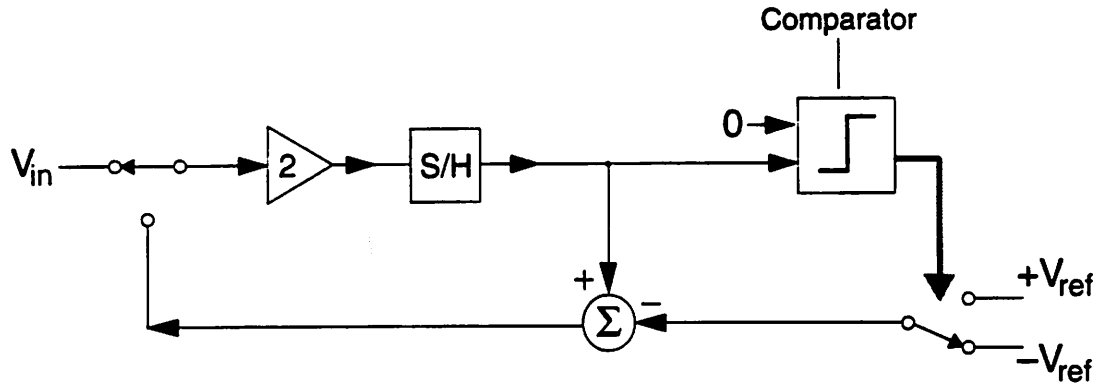


Fig. 4.5 Block diagram of an algorithmic A/D converter with bipolar signal range; the comparator threshold level is at zero.

Algorithm Example 3: 1 bit per stage with redundancy

In this example, the stage has a gain of two but has *two* threshold levels, resulting in a factor of two of redundancy. The transfer characteristic is shown below in Fig. 4.6.

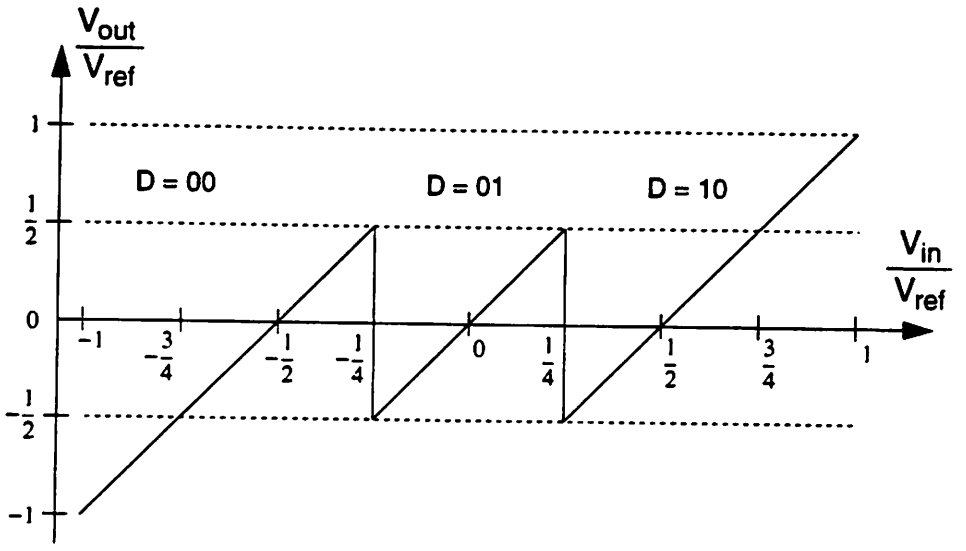


Fig. 4.6 Static transfer characteristic from stage input to gained-up residue of a pipeline stage with interstage gain of two, two threshold levels, and three DAC levels.

The static parameters are listed below:

Input range:	[–1, +1] (normalized to V_{ref})
2 ADC threshold levels:	{–1/4, 1/4}
3 digital codes:	{0, 1, 2} or {00, 01, 10}
3 corresponding DAC levels:	{–1/2, 0, 1/2}
Interstage gain:	G = 2
Number of bits:	k = 2 (net, after digital correction: k = 1)

This scheme has been employed in a high-speed CMOS pipeline implementation [74], [75] described earlier in Implementation Example 4 in Section 3.4.1. Closely-related schemes have been used in various other ADC topologies including a number of CMOS algorithmic/recirculating architectures [56], [33]–[35], and a bipolar ripple-through scheme [54] described in Implementation Example 6 in Section 3.4.1. (The terminology "1.5 bits per stage" is sometimes used to describe this stage [75]; note, however, that with N_S stages, the overall converter resolution is $N_S + 1$ bits.)

The individual stage input-output relation is

$$v_{out} = 2\,(v_{in} - v_{DAC,D}) \tag{4.51}$$

As in the previous example, the expression for the quantized analog output is

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{2} + \frac{v_{DAC,D_3}}{2^2} + \frac{v_{DAC,D_4}}{2^3} + \ldots + \frac{v_{DAC,D_{N_S}}}{2^{N_S-1}} \tag{4.52}$$

$$= \sum_{i=1}^{N_S} \frac{v_{DAC,D_i}}{2^{i-1}} \tag{4.53}$$

Next, the relation between the digital codes and the corresponding DAC levels is needed. By looking at the codes $\{D_i\}$ and the corresponding DAC levels $\{v_{DAC,D_i}\}$, it may be seen that the appropriate relation is

$$v_{DAC,D_i} = -\frac{1}{2} + \frac{D_i}{2} = \frac{1}{2}(-1 + D_i) \tag{4.54}$$

where $D_i = 0, 1, 2$.

Substituting this expression into (4.53) gives

$$\tilde{v}_{in} = \sum_{i=1}^{N_S} \frac{-\frac{1}{2} + \frac{D_i}{2}}{2^{i-1}} \tag{4.55}$$

$$= \sum_{i=1}^{N_S} \frac{-\frac{1}{2}}{2^{i-1}} + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.56}$$

$$= -\sum_{i=1}^{N_S} \frac{1}{2^i} + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.57}$$

$$= -\left(1 - \frac{1}{2^{N_S}}\right) + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.58}$$

Thus, the relation between $\tilde{v}_{in}$, the amplitude-quantized output of the ADC, and the digital codes from each stage is given by

$$\tilde{v}_{in} = -1 + \frac{1}{2^{N_S}} + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.59}$$

As before, the $1/2^{N_S}$ term in (4.59) is simply a shift of 1/2 LSB. Note that although this equation appears similar to the standard binary representation in (4.50) of a bipolar number between −1 and +1, there are a number of key differences. First, the above equation does not have a multiplier of 2 in front of the summation, as is present in, for example, (4.50). Second, the $\{D_i\}$ in (4.59) are not simply binary bits, since each may assume *three* possible values: $\{0, 1, 2\}$: thus, the $\{D_i\}$ may be considered *2-bit* numbers. Therefore, the summation in (4.59) is seen to correspond to *addition with one bit overlapped between adjacent stages.* This differs from the summations in (4.39) in Algorithm Example 1, and (4.50) in Algorithm Example 2, respectively, as follows. In those cases, which employ standard binary coding with no redundancy, and in which the $\{D_i\}$ can have only 2 possible values, 0 or 1, the "addition" operation is trivial: the digital outputs from successive stages are simply placed to the right of each other — i.e., corresponding to terms of the form $D_i/2^i$ — and there is no interaction between the digital output codes of different stages. Here, however, there *is* interaction between the output codes of adjacent stages.

The usefulness of the generalized formulation presented in Section 4.1 is now apparent: for a given stage transfer characteristic and given dc parameters, the approach of Section 4.1 is systematic way to derive precisely *how* to combine the stage output codes $\{D_i\}$ in order to generate the correct amplitude-quantized ADC output. Furthermore, it facilitates investigation of the use of different static parameters — for example, shifted ADC thresholds or DAC levels — or different code assignment schemes, and it allows examination of how such possible alternative schemes affect the complexity of the precise way the stage outputs codes are combined to produce the final ADC output.

Algorithm Example 4: Bipolar, 2 bits per stage, redundancy with "naive" coding

The previous example showed that by reducing the interstage gain by a factor of two from the classical multistage A/D conversion scheme, and by choosing the threshold values, DAC values, and code assignment according to Fig. 4.6, a relatively straightforward "overlapped addition" gives the final ADC output code. The question now arises: what about the 2-bit stage with the "naive" coding scheme described earlier in Section 3.2.2? The stage transfer characteristic is illustrated again in Fig. 4.7 below; it was shown in Section 3.2.2 that this coding scheme and dc input-output relation did not work "nicely".

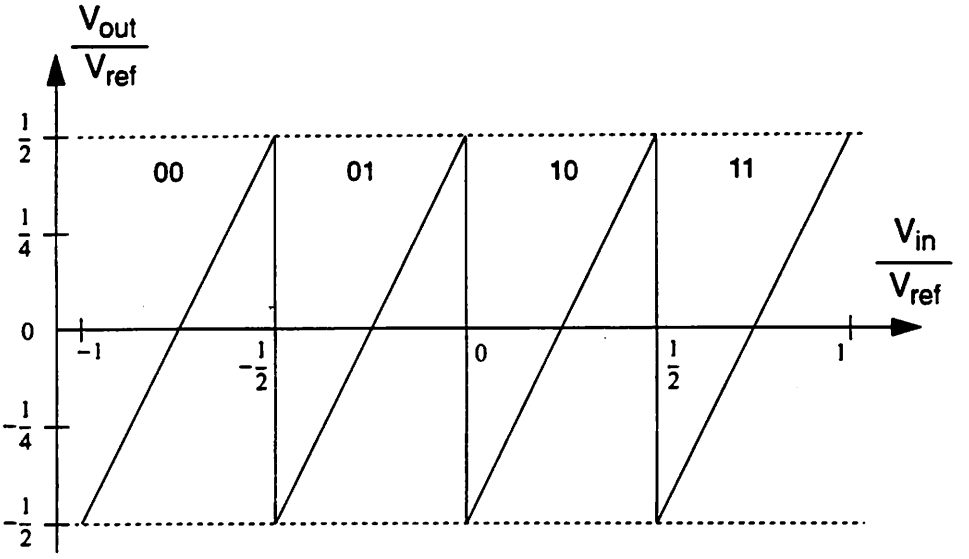


Fig. 4.7 Transfer curve of a 2-bit stage with reduced interstage gain and "naive" coding.

The static parameters associated with this transfer characteristic are listed below.

Input range:	[–1, +1] (normalized to V_{ref})
3 ADC threshold levels:	{–1/2, 0, 1/2}
4 digital codes:	{0, 1, 2, 3} or {00, 01, 10, 11}
4 corresponding DAC levels:	{–3/4, –1/4, 1/4, 3/4}
Interstage gain:	G = 2
Number of bits:	k = 2 (net, after digital correction: k = 1)

Following the same procedure as in the previous Algorithm Examples, the relationship between the stage output codes $\{D_i\}$ and the final ADC output $\tilde{v}_{in}$ may be derived. The individual stage input-output relation is given by

$$v_{out} = 2\,(v_{in} - v_{DAC,D}) \tag{4.60}$$

As in the previous case, writing the expression for the amplitude-quantized analog output voltage gives

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{2} + \frac{v_{DAC,D_3}}{2^2} + \frac{v_{DAC,D_4}}{2^3} + \ldots + \frac{v_{DAC,D_{N_S}}}{2^{N_S-1}} \tag{4.61}$$

$$= \sum_{i=1}^{N_S} \frac{v_{DAC,D_i}}{2^{i-1}} \tag{4.62}$$

Next, the relation between each digital code D_i and the corresponding DAC level is needed. By looking at the codes $\{D_i\}$ and the corresponding DAC levels $\{v_{DAC,D_i}\}$, it may be seen that the relation is given by

$$v_{DAC,D_i} = -\frac{3}{4} + \frac{D_i}{2} = \frac{1}{2}\left(-\frac{3}{2} + D_i\right) \tag{4.63}$$

where $D_i = 0, 1, 2, 3$. Substituting this expression for v_{DAC,D_i} into (4.62) yields

$$\tilde{v}_{in} = \sum_{i=1}^{N_S} \frac{\frac{1}{2}\left(-\frac{3}{2} + D_i\right)}{2^{i-1}} \tag{4.64}$$

$$= -\frac{3}{2}\sum_{i=1}^{N_S} \frac{1}{2^i} + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.65}$$

$$= -\frac{3}{2}\left(1 - \frac{1}{2^{N_S}}\right) + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.66}$$

It is clear that the first term on the right-hand side in the above equation (4.66) is a substantial offset — not simply a 1/2 LSB shift of the transfer characteristic as in some previous examples. However, it is a well-defined *digital* offset — and at least potentially

could be accounted for in subsequent digital processing after the ADC. The second term is the conventional expression. So, the conclusion here is that the scheme of Fig. 4.7 *can* be made to work, but at the expense of messy and non-intuitive coding involving a somewhat complicated (digital) offset.

Focussing now on the value of the "offset" term in (4.66), it is of interest to determine what its *digital binary* representation is. Rearranging (4.66) gives

$$\tilde{v}_{in} = -1 - \frac{1}{2} + \frac{3}{2^{N_S+1}} + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.67}$$

$$= -1 - \frac{1}{2} + \left(\frac{1}{2^{N_S+1}} + \frac{2}{2^{N_S+1}}\right) + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.68}$$

$$= -1 - \frac{2}{4} + \frac{1}{2} \cdot \left(\frac{2}{2^{N_S+1}}\right) + \left(\frac{2}{2^{N_S+1}}\right) + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.69}$$

The -1 on the right-hand side of (4.69) is expected, since this is a bipolar system. The next three terms on the right-hand side constitute the offset. These terms correspond to $-1/4$ Full Scale + 1/2 LSB + 1 LSB. (Note that the overall resolution is $N_S + 1$.) Thus, for $N_S = 3$, that is, 4-bit resolution, the binary representation of these three terms is

$$-0100 + (1/2)\,\text{LSB} + 0001 = -0011 + (1/2)\,\text{LSB}$$

The example given earlier in Section 3.2.2 as part of the discussion of "naive" coding is now recalled. Consider a multistage ADC composed of 3 of these stages; let the input be $V_{in} = V_1 = +\varepsilon$, where ε is a positive voltage of magnitude much smaller than 1 LSB at the overall ADC resolution. Referring to Fig. 4.7, the outputs from the three stages are

$D_1 = 2 = 10 \Rightarrow V_{DAC,1} = 0.25 \Rightarrow V_2 = 2(\varepsilon - 0.25) = -0.5 + 2\varepsilon,$

$D_2 = 1 = 01 \Rightarrow V_{DAC,2} = -0.25 \Rightarrow V_3 = 2[(-0.5 + 2\varepsilon) - (-0.25)] = -0.5 + 4\varepsilon,$

$D_3 = 1 = 01.$

The raw output from the 3-stage ADC is obtained from the addition:

$$\begin{array}{r} 10 \\ 01 \\ \quad 01 \\ \hline 1011 \end{array}$$

Now, combining this raw output, 1011, with the offset calculated above gives

$$-0011 + (1/2)\,\mathrm{LSB} + 1011 = 1000 + (1/2)\,\mathrm{LSB}$$

which *is* the answer expected — i.e., for $V_{in} = +\varepsilon$ on a bipolar scale, the digital output code would normally be expected to be within 1/2 LSB of 1000 in a 4-bit system. So, for this 3-stage case, the algorithm may be viewed as taking the raw output from the overlapped addition, and then adding a digital offset of −0011.

This may be verified also by working directly from (4.66). For example, if the number of stages N_S is 3, the offset term is given by $-\frac{3}{2}\left(1-\frac{1}{2^3}\right) = -\frac{21}{16}$. Thus, in the above example, the overall amplitude-quantized output from the 3-stage ADC can be obtained directly from the general expression (4.66), that is,

$$\tilde{v}_{in} = -\frac{3}{2}\left(1-\frac{1}{2^{N_S}}\right) + \sum_{i=1}^{N_S} \frac{D_i}{2^i} \tag{4.70}$$

$$= -\frac{21}{16} + \frac{2}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} = \frac{1}{16} \tag{4.71}$$

which is +1/2 LSB at 4-bits on a bipolar scale, as expected.

Of course, other more elegant coding schemes may exist. However, the motivation for presenting the above analysis is to show that there is nothing inherently "wrong" with the stage transfer characteristic shown in Fig. 4.7. It can be made to work and has all the usual advantages of reduced sensitivity to threshold and subtractor offsets. However, it does not satisfy the practical requirement for a "nice" coding scheme.

Next a number of examples are discussed showing different coding possibilities for a multibit (k>1) stage.

Algorithm Example 5: Unipolar with multiple (k) bits per stage

This example considers the simple case of a multistage ADC with a unipolar input signal, k bits per stage, and no overrange detection.

The corresponding dc transfer characteristic is shown in Fig. 4.8 below for k = 3.

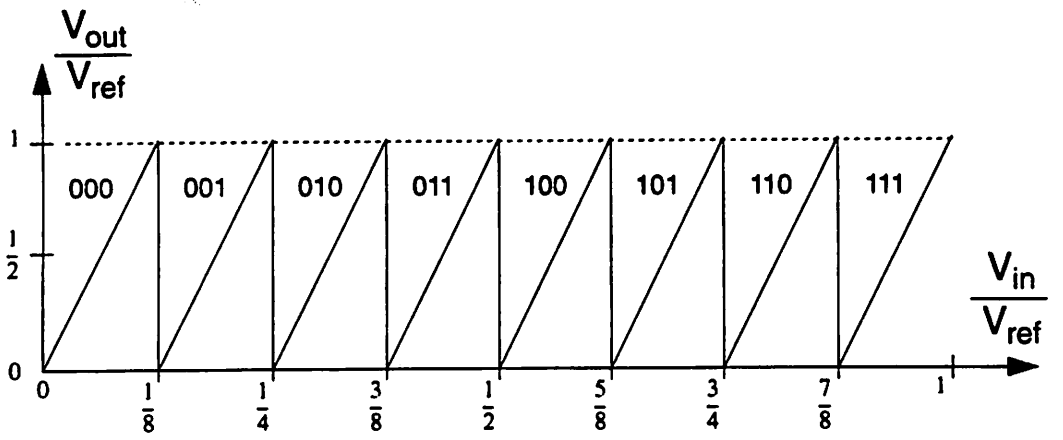


Fig. 4.8 Static transfer characteristic from stage input to gained-up residue of a 3-bit stage with no redundancy; interstage gain is 8.

The following table gives the values of the static parameters for the case k = 3.

Input range:	[0, +1] (normalized to V_{ref})
7 ADC threshold levels:	{1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8}
8 digital codes:	{0, 1, 2, 3, 4, 5, 6, 7} or
	{000, 001, 010, 011, 100, 101, 110, 111}
8 corresponding DAC levels:	{0, 1/8, 1/4, 3/8, 1/2, 5/8, 3/4, 7/8}
Interstage gain:	G = 8
Number of bits:	k = 3

In the following analysis, for generality, it is assumed that the ADC consists of N_S k-bit (rather than 3-bit) stages. The stage input-output relation is therefore given by

$$v_{out} = 2^k (v_{in} - v_{DAC,D}) \tag{4.72}$$

Substituting the interstage gains into the general equation — i.e., (4.34) — yields

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{2^k} + \frac{v_{DAC,D_3}}{2^k \cdot 2^k} + \frac{v_{DAC,D_4}}{2^k \cdot 2^k \cdot 2^k} + \ldots + \frac{v_{DAC,D_{N_S}}}{(2^k)^{N_S - 1}} \tag{4.73}$$

$$= \sum_{i=1}^{N_S} \frac{v_{DAC,D_i}}{(2^k)^{i-1}} \tag{4.74}$$

By examining the codes $\{D_i\}$ and the corresponding DAC levels $\{v_{DAC,D_i}\}$, it is clear that the relation between them is given by

$$v_{DAC,D_i} = \frac{D_i}{2^k} \tag{4.75}$$

where $D_i = 0, 1, \ldots, 2^k - 1$.

Substituting the above expression for v_{DAC,D_i} into (4.74) yields

$$\tilde{v}_{in} = \sum_{i=1}^{N_S} \frac{D_i \cdot \frac{1}{2^k}}{(2^k)^{i-1}} \tag{4.76}$$

$$= \sum_{i=1}^{N_S} \frac{D_i}{(2^k)^i} \tag{4.77}$$

Note that in the above equation, the codes $\{D_i\}$ are k-bit numbers, and the codes from successive stages are scaled down by 2^k: *this simply corresponds to string concatenation — exactly as expected, i.e., the digital codes from successive stages are placed to the right of each other.* This example emphasizes that the "normal" or "classical" multistage A/D conversion algorithm (no redundancy) may be regarded as a special case of the more general formulation presented above and, furthermore, a special case in which the mapping from digital codes to the underlying quantized analog voltage is trivial.

Algorithm Example 6: Bipolar, k-bit stage, 1 bit of redundancy, 2^k-2 thresholds

The transfer characteristic of a 3-bit stage with 6 decision levels is shown below.

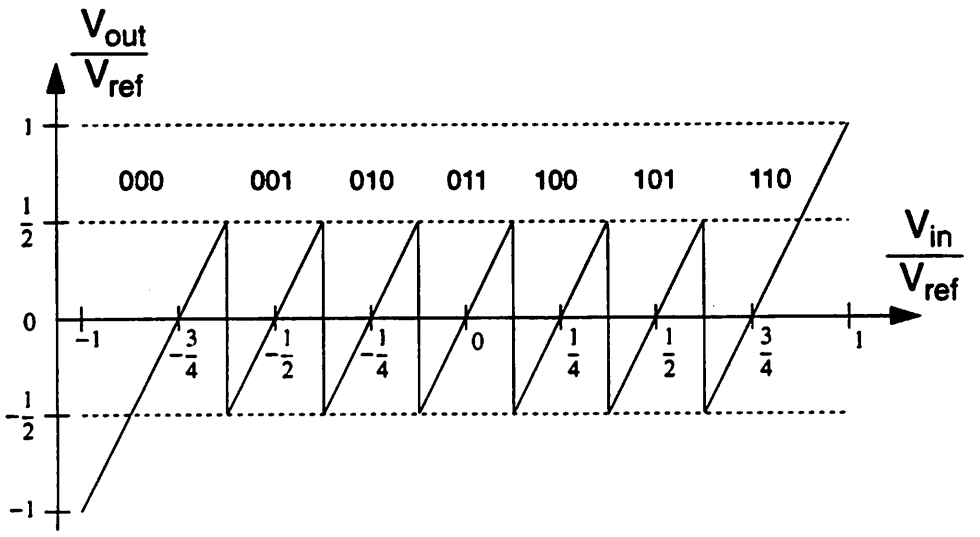


Fig. 4.9 Static transfer characteristic from stage input to amplified residue of a 3-bit pipeline stage with 6 (= 2^3-2) threshold levels, and interstage gain of 4.

This stage has the following parameters.

Input range:	[−1, +1] (normalized to V_{ref})
6 ADC threshold levels:	{−5/8, −3/8, −1/8, 1/8, 3/8, 5/8}
7 digital codes:	{0, 1, 2, 3, 4, 5, 6} or
	{000, 001, 010, 011, 100, 101, 110}
7 corresponding DAC levels:	{−3/4, −1/2, −1/4, 0, 1/4, 1/2, 3/4}
Interstage gain:	G = 4
Number of bits:	k = 3 (net, after digital correction: k = 2)

As discussed later, the input-output relationship shown in Fig. 4.9 is an example of a more general set of transfer characteristics with k bits per stage and an interstage gain of 2^{k-1}, i.e., one bit of redundancy. Accordingly, in this case, the complete derivation is given, assuming for generality that the ADC consists of N_S k-bits stages.

The individual stage input-output relation is

$$v_{out} = 2^{k-1}(v_{in} - v_{DAC,D}) \tag{4.78}$$

Once again the starting point is

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{DAC,D_3}}{G_1G_2} + \ldots + \frac{v_{DAC,D_{N_S}}}{G_1G_2\ldots G_{N_S-1}} \tag{4.79}$$

Here, this equation becomes

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{2^{k-1}} + \frac{v_{DAC,D_3}}{2^{k-1}\cdot 2^{k-1}} + \frac{v_{DAC,D_4}}{2^{k-1}\cdot 2^{k-1}\cdot 2^{k-1}} + \ldots + \frac{v_{DAC,D_{N_S}}}{(2^{k-1})^{N_S-1}} \tag{4.80}$$

$$= \sum_{i=1}^{N_S} \frac{v_{DAC,D_i}}{(2^{k-1})^{i-1}} \tag{4.81}$$

By looking at the codes $\{D_i\}$ and the corresponding DAC levels $\{v_{DAC,D_i}\}$, it may be seen that for this case, the appropriate relation between them is given by

$$v_{DAC,D_i} = -\left(\frac{2^{k-1}-1}{2^{k-1}}\right) + \frac{D_i}{2^{k-1}} \tag{4.82}$$

where

$$D_i = 0, 1, \ldots, 2^k - 2 \tag{4.83}$$

This yields

$$\tilde{v}_{in} = \sum_{i=1}^{N_S} \frac{v_{DAC,D_i}}{(2^{k-1})^{i-1}} \tag{4.84}$$

$$= \sum_{i=1}^{N_S} \frac{-\left(\frac{2^{k-1}-1}{2^{k-1}}\right) + \frac{D_i}{2^{k-1}}}{(2^{k-1})^{i-1}} \tag{4.85}$$

$$= \sum_{i=1}^{N_S} \frac{-(2^{k-1}-1)}{(2^{k-1})^i} + \sum_{i=1}^{N_S} \frac{D_i}{(2^{k-1})^i} \tag{4.86}$$

$$= -\left(1 - \frac{1}{(2^{k-1})^{N_S}}\right) + \sum_{i=1}^{N_S} \frac{D_i}{(2^{k-1})^i} \tag{4.87}$$

i.e.,

$$\tilde{v}_{in} = -1 + \frac{1}{(2^{k-1})^{N_S}} + \sum_{i=1}^{N_S} \frac{D_i}{(2^{k-1})^i} \tag{4.88}$$

Since the stage digital outputs $\{D_i\}$ are k-bit numbers and the interstage gain is 2^{k-1}, the summation in (4.88) corresponds to *addition with one bit overlapped between adjacent stages*. The overall resolution n_{tot} of the multistage ADC therefore corresponds to $N_S - 1$ stages having k-bit resolution, but with one bit used for redundancy/digital correction, followed by one stage (the last stage) that cannot be corrected and so contributes k non-redundant bits. This gives

$$n_{tot} = (N_S - 1)(k-1) + k = N_S(k-1) + 1 \tag{4.89}$$

(Clearly, if the stages are not all identical, a different total resolution can be obtained.) It is worthwhile to reemphasize that, in some sense, the above analysis defines a family of multistep A/D conversion algorithms, of which Fig. 4.6 and Fig. 4.9 are examples [75]. These algorithms have the following key characteristics:

- k bits per stage, with one bit of redundancy;
- $2^k - 2$ threshold levels (and hence $2^k - 2$ comparators) in each stage;
- interstage gain equal to 2^{k-1}.

Furthermore, all the usual advantages of digital correction / overrange detection are present, and implementation is relatively straightforward.

The transfer characteristics from analog input v_{in} to amplified residue output v_{out} for the two cases corresponding to k = 2 and k = 3 are illustrated in Fig. 4.10 below.

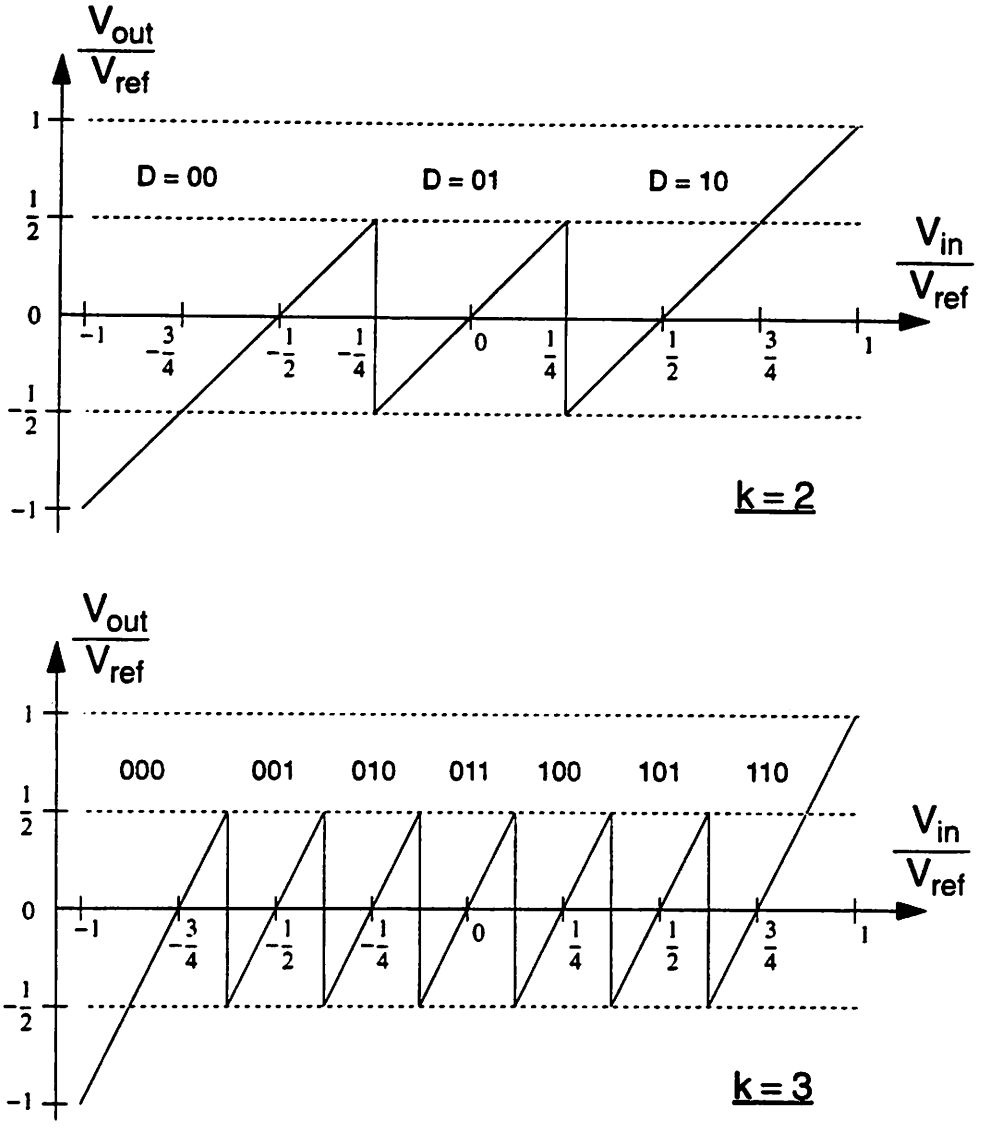


Fig. 4.10 Examples of transfer characteristics from stage input to gained-up residue with k bits per stage, interstage gain of 2^{k-1} (implying a factor of 2 of redundancy), and 2^k-2 thresholds.

In this architecture, an offset error in any threshold of up to $\pm (V_{ref}/2^k)$ does not cause an overrange. Note that for the two threshold levels closest to zero, this implies that an offset in the threshold level *up to the value of the threshold itself* is tolerated. (See also Fig. 3.25 and the accompanying discussion.)

Thus, in summary, the offset requirements on the combination of comparators and interstage gain block have been relaxed to a k-bit level — i.e., the *stage* resolution, which is typically 2–4 bits. Clearly, this is a significantly less stringent accuracy requirement than the *overall* resolution of the multistage ADC.

In order to obtain a digital output code of *all ones*, and so resolve the uppermost portion of the range (which may be necessary in some applications), an extra threshold level is required in the last stage. The last stage therefore has $2^k - 1$ comparators and decision levels, as described in [75].

Note that the above stage transfer characteristics are almost exactly identical to those described earlier in Section 3.2.3, specifically, the 2-bit example in Fig. 3.22 and the 3-bit case in Fig. 3.23. The difference is that the top decision level has been omitted. This last threshold is not necessary to guarantee that the no-overrange condition holds. In order to get an all-ones digital code from the ADC, however, the *last* stage in the pipeline should include the top comparator.

The prototype chip described in Chapter 8 uses 4 stages with $k = 3$ bits per stage. For all stages except the last, there are 6 comparators, the interstage gain is 4, and the static transfer characteristic from stage input to gained-up residue is as shown in Fig. 4.9. Note that the DAC reference levels and ADC threshold levels are offset from each other by 1/2 LSB of the stage — i.e., by $V_{ref}/8$. The last stage does not require a DAC or interstage gain function — it consists only of a 3-bit ADC with 7 comparators — the threshold level of the extra comparator is at $7/8V_{ref}$ and the digital output code when V_{in} is greater than $7/8V_{ref}$ is 111.

Summary of Preceding Examples

The key message from the preceding examples and discussion is that the details of coding assignment in multistage A/D converters are important and also sometimes quite subtle — especially for schemes involving redundancy/overrange detection. However, when considered from the general viewpoint of Section 4.1, any scheme may be examined systematically and verified. A number of input-output transfer characteristics for an individual stage within a multistage ADC have been presented. In these cases, the following procedure has been followed.

1. Given the stage input-output relation, substitute the interstage gain values into the general equation, which, for the usual case of nominally zero offsets, is given by

$$\tilde{v}_{in} = v_{DAC,D_1} + \frac{v_{DAC,D_2}}{G_1} + \frac{v_{DAC,D_3}}{G_1G_2} + \ldots + \frac{v_{DAC,D_{N_S}}}{G_1G_2\ldots G_{N_S-1}}$$

2. Write down a relation between the DAC value v_{DAC,D_i} and code D_i.

3. Substitute back into the equation for $\tilde{v}_{in}$, and simplify to yield an equation relating the analog output to the digital codes from each stage.

Examples of Non-Binary Radix Schemes

So far in this chapter, the focus has been exclusively on binary codes and binary representations with interstage gains being powers of 2. In those cases, the relationships between the code and the DAC value are quite straightforward and result in (at least in most cases) quite simple expressions. However, it is also feasible not to use a binary representation: in such cases, the DAC levels do not have a simple binary equivalent. The following are two examples in which the underlying decomposition kernel is not binary.

Algorithm Example 7: Bipolar, 1 bit per stage, gain less than 2

This example is perhaps of theoretical interest only but is conceptually possible. It shows the application of reduced interstage gain to the 1-comparator per stage case. The transfer characteristic is shown in Fig. 4.11 below.

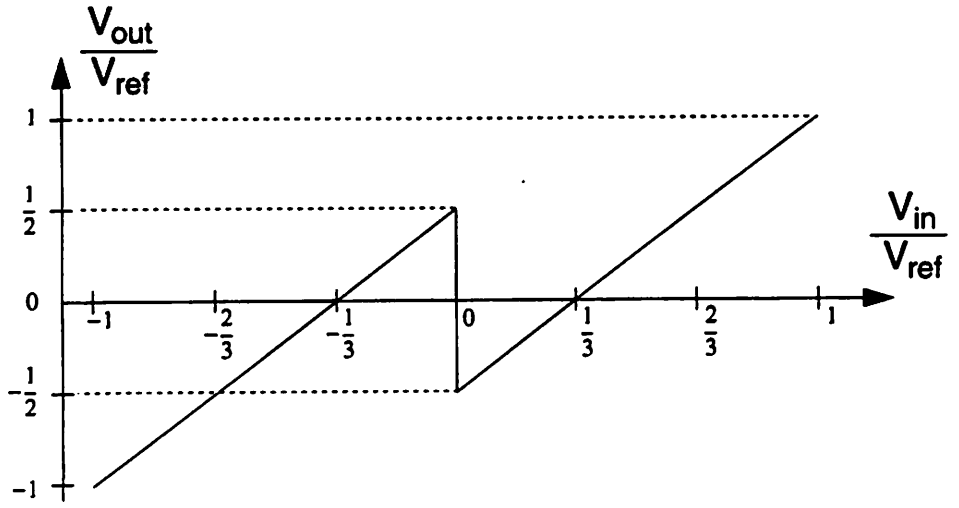


Fig. 4.11 Static transfer characteristic from stage input to gained-up residue of pipeline stage using a single threshold level and a non-binary interstage gain.

Input range:	[−1, +1] (normalized to V_{ref})
1 ADC threshold level:	{0}
2 digital codes:	{00, 01}
2 corresponding DAC levels:	{−1/3, 1/3}
Interstage gain:	G = 3/2 = 1.5
Number of bits:	See text

It is clear from the transfer curve that the threshold level and subtractor offset requirements are not stringent. Specifically, the threshold can vàry up to $\pm V_{ref}/3$. The mapping of digital codes to quantized analog voltage is complicated and has to be determined by using (4.21) directly. The equivalent resolution can be computed using the approach mentioned in Note 7 in Section 4.1.

Algorithm Example 8: Digitally self-calibrated — 1 bit per stage, non-binary radix

Finally in this selection of multistage ADC algorithms, a non-binary scheme used in a digitally self-calibrated architecture [57] is briefly described. This also uses one decision level and a gain less than 2. The input-output relationship of an individual stage is shown below in Fig. 4.12.

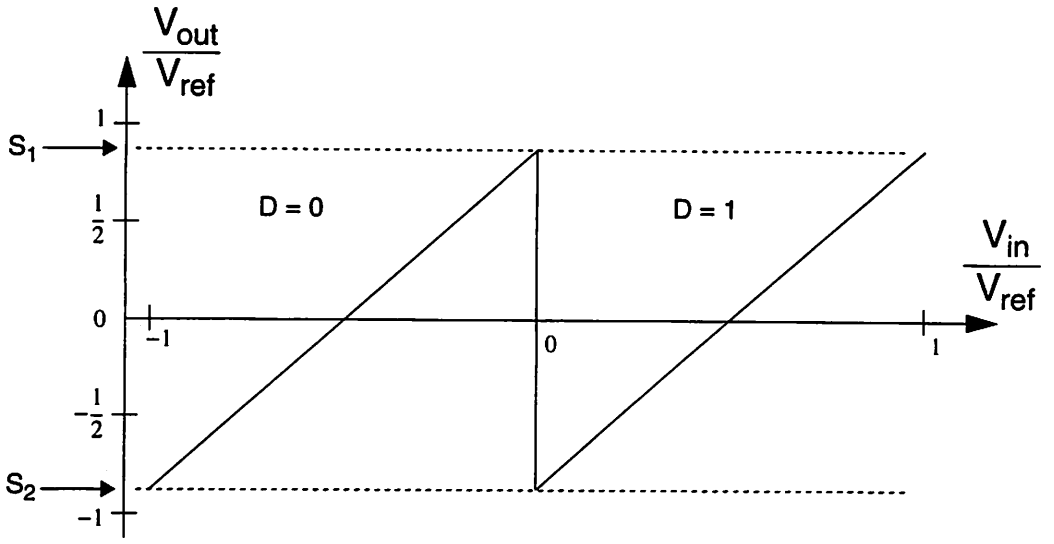


Fig. 4.12 Static transfer characteristic from input to amplified residue of non-binary stage [57].

Here, the interstage gain and DAC levels are implicitly defined in terms of the quantities S_1 and S_2 shown in Fig. 4.12. In order to determine the effective gain and DAC levels, the transfer characteristic is examined as follows.

When $v_{in} < 0$, the digital code is $D = 0$, and

$$v_{out} = (S_1 - S_2)\, v_{in} + S_1 \tag{4.90}$$

When $v_{in} > 0$, the digital code is $D = 1$, and

$$v_{out} = (S_1 - S_2)\, v_{in} + S_2 \tag{4.91}$$

Rearranging gives, for $D = 0$,

$$v_{out} = (S_1 - S_2)\, (v_{in} - (\frac{-S_1}{S_1 - S_2})) \tag{4.92}$$

and, for D = 1

$$v_{out} = (S_1 - S_2)\,(v_{in} - (\frac{-S_2}{S_1 - S_2})) \tag{4.93}$$

It is apparent that the interstage gain is given by $G = (S_1 - S_2)$ and the mapping from codes to DAC levels is as follows:

$$D = 0 \Rightarrow v_{DAC} = (-S_1)/(S_1 - S_2)$$

$$D = 1 \Rightarrow v_{DAC} = (-S_2)/(S_1 - S_2)$$

From this, the table listing of dc parameters can be written:

Input range:	[–1, +1] (normalized to V_{ref})
1 ADC threshold level:	{0}
2 digital codes:	{0, 1}
2 corresponding DAC levels:	$\{-S_1/(S_1-S_2), -S_2/(S_1-S_2)\}$
Interstage gain:	$G = (S_1-S_2)$
Number of bits:	See text

Note that if $S_1 = -S_2 = (1-\gamma)$, then $v_{DAC,1} = -v_{DAC,0} = 1/2$, and $G = 2(1-\gamma)$. Thus, the interstage gain is no longer a power of two: the A/D conversion process is using a non-binary radix.

Note that the relationship between DAC levels and the digital code of the stage is

$$v_{DAC,D} = \frac{-S_1 + D\,(S_1 - S_2)}{(S_1 - S_2)} \tag{4.94}$$

which gives

$$(S_1 - S_2)\,v_{DAC,D} = -S_1 + D\,(S_1 - S_2) \tag{4.95}$$

i.e.,

$$G v_{DAC,D} = -S_1 + DG \tag{4.96}$$

As mentioned earlier in Note 6 in Section 4.1, multiplication operations are not

necessary in the calibration — the relative weights are accounted for automatically. Note that the comparator requirements here are quite stringent. In the implementation described in [57] the comparator is quite complex.

4.3 SUMMARY

At this point, having gone through all the details of the generalized analysis and the examples, it is worthwhile to restate the key motivation. In summary, the aim of this chapter is to present a useful and quite general way to analyze and think about multistage A/D conversion algorithms. This systematic approach allows detailed examination and verification of particular schemes, and also removes the "magic" or "mystery" from the digital error correction and non-binary radix schemes. Furthermore, it permits analysis of errors and calibration techniques. The above is an analytical framework — a systematic way to unify and compare multistage A/D conversion schemes.

APPENDIX

4.A FURTHER EXAMPLES OF MULTISTAGE ALGORITHMS

Some more examples are given in this appendix of multistage A/D converter algorithms from the point of view of the generalized approach presented in Section 4.1.

Algorithm Example 9: Bipolar with multiple (k) bits per stage

Shown below is the transfer characteristic for a stage that resolves k = 3 bits and uses no digital correction or redundancy — i.e., a bipolar version of Algorithm Example 5.

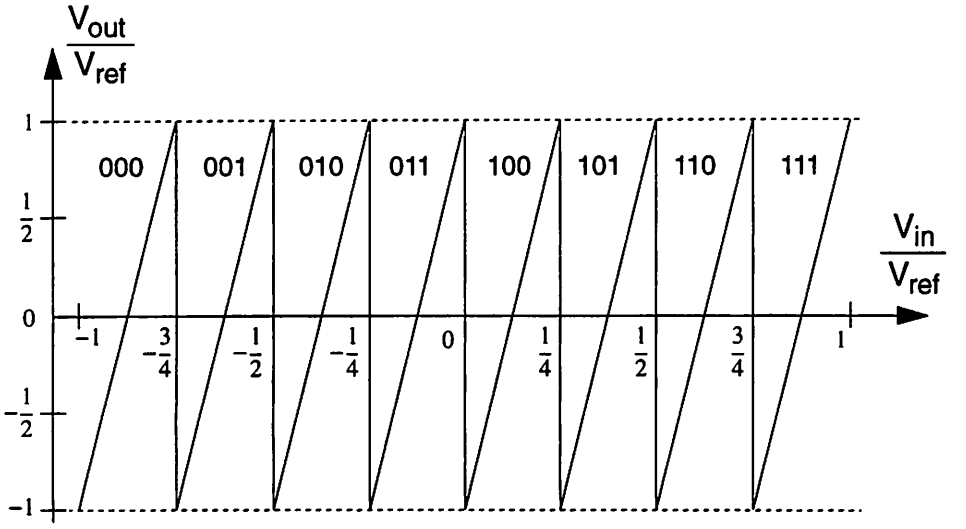


Fig. 4.13 Static transfer characteristic from input to amplified residue output of a bipolar, 3-bit stage with interstage gain of 8.

This stage has the following parameters.

Input range:	[–1, +1] (normalized to V_{ref})
7 ADC threshold levels:	{–3/4, –1/2, –1/4, 0, 1/4, 1/2, 3/4}
8 digital codes:	{0, 1, 2, 3, 4, 5, 6, 7} or
	{000, 001, 010, 011, 100, 101, 110, 111}
8 corresponding DAC levels:	{–7/8, –5/8, –3/8, –1/8, 1/8, 3/8, 5/8, 7/8}
Interstage gain:	G = 8
Number of bits:	k = 3

The derivation in terms of the generalized model is similar to previous cases. From the transfer characteristic and parameter listing the relationship between the DAC voltages and the digital codes is seen to be

$$v_{DAC,D_i} = -1 + \frac{(2D_i + 1)}{2^k} \tag{4.97}$$

where $D_i = 0, 1, \ldots, 2^k - 1$, and the final expression for the amplitude-quantized analog output is

$$\tilde{v}_{in} = -1 + \frac{1}{(2^k)^{N_S}} + 2\sum_{i=1}^{N_S} \frac{D_i}{(2^k)^i} \tag{4.98}$$

Algorithm Example 10: Bipolar, 3-bit stage, 1 bit of redundancy, 7 threshold levels

The input-output relationship for the 3-bit stage with interstage gain reduced from 8 to 4 and shifted ADC and DAC levels previously mentioned in Section 3.2.3 is shown below.

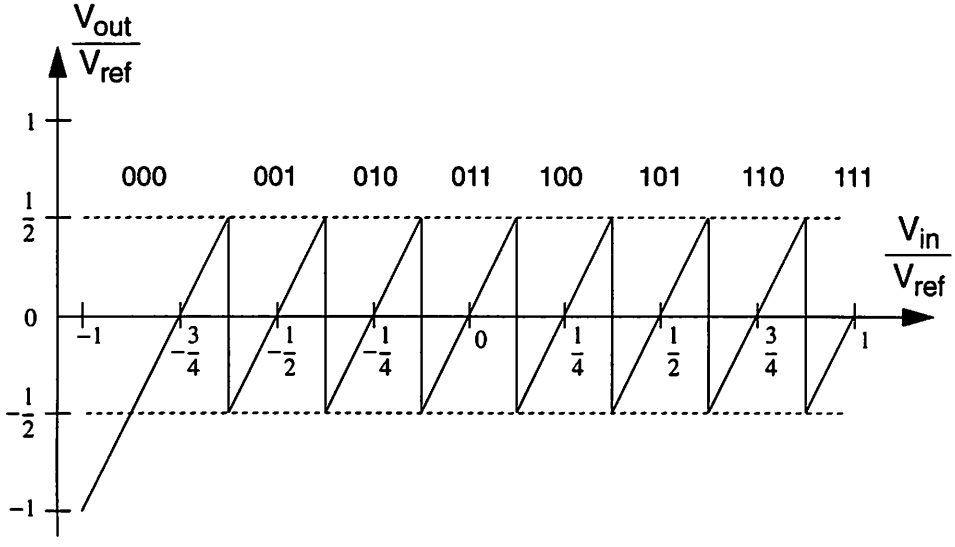


Fig. 4.14 Static transfer characteristic from input to amplified residue output of a bipolar, 3-bit stage with digital error correction — used in [71], [73].

For this example, the stage has the following parameters.

Input range:	[–1, +1] (normalized to V_{ref})
7 ADC threshold levels:	{–5/8, –3/8, –1/8, 1/8, 3/8, 5/8, 7/8}
8 digital codes:	{0, 1, 2, 3, 4, 5, 6, 7} or
	{000, 001, 010, 011, 100, 101, 110, 111}
8 corresponding DAC levels:	{–3/4, –1/2, –1/4, 0, 1/4, 1/2, 3/4, 1}
Interstage gain:	G = 4
Number of bits:	k = 3 (net, after digital correction: k = 2)

As mentioned previously, the coding for this case "works" nicely — a simple overlapped addition.

Algorithm Example 11: Bipolar, 3-bit stage, 1 bit of redundancy, 8 threshold levels

A slight modification to the scheme of Algorithm Example 10 is shown in Fig. 4.15 below. This is similar to the scheme used in [78]–[80].

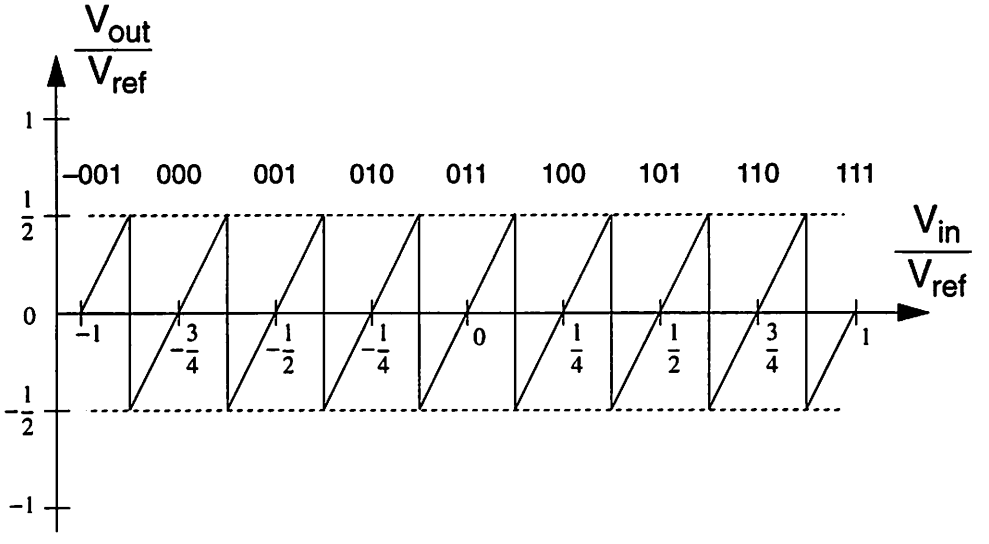


Fig. 4.15 Static transfer characteristic from stage input to gained-up residue of 3-bit pipeline stage using 8 decision levels.

In this case an extra threshold level has been added at $(-7/8)\ V_{ref}$ giving a total of 8 decision levels (and hence comparators). The static parameters are listed below.

Input range:	[–1, +1] (normalized to V_{ref})
8 ADC threshold levels:	{–7/8, –5/8, –3/8, –1/8, 1/8, 3/8, 5/8, 7/8}
9 digital codes:	{–1, 0, 1, 2, 3, 4, 5, 6, 7} or
	{–001, 000, 001, 010, 011, 100, 101, 110, 111}
9 corresponding DAC levels:	{–1, –3/4, –1/2, –1/4, 0, 1/4, 1/2, 3/4, 1}
Interstage gain:	G = 4
Number of bits:	k = 3 (net, after digital correction: k = 2)

The above scheme employs an overlapped addition to generate the final digital code. The presence of a negative number (–1) complicates things a little. It does mean, however, that the range is extended slightly, since a code is now possible that is more negative than all 0's.

To see this, consider a multistage ADC composed of a cascade of 3 of these stages. Let the input be $v_{in} = v_1 = -1 + \varepsilon$, where ε is a positive voltage of magnitude much smaller than 1 LSB at the overall ADC resolution. (All voltages are normalized to V_{ref}). The outputs from the three stages are as follows:

$$D_1 = -1 = -001 \Rightarrow v_{DAC,1} = -1 \quad \Rightarrow v_2 = 4[-1+\varepsilon-(-1)] = 4\varepsilon,$$

$$D_2 = 3 = 011 \quad \Rightarrow v_{DAC,2} = 0 \quad \Rightarrow v_3 = 4(4\varepsilon) = 16\varepsilon,$$

$$D_3 = 3 = 011.$$

The final output is then obtained from an overlapped addition as given below:

$$\begin{array}{r} -001 \\ \quad 011 \\ \qquad 011 \\ \hline -0000001 \end{array}$$

i.e., the output code is –1. This could be used as an underrange signal.

Algorithm Example 12: Bipolar, 3-bit stage, 1 bit of redundancy, 8 threshold levels

Note that if the coding in the above example is modified to that shown below in Fig. 4.16 below, a significant digital offset results.

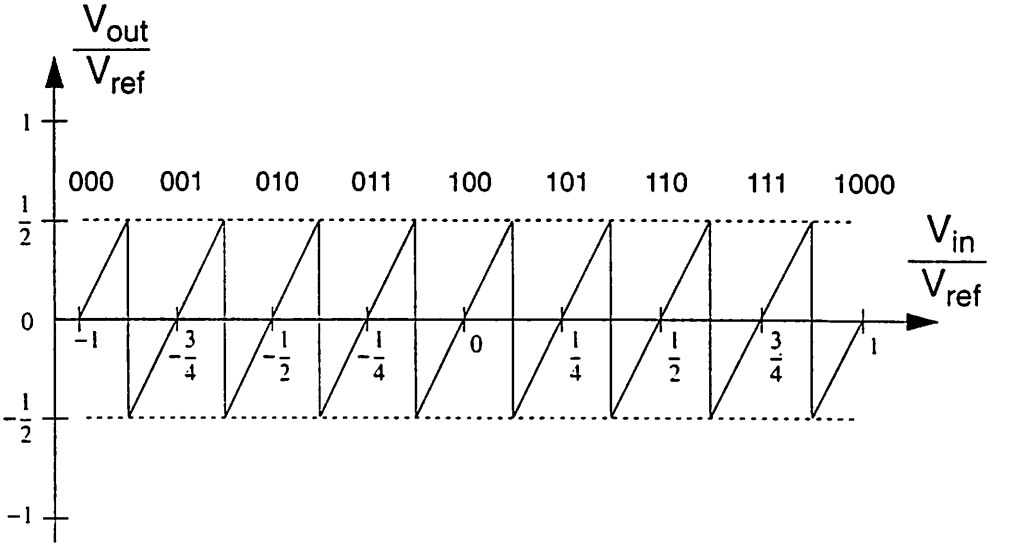


Fig. 4.16 Static transfer characteristic from stage input to gained-up residue of pipeline stage.

In this case, the stage has the following parameters.

Input range:	[−1, +1] (normalized to V_{ref})
8 ADC threshold levels:	{−7/8, −5/8, −3/8, −1/8, 1/8, 3/8, 5/8, 7/8}
9 digital codes:	{0, 1, 2, 3, 4, 5, 6, 7, 8} or
	{000, 001, 010, 011, 100, 101, 110, 111, 1000}
9 corresponding DAC levels:	{−1, −3/4, −1/2, −1/4, 0, 1/4, 1/2, 3/4, 1}
Interstage gain:	G = 4
Number of bits:	k = 3 (net, after digital correction: k = 2)

The presence of an offset is clear if the behavior is examined when the input is 0. As in Algorithm Example 4, the offset is well-defined and may be computed systematically, but in general it is undesirable since it adds unnecessary complexity to the digital coding.

Chapter 5 Optimization of Settling Time in Switched-Capacitor Gain Stages

5.0 INTRODUCTION AND MOTIVATION

This chapter considers the settling time optimization of a signal processing building block used extensively in mixed-signal CMOS IC's: the switched-capacitor (SC) gain stage. In order to motivate the analysis, a number of common configurations of the SC gain block are reviewed in this section. Then, in Section 5.1, a general analysis and optimization of settling time is presented for the case of a SC gain stage employing a single-stage op amp topology. Some specific examples are presented in Section 5.2, and finally, in Section 5.3, the results and discussion are summarized, and some conclusions are drawn.

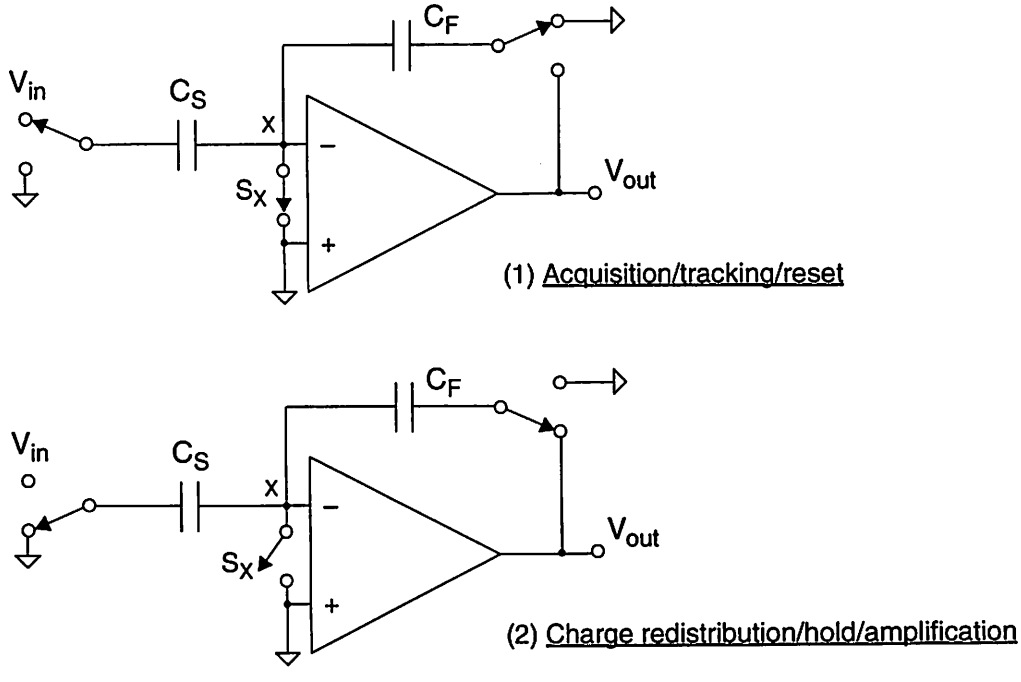


Fig. 5.0 Basic schematic for a switched-capacitor gain stage. The sampling capacitance (C_S) and feedback capacitance (C_F) are indicated. Closed-loop gain (G) is given by (C_S/C_F). The op amp is assumed to be an operational transconductance amplifier (OTA).

The basic schematic of a SC gain stage is shown above in Fig. 5.0. As described in

Chapter 3, this block and variations of it are used widely in pipeline and algorithmic A/D converters. Three remarks are now made that apply in general to the discussion and analysis throughout this chapter.

1. The op amps are assumed to be operational transconductance amplifiers (OTA's).
2. All of these circuit blocks operate using two-phase nonoverlapping clocks; these clocks need not necessarily have a 50% duty cycle.
3. For simplicity, the figures usually indicate single-ended implementations; however, it is to be understood that very often fully-differential signal paths are used in actual IC implementations.

Another configuration of the SC gain block, which is an alternative to that shown in Fig. 5.0, is illustrated in Fig. 5.1 below. This circuit is also used widely.

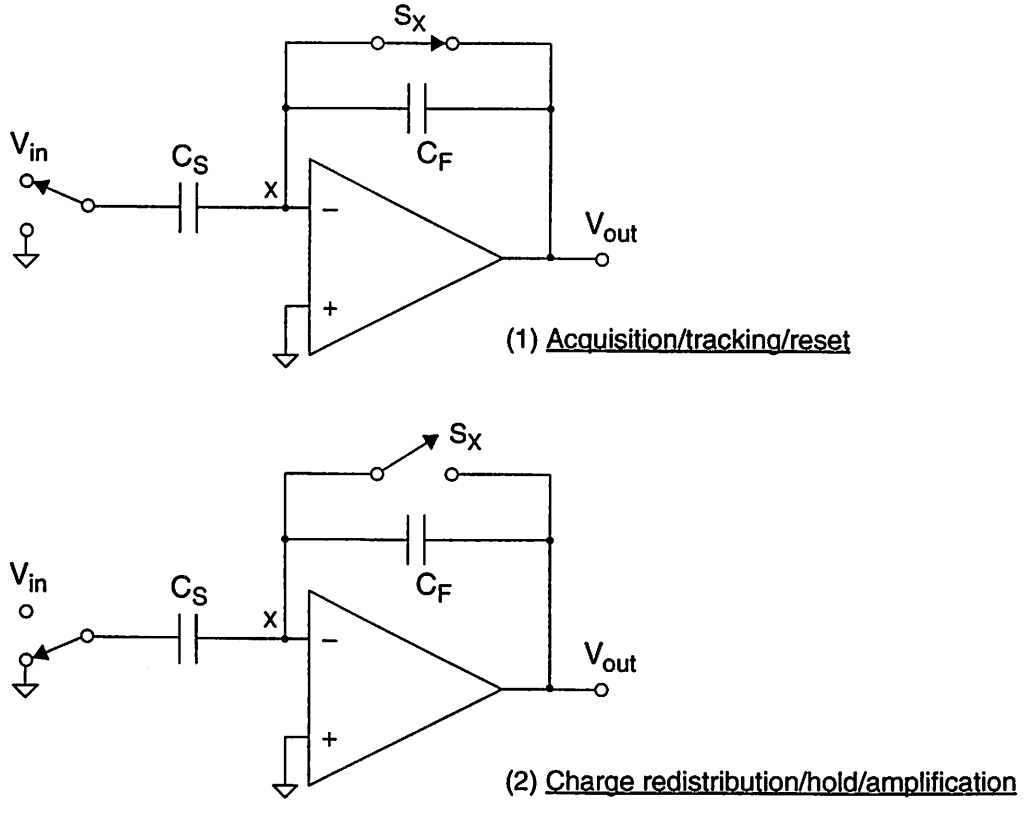


Fig. 5.1 Schematic of an alternative topology for a switched-capacitor gain stage. Closed-loop gain (G) is again given by (C_S/C_F).

The basic operation of both of the blocks shown in Fig. 5.0 and Fig. 5.1 is identical. Assume that the op amps have infinite open-loop gain and zero offset. First, consider the reset/signal acquisition/track mode — phase 1. In the top diagram of Fig. 5.0 the op amp inputs are shorted out, and in the top diagram of Fig. 5.1 the op amp is in unity gain feedback: both configurations cause the voltage V_X at the summing node X to be zero. Note that, as mentioned also in Section 3.5.0, the output node reset circuitry is omitted in Fig. 5.0. The charge Q_X at the summing node is

$$Q_X = C_S(0 - V_{in}) = -C_S V_{in} \tag{5.0}$$

After the sampling switch S_X opens, charge is conserved at node X. The switch configuration during the amplification or charge redistribution mode — phase 2 — is shown in the lower diagrams in Fig. 5.0 and Fig. 5.1. The ideal op amp forces a virtual ground at node X, and the charge, which has now been redistributed to capacitor C_F, can be written as

$$Q_X = C_F(0 - V_{out}) = -C_F V_{out} \tag{5.1}$$

Equating (5.0) and (5.1) gives the input-output relationship for both gain stage topologies:

$$V_{out} = \left(\frac{C_S}{C_F}\right) V_{in} \tag{5.2}$$

However, although the basic functionality is the same, there are some important circuit-level differences between these two blocks. The configuration of Fig. 5.1 requires that the amplifier is stable in unity-gain feedback, since the input and output are shorted together during the reset phase. This is a much more stringent stability constraint than applies for the circuit in Fig. 5.0, since the loop gain in the closed-loop configuration of Fig. 5.0 is reduced from the op amp open-loop gain by the feedback factor (f) due to the capacitive divider effect. Specifically, the feedback factor in the circuit of Fig. 5.0 is given by $f = C_F / (C_F + C_{IP} + C_S)$, where C_{IP} is the parasitic input capacitance of the amplifier.

Another constraint present in the configuration of Fig. 5.1 is that the input and output common-mode levels of the op amp are the same. However, in applications where voltage swing is critical, it is often advantageous to have the flexibility to employ different common-mode levels at the input and output. Of course, some power is consumed in order to generate the input and output common-mode reference levels.

The traditional advantage of the circuit in Fig. 5.1 is that the op amp dc offset is cancelled. However, the offset or *pedestal error* caused by charge injection from the sampling switch S_X is not cancelled, and in high-speed SC circuits that require switches significantly larger than minimum size for bandwidth reasons, this charge-injection offset tends to be the dominant component in the overall stage input-referred offset. From now on, the focus is on the technique of Fig. 5.0.

Note that throughout this chapter, and elsewhere in this dissertation also, it is assumed that the appropriate *delayed* clock edges are used such that signal-dependent charge injection is eliminated — i.e., it is assumed the technique of "bottom-plate sampling" is employed [77], [80]. To further motivate this discussion, an example of a CMOS implementation of a differential SC gain stage is shown in Fig. 5.2 with appropriate clock waveforms indicated.

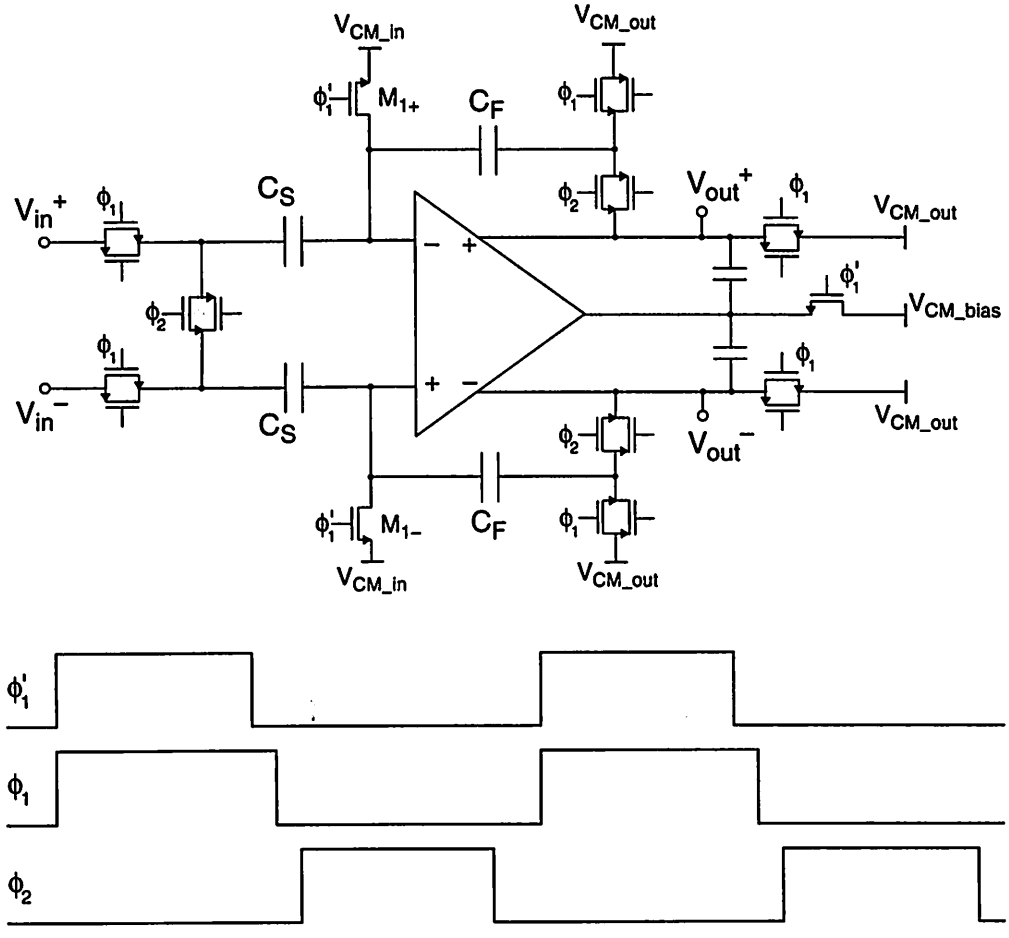


Fig. 5.2 Differential SC gain stage with clock phases indicated. V_{CM_in}, V_{CM_out}, and V_{CM_bias} are dc bias levels generated elsewhere. Sampling occurs at the instant that transistors M_{1+} and M_{1-} are turned off — i.e., on the high-to-low transition of ϕ_1'. (Note that whenever CMOS switches are shown, it is implied that if clock ϕ is driving the NMOS device, then the PMOS device is driven by the inverse of that clock — i.e., $\overline{\phi}$.)

A typical choice for the op amp is a fully-differential, folded-cascode or non-folded (or unfolded) cascode, as shown in Fig. 5.3 below.

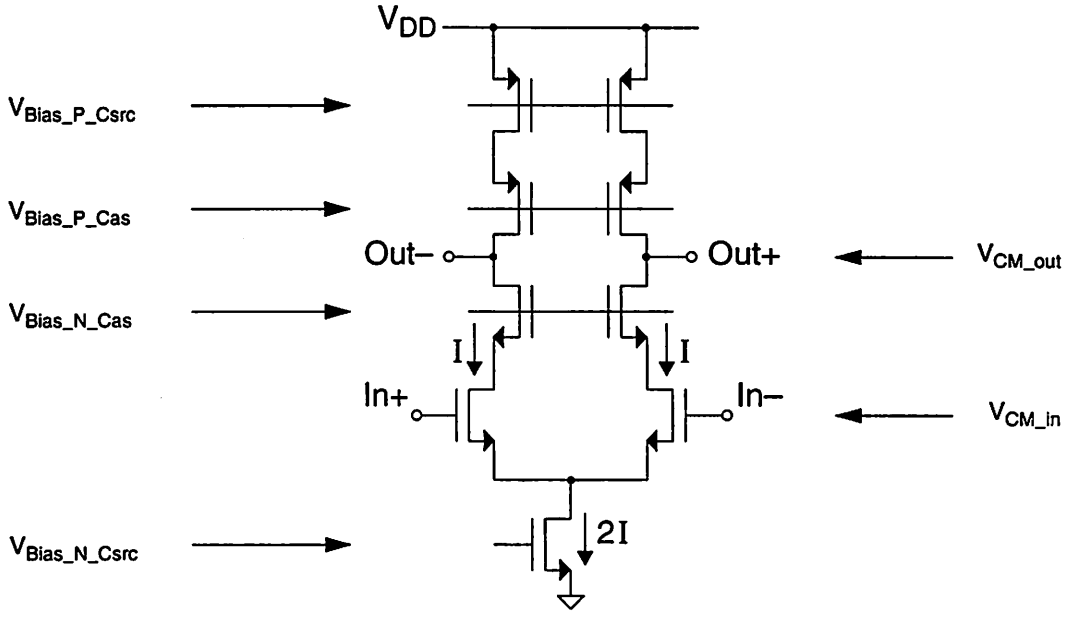


Fig. 5.3 Basic schematic of a fully-differential non-folded cascode op amp with bias voltages indicated.

Some other SC building blocks are now described. A *unity-gain* sample-and-hold (S/H) using a single capacitor is shown in Fig. 5.4.

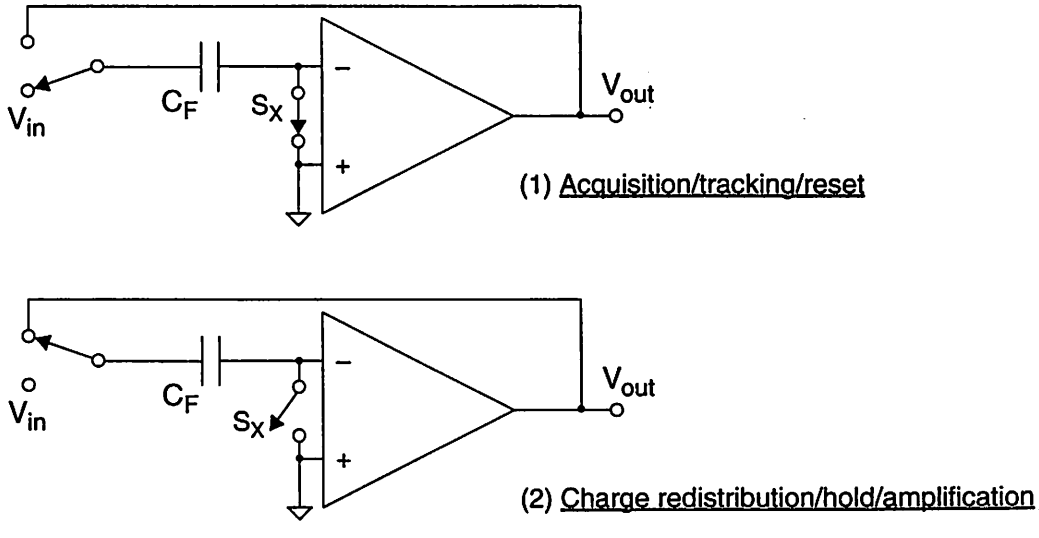


Fig. 5.4 Switched-capacitor unity-gain S/H stage. A single capacitor C_F is used for both acquisition/sample mode and hold mode. The stage gain is 1.

As mentioned in Section 3.4, and also discussed in [135] and [75], it is possible to *share* the feedback capacitor — i.e., use it both for the sampling *and* feedback functions — in a more general way, as shown in Fig. 5.5.

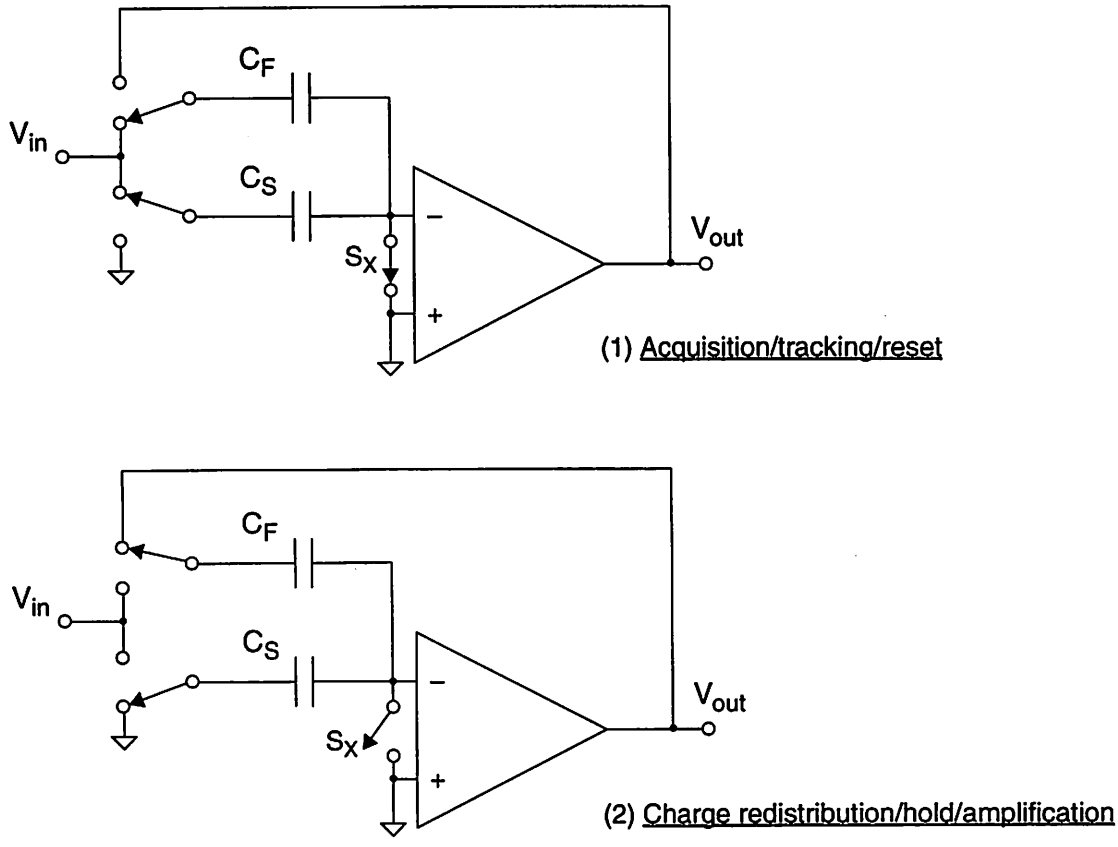


Fig. 5.5 Schematic for a switched-capacitor gain stage with shared feedback capacitance C_F. Closed-loop gain is given by $G = (C_S+C_F)/C_F = 1+C_S/C_F$.

The single-capacitor S/H of Fig. 5.4 is seen to be a special case of the configuration of Fig. 5.5 with $C_S = 0$.

Finally, shown in Fig. 5.6, is the schematic of a SC integrator stage, which is a key component in filters and Δ-Σ converters.

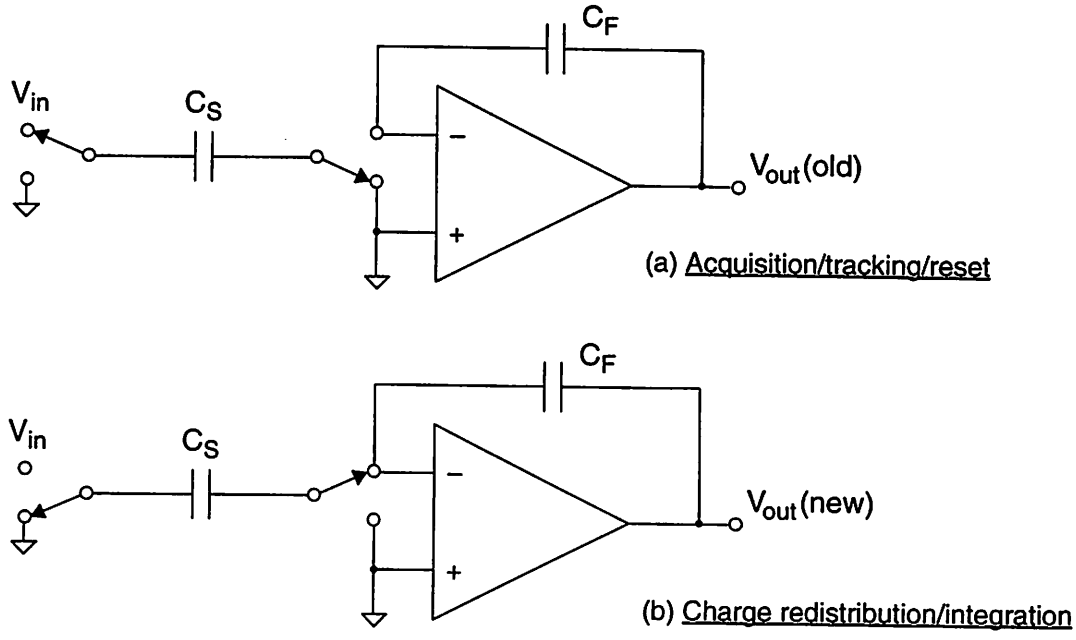


Fig. 5.6 Basic schematic of switched-capacitor integrator stage showing sampling capacitance C_S and integrating/feedback capacitance C_F. The input-output relationship can be expressed as $V_{out}(new) = V_{out}(old) + (C_S/C_F)V_{in}$.

Previously, in Sections 3.5 and 3.6, the static characteristics of these blocks were considered in the context of performing the interstage gain function — and, in some cases, also the DAC function — in pipelined multistage ADC's. Here, the focus is on speed: in all of these circuits, it is required to minimize the settling time in closed-loop configuration (i.e., the integrator or S/H "hold" time), since this block limits the throughput of the pipeline A/D converter. Specifically, the amplifier output has to settle to an accuracy of 1/2 LSB at the system resolution. In a first-order system, this requires (n+1)ln(2) time constants. The design entails first choosing the op amp topology, and then choosing the capacitor sizes, op amp currents, and transistor sizes. There are at least three obvious choices for amplifier topology: (i) single-stage, (ii) single-stage with low-gain, wideband preamp, and (iii) two-stage. The focus of this chapter is analysis of the single-stage op amp — for example, a folded cascode or non-folded cascode (telescopic).

5.1 ANALYSIS OF SC GAIN BLOCK WITH ONE-STAGE OP AMP

A high-frequency equivalent circuit, including parasitic capacitances, for the above SC gain and integrator stages in their closed-loop configuration — i.e., the "hold" or "integrate" phase — is shown in Fig. 5.7. In this figure, v_i is the summing node voltage.

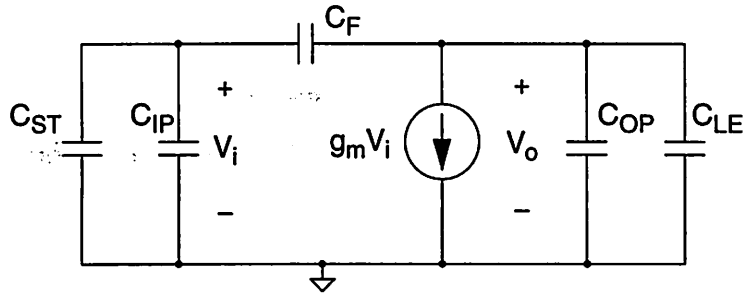


Fig. 5.7 Small-signal equivalent model for high-frequency analysis of the circuits shown in Fig. 5.0–Fig. 5.6 in their closed-loop configuration, assuming a single-stage OTA.

The notation used is as follows.

τ	closed-loop time constant of amplifier
ω_{CL}	closed-loop bandwidth of amplifier (= $1/\tau$)
g_m	transconductance of op amp input device
C_{IP}	input (parasitic) capacitance of op amp
C_{ST}	total source capacitance
C_F	feedback capacitance
C_{OP}	output parasitic capacitance of amplifier
α	ratio of output parasitic capacitance of amplifier to input parasitic capacitance (= C_{OP}/C_{IP})
f	feedback factor of closed-loop system due to capacitive divider (= $C_F/[C_F + C_{IP} + C_{ST}]$)
C_{LE}	extrinsic load capacitance — i.e., external constant loads — *excludes* C_{OP}
ω_T	unity-gain angular frequency of op amp input device (= g_m/C_{IP})

Output resistances are omitted from the above equivalent circuit: clearly they are essential for dc analysis, but do not affect settling time calculation. It is important to note that the C_{ST} indicated in Fig. 5.7 corresponds to the notation $C_{S2T} = \sum_j C_{S2_j}$ introduced in Section 3.5.0, that is, C_{ST} is the total capacitance switched to input sources or reference voltage sources *during the charge redistribution or amplification phase (phase 2), which therefore is the total effective source capacitance seen by the amplifier in its closed-loop configuration*; C_{ST} may or may not equal the *sampling* capacitance of the *acquisition* phase (phase 1).

Note the circuits in Fig. 5.0–Fig. 5.6 perform *discrete-time* signal processing. There is never a continuous path from a time-varying input to the output, although it is sometimes useful to consider an overall closed-loop input-output relationship in the s-domain or ω-domain. Essentially, at the end of the acquisition/tracking phase, the opening of the sampling switch causes charge to be stored at the summing node and thereby defines the initial conditions in the network. In the closed-loop phase, the system starts from those initial conditions, undergoes a transient response, and settles to its final state. For the specific case of Fig. 5.0, a useful view is to consider the switching process as the application of a negative step to the closed loop system, i.e., initially the voltage applied to the left-hand side of C_S is V_{in}, and then it changes to zero.

The basic relation for the equivalent closed-loop –3 dB bandwidth ω_{CL} and the corresponding time constant τ of the system shown in Fig. 5.7 is

$$\omega_{CL} = \frac{1}{\tau} = f \cdot \frac{g_m}{C_{Ltot}} \tag{5.3}$$

where g_m is the transconductance, C_{Ltot} is the total capacitive loading at the output node, g_m/C_{Ltot} is the unity-gain frequency of the closed-loop system, and the feedback factor f is given by

$$f = \frac{C_F}{C_F + C_{ST} + C_{IP}} \tag{5.4}$$

Referring to Fig. 5.7, the feedback factor f corresponds to the capacitive divider seen looking back from the output node to the summing node. Clearly, it is closely related to the feedback factor defined in Section 3.5 in the context of analyzing dc gain errors in a SC gain stage. A key difference, however, is that for dc analysis purposes, there may be a significant "Miller" capacitance seen at the op amp input, since, for example, the low-frequency gain from gate to drain of the input devices in Fig. 5.3 is on the order of $g_m r_o$, whereas the high-frequency gain is the ratio of the transconductance of the input device to that of the cascode device, typically equal to 1.

Note that assuming equal input capacitance C_{IP}, the configurations of Fig. 5.0 and Fig. 5.5 have the same feedback factor

$$f = \frac{C_S}{C_S + C_F + C_{IP}} \tag{5.5}$$

although in the configuration of Fig. 5.5, the closed-loop gain is $1 + C_S/C_F$, whereas in Fig. 5.0 it is C_S/C_F. Alternatively stated, for a given closed-loop gain, the configuration of Fig. 5.5 has a higher feedback factor and hence a faster response.

Throughout this analysis it is assumed that the system can be adequately modeled as a first-order, single-pole network and that minimization of the closed-loop time constant τ implies minimization of the system settling time: the justification of this assertion is given later.

There is a feedforward zero in the transfer function at $s = +g_m/C_F$, but that is at a much higher frequency than the overall closed-loop bandwidth and so is ignored here. Note that this zero corresponds to the frequency at which all the current through capacitor C_F is supplied by the g_m generator, i.e.,

$$sC_F v_i = g_m v_i \Rightarrow s = \frac{g_m}{C_f} \tag{5.6}$$

Referring again to Fig. 5.7, C_{Ltot}, the total load capacitance including the loading effect of the feedback network, is the sum of the load capacitances at the output node $(C_{LE} + C_{OP})$ plus the series combination of the feedback capacitance (C_F) and the capacitance to ground at the summing node $(C_{ST} + C_{IP})$, i.e.,

$$C_{Ltot} = C_{LE} + C_{OP} + \frac{C_F (C_{ST} + C_{IP})}{C_F + C_{ST} + C_{IP}} \tag{5.7}$$

In this analysis, it is assumed that from, for example, open-loop gain or output voltage swing constraints, the bias conditions on the op amp transistors are fixed, i.e., the $(V_{GS} - V_T)$ values. Thus, the current (and transconductance) per unit width of the devices is fixed, and hence the intrinsic speed of the devices as given by ω_T is constant. (Some further discussion of MOS ω_T is given in Appendix 5.A at the end of this chapter.) There is assumed to be no constraint on the maximum current or area. Therefore, the design problem essentially consists of choosing the size of this scalable core block. The transconductance and input parasitic capacitance are related by

$$g_m = \omega_T C_{IP} \tag{5.8}$$

Note that this relation holds for a large class of active devices — including bipolar transistors.

From (5.3), (5.4), and (5.7), the closed-loop time constant is given by

$$\tau = \frac{C_F + C_{ST} + C_{IP}}{C_F} \cdot \frac{C_{LE} + C_{OP} + \dfrac{C_F (C_{ST} + C_{IP})}{C_F + C_{ST} + C_{IP}}}{g_m} \tag{5.9}$$

Rearranging (5.9), and substituting for g_m from (5.8), yields

$$\tau = \frac{1}{\omega_T} \left(\frac{C_{LE}C_F + C_{LE}C_{ST} + C_F C_{ST}}{C_{IP}C_F} + \frac{C_F + C_{LE} + (C_{OP}/C_{IP})\,(C_F + C_{ST})}{C_F} + \frac{C_{OP}}{C_F} \right) \tag{5.10}$$

At this point, the partitioning in Fig. 5.7 of the total load capacitance into two components — C_{OP}, the output parasitic capacitance of the op amp, and C_{LE}, the external or extrinsic load capacitance — is explained. The output parasitic capacitance of the amplifier C_{OP} is due to drain-bulk junction capacitances and overlap capacitance, etc., and *scales* in direct proportion to the size of the op amp device widths, and hence with the input parasitic capacitance — as shown schematically below for a simple device layout.

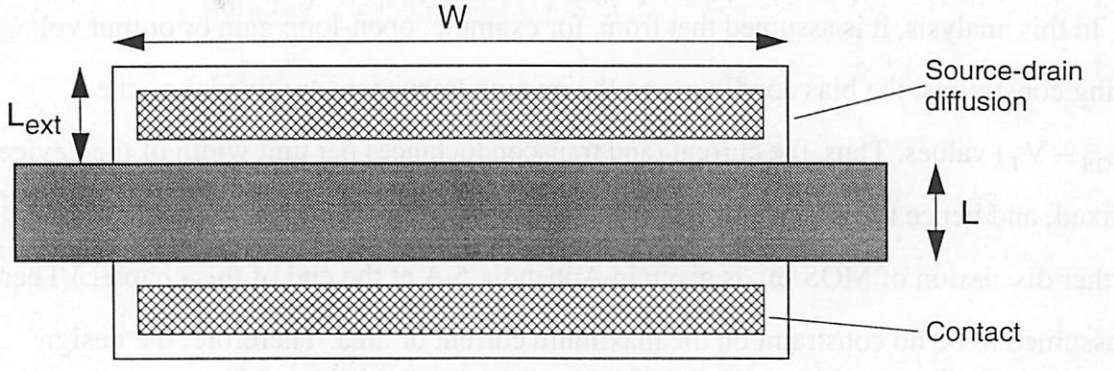


Fig. 5.8 Layout of a MOSFET showing intrinsic gate capacitance and parasitic source-drain capacitance.

In Fig. 5.8, the intrinsic gate capacitance of the device, which corresponds to the input parasitic C_{IP} defined above, is given by (ignoring overlap)

$$C_G = WLC'_{OX} \tag{5.11}$$

where C'_{OX} is the gate oxide capacitance per unit area, and the junction capacitance due to the source/drain diffusions is

$$C_{SD} = WL_{ext}C'_J \tag{5.12}$$

where C'_J is the junction capacitance per unit area.

It is clear that as the width of the device varies, the junction capacitance scales in direct proportion (the overlap capacitance behaves similarly); therefore, more generally, in the case of a SC gain block using a single-stage op amp, the portion of the load capacitance associated with the device junction capacitance and overlap capacitance

scales with the gate capacitance. Thus, the load capacitance at the output node of the stage has been partitioned into (i) a component C_{OP} that scales with the input device size, and (ii) a component C_{LE} that does not scale, and so can be regarded as fixed, i.e., previously specified or known. This effect has not been included systematically in previous analysis and optimization for amplifiers [158]. Note that the above has ignored the perimeter or sidewall capacitance; usually, however, in typical op amp implementations, the MOS devices are laid out using multiple "fingers" — i.e., gates connected in parallel. It is usually possible to avoid using "fingers" that are at the ends of this structure for the most critical nodes, and there is negligible capacitance contribution from the perimeter of the "fingers" that are *not* at the ends. If necessary, this contribution could be estimated and lumped with C_{LE} without significant error.

The constant of proportionality α is defined as the ratio of the (scalable) parasitic capacitance at the output node to the input capacitance, that is,

$$\alpha = \frac{C_{OP}}{C_{IP}} \tag{5.13}$$

For a single MOS device, this quantity is process-dependent, but typically lies in the range 0.5–1. For MOS op amps, α can sometimes be as large as 2 or 3, since the parasitic capacitance at the output node is usually due to contributions from two devices, an NMOS and a PMOS. Note that for a bipolar device, α can easily be less than 0.1.

Substituting for α into (5.10) gives an expression for the closed-loop time constant:

$$\tau = \frac{1}{\omega_T}\left(\frac{C_{LE}C_F + C_{LE}C_{ST} + C_FC_{ST}}{C_{IP}C_F} + \frac{C_F + C_{LE} + \alpha\,(C_F + C_{ST})}{C_F} + \alpha\frac{C_{IP}}{C_F} \right) \tag{5.14}$$

$$= \frac{1}{\omega_T}\left(\left[\frac{C_{LE}C_F + C_{LE}C_{ST} + C_FC_{ST}}{C_{IP}C_F}\right] + \left[1 + \frac{C_{LE}}{C_F} + \alpha\,(1 + \frac{C_{ST}}{C_F})\right] + \left[\alpha\frac{C_{IP}}{C_F}\right] \right) \tag{5.15}$$

The above is a general expression: how can it be applied in a practical design situation? In order to proceed further, some assumptions are needed. First, it is assumed

that the source and feedback capacitances are known — e.g., for matching or area or noise reasons. Second, it is assumed that the extrinsic load capacitance C_{LE} is known. In pipeline ADC's, C_{LE} consists of the sampling capacitance of the next stage, capacitances associated with comparators, parasitic switch capacitances, and any other parasitics such as interconnect capacitance, etc. Given these two assumptions, it is apparent from (5.15) that the closed-loop time constant is a function of only one parameter: C_{IP}, the op amp input capacitance, which in turn directly determines the device widths and currents.

Next, note that the expression in (5.15) has three main terms: (i) a term inversely proportional to C_{IP}, (ii) a constant term, and (iii) a term directly proportional to C_{IP} (corresponding to the effect of the output parasitic capacitance, as discussed above). Thus, it has the form $(\frac{a}{x} + b + cx)$. This expression has an optimum (i.e., minimum) value; by differentiating, this minimum is found to occur at a value of C_{IP} given by

$$C_{IPopt} = \sqrt{\frac{C_{LE}C_F + C_{LE}C_{ST} + C_FC_{ST}}{\alpha}} \tag{5.16}$$

The corresponding values of τ and ω_{CL} are

$$\tau_{opt} = \frac{1}{\omega_{CLopt}} = \frac{1}{\omega_T}\left(2\sqrt{\alpha}\sqrt{\frac{C_{LE}C_F + C_{LE}C_{ST} + C_FC_{ST}}{C_F^2}} + 1 + \frac{C_{LE}}{C_F} + \alpha\left(1 + \frac{C_{ST}}{C_F}\right)\right) \tag{5.17}$$

The following points are pertinent.

Note 1: The above equation is a quite general expression for the highest achievable bandwidth of a switched-capacitor gain stage, given the technology, the external load capacitance, and the source and feedback capacitances.

Note 2: Physically the reason an optimum exists is as follows. Referring to (5.3), clearly when $C_{IP} = 0$, $g_m = 0$, and the bandwidth is zero. As g_m increases, ω_{CL} increases roughly in direct proportion. However, at some point, the effects of C_{IP} on *both* (a) the

feedback factor and (b) the output parasitic capacitance become important, and so the eventual effect is that the bandwidth decreases.

Note 3: It is important that ω_{CLopt} is directly proportional to ω_T, and depends only on *ratios* of the sampling, feedback, and extrinsic load capacitances — which are assumed to fixed. This implies that ω_{CLopt} is a *technology-limited optimum:* as ω_T increases, the achievable ω_{CLopt} increases in direct proportion. The value of α depends weakly on the particular IC process technology. It is also affected by the widths of the devices at the output node relative to the width of the input device; for example, it is possible, for gain or swing reasons, to bias the input and cascode transistors at different nominal $(V_{GS} - V_T)$ values: in such a case, the input and cascode devices have different widths.

Note 4: The optimum value of the input capacitance is a function of the ratios of the source, feedback, and external load capacitances, and also *depends strongly* on the value of α — i.e., the relative contribution to the load of the amplifier parasitic capacitance at the output terminal. This indicates that for a given minimum gate length, the process technology design rules should permit L_{ext} in Fig. 5.8 to be minimized as much as possible in order to minimize the source-drain area.

Note 5: Under certain circumstances C_{IPopt} may be approximately equal to the sampling capacitance, but in general this is not strictly true.

5.2 SOME PARTICULAR CASES

As mentioned already, a very common application of this circuit block is as a gain stage, in particular, to perform the interstage residue amplification function in a multistage pipeline A/D converter. It is useful to examine the maximum achievable speed and corresponding optimum value of C_{IP} as a function of the ratio (C_{ST}/C_F), since for circuit configurations such as those of Fig. 5.0–Fig. 5.2, the closed-loop gain is given by $G = (C_{ST}/C_F)$, and in the case of Fig. 5.4–Fig. 5.5, in which the feedback capacitor is shared and also used for sampling, the closed-loop gain is $1 + (C_{ST}/C_F)$. One of the

reasons why this is important is that, as discussed throughout Chapter 3, in the design of multistage ADC's, the number of bits per stage and the gain per stage are key design choices at the architecture level; consequently, it is important to know the achievable speed as a function of the interstage gain. Accordingly, using the general equations (5.16) and (5.17), the optimum τ is now calculated explicitly for three representative values of the extrinsic load capacitance: (i) $C_{LE} = 0$, (ii) $C_{LE} = C_F$, and (iii) $C_{LE} = C_{ST}$. It is assumed that $\alpha = 1$; for MOS op amps the assumption $\alpha = 1$ is usually the best that is practically achievable.

Case 1: $\alpha = 1$ and $C_{LE} = 0$

These conditions give upper bounds on the achievable speed with single-stage MOS op amps, since specifying $C_{LE} = 0$ is equivalent to stating that the sampling and feedback capacitors, and also the amplifier device widths and currents, are made large enough so that the extrinsic load capacitance C_{LE} becomes negligible. These assumptions give

$$C_{IPopt} = \sqrt{C_F C_{ST}} = C_F\sqrt{\frac{C_{ST}}{C_F}} = C_{ST}\sqrt{\frac{C_F}{C_{ST}}} \tag{5.18}$$

and

$$\tau_{opt} = \frac{1}{\omega_{CLopt}} = \frac{1}{\omega_T}\left[2\sqrt{\frac{C_{ST}}{C_F}} + 2 + \left(\frac{C_{ST}}{C_F}\right)\right] = \frac{1}{\omega_T}\left[\left(1 + \sqrt{\frac{C_{ST}}{C_F}}\right)^2 + 1\right] \tag{5.19}$$

The meaning of (5.18) when $C_{ST} = 0$ needs some elucidation: this case corresponds to a shared sampling and feedback capacitance *and* zero extrinsic load. Since the amplifier is then driving only itself, the smaller C_{IP} the better. In an actual implementation, the presence of a parasitic capacitance to ground at the summing node associated with C_{ST} implies that C_{IPopt} has some non-zero value.

Case 2: $\alpha = 1$ and $C_{LE} = C_F$

These conditions give

$$C_{IPopt} = C_F\sqrt{1+2\left(\frac{C_{ST}}{C_F}\right)} = C_{ST}\sqrt{\left(\frac{C_F}{C_{ST}}\right)^2+2\left(\frac{C_F}{C_{ST}}\right)} \tag{5.20}$$

and

$$\tau_{opt} = \frac{1}{\omega_{CLopt}} = \frac{1}{\omega_T}\left(2\sqrt{1+2\left(\frac{C_{ST}}{C_F}\right)}+3+\left(\frac{C_{ST}}{C_F}\right)\right) \tag{5.21}$$

Case 3: $\alpha = 1$ and $C_{LE} = C_{ST}$

The condition $C_{LE} = C_{ST}$ is typical of what occurs in real implementations — e.g., in a pipeline ADC, often the extrinsic capacitive load consists primarily of the sampling capacitance of the next stage. The optimal values in this case are

$$C_{IPopt} = C_F\sqrt{2\left(\frac{C_{ST}}{C_F}\right)+\left(\frac{C_{ST}}{C_F}\right)^2} = C_{ST}\sqrt{2\left(\frac{C_F}{C_{ST}}\right)+1} \tag{5.22}$$

and

$$\tau_{opt} = \frac{1}{\omega_{CLopt}} = \frac{1}{\omega_T}\left(2\sqrt{2\left(\frac{C_{ST}}{C_F}\right)+\left(\frac{C_{ST}}{C_F}\right)^2}+2+2\left(\frac{C_{ST}}{C_F}\right)\right) \tag{5.23}$$

Note that the relationship between closed-loop "gain" and "bandwidth" is quite nonlinear due to the interdependence of the amplifier forward path and the feedback path. Specifically, both the transconductance g_m and the feedback factor f depend on the input capacitance C_{IP}. In the limit of large C_{ST}/C_F — i.e., large closed-loop gain — and the extrinsic load C_{LE} small or at most comparable to C_F, then this interdependence diminishes. In that case, the quantity under the square root in (5.17) becomes negligible, and the expression for the closed-loop time constant and bandwidth approximately reduces to

$$\tau_{opt} = \frac{1}{\omega_{CLopt}} \approx \frac{1}{\omega_T}\alpha\left(\frac{C_{ST}}{C_F}\right) = \alpha\frac{(C_{ST}/C_F)}{\omega_T} \tag{5.24}$$

which is more consistent with the notion of a constant "gain-bandwidth product".

The above equations for C_{IPopt} and τ_{opt} are now plotted for the three cases mentioned above: $C_{LE} = 0$, $C_{LE} = C_F$, and $C_{LE} = C_{ST}$. The first two plots show C_{IPopt} — the value of the input capacitance needed to achieve the fastest speed — plotted in two ways. First, in Fig. 5.9 the ratio of the optimum input capacitance to the source capacitance is plotted as a function of C_{ST}/C_F.

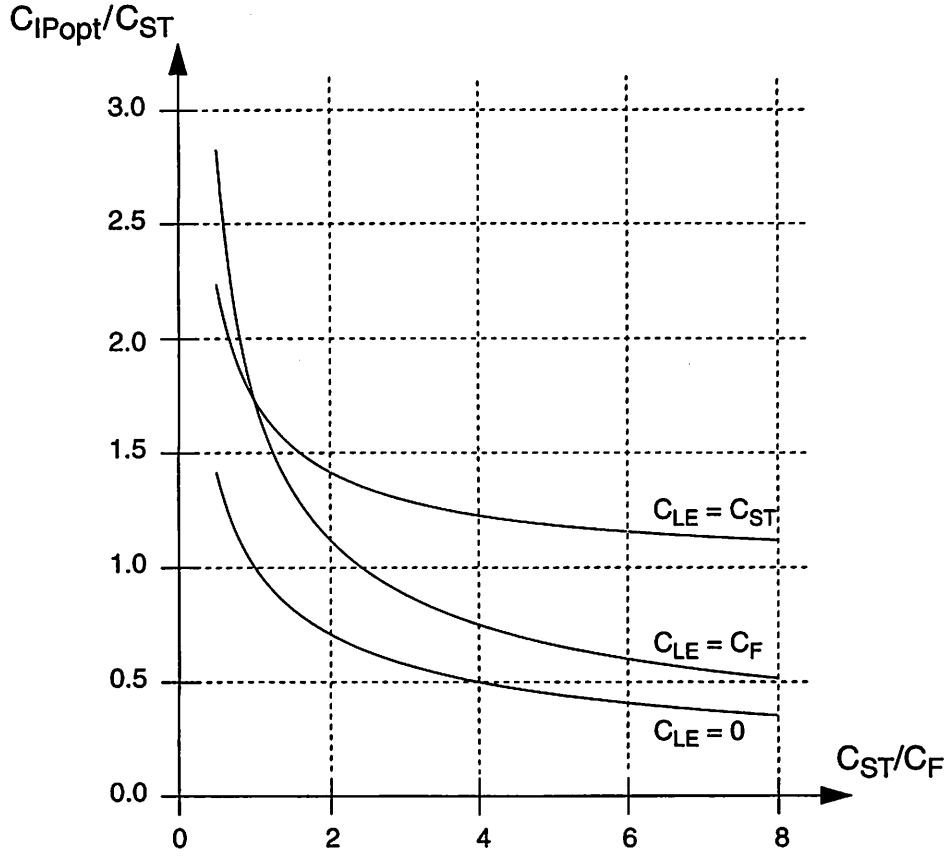


Fig. 5.9 Plot of C_{IPopt}/C_{ST} as a function of C_{ST}/C_F.

It is apparent from the graphs that, especially for values of C_{ST}/C_F in the range 1–4, which is particularly relevant for pipeline ADC's, the optimum value of C_{IP} is quite a strong function of C_{ST}/C_F. Note that the often-quoted condition $C_{IP} = C_{ST}$ does *not* necessarily give the optimum speed.

A second way of looking at this is to plot C_{IPopt} normalized relative to the feedback capacitance C_F — this is shown in Fig. 5.10.

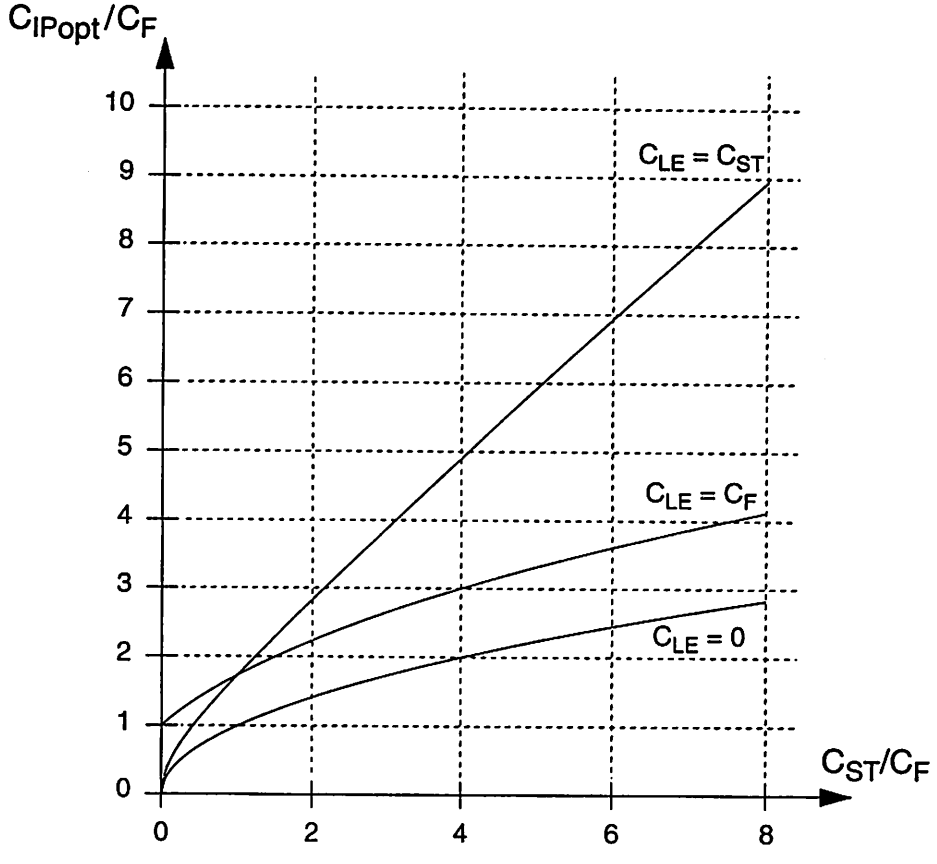


Fig. 5.10 Plot of C_{IPopt}/C_F as a function of C_{ST}/C_F.

Finally, the plot in Fig. 5.11 below is the optimum time constant τ_{opt} as a function of C_{ST}/C_F. This minimum time constant τ_{opt} is normalized with respect to the intrinsic time constant of the transistor, τ_T, where $\tau_T = 1/\omega_T$, thus giving a technology-independent result. This optimum time constant directly determines the settling time to a given accuracy of a SC gain stage, and hence the maximum achievable sample rate of a pipeline A/D converter,

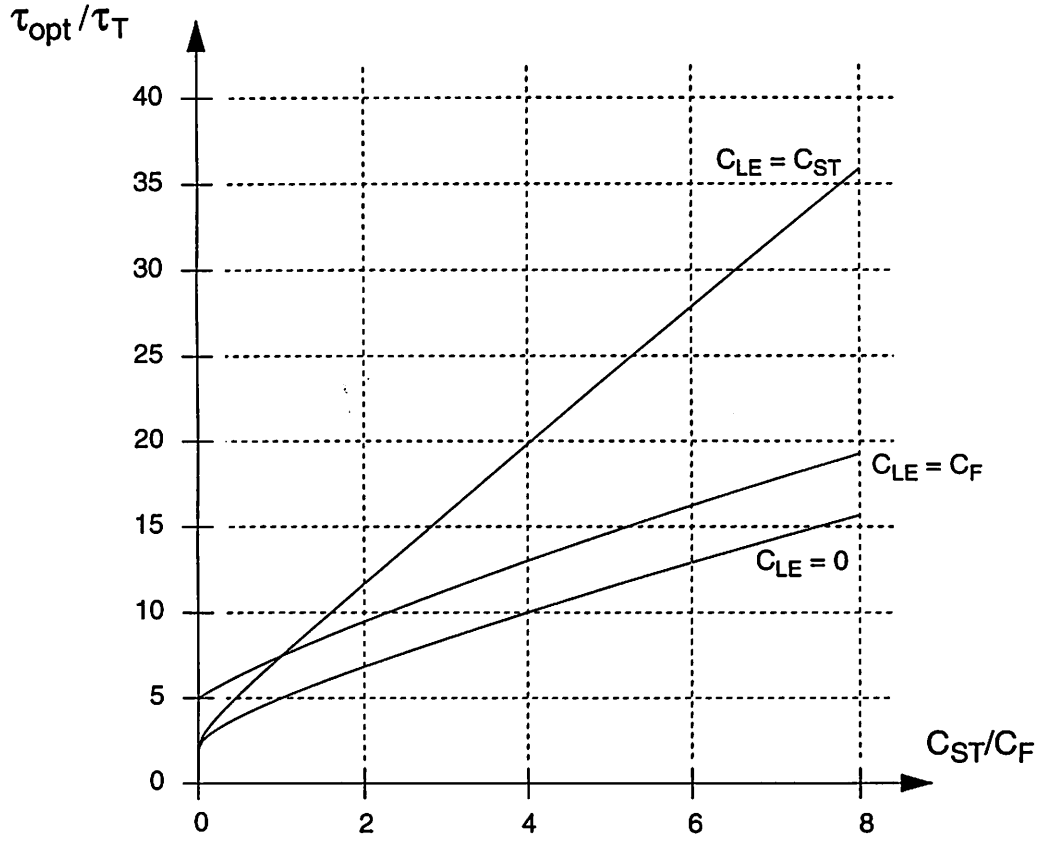


Fig. 5.11 Plot of the (technology-normalized) fastest achievable closed-loop time constant τ_{opt}/τ_T, as a function of C_{ST}/C_F.

Again note that the value of τ_{opt} depends strongly on C_{ST}/C_F — especially for small values of C_{ST}/C_F — emphasizing that careful optimization is worthwhile. Furthermore, this plot clearly demonstrates the advantage of sharing the feedback capacitor in order to minimize C_{ST}/C_F. For example, in a gain-of-two stage where the feedback capacitor is not shared, $C_{ST}/C_F = 2$, whereas if the feedback capacitor is shared, then $C_{ST}/C_F = 1$.

A number of other special cases and limiting cases are of interest as follows.

Case 4: Constant — i.e., *non-scaling* — output parasitic capacitance: $\alpha = 0$

Consider the case where the output parasitic capacitance of the amplifier does *not* scale with C_{IP}. As mentioned already, this is highly unrealistic for MOS op amps. For bipolar devices it is more reasonable.

Since all the load capacitance is lumped with the fixed term C_{LE}, this implies

$$\alpha = 0 \tag{5.25}$$

Equation (5.16) then yields

$$C_{IPopt} \rightarrow \infty \tag{5.26}$$

and (5.17) gives

$$\tau_{opt} = \frac{1}{\omega_{CLopt}} = \frac{1}{\omega_T}\left(1 + \frac{C_{LE}}{C_F}\right) \tag{5.27}$$

This result also follows from a more direct analysis as follows. As C_{IP} becomes large (with C_S and C_F remaining constant), eventually the feedback factor becomes close to (C_F/C_{IP}), and the total load capacitance approaches $(C_{LE} + C_F)$. Therefore, the closed-loop bandwidth becomes

$$\omega_{CL} = f\frac{g_m}{C_{Ltot}} \rightarrow \frac{C_F}{C_{IP}} \cdot \frac{g_m}{C_{LE} + C_F} = \frac{\omega_T}{1 + \dfrac{C_{LE}}{C_F}} \tag{5.28}$$

Case 5: Unity-gain stage using a single capacitor C_F

Another particular case of interest is the unity-gain S/H stage using a single capacitor, which is shown in Fig. 5.4. In terms of the equivalent circuit model illustrated in Fig. 5.7, the configuration of Fig. 5.4 corresponds to setting $C_{ST} = 0$, i.e., the feedback capacitor C_F also functions as the "sampling" or "acquisition" capacitance, and therefore there is no "source capacitance" in Fig. 5.7. (Any additional parasitic capacitance to ground at the

summing node is ignored here also.)

Substituting the condition $C_{ST} = 0$ into (5.16) and (5.17) gives

$$C_{IPopt} = \sqrt{\frac{C_{LE}C_F}{\alpha}} \tag{5.29}$$

and

$$\tau_{opt} = \frac{1}{\omega_{CLopt}} = \frac{1}{\omega_T}\left(2\sqrt{\alpha\frac{C_{LE}}{C_F}} + 1 + \frac{C_{LE}}{C_F} + \alpha\right) \tag{5.30}$$

Again, note the strong dependence on α. One circuit implication of this is that it is usually worthwhile to make the cascode devices (which contribute to the output parasitic capacitance and scale with the input device) narrower than the input device.

5.3 DISCUSSION AND SUMMARY

From the above equations and especially from the plot in Fig. 5.11, it is clear that in almost all cases the closed-loop system –3db bandwidth is significantly less than ω_T. Since in single-stage op amps the non-dominant pole (at the cascode node) is approximately at ω_T, the assumption of single-pole behavior is seen to be reasonable — i.e., the non-dominant pole does not affect the speed significantly.

A key issue is technology choice: for example, does BiCMOS provide a performance improvement in this circuit? One possibility is to use npn devices to replace the NMOS cascodes — this reduces the parasitic capacitance at the output node and provides higher dc gain. The PMOS side of course is still a limitation with respect to both gain and speed. Note that using an npn transistor for the cascode device *does* require a contact at the cascode node: a contact is *not* required if a non-folded cascode op amp is used with input and cascode devices of the same width. More generally, npn emitter followers devices could be used as buffers throughout a mixed-signal chip.

Finally, note that in all of the above examples the parasitic capacitance to ground associated with the sampling and feedback capacitors has not been explicitly discussed. In

practice, using precision capacitors having capacitance typically around 0.5 fF/μm^2, the value of this parasitic is 10–15% of the nominal capacitor value. The capacitor layout is such that this parasitic is physically at the input or output nodes of the overall stage — *not at the summing node.* Shown in Fig. 5.12 below is a simplified diagram of a capacitor structure in a typical CMOS technology and the configuration in an actual circuit.

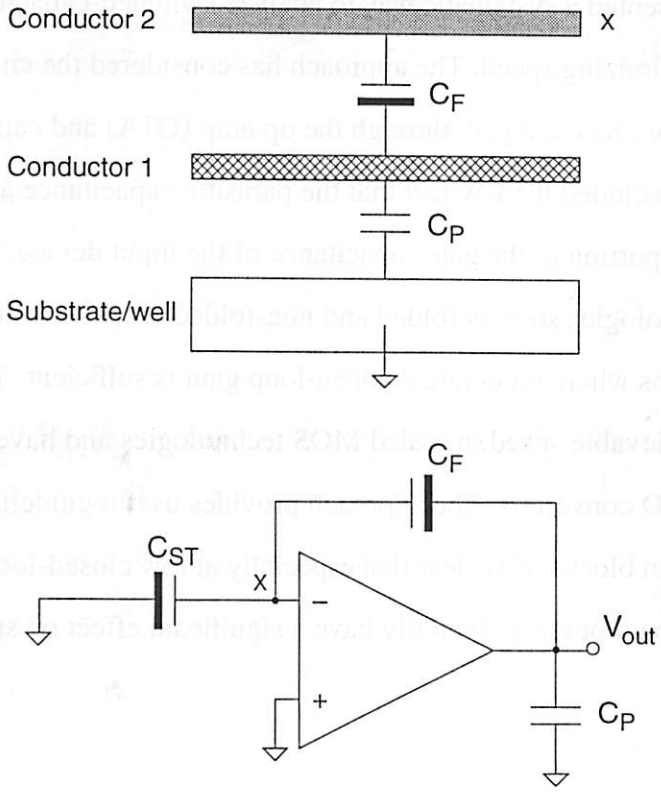


Fig. 5.12 Schematic of switched-capacitor stage in closed-loop configuration showing the parasitic capacitance to ground associated with the feedback capacitor.

Typically, for a high-precision analog capacitor such as a Poly1-Poly2 structure, the parasitic C_P is about 10–15% of C_F, the value of the floating capacitor. However, there is some cost motivation to using the capacitance between metal interconnect layers for the sampling and feedback capacitors in order to realize the circuits in standard digital CMOS processes, which usually do not possess special high-quality capacitors (e.g., poly-poly capacitors). The drawback of course is that if, for example, Metal2-Metal1 capacitors are used, then the parasitic has a much higher relative value ($\approx$ 50–100%), which clearly has a

direct impact on the achievable speed. (Analytically, this is apparent from (5.17) by considering the effect of increased C_{LE}.) The effect of this parasitic could be included in the analysis described in this chapter by replacing C_{LE} — the external/extrinsic load capacitance — by $C_{LE} + \alpha_2 C_F$, where $\alpha_2 C_F$ is the parasitic capacitance C_P shown in Fig. 5.12 and is directly proportional to C_F.

This chapter has presented a systematic way to analyze switched-capacitor gain stages with the objective of optimizing speed. The approach has considered the entire closed-loop system — i.e., forward path through the op amp (OTA) and capacitive feedback factor — and included the key fact that the parasitic capacitance at the output node scales in direct proportion to the gate capacitance of the input device. The focus has been on single-stage topologies such as folded and non-folded cascode, which are important for applications where moderate dc open-loop gain is sufficient. The results represent the fastest achievable speed in scaled MOS technologies and have direct relevance to pipeline A/D converters. The approach provides useful guidelines for choice of device sizes in SC gain blocks. It is clear that especially at low closed-loop gains (1–4), sharing the feedback capacitor can potentially have a significant effect on speed.

APPENDIX

5.A ANALYSIS OF MOSFET f_T

5.A.0 Introduction and Motivation

The small-signal performance of MOS devices at high frequencies is critical for many analog IC applications. Although much work has been reported in the area of generalized high-frequency ac MOS modeling [143], very often a designer requires a single figure of merit that can summarize the high-speed capabilities of the device as a function of the design variables. Traditionally, the device f_T, defined as the unity ac current gain frequency or current gain-bandwidth product, has fulfilled this requirement for the bipolar transistor. For the MOSFET, however, f_T is a strong function of the effective gate length L and the bias condition $V_{GS} - V_T$. This appendix gives a brief analysis of MOS including two key second-order effects; namely, velocity saturation and transverse field dependent mobility. A conceptual view of a single-MOSFET current amplifier is shown in Fig. 5.13.

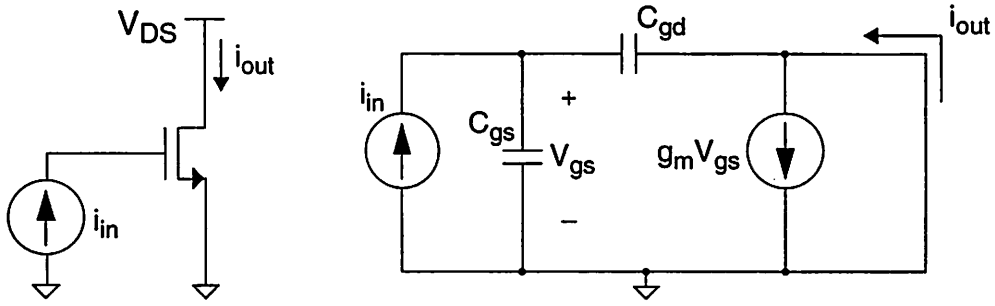


Fig. 5.13 Conceptual schematic of a MOSFET current amplifier and small-signal equivalent circuit.

From analysis of the small-signal circuit of Fig. 5.13, it is easily shown that the MOSFET unity current gain frequency f_T is given by

$$f_T = \frac{1}{2\pi} \cdot \frac{g_m}{C_{gs}+C_{gd}} = \frac{1}{2\pi} \cdot \frac{g_m}{C_{gs}} \cdot \frac{1}{(1+C_{gd}/C_{gs})} = f_{Tintr} \cdot \frac{1}{(1+C_{gd}/C_{gs})} \qquad (5.31)$$

where f_{Tintr} may be thought of as the *intrinsic* f_T. In the saturation region, the gate capacitance C_{gs} is equal to $(2/3)\,WLC_{OX}$ [143].

5.A.1 Analysis

For most analog circuit applications MOSFET's are biased in the saturation region. Thus, following the high-field model described in [124] and using the same notation, the current in saturation (for an NMOS transistor) is given by

$$I_{Dsat} = WC_{OX}(V_G' - V_{Dsat})\,v_{sat} \tag{5.32}$$

where W is the device width, C_{OX} is the gate oxide capacitance per unit area, $V_G' = V_{GS} - V_T$ is the effective gate drive, and v_{sat} is the electron saturation velocity (1.1×10^7cm/s). V_{Dsat} includes the second-order effects mentioned above and is given by

$$V_{Dsat} = \left[\frac{1}{V_G'} + \frac{1}{E_C L}\right]^{-1} \tag{5.33}$$

where $v_{sat} = \frac{\mu_{eff}E_C}{2}$ and $\mu_{eff} = \frac{\mu_0}{1+\theta V_G'}$, which implies that $E_C = \frac{2v_{sat}}{\mu_0}(1+\theta V_G')$.

The parameter θ determines the inversion layer mobility dependence on transverse field (i.e., *perpendicular* to the gate) and is inversely proportional to the gate oxide thickness.

By combining the above equations, an expression for I_{Dsat} may be derived in terms of W, μ_0, C_{OX}, L, V_G', and v_{sat}. Differentiation gives the saturation region transconductance $g_{msat} = \partial I_{Dsat}/\partial V_G$, and it is straightforward to show that the intrinsic f_T is given by

$$f_{Tintr} = \frac{1}{2\pi}\cdot\frac{g_{msat}}{\frac{2}{3}WLC_{OX}} = \frac{3\mu_0 V_G'}{8\pi L^2}\cdot\frac{2+V_G'(\theta+\frac{\mu_0}{2Lv_{sat}})}{\left[2+V_G'(\theta+\frac{\mu_0}{2Lv_{sat}})\right]^2} \tag{5.34}$$

Therefore, f_{Tintr}, MOSFET intrinsic f_T, depends on (i) bias voltage V_G' (but in a *sublinear* manner), (ii) effective channel length L, (iii) process/technology parameters μ_0 and θ, and (iv) v_{sat}, a constant for silicon.

Chapter 6 Path Mismatch Issues in Parallel Time-Interleaved ADC's

6.0 INTRODUCTION

A fundamental problem associated with parallelism in the analog domain, which does not arise in the digital domain, is the issue of mismatches between the parallel *signal paths* or *channels*. In particular, the following mismatches are present in time-interleaved A/D converters: (i) gain mismatch, (ii) offset mismatch, and (iii) timing mismatch, which is also referred to as *nonuniform timing*, or *sampling skew*. Previous work — in particular, by Black and Hodges [8], Jenq [53], Messerschmitt [94], and Petraglia and Mitra [105] — has considered some of the circuit and system issues associated with mismatches in parallel channels. This chapter is, in some sense, a review of this existing literature; however, it also aims to present a unified interpretation and to convey some intuition on the underlying signal processing issues.

The chapter is structured as follows. In Section 6.1, gain and offset mismatches are considered as discrete-time sequences and are analyzed using an approach based on the Discrete Fourier Series (DFS). Another approach to the analysis of gain mismatch is to examine the system in terms of decimation and interpolation operations: this is the viewpoint of Section 6.2. Section 6.3 considers timing mismatches, which have some qualitatively different characteristics from offset and gain mismatches; the important special case of a sinusoidal input is dealt with and appropriate expressions are derived. Finally, in Section 6.4, some issues regarding the statistics of mismatches are discussed and expressions for error power, total distortion power, and signal-to-noise ratio (SNR) are derived.

The following notation is used throughout the chapter.

M	number of parallel channels
T	system sampling interval
F_s	system sampling rate ($= 1/T$)
F_s/M	individual channel sampling rate ($= 1/MT$)
ω	absolute angular frequency (i.e., *not* normalized to F_s)
f	absolute frequency, i.e., cycles per second ($f = \omega/2\pi$)
$x(t)$, $X(j\omega)$	analog input and its Fourier Transform
$p(t)$, $P(j\omega)$	sampling impulse train and its Fourier Transform
$y(t)$, $Y(j\omega)$	time-interleaved A/D converter output and its Fourier Transform
i	index for successive samples of the A/D converter
r	index for successive cycles through all the parallel channels
k	channel index: $k = 0, 1, \ldots, M-1$
a_k	gain of channel k (ideally, $a_k = 1$)
A_n	Discrete Fourier Series coefficient of channel gains
b_k	offset of channel k (ideally, $b_k = 0$)
B_n	Discrete Fourier Series coefficient of channel offsets
Δt_k	sampling time error of channel k (ideally, $\Delta t_k = 0$)
t_k	sampling instant of channel k ($t_k = kT + \Delta t_k$; ideally, $t_k = kT$)
$\phi_{nn}(\omega)$	frequency-domain coefficient for channel timing mismatches
ω_0, f_0	frequency, in radians/second and cycles/second respectively, of input sinusoid or complex exponential

The symbol $\leftrightarrow$ denotes a Fourier Transform relationship, i.e.,

$$f(t) \leftrightarrow F(j\omega)$$

indicates that the frequency-domain quantity on the right-hand side is the Fourier Transform (FT) — or Discrete-Time Fourier Transform, as appropriate — of the time-domain quantity on the left-hand side.

The basic general model for a time-interleaved A/D converter (ADC) including mismatches is shown below in Fig. 6.0.

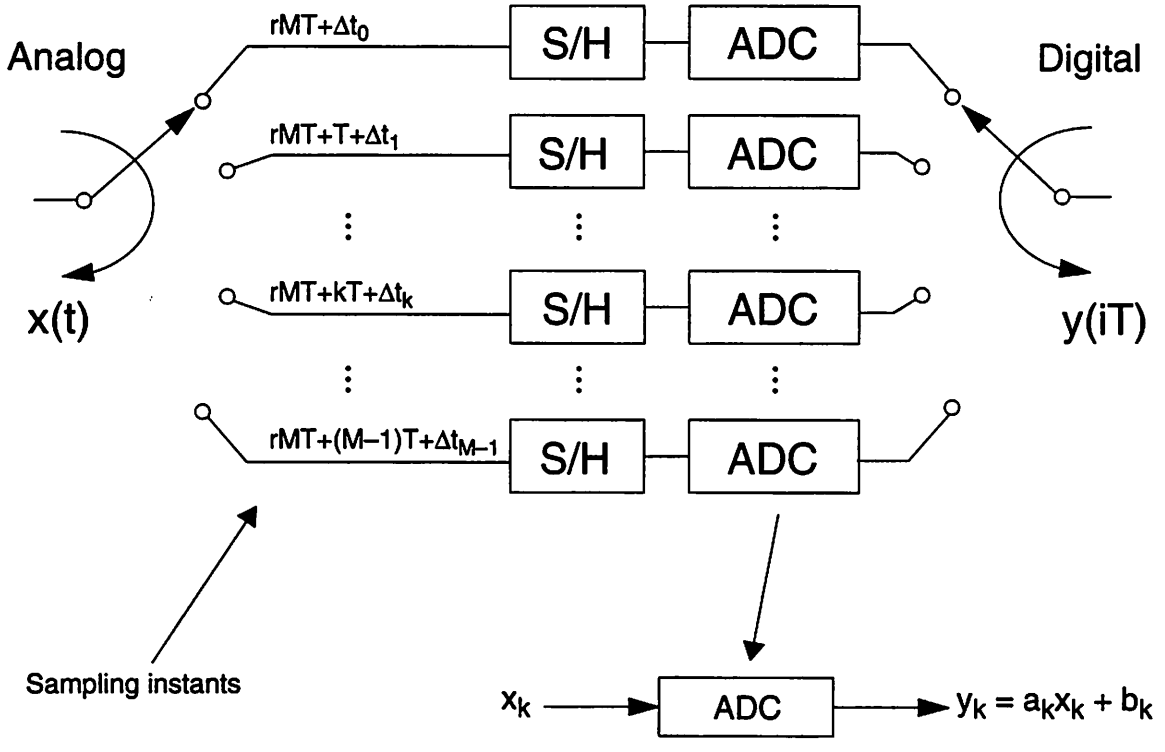


Fig. 6.0 Basic general model for a time-interleaved ADC consisting of M parallel channels with overall sampling rate $F_s = 1/T$; each channel or path operates at the rate $F_s/M = 1/MT$.

The input x(t) is a continuous-time signal; its Fourier Transform is X(jω). The system consists of M parallel channels. The diagram in Fig. 6.0 shows each channel k consisting of a sample-and-hold (S/H) block and an ADC block. The S/H block performs no function in the analysis — it is present in the figures simply to emphasize the fact that sampling is taking place; all sampling is assumed to be ideal impulse sampling. (In next chapter there

is some discussion of input bandwidth issues in the S/H.) The overall system sampling rate is $F_s = 1/T$; each channel samples at a rate of $1/MT$. The sampling instant of channel k occurs nominally at time T after channel k–1 has sampled and in general has some timing offset Δt_k. The ADC block in each channel k is modeled as a memoryless block having linear gain a_k (nominally 1) and offset b_k (nominally 0). Note that more generally each of the parallel signal paths may have some frequency dependence, thus giving rise to parallel filters [105], [147]; however, that level of generality is not considered here. The final output is a discrete-time signal y(iT) (sometimes, the notation y_i is used); y(iT) has Fourier Transform $Y(j\omega)$. The output rate of the entire system is identical to the sampling rate, i.e., $F_s = 1/T$. Clearly, in a real A/D converter system, x(t) is continuous-valued and y(iT) is discrete-valued: this fact is ignored throughout most of this chapter — i.e., quantization effects in the ADC's are ignored and it is assumed that the quantization has infinite resolution.

The terms *distributor* and *commutator* are used as indicated in Fig. 6.1 below.

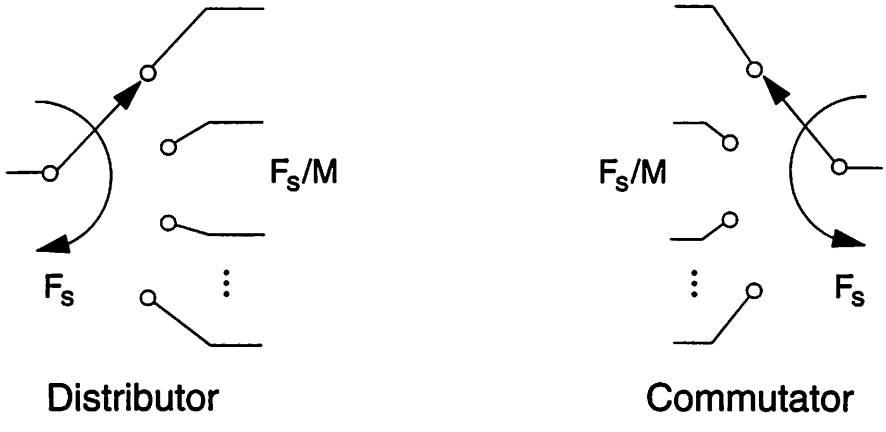


Fig. 6.1 Definition of terms: distributor and commutator.

These blocks are defined as follows [7]:

a *distributor* takes one input stream and synchronously splits it into M output streams;

a *commutator* takes M input streams and combines them to form one output stream.

In this chapter, however, it is often assumed that the distributor symbol in Fig. 6.1 also incorporates *ideal impulse sampling* of a continuous-time input. Note that in the literature and also in other chapter of this dissertation, the more familiar terms demultiplexer and multiplexer are often used for distributor and commutator respectively — i.e., an *analog demultiplexer* at the input and a *digital multiplexer* at the output. Strictly, however, the terms demultiplexer and multiplexer imply the existence of a "control" or "select" input.

6.1 GAIN AND OFFSET MISMATCHES AS DISCRETE-TIME SEQUENCES

This section is organized as follows. Section 6.1.0 presents an analysis of gain mismatches; some examples are given in Section 6.1.1, and Section 6.1.2 discusses some intuitive interpretations. Using a similar method, offset mismatches are analyzed in Section 6.1.3. Finally, the approach and the key results are summarized in Section 6.1.4.

6.1.0 Gain Mismatches in Parallel Paths: Fourier Analysis

The system being examined is illustrated in Fig. 6.2 below for the case of 4 channels.

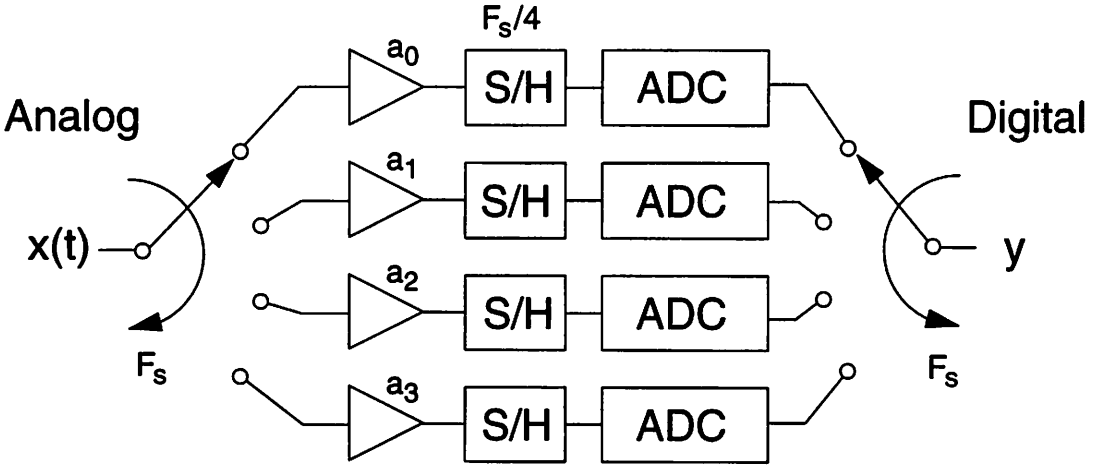


Fig. 6.2 Model for 4-channel time-interleaved A/D converter system indicating gain mismatches. Ideally, the $\{a_k\}$ are all equal (to 1); in general, however, the $\{a_k\}$ are not all identical.

First, as reference, the behavior in the ideal case (no path mismatches) is considered. Then two methods are presented for analysis of the case of mismatched gains.

The basic approach here consists of the following three steps.

1. Consider the sampling operation as the multiplication of the input signal x(t) by an impulse train p(t).
2. Calculate the Fourier Transform P(jω) of the sampling impulse train p(t).
3. Convolve P(jω) with the Fourier Transform X(jω) of the analog input signal x(t).

Sampling with Ideal Matched Channel Gains

Clearly, in the degenerate case of no mismatches, all the results should reduce to what would be expected for a single channel. In particular, an M-channel system with overall sample rate F_s, and in which the channel gains are equal, behaves identically to a single-channel system sampling at F_s. An ideal impulse train is shown in Fig. 6.3.

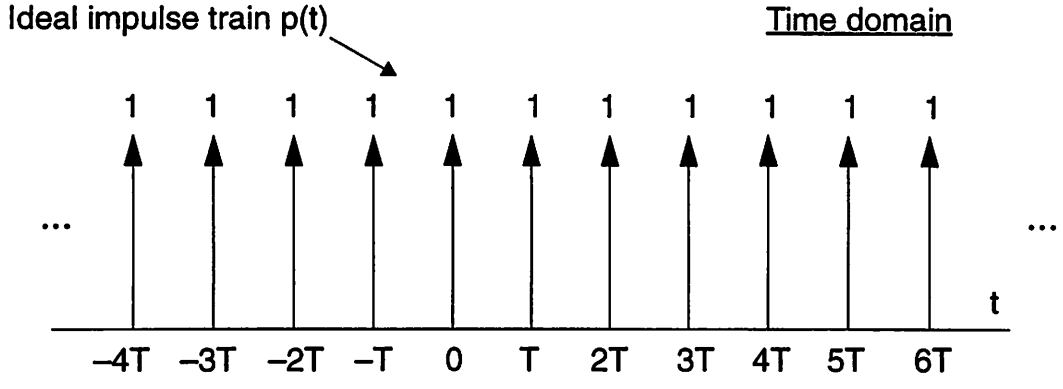


Fig. 6.3 Ideal periodic impulse train p(t).

The basic well-known analytical relation for the FT of an ideal impulse train is

$$p(t) = \sum_{m=-\infty}^{\infty} \delta(t - mT) \leftrightarrow \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} \delta(\omega - \frac{2\pi n}{T}) = P(j\omega) \qquad (6.0)$$

This relationship is illustrated in Fig. 6.4 below.

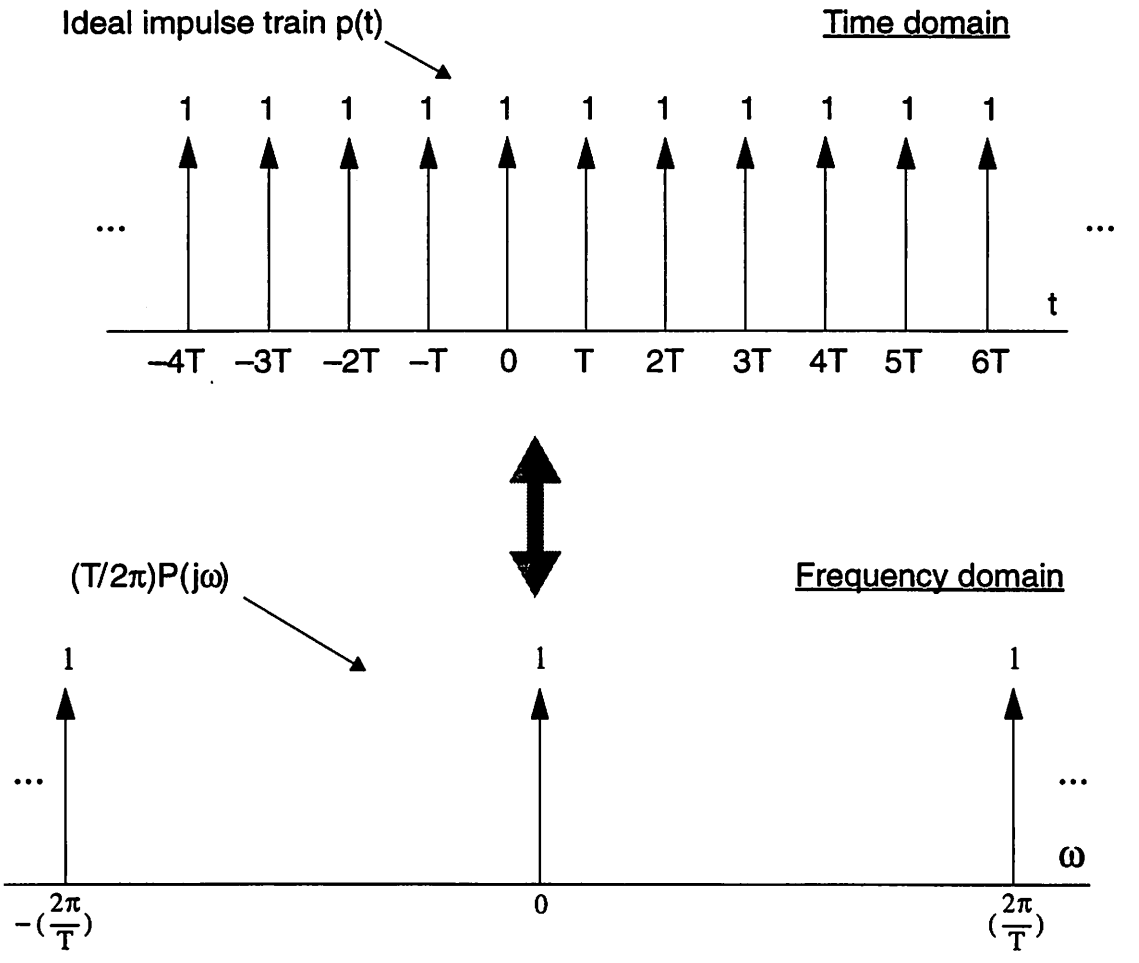


Fig. 6.4 Ideal impulse train p(t) and its Fourier Transform P(jω).

Aside

The following three remarks are generally relevant to the presentation in this chapter.

- To minimize clutter on the graphs of spectra, constant factors such as $1/T$ or 2π, etc., are usually absorbed into the quantity being plotted.
- For correctness, the factor of (1/T) is retained in the Fourier Transform expressions for P(jω) and Y(jω) throughout most of this chapter, but it should be realized that this factor disappears when the ADC output is passed through a reconstruction filter, specifically, an ideal low-pass filter with gain T.

- Whenever spectra of complex quantities are displayed, the plots in general indicate the magnitudes, even though the labelling and discussion refer to the entire quantities, which in general may be complex-valued.

The familiar effect in the frequency domain of sampling by an ideal impulse train — i.e., as shown in Fig. 6.4 — is illustrated in Fig. 6.5 below: the spectrum of the input $X(j\omega)$ is replicated over the entire angular frequency axis at intervals of $2\pi/T$.

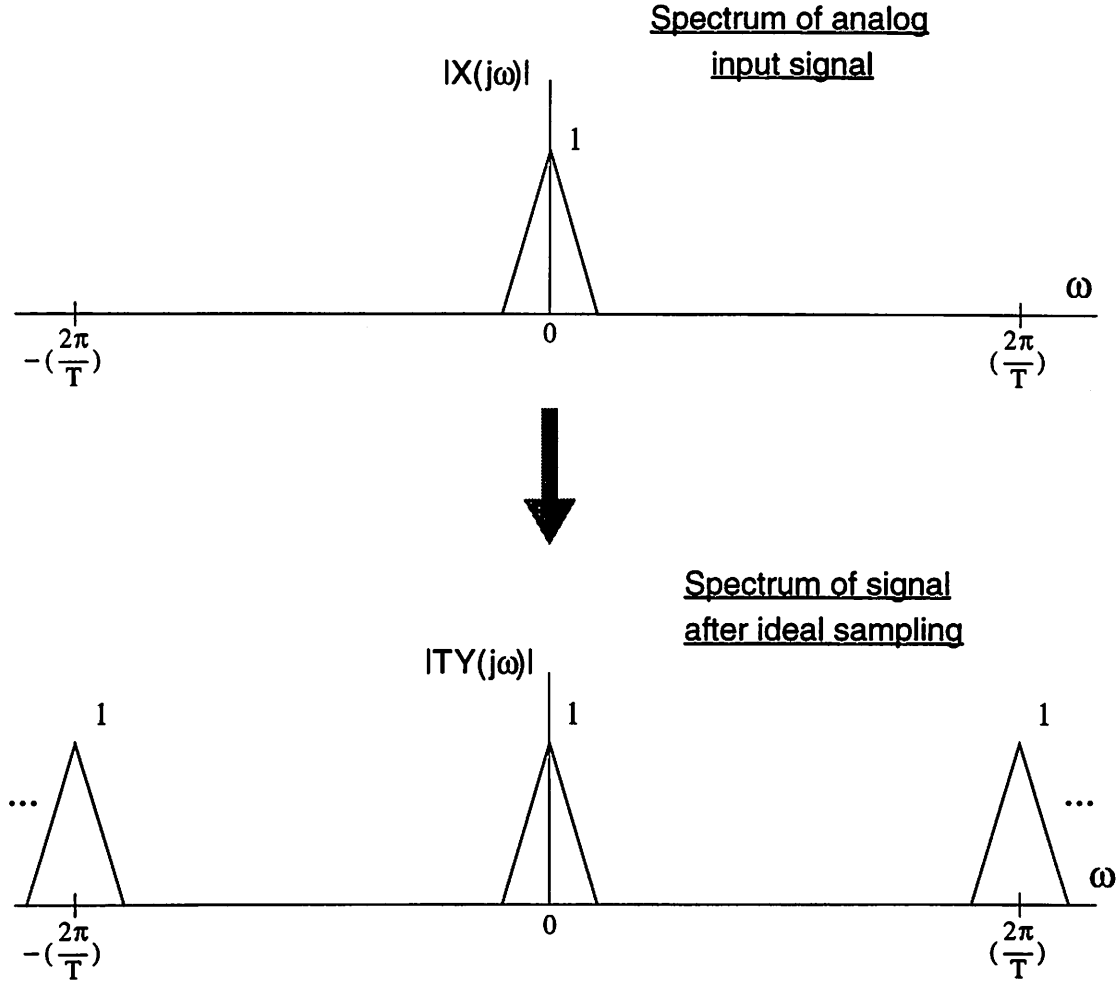


Fig. 6.5 Spectrum of a continuous-time input signal and the spectrum after sampling by an ideal impulse train.

Next consider a system with the same overall sampling frequency F_s, but comprised of M time-interleaved ADC's. Let each channel k have gain a_k, for $k = 0, \ldots, M-1$. In this case, the sampling waveform is the periodic impulse train

$$\ldots, a_0, a_1, a_2, \ldots, a_{M-1}, a_0, a_1, a_2, \ldots, a_{M-1}, a_0, a_1, a_2, \ldots, a_{M-1}, a_0, \ldots$$

as shown in Fig. 6.6 for a 4-channel case.

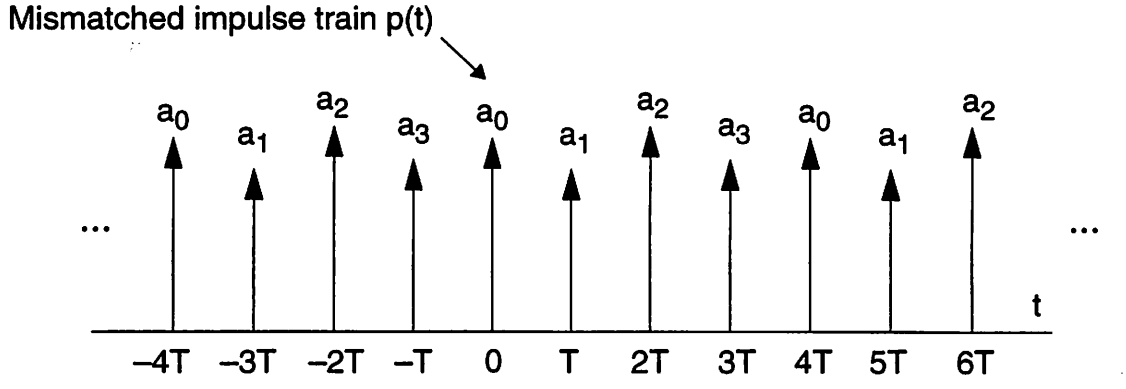


Fig. 6.6 Sampling impulse train in the presence of channel gain mismatches — here shown for case of M=4 channels.

Two methods are now presented for deriving the Fourier Transform of the modified sampling impulse train, in the presence of mismatches.

Sampling with Nonideal Mismatched Channel Gains: Analysis Method 1

The key point here is that the sampling impulse train of Fig. 6.6 may be regarded as the superposition of M sequences, each having period MT, and delayed by T relative to each other, and so can be written as

$$p(t) = \sum_{k=0}^{M-1} a_k \sum_{n=-\infty}^{\infty} \delta(t - kT - nMT) \tag{6.1}$$

The corresponding Fourier Transform $P(j\omega)$ thus becomes

$$P(j\omega) = \sum_{k=0}^{M-1} a_k e^{-j\omega kT} \frac{2\pi}{MT} \sum_{n=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi n}{MT}\right) \tag{6.2}$$

which may be simplified as follows:

$$P(j\omega) = \sum_{k=0}^{M-1} a_k e^{-j\omega kT} \frac{2\pi}{MT} \sum_{n=-\infty}^{\infty} \delta(\omega - \frac{2\pi n}{MT}) \tag{6.3}$$

$$= \sum_{k=0}^{M-1} a_k \frac{2\pi}{MT} \sum_{n=-\infty}^{\infty} e^{-j\frac{2\pi n}{MT}kT} \delta(\omega - \frac{2\pi n}{MT}) \tag{6.4}$$

$$= \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} \left[\frac{1}{M} \sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk} \right] \delta(\omega - \frac{2\pi n}{MT}) \tag{6.5}$$

In summary, for the case of M parallel, time-interleaved channels, with each path k having gain a_k, the Fourier Transform of the resulting impulse sampling train is given by

$$P(j\omega) = \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} A_n \delta(\omega - \frac{2\pi n}{MT}) \tag{6.6}$$

where

$$A_n = \frac{1}{M} \sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk} \tag{6.7}$$

The following points are pertinent regarding the $\{A_n\}$.

Note 1: The coefficients $\{A_n\}$ are the *Discrete Fourier Transform* (DFT) coefficients of the path gains $\{a_k\}$ when the $\{a_k\}$ are considered as an aperiodic sequence, or the *Discrete Fourier Series* (DFS) coefficients of the $\{a_k\}$ when the $\{a_k\}$ are considered as a periodic sequence with period M. In this chapter, the latter interpretation is emphasized.

Note 2: The DFS coefficients $\{A_n\}$ have period M: $A_{M+n} = A_n$. Specifically, for the case of a 4-channel system, i.e., $M = 4$,

$$\ldots = A_{-4} = A_0 = A_4 = \ldots$$

$$\ldots = A_{-3} = A_1 = A_5 = \ldots$$

$$\ldots = A_{-2} = A_2 = A_6 = \ldots$$

$$\ldots = A_{-1} = A_3 = A_7 = \ldots$$

Note 3: For real sequence values, the DFS/DFT coefficients exhibit complex conjugate symmetry. Here, since the path gains $\{a_k\}$ are real, this property always holds. Mathematically,

$$A_{M-n} = \frac{1}{M}\sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)(M-n)k} \tag{6.8}$$

$$= \frac{1}{M}\sum_{k=0}^{M-1} a_k e^{-j2\pi k} e^{j(2\pi/M)nk} \tag{6.9}$$

$$= \frac{1}{M}\sum_{k=0}^{M-1} a_k e^{j(2\pi/M)nk} \tag{6.10}$$

$$= A_n^* \tag{6.11}$$

In particular, the magnitudes are symmetric about the mid-sequence point, i.e.,

$$|A_{M-n}| = |A_n| \tag{6.12}$$

Specifically, for the case of a 4-channel system, i.e., $M = 4$, this implies $|A_1| = |A_3|$.

Using some of the properties mentioned above, the Fourier Transform pair relationship for the mismatched impulse train is illustrated graphically below.

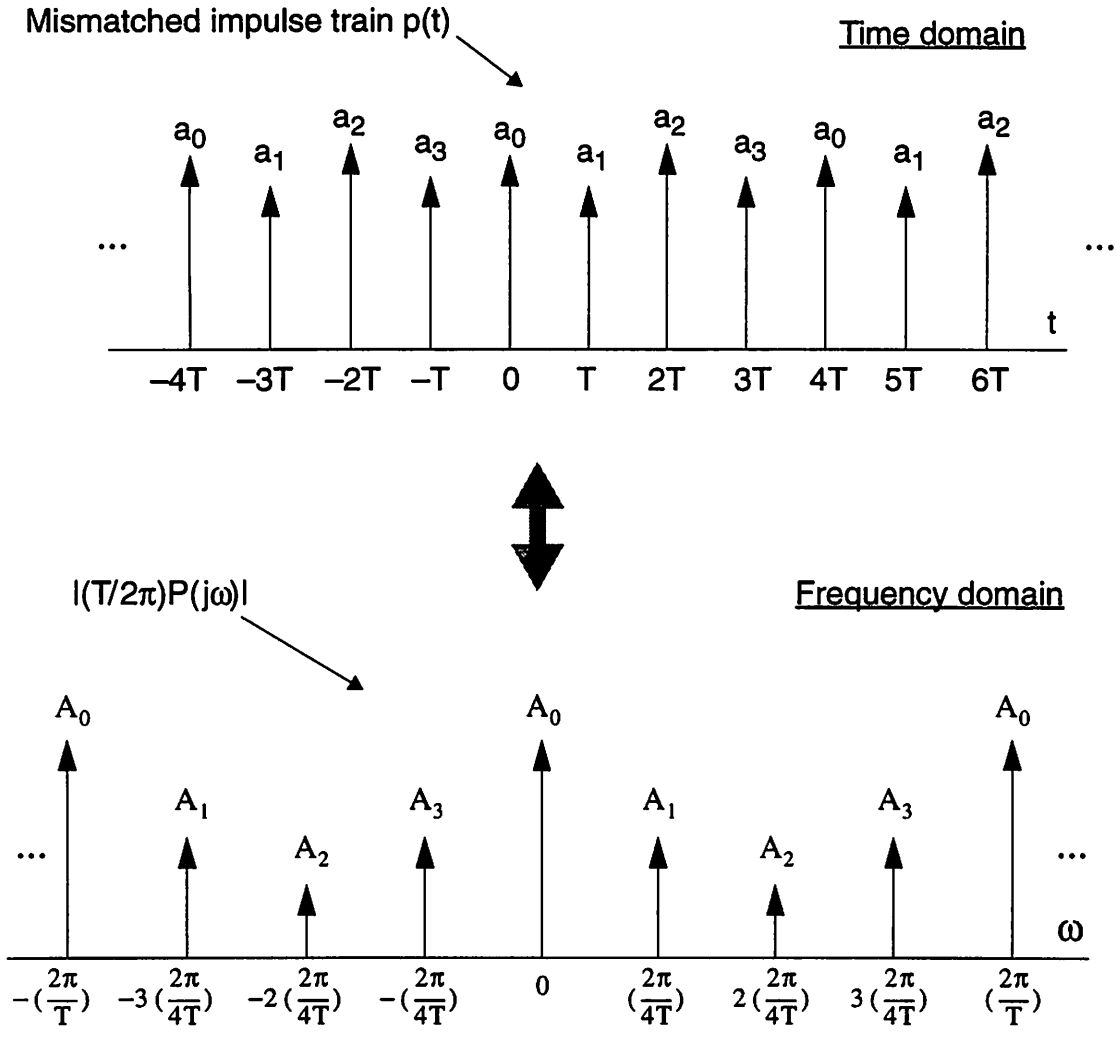


Fig. 6.7 Modified sampling impulse train p(t) in the presence of channel mismatches and the corresponding Fourier Transform P(jω). The output is the convolution of P(jω) and X(jω).

The output spectrum is obtained by convolving the input spectrum with the impulse train P(jω) as given in (6.6) rather than in (6.0): the result is shown below in Fig. 6.8.

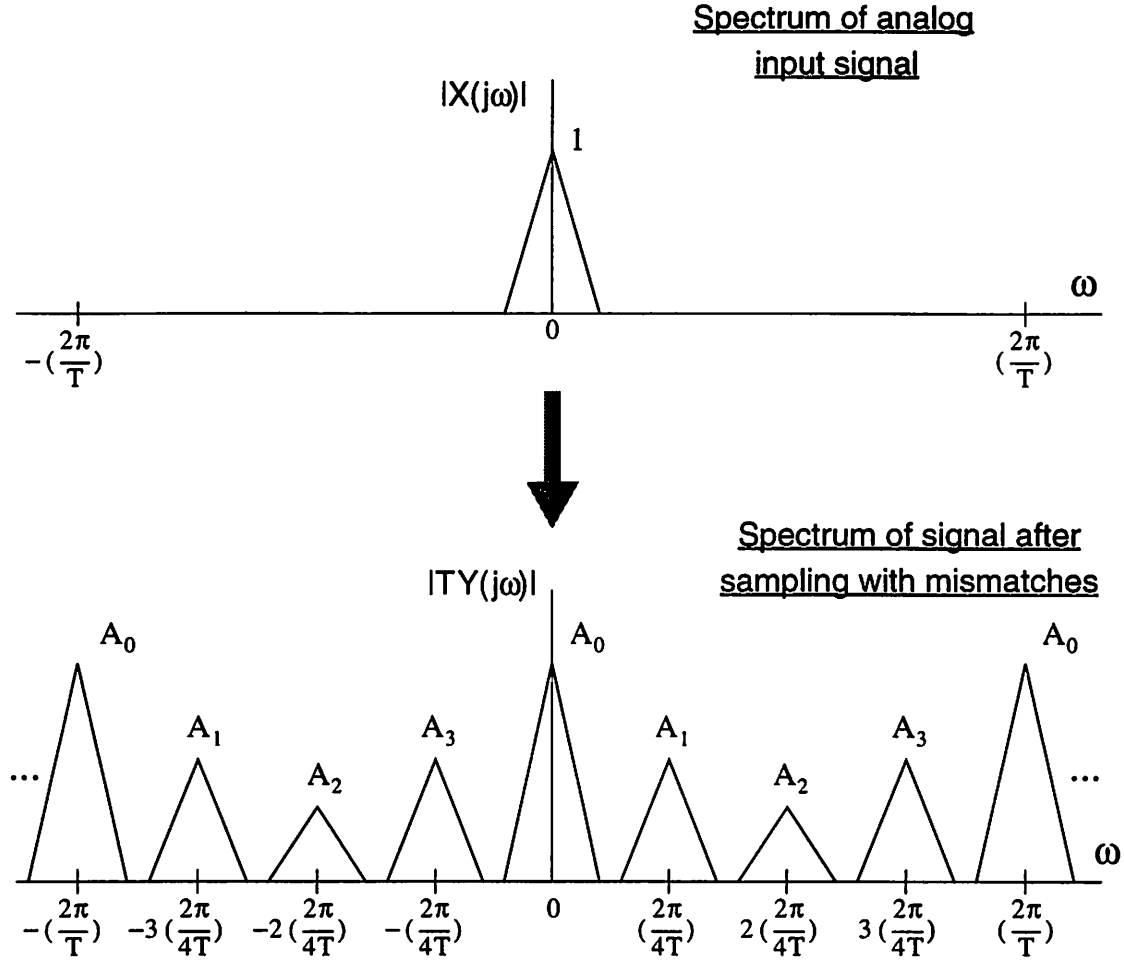


Fig. 6.8 Spectra of continuous-time input signal and discrete-time output signal after sampling by an impulse train from M mismatched parallel channels; Y(jω) is the result of the convolution of X(jω) with the Fourier Transform of the impulse train P(jω) shown in Fig. 6.7.

Comparing the output spectrum shown in Fig. 6.8 with that shown Fig. 6.5 for the ideal case, it is clear that there is a fundamental difference: in the case of mismatched channels, aliases of the input spectrum are present *at multiples of the individual channel sampling rate.*

For simplicity, throughout most of this chapter, the figures show X(jω), the spectrum of the input signal, as having relatively low bandwidth. This is primarily for graphical convenience — to avoid the diagrams becoming cluttered with overlapping aliases. However, it should be borne in mind throughout that, in general, for an ADC sampling at F_s, the input signal spectrum may occupy the entire band from 0 to $F_s/2$. In that case, the spectrum at the output due to sampling with mismatched path gains resembles that shown in Fig. 6.9 below. Note that the channel mismatch aliases overlap in the band of interest, and cannot simply be filtered out, as might appear to be possible from Fig. 6.8.

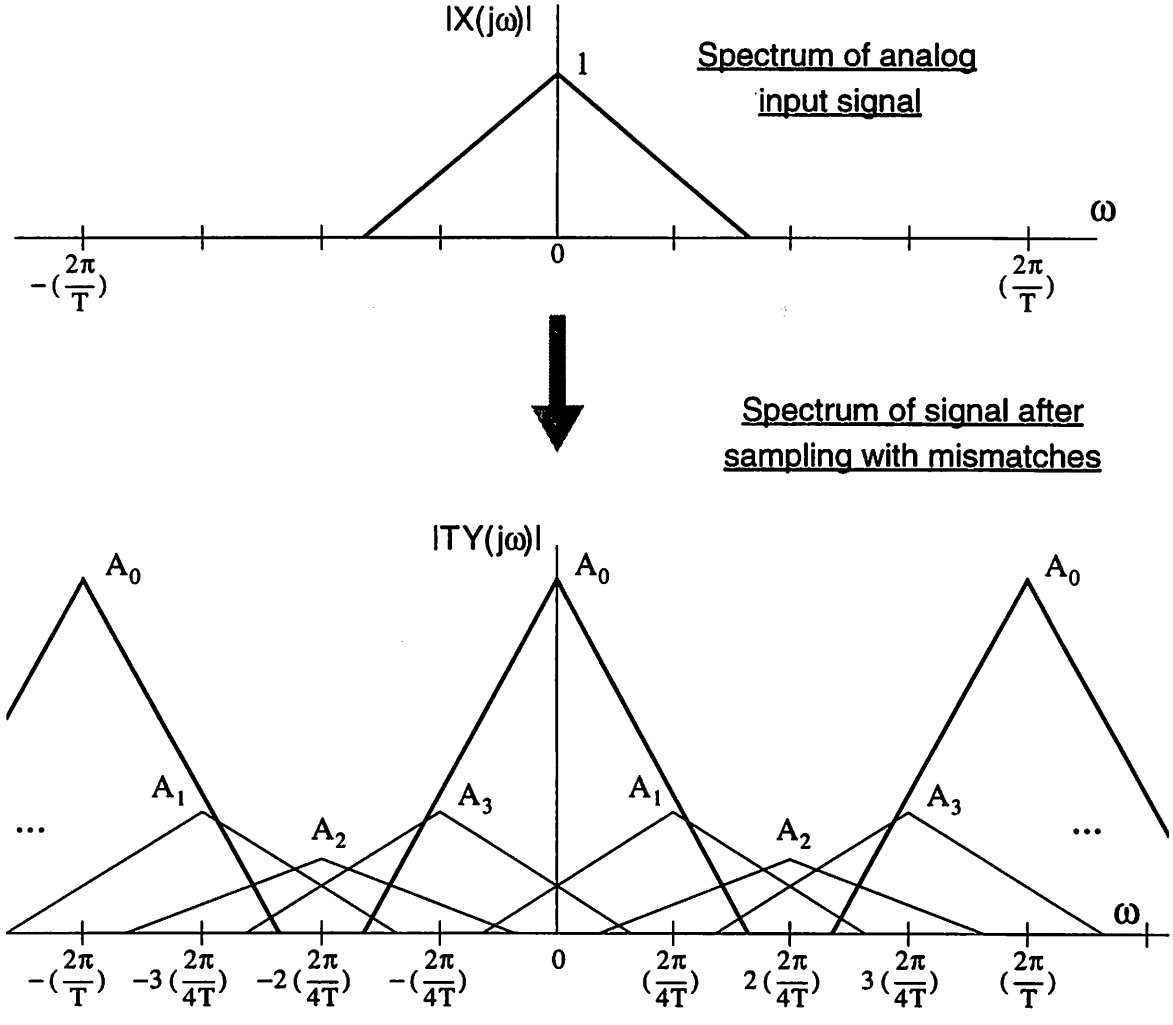


Fig. 6.9 Spectra of continuous-time input signal and discrete-time output signal after sampling by an impulse train from M parallel channels with mismatched gains. Here, X(jω) occupies most of the band from 0 to $F_s/2$ and thus the mismatch aliases overlap *in-band.*

Sampling with Nonideal Mismatched Channel Gains: Analysis Method 2

An alternative way to obtain the expression for $P(j\omega)$ — the spectrum of the sampling impulse train in a system with gain mismatches — is now given. The following well-known Fourier Transform relation is needed.

$$e^{j\omega_0 kT} \leftrightarrow \sum_{n=-\infty}^{\infty} \frac{2\pi}{T}\delta(\omega - \omega_0 + \frac{2\pi n}{T}) \tag{6.13}$$

The key point here is that the channel gains $\{a_k\}$ may be viewed as a periodic discrete-time sequence with period M, and consequently may be represented by a Discrete Fourier Series (DFS), with coefficients $\{A_n\}$, as follows:

$$a_k = \sum_{n=0}^{M-1} A_n e^{j\frac{2\pi k}{M}n} = \sum_{n=0}^{M-1} A_n e^{j(\frac{2\pi n}{MT})kT} \tag{6.14}$$

Using (6.13), the Fourier Transform of the rightmost expression in (6.14) is given by

$$\sum_{n=0}^{M-1} A_n \sum_{r=-\infty}^{\infty} \frac{2\pi}{T}\delta(\omega - \frac{2\pi n}{MT} + \frac{2\pi r}{T}) \tag{6.15}$$

$$= \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} A_n \delta(\omega - \frac{2\pi n}{MT}) \tag{6.16}$$

which is identical to the expression for $P(j\omega)$ obtained before. This approach is discussed further in a qualitative, intuitive way in Section 6.1.2.

Summary

In the case of "ideal" sampling, the Fourier Transform of the sampling impulse train is

$$P(j\omega) = \frac{2\pi}{T}\sum_{n=-\infty}^{\infty}\delta(\omega-\frac{2\pi n}{T}) \tag{6.17}$$

and the Fourier Transform of the sampled waveform is (with $\otimes$ denoting convolution)

$$Y(j\omega) = \frac{1}{2\pi}X(j\omega)\otimes P(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty}X\left[j(\omega-\frac{2\pi n}{T})\right] \tag{6.18}$$

In the case of "mismatched" sampling, the FT of the sampling impulse train is

$$P(j\omega) = \frac{2\pi}{T}\sum_{n=-\infty}^{\infty}A_n\delta(\omega-\frac{2\pi n}{MT}) \tag{6.19}$$

where

$$A_n = \frac{1}{M}\sum_{k=0}^{M-1}a_k e^{-j(2\pi/M)nk} \tag{6.20}$$

and so the FT of the sampled waveform is

$$Y(j\omega) = \frac{1}{2\pi}X(j\omega)\otimes P(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty}A_nX\left[j(\omega-\frac{2\pi n}{MT})\right] \tag{6.21}$$

Usually in practical applications the spectra are plotted versus frequency f (i.e., not angular frequency ω). In that case, if the input spectrum is written as X(f), the above expressions for the output spectrum become the following.

For one channel or multiple channels with ideal matching, the output spectrum is

$$Y(f) = \frac{1}{T}\sum_{n=-\infty}^{\infty}X(f-\frac{n}{T}) = \frac{1}{T}\sum_{n=-\infty}^{\infty}X(f-nF_s) \tag{6.22}$$

For M time-interleaved channels with mismatches, the output spectrum is

$$Y(f) = \frac{1}{T}\sum_{n=-\infty}^{\infty}A_nX(f-\frac{n}{MT}) = \frac{1}{T}\sum_{n=-\infty}^{\infty}A_nX(f-n\frac{F_s}{M}) \tag{6.23}$$

6.1.1 Examples of Gain Mismatch

Some specific examples are now given to illustrate the effects of gain mismatch.

Example 1: 4 channels (M = 4) and 4 mismatched path gains a_0, a_1, a_2, a_3

Using the general expression in (6.21), the output spectrum — i.e., the FT of the sampled waveform — is given by

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty} A_n X\left[j\left(\omega - \frac{2\pi n}{4T}\right)\right] \tag{6.24}$$

where

$$A_n = \frac{1}{4}\sum_{k=0}^{3} a_k e^{-j(2\pi/4)nk} \tag{6.25}$$

Explicitly writing out the expressions for the $\{A_n\}$ yields

$$A_0 = \frac{1}{4}[a_0 + a_1 + a_2 + a_3] \tag{6.26}$$

$$A_1 = \frac{1}{4}[(a_0 - a_2) - j(a_1 - a_3)] \tag{6.27}$$

$$A_2 = \frac{1}{4}[(a_0 - a_1) + (a_2 - a_3)] \tag{6.28}$$

$$A_3 = \frac{1}{4}[(a_0 - a_2) + j(a_1 - a_3)] \tag{6.29}$$

The coefficients $\{A_n\}$ may be interpreted as follows: A_0 is the average gain — expected to be approximately 1; A_1 and A_3 are related to the mismatch between pairs of channels that are 180° out of phase; A_2 depends on the average mismatch between adjacent channels.

Next, in the following cases, specific value are assigned to the $\{A_n\}$.

Example 1a: $a_0 = a_1 = a_2 = 1,\ a_3 = 0$ (Channel 3 totally absent.)

For this case, the Fourier coefficients are

$$A_0 = \frac{3}{4},\ A_1 = -j\frac{1}{4},\ A_2 = -\frac{1}{4},\ A_3 = j\frac{1}{4} \tag{6.30}$$

The time-domain and frequency-domain representations of the sampling impulse train are shown below.

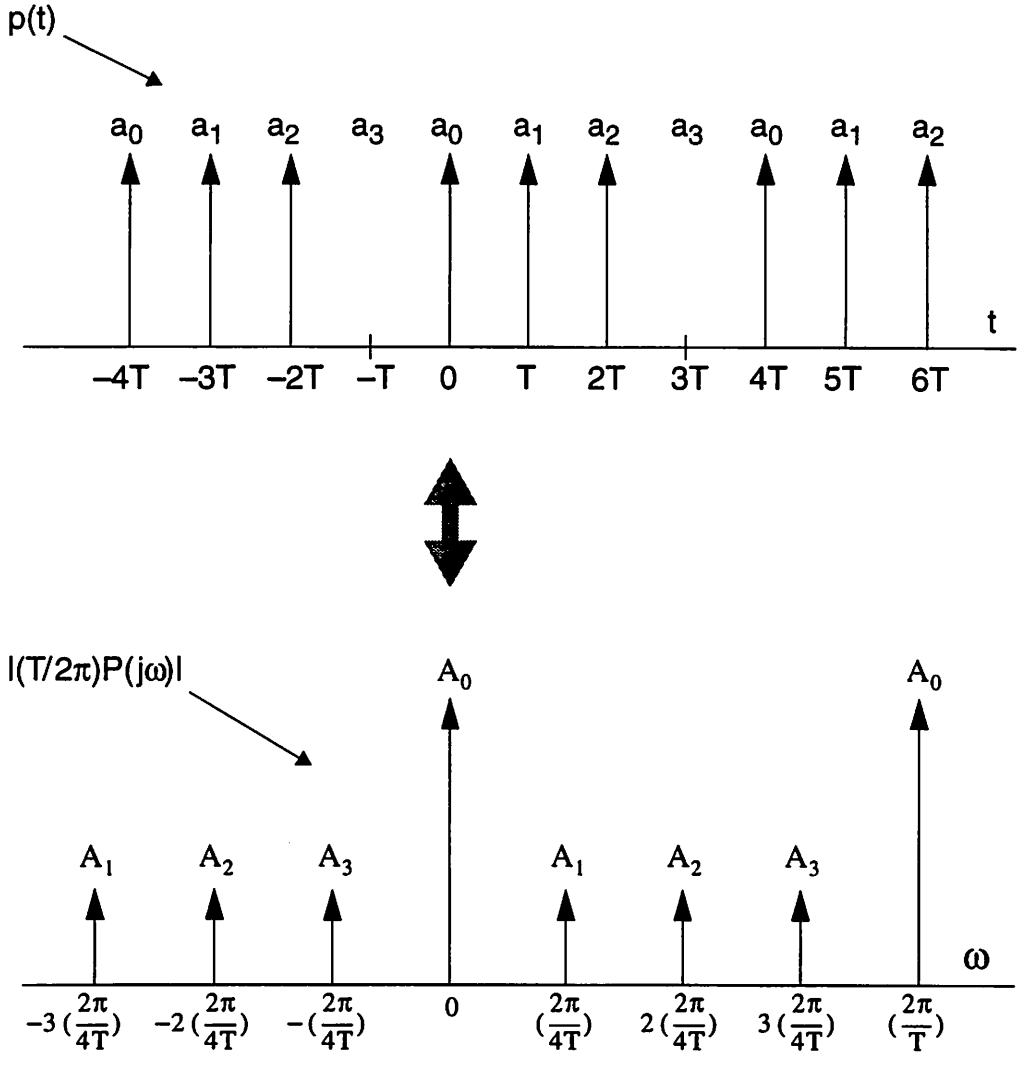


Fig. 6.10 Sampling impulse train p(t) and its Fourier Transform with channel 3 having gain of zero.

In this extreme case, the mismatch aliases have magnitude one third of the signal magnitude.

Example 1b: $a_0 = a_2 = a_3 = 1,\ a_1 = 1 + \varepsilon$ (Channel 1 has gain error ε.)

For this case, the coefficients are

$$A_0 = 1 + \frac{\varepsilon}{4},\ A_1 = -j\frac{\varepsilon}{4},\ A_2 = -\frac{\varepsilon}{4},\ A_3 = j\frac{\varepsilon}{4} \tag{6.31}$$

The time-domain and frequency-domain representations are shown below.

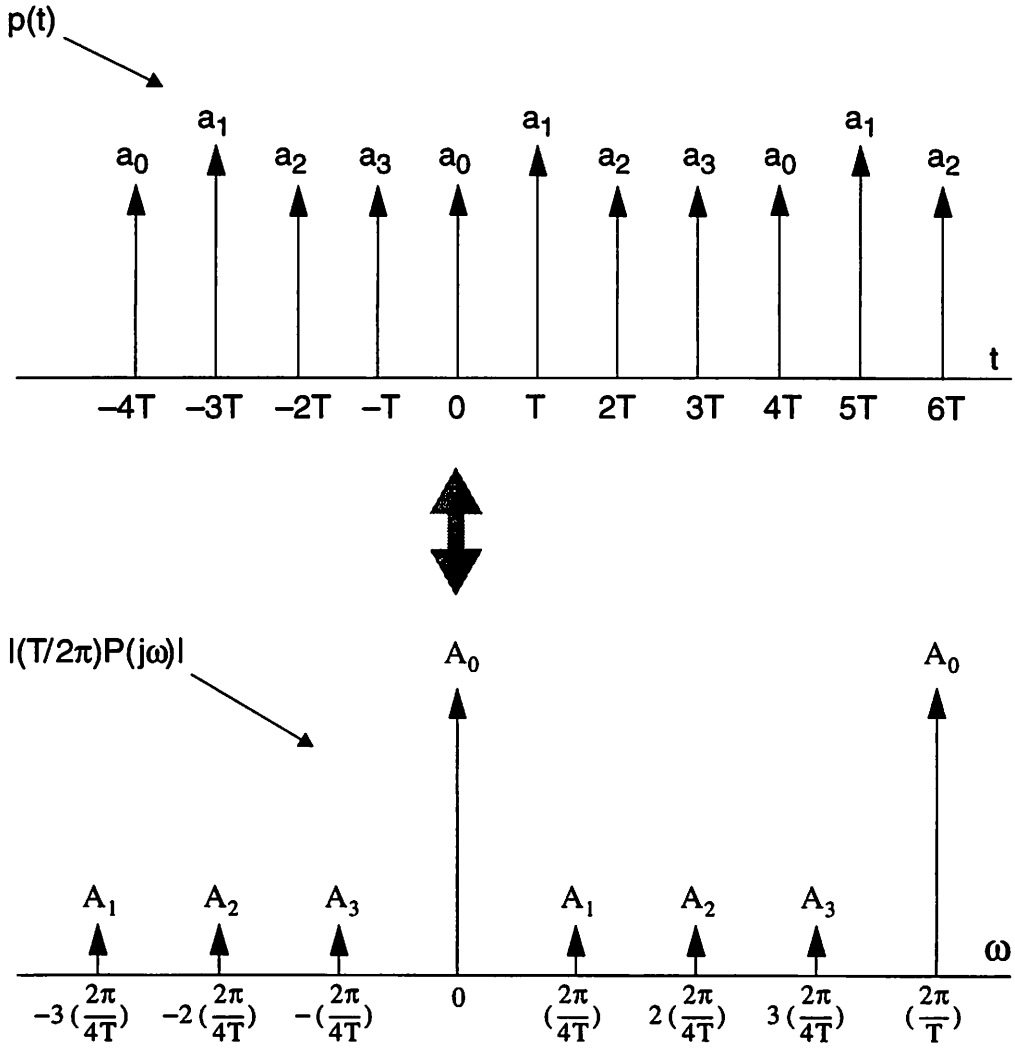


Fig. 6.11 Sampling impulse train with a gain error ε (shown here as positive) in channel 1.

Example 1c: $a_0 = 1, a_1 = 1 + \varepsilon, a_2 = 1, a_3 = 1 - \varepsilon$ (Symmetric gain errors in channels 1 and 3.)

For this case the coefficients $\{A_n\}$ have values

$$A_0 = 1, A_1 = -j\frac{\varepsilon}{2}, A_2 = 0, A_3 = j\frac{\varepsilon}{2} \tag{6.32}$$

The time-domain and frequency-domain representations are shown below.

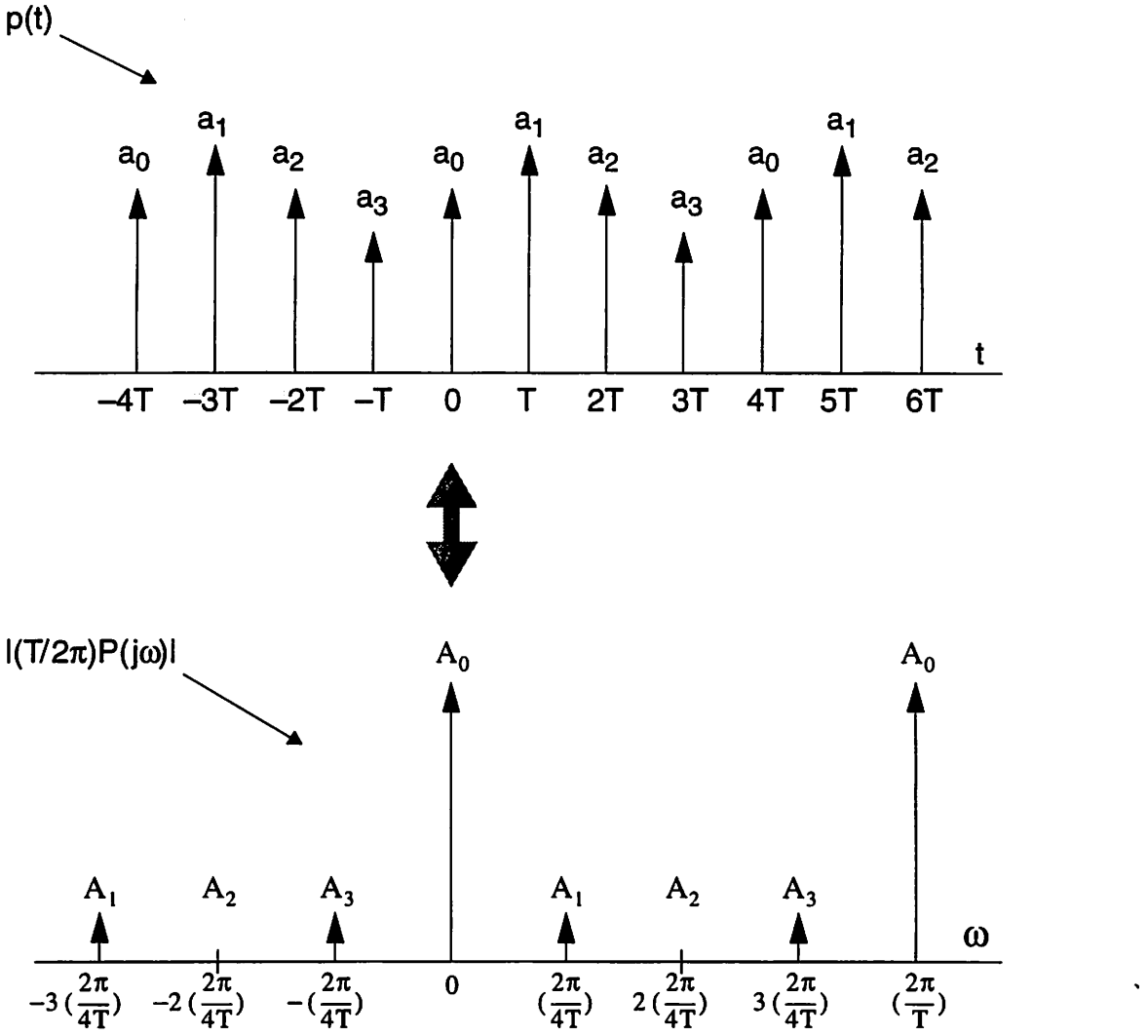


Fig. 6.12 Path gain sequence with errors of equal magnitude but opposite sign in channels 1, 3.

Note that in this case, because of the odd symmetry of the gain sequence, the coefficient A_2 is zero.

Example 2: Special Case — Sinusoidal Input

For the case of a sinusoidal input, the input is of the form

$$x(t) = \sin(\omega_0 t) = \frac{e^{j\omega_0 t} - e^{-j\omega_0 t}}{2j} \tag{6.33}$$

and its spectrum is given by

$$X(j\omega) = 2\pi\left(\frac{\delta(\omega-\omega_0) - \delta(\omega+\omega_0)}{2j}\right) \tag{6.34}$$

From (6.21), the spectrum at the output is

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty} A_n X\left[j\left(\omega - \frac{2\pi n}{MT}\right)\right] \tag{6.35}$$

$$= \frac{2\pi}{T}\sum_{n=-\infty}^{\infty} \frac{A_n}{2j}\left[\delta\left(\omega-\omega_0-\frac{2\pi n}{MT}\right) - \delta\left(\omega+\omega_0-\frac{2\pi n}{MT}\right)\right] \tag{6.36}$$

Typical input and output spectra are illustrated in Fig. 6.13.

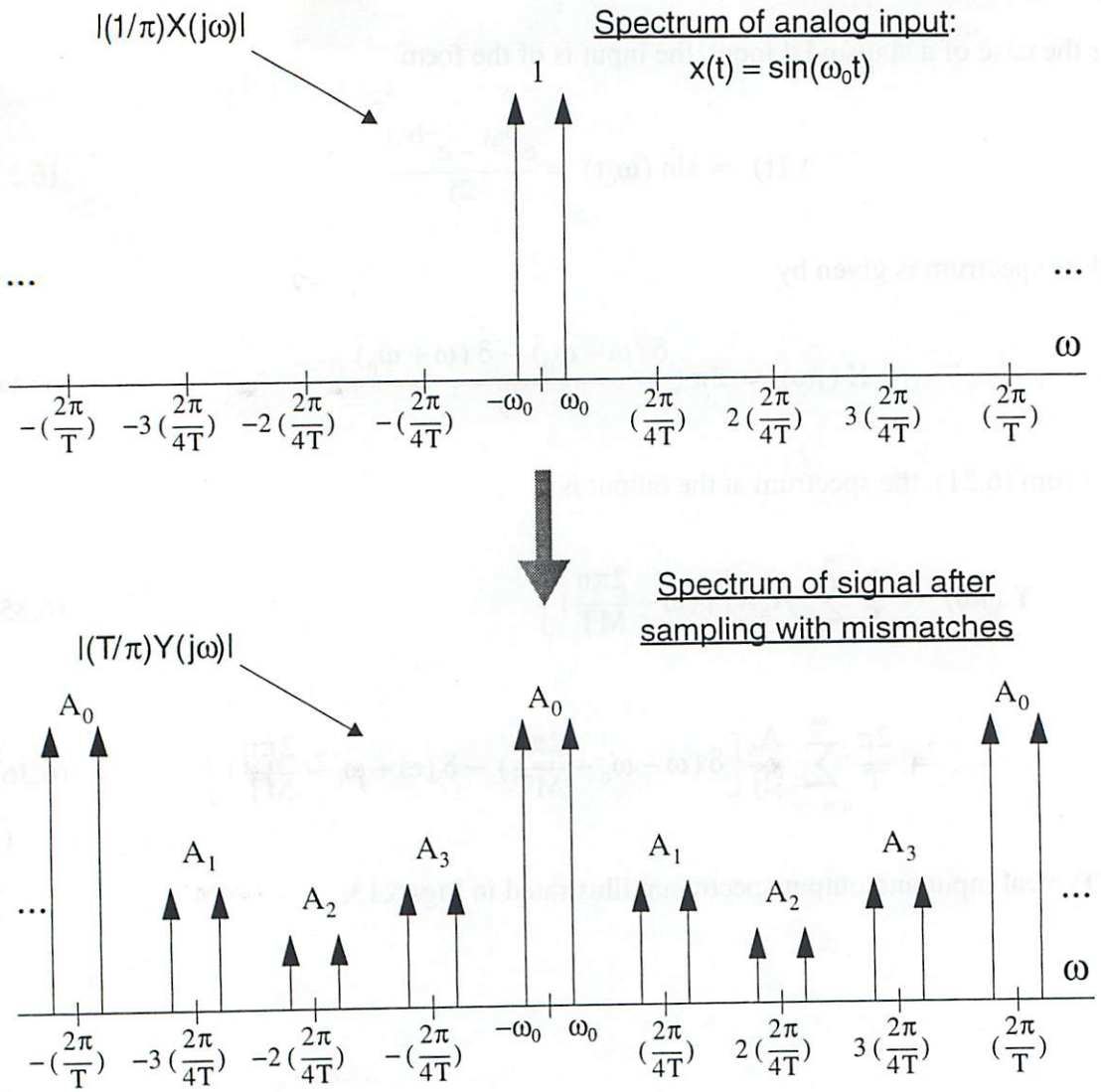


Fig. 6.13 Frequency-domain view when a sinusoid is sampled by a 4-channel time-interleaved ADC with gain mismatches: the output spectrum has sidebands around multiples of $F_s/4$.

Note that the effect of the gain mismatches in the parallel paths is to produce sidebands around multiples of $2\pi/MT$. In terms of the frequency variable f, if the input signal is a sinusoid at frequency f_0, then the output spectrum has additional tones at the frequencies

$$\frac{F_s}{M} \pm f_0,\ 2\frac{F_s}{M} \pm f_0,\ 3\frac{F_s}{M} \pm f_0,\ \ldots,\ (M-1)\frac{F_s}{M} \pm f_0$$

i.e., sidebands around multiples of the individual channel sampling rate. Note that, since in general the input goes from 0 to $F_s/2$, these frequency components are in-band.

6.1.2 Interpretation of Gain Mismatch Effects

This section discusses the preceding analysis and examples: the aim is to present some qualitative explanations and convey some intuition.

Interpretation 1:

If the path gains are close to ideal, i.e.,

$$a_0, a_1, a_2, \ldots, a_{M-1} \approx 1,$$

then

$$|A_0| \approx 1 \text{ and } |A_1| \approx 0, |A_2| \approx 0, \ldots, |A_{M-1}| \approx 0,$$

i.e., the Fourier Transform of the sampling impulse train is close to ideal, and has non-zero values only at the angular frequencies $\omega = 0, \pm 2\pi/T, \pm 4\pi/T, \ldots, \pm 2\pi n/T, \ldots$

Interpretation 2:

Consider the gains $\{a_k\}$ as a periodic discrete-time sequence: all the $a_k \approx 1$ corresponds to *no harmonics in that sequence* — just a DC term, and so

$$|A_0| \approx 1 \text{ and } |A_1| \approx 0, |A_2| \approx 0, \ldots, |A_{M-1}| \approx 0$$

Interpretation 3:

The periodic discrete-time sequence $\{a_k\}$ has a Discrete Fourier Series representation

$$a_k = \sum_{n=0}^{M-1} A_n e^{j(2\pi/M)kn} = \sum_{n=0}^{M-1} A_n e^{j\left(\frac{2\pi n}{MT}\right)(kT)} \tag{6.37}$$

i.e., the periodic sequence is expressible in the form

$$a_k = \sum_{n=0}^{M-1} A_n e^{j\omega_n(kT)} \tag{6.38}$$

where $\omega_n = 2\pi n/MT$.

In this representation, A_0 is the dc term, and the $\{A_n\}$ (for $n > 0$) represent the harmonic content of the mismatched gain sequence. Thus, the sequence $\{a_k\}$ may be viewed as a linear combination of complex exponentials sampled at times kT. Now, since multiplication by a complex exponential in the time domain is equivalent to frequency translation in the frequency domain, the net result of the operation of multiplying the input by the waveform given in (6.38) is expected to be of the form

$$\sum_{n=0}^{M-1} A_n X\left[j(\omega - \frac{2\pi n}{MT})\right] \tag{6.39}$$

In fact, because of the discrete-time nature of the system, the above expression is repeated periodically in frequency:

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty} A_n X\left[j(\omega - \frac{2\pi n}{MT})\right] \tag{6.40}$$

Interpretation 4:

The process of sampling may always be viewed as modulation. Here, x(t) is the "message" and has spectrum $X(j\omega)$; the frequencies

$$\omega_n = 0, \pm 2\pi/MT, \pm 4\pi/MT, \ldots, \pm 2\pi n/MT, \ldots$$

correspond to the carrier frequencies. Associated with each modulation frequency is a gain or "strength" A_n, where the $\{A_n\}$ are the DFS coefficients of the gains $\{a_k\}$. This leads naturally to the output spectrum shown in Fig. 6.8, i.e., the baseband spectrum is *shifted* or *translated* to the carrier frequencies $\omega_n = 0, \pm 2\pi/MT, \pm 4\pi/MT, \ldots, \pm 2\pi n/MT, \ldots$

6.1.3 Offset Mismatch

A schematic representation of nonideal channel offsets is shown below.

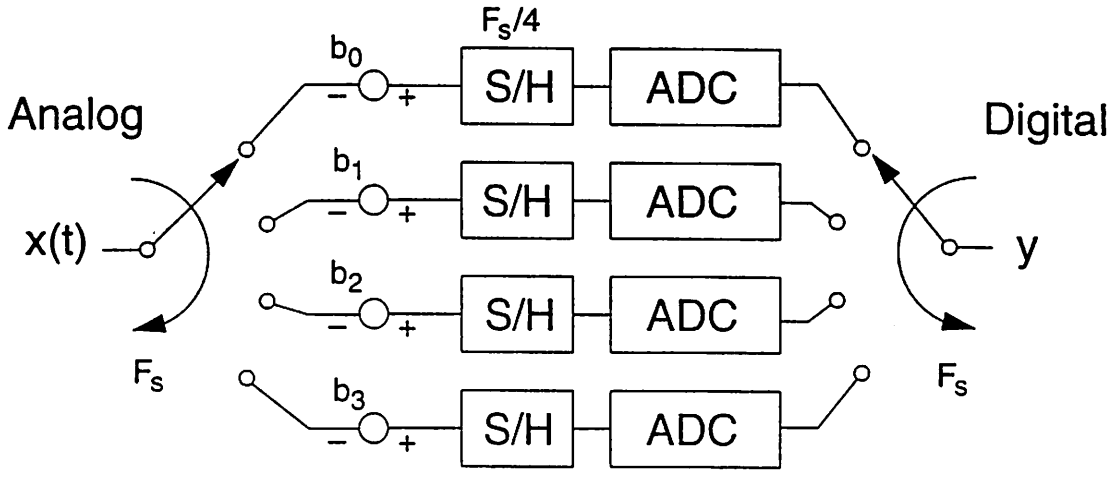


Fig. 6.14 Basic model for offset mismatches in a 4-channel parallel time-interleaved ADC.

In this case, a discrete-time periodic sequence of period M given by

$$\ldots, b_0, b_1, b_2, \ldots, b_{M-1}, b_0, b_1, b_2, \ldots, b_{M-1}, b_0, b_1, b_2, \ldots, b_{M-1}, b_0, \ldots$$

is *added* to input signal, as indicated schematically below for the case of four channels.

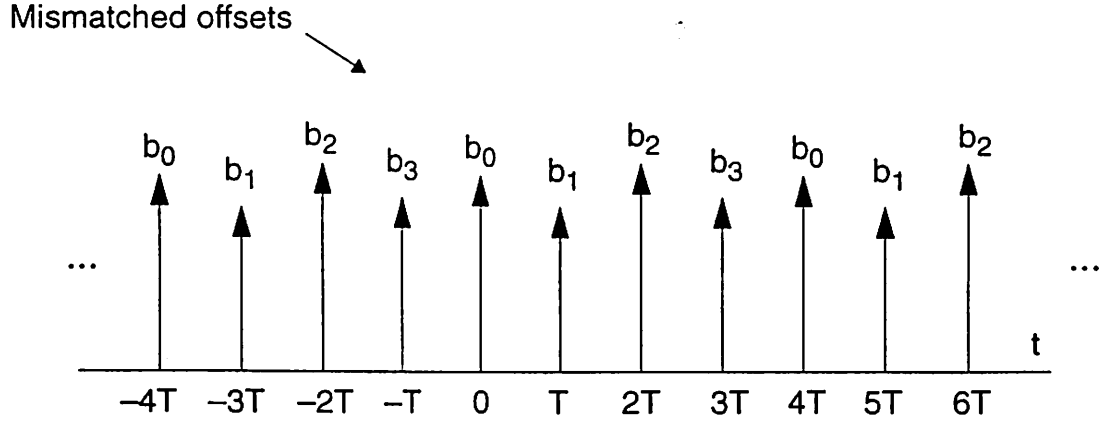


Fig. 6.15 Offset mismatch sequence in a 4-channel parallel time-interleaved analog system.

Calculation of the Fourier Transform of the offset sequence is identical to that of the gain mismatch sequence given previously. Since the mismatched offset sequence is *added* to the input, its spectrum is added to the output spectrum; this is in contrast to the case of

gain mismatch, in which the FT of the gain mismatch sequence is *convolved* with the input spectrum. The FT of the output resulting from the offset mismatch is therefore given by

$$Y(j\omega) = \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} B_n \delta\left(\omega - \frac{2\pi n}{MT}\right) \tag{6.41}$$

where

$$B_n = \frac{1}{M} \sum_{k=0}^{M-1} b_k e^{-j(2\pi/M)nk} \tag{6.42}$$

This relationship is shown below in Fig. 6.16.

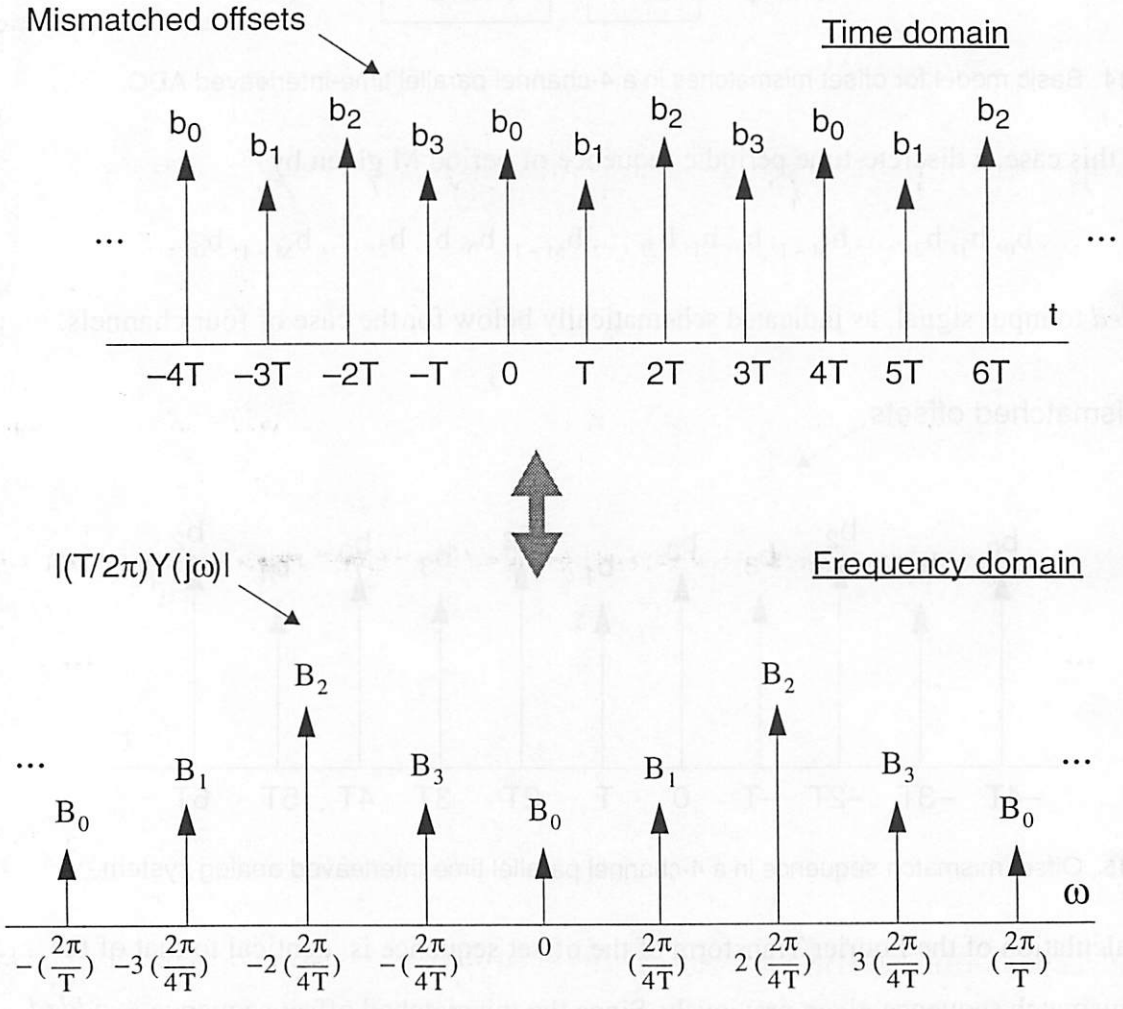


Fig. 6.16 Time-domain and frequency-domain views of offset mismatch in parallel channels.

The overall spectrum at the output, including the input signal, is shown in Fig. 6.8.

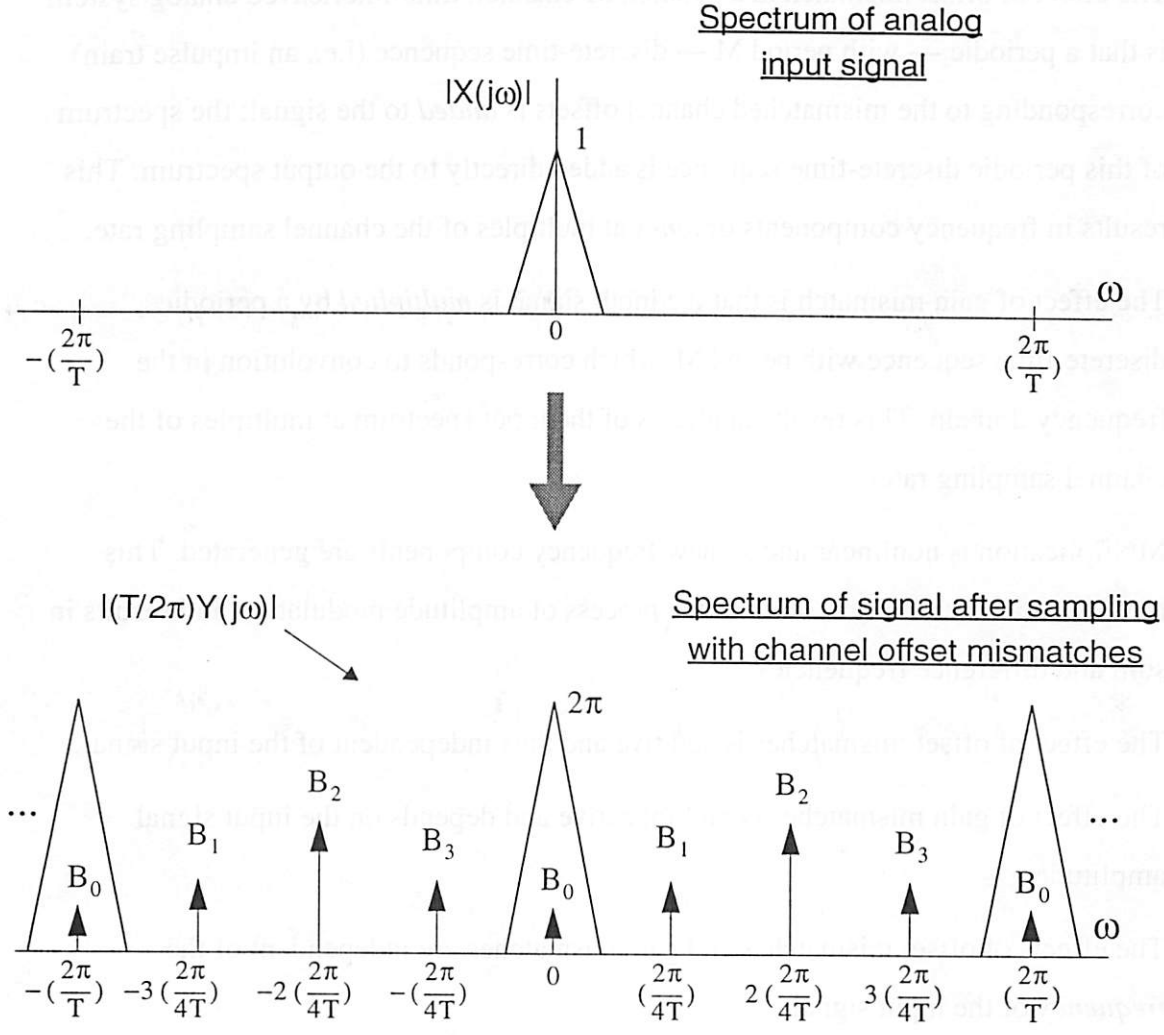


Fig. 6.17 Input spectrum X(jω) and output spectrum Y(jω) that results from sampling by an impulse train from M = 4 parallel channels with offset mismatches. No gain mismatches are present, i.e., all the path gains $a_k = 1$.

Writing the output spectrum as a function of f rather than ω gives

$$Y(f) = \frac{1}{T}\sum_{n=-\infty}^{\infty} B_n \delta\left(f - \frac{n}{MT}\right) = \frac{1}{T}\sum_{n=-\infty}^{\infty} B_n \delta\left(f - n\frac{F_s}{M}\right) \tag{6.43}$$

It is clear that the effect of offset mismatch is to produce frequency components at integer multiples of the channel sampling rate.

6.1.4 Summary

- The effect of offset mismatch in a parallel, M-channel, time-interleaved analog system is that a periodic — with period M — discrete-time sequence (i.e., an impulse train) corresponding to the mismatched channel offsets is *added* to the signal; the spectrum of this periodic discrete-time sequence is added directly to the output spectrum. This results in frequency components or *tones* at multiples of the channel sampling rate.
- The effect of gain mismatch is that the input signal is *multiplied* by a periodic discrete-time sequence with period M, which corresponds to convolution in the frequency domain. This results in aliases of the input spectrum at multiples of the channel sampling rate.
- Multiplication is nonlinear and so new frequency components are generated. This multiplication or *mixing* resembles the process of amplitude modulation and results in sum and difference frequencies.
- The effect of offset mismatches is additive and thus independent of the input signal.
- The effect of gain mismatches is multiplicative and depends on the input signal amplitude.
- The effects of offset mismatches and gain mismatches are independent of the *frequency* of the input signal.
- Exact expressions have been derived for the values of the aliases and tones caused by mismatches — these values are seen to be the DFS coefficients of the path gains and offsets.
- For simplicity, most of the *figures* in this section have shown the input spectrum as occupying significantly less than the full Nyquist band $[0, F_s/2]$. However, all the analysis, results, and discussion apply to time-interleaved ADC systems with input signals occupying the full Nyquist band.
- Finally, note that the treatment so far has been deterministic: later, in Section 6.4, the statistical distributions of the gains and offsets are considered.

6.2 MULTIRATE VIEW OF CHANNEL GAIN MISMATCHES: DECIMATION/INTERPOLATION ANALYSIS

6.2.0 Introduction: Time-Interleaved ADC as a Multirate System

This section considers a time-interleaved A/D converter from the point of view of multirate signal processing. First, note that a key assumption necessary for this analysis is that the system timing is uniform — i.e., the analysis deals with the input samples {x(iT)} as if they are sampled uniformly. Some background on multirate signal processing is now reviewed [24]. Two key multirate blocks used in the following analysis are the *downsampler* and *upsampler*, shown in Fig. 6.18 below. These two blocks are defined as follows.

- A *downsampler* with parameter M takes an input sequence and outputs every Mth sample at a rate 1/M of the input rate.
- An *upsampler* with parameter L takes an input sequence and outputs a sequence at L times the input rate with L–1 zeros inserted between each pair of input samples.

Downsampler: chooses every Mth sample

Upsampler: fills in (L–1) zeros

Fig. 6.18 Key multirate blocks: downsampler and upsampler.

Using these blocks, a time-interleaved ADC may be represented as shown in Fig. 6.19.

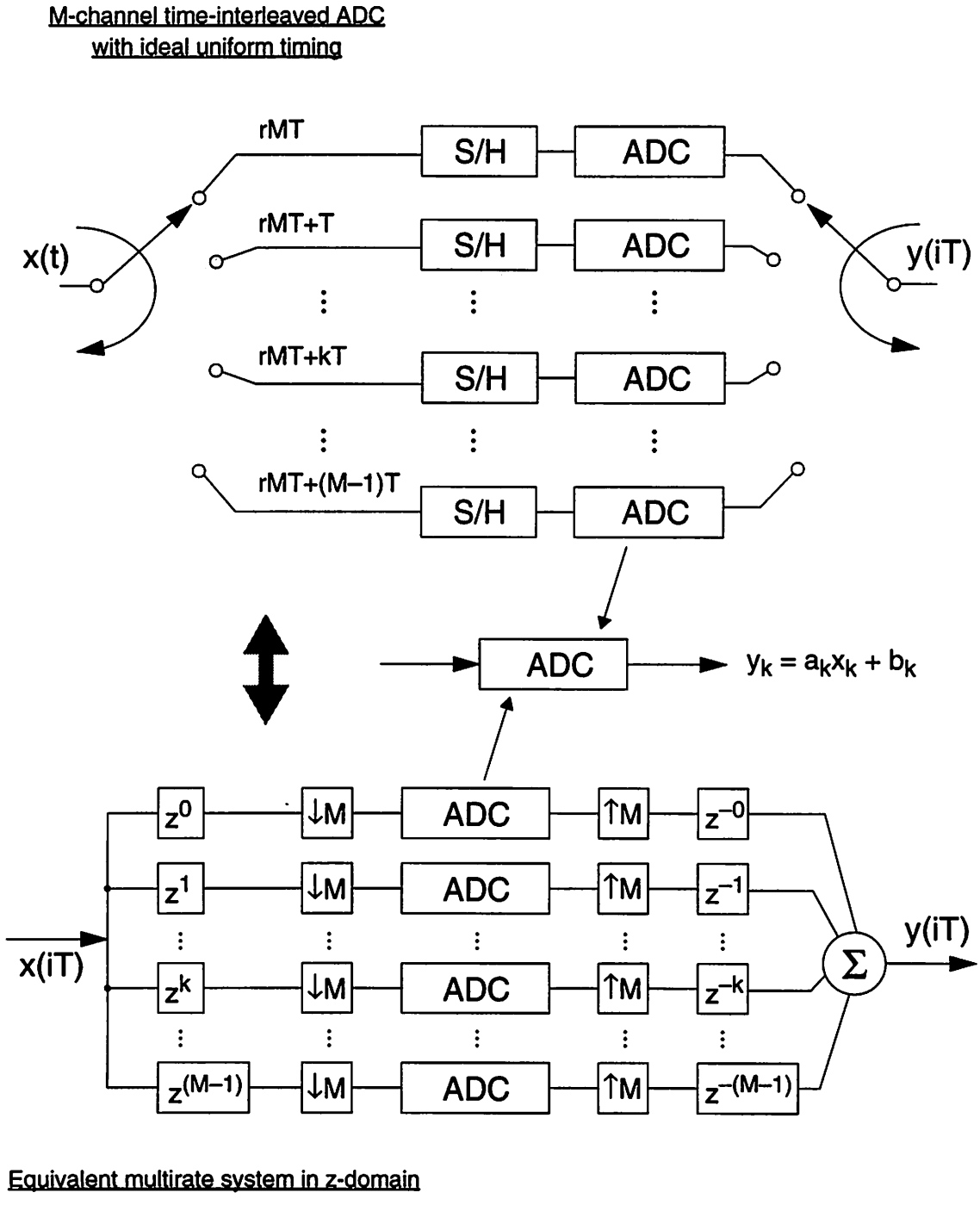


Fig. 6.19 Time-interleaved A/D converter (with ideal uniform sampling) and z-domain model.

6.2.1 Frequency-Domain Analysis

The basic approach of this analysis is to obtain the transfer function for a single path in the parallel system, and then to combine the terms for all channels, taking appropriate account of phase. Before presenting the full analytical expressions for an M-channel case, the behavior in the frequency domain is illustrated graphically for a 4-channel case in Fig. 6.20 on the next page. At the top of the figure, a model for one path (path k) of the time-interleaved ADC is indicated; below that, the spectra at various points are shown. A brief explanation of the spectra at the various points is now given.

(0): This is the continuous-time input $X(j\omega)$; for convenience on the graphs, the magnitude of the baseband spectrum at dc is chosen to be T.

(1): This shows the spectrum after uniform sampling at rate $F_s = 1/T$: $X(j\omega)$ is replicated over the entire angular frequency axis at intervals of $2\pi/T$.

(2): The magnitude spectrum here is identical to that at point **(1)**, since the z^k block — an ideal advance — only affects the phase.

(3): Two interpretations are possible for the spectrum here: (a) it may be viewed as the discrete-time signal at point **(2)** downsampled by 4, giving aliases spaced $1/4(2\pi/T)$ apart, or (b) it may be viewed as the original continuous-time input spectrum sampled at the rate $F_s/4$.

(4): For this analysis, the ADC is modeled simply as a gain block with gain a_k — so this is the spectrum at point **(3)** multiplied by a_k.

(5), (6): The magnitude spectra here are unchanged from point **(4)**.

(Combined): This plot illustrates the final magnitude spectrum obtained by combining the spectra at point **(6)** for all 4 channels. In the ideal case, with all channel gains equal, the undesired aliases cancel.

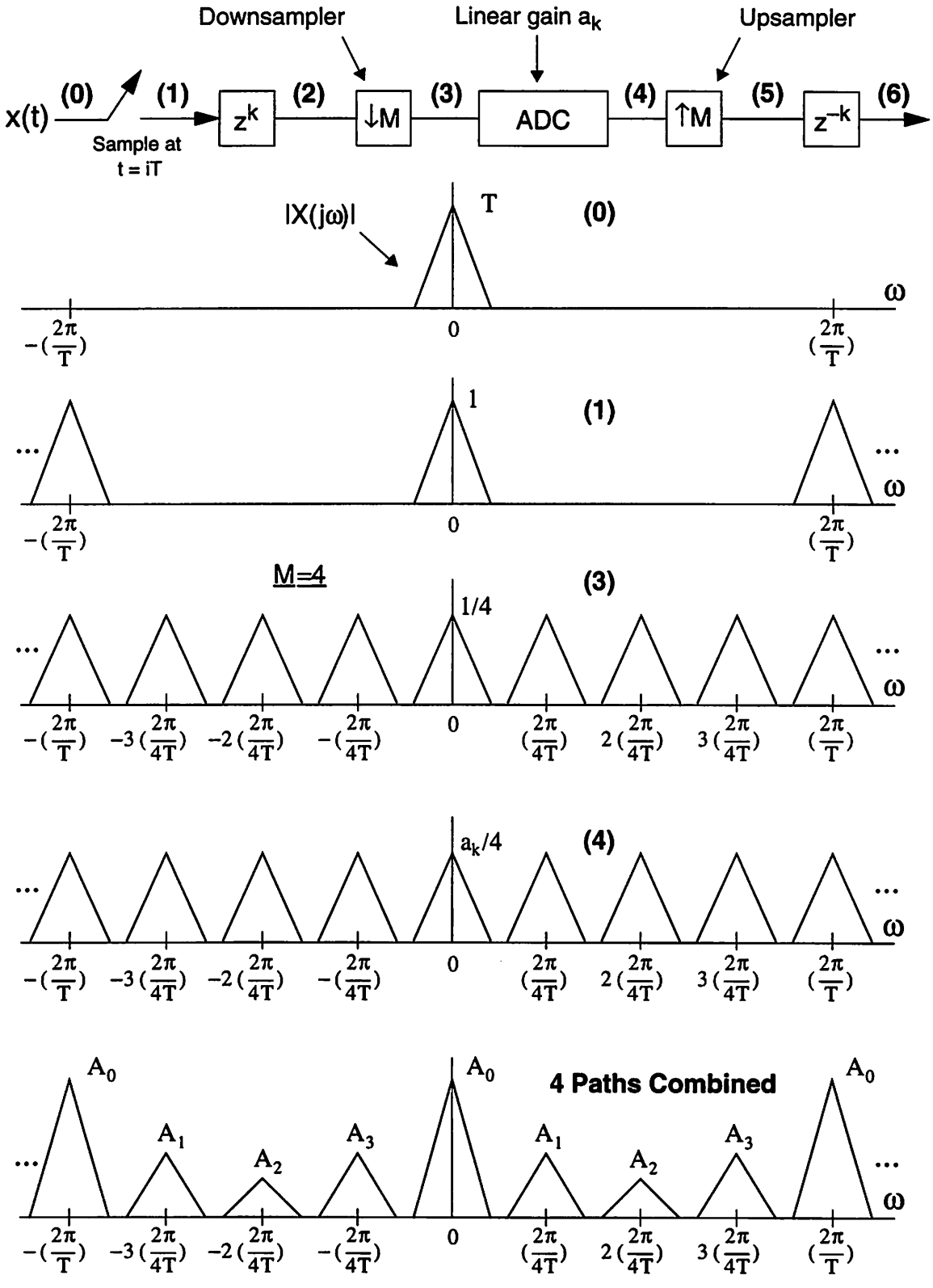


Fig. 6.20 Spectra at various points in one path of a time-interleaved ADC and at the overall output.

The block diagram for path k of the multichannel system is repeated below.

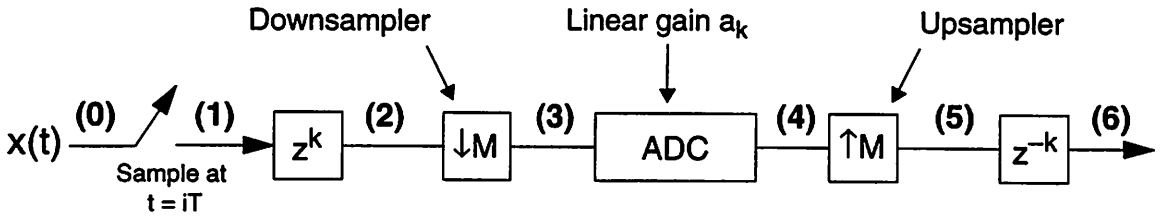


Fig. 6.21 Model for a single path within an M-channel time-interleaved ADC.

The analytical expressions corresponding to indicated points on Fig. 6.21 and corresponding to the spectra of Fig. 6.20 are now given.

(0): $x(t) \leftrightarrow X(j\omega)$

(1): $$x(iT) \leftrightarrow \frac{1}{T}\sum_{n=-\infty}^{\infty} X\left[j\left(\omega - \frac{2\pi n}{T}\right)\right]$$

(2): $$x(iT+kT) \leftrightarrow \frac{1}{T}\sum_{n=-\infty}^{\infty} e^{j\left(\omega-\frac{2\pi n}{T}\right)kT} X\left[j\left(\omega - \frac{2\pi n}{T}\right)\right]$$

(3): $$\frac{1}{T}\sum_{n=-\infty}^{\infty} \frac{1}{M} e^{j\left(\omega-\frac{2\pi n}{MT}\right)kT} X\left[j\left(\omega - \frac{2\pi n}{MT}\right)\right]$$

(4), (5): $$\frac{1}{T}\sum_{n=-\infty}^{\infty} \left[\frac{a_k}{M} e^{j\left(\omega-\frac{2\pi n}{MT}\right)kT}\right] X\left[j\left(\omega - \frac{2\pi n}{MT}\right)\right]$$

(6): $$e^{-j\omega kT} \cdot \frac{1}{T}\sum_{n=-\infty}^{\infty} \left[\frac{a_k}{M} e^{j\left(\omega-\frac{2\pi n}{MT}\right)kT}\right] X\left[j\left(\omega - \frac{2\pi n}{MT}\right)\right]$$

$$= \frac{1}{T}\sum_{n=-\infty}^{\infty} \left[\frac{a_k}{M} e^{-j\left(\frac{2\pi n}{MT}\right)kT}\right] X\left[j\left(\omega - \frac{2\pi n}{MT}\right)\right]$$

The Fourier Transform of the signal at point **(3)** — i.e., after the downsampler — can also be obtained directly by considering channel k to sample the advanced signal $x(t+kT)$ at rate 1/MT. Note that since the upsampler only inserts zeros, the spectrum (with respect to absolute frequency) is the same at points **(4)** and **(5)**.

Finally, from Fig. 6.19, it is clear that the final output is obtained by adding the terms at **(6)** for all the channels; this yields the same result as in (6.20) and (6.21), i.e.,

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty} A_n X\left[j\left(\omega - \frac{2\pi n}{MT}\right)\right] \tag{6.44}$$

where

$$A_n = \frac{1}{M}\sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk} \tag{6.45}$$

Interpretation:

The above discussion leads to the following interpretation of the frequency-domain behavior when gain mismatches are present. If the path gains are close to ideal, i.e., if $a_0, a_1, a_2, \ldots, a_{M-1} \approx 1$, then the aliased components of $X(j\omega)$ combine *so as to cancel, except at multiples of the system sampling frequency,* that is, only at angular frequencies $\omega = 0, \pm 2\pi/T, \pm 4\pi/T, \ldots, \pm 2\pi n/T, \ldots$ is any energy present in the output spectrum.

It is apparent that the terms $(a_k/M)\,e^{-j(\frac{2\pi n}{MT})kT}$ consist of both a gain contribution and a phase contribution. Assume *equal* path gains $\{a_k\}$: note that when added together over all channels k, the phase terms $(1/M)\,e^{-j(\frac{2\pi n}{MT})kT}$ sum to 1 when n is a multiple of M, and sum to zero when n is not a multiple of M. However, when the $\{a_k\}$ are *not equal*, perfect cancellation of the "gain-phase" product terms $a_k e^{-j(\frac{2\pi n}{MT})kT}$ does not occur, and thus some residual portion of the aliased components at multiples of $2\pi/MT$ remains in the output spectrum.

6.2.2 Z-Domain Analysis

In this section the analysis is performed completely in the z-domain [105]. First, the decimator and interpolator are introduced and their behavior in the z-domain described analytically [24]. Decimation and interpolation are illustrated schematically in Fig. 6.22.

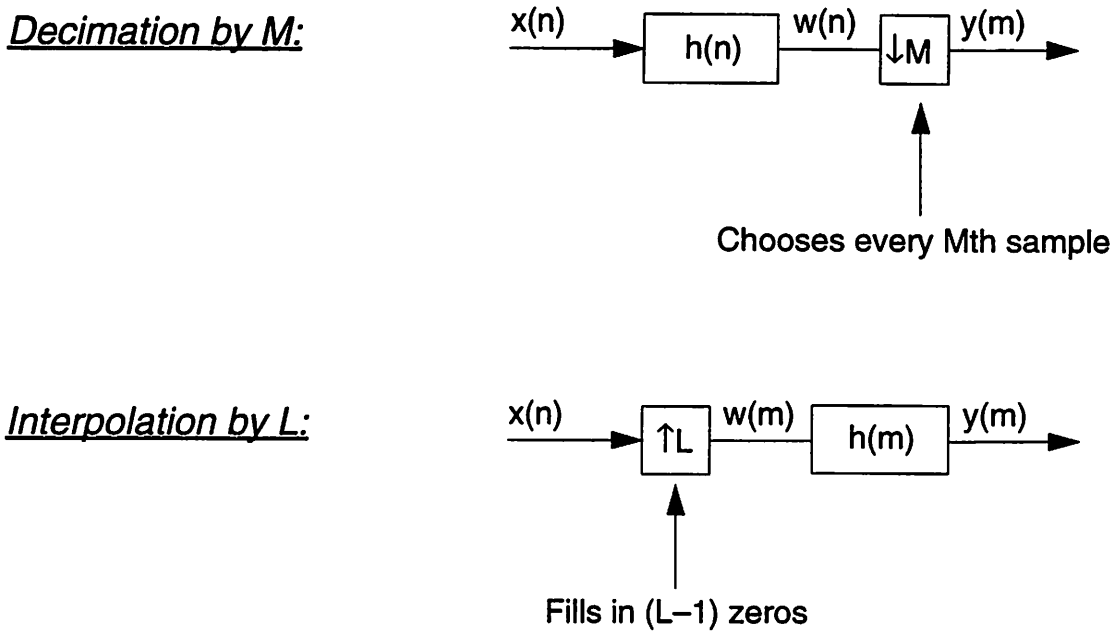


Fig. 6.22 Basic schematic and definitions for decimator and interpolator blocks

The input-output relationships for these block in the z domain are as follows [24].

Decimation by M:

$$Y(z) = \frac{1}{M}\sum_{n=0}^{M-1} H[e^{-j(2\pi n/M)}z^{1/M}]\,X[e^{-j(2\pi n/M)}z^{1/M}] \tag{6.46}$$

Interpolation by L:

$$Y(z) = H(z)\,X(z^{L}) \tag{6.47}$$

The basic approach of this analysis is similar to earlier in this section: first, the transfer function for a single path in the parallel system is obtained, and then the terms for all channels are combined, taking appropriate account of phase. An individual channel k is shown in Fig. 6.23 below.

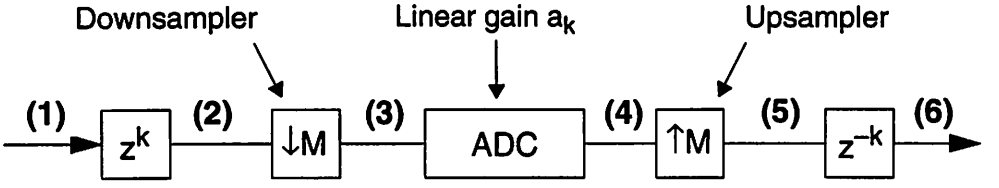


Fig. 6.23 Z-domain model of a single path (path k) within an M-channel time-interleaved ADC.

The z-domain representation of the signal at the various points indicated on the above block diagram is given below.

(1): $x(iT) \leftrightarrow X(z)$

(2): $x(iT + kT) \leftrightarrow z^k X(z)$

(3): $$\frac{1}{M}\sum_{n=0}^{M-1} e^{-j(2\pi kn/M)} z^{k/M} X\left[e^{-j(2\pi n/M)} z^{1/M}\right]$$

(4): $$a_k\left(\frac{1}{M}\sum_{n=0}^{M-1} e^{-j(2\pi kn/M)} z^{k/M} X\left[e^{-j(2\pi n/M)} z^{1/M}\right]\right)$$

(5): $$\frac{a_k z^k}{M}\sum_{n=0}^{M-1} e^{-j(2\pi kn/M)} X\left[e^{-j(2\pi n/M)} z\right]$$

(6): $$\frac{a_k}{M}\sum_{n=0}^{M-1} e^{-j(2\pi kn/M)} X\left[e^{-j(2\pi n/M)} z\right]$$

The z-transform expression at point **(3)** above — i.e., after the downsampler — is obtained by applying the relation in (6.46), with $H(z) = z^k$; expression **(5)** is obtained by using (6.47) with $H(z) = 1$ and $L = M$.

Following the same procedure as in Section 6.2.1, the final output of the time-interleaved system is obtained by combining terms at point **(6)** from all M channels.

Therefore

$$Y(z) = \sum_{n=0}^{M-1}\left[\frac{1}{M}\sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk}\right] X\left[e^{-j(2\pi n/M)} z\right] \tag{6.48}$$

which can be written in the form

$$Y(z) = \sum_{n=0}^{M-1} A_n X\left[e^{-j(2\pi n/M)} z\right] \tag{6.49}$$

where

$$A_n = \frac{1}{M}\sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk} \tag{6.50}$$

The spectrum is obtained by setting $z = e^{j\omega T}$, which yields

$$Y(e^{j\omega T}) = \sum_{n=0}^{M-1} A_n X\left(e^{j(\omega - \frac{2\pi n}{MT})T}\right) \tag{6.51}$$

To relate this back to continuous time, recall that a basic relation for the Discrete-time Fourier Transform is

$$X(e^{j\omega T}) = \frac{1}{T}\sum_{n=-\infty}^{\infty} X\left[j(\omega - \frac{2\pi n}{T})\right] \tag{6.52}$$

and substituting this back into the right-hand side of (6.51) gives

$$Y(j\omega) = \sum_{n=0}^{M-1} A_n X\left(e^{j(\omega - \frac{2\pi n}{MT})T}\right) = \frac{1}{T}\sum_{n=-\infty}^{\infty} A_n X\left[j(\omega - \frac{2\pi n}{MT})\right] \tag{6.53}$$

as before.

A Remark on Distributor/Commutator Notation: There are two possible notational conventions to describe the sampling sequence from one channel to the next in a time-interleaved ADC. All the system diagrams so far in this chapter have employed a distributor at the input going *clockwise* from channel 0 to channel M–1. In this notation, which is used (implicitly) in [8] and [53], at any instant, channel 0 has the oldest sample and channel M–1 has the most recent, and unrealizable advances are used in the discrete-time model. However, it is also possible for the distributor at the input to go *counterclockwise* from channel M–1 to channel 0, as in [105]. In this approach, at any given time, channel 0 is processing the most recent sample and channel (M–1) is processing the oldest, and realizable delays are used in the discrete-time model.

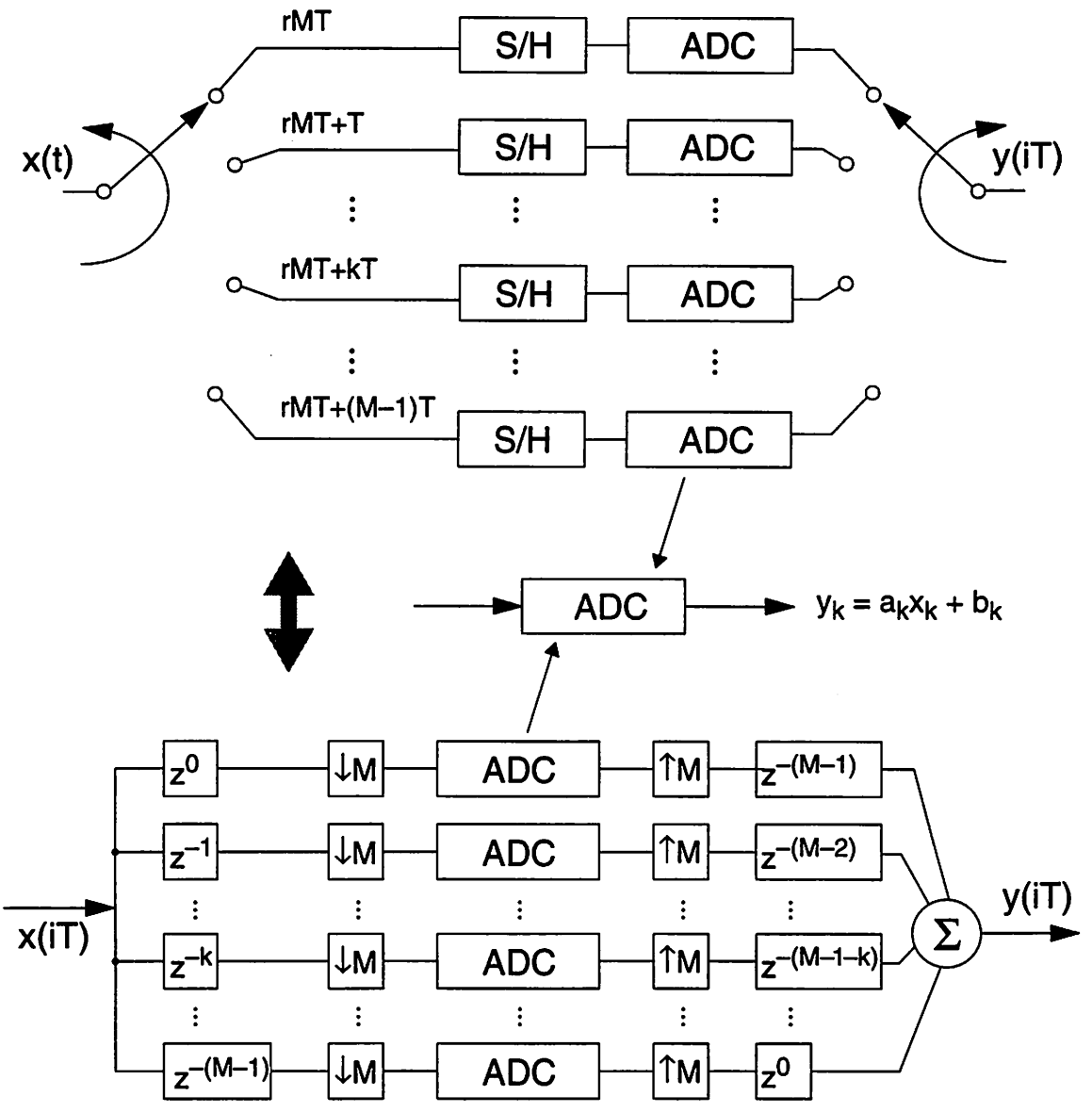


Fig. 6.24 Z-domain model for an M-channel time-interleaved ADC with ideal uniform sample timing [105]. This model employs a counterclockwise distributor and realizable delays instead of advances.

The above diagrams should be compared with those in Fig. 6.19: the point here is that the results are the same in both cases (ignoring fixed delays) but that in the literature both approaches are used. (The details of the analysis are somewhat more tractable with the model of Fig. 6.19, since when the outputs from all the channels are combined, the general case involves terms of the form $e^{j\omega kT}$ rather than $e^{-j\omega(M-1-k)T}$.)

6.3 TIMING MISMATCH

6.3.0 Introduction

Timing mismatches are a key issue in any signal processing system that involves sampling of a continuous-time input by a number of time-interleaved parallel channels. A model for a time-interleaved ADC with nonuniform sampling is shown below.

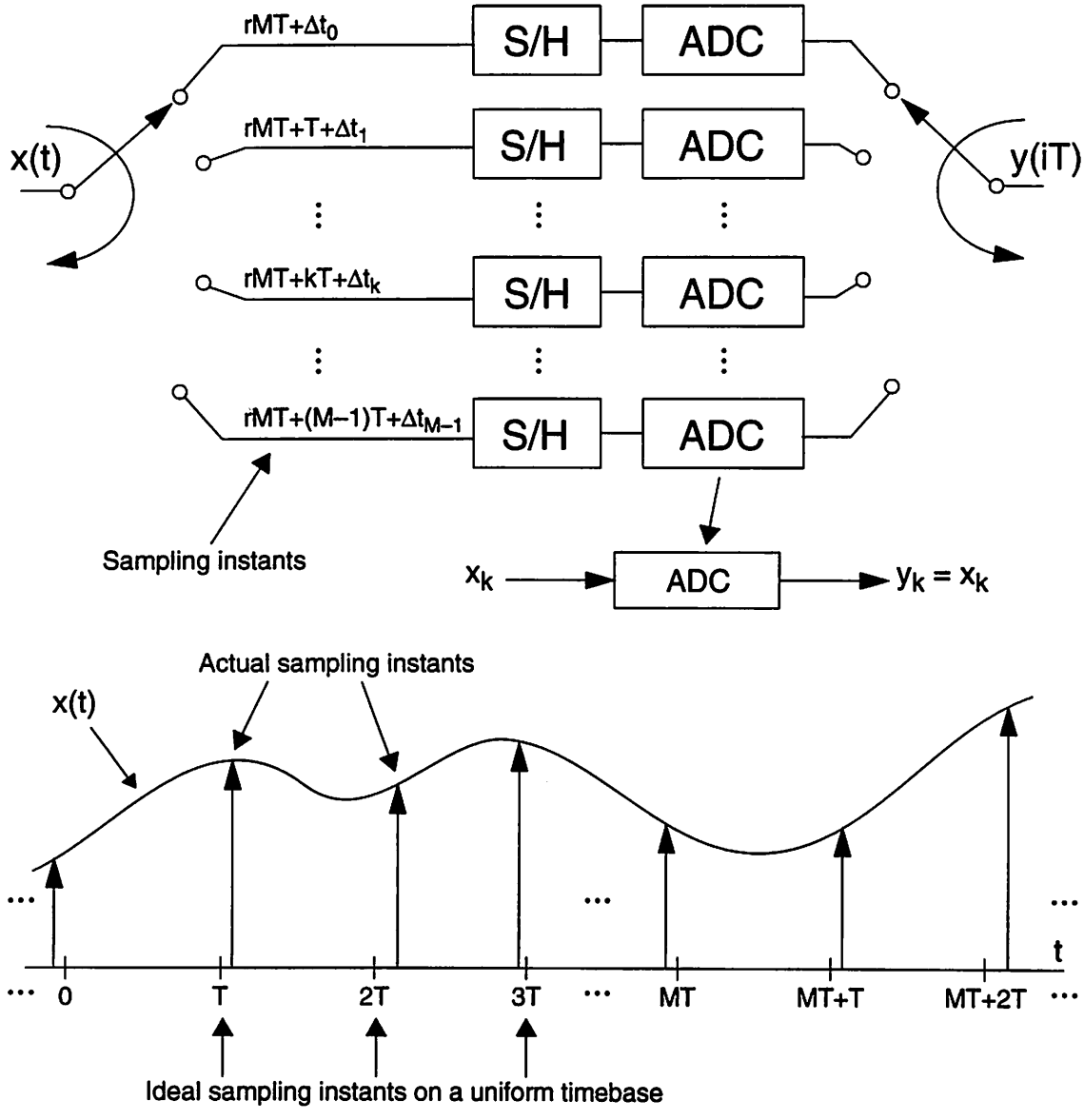


Fig. 6.25 Model for an M-channel time-interleaved ADC system showing a signal x(t) undergoing nonuniform sampling; the ticks on the time axis are the ideal sampling times. The ADC blocks *within* each path are assumed to have ideal gain and offset, i.e., $a_k=1$ and $b_k=0$.

The ADC blocks within each channel are assumed to have ideal unity gain and zero offset — i.e., $a_k = 1$ and $b_k = 0$. Each channel k has a fixed timing offset Δt_k.

The actual sampling instants are

$$
\begin{array}{lllll}
 & & & \cdots & \\
\Delta t_0, & T+\Delta t_1, & 2T+\Delta t_2, & \ldots, kT+\Delta t_k, & \ldots, (M-1)T+\Delta t_{M-1}, \\
 & & & \cdots & \\
rMT+\Delta t_0, & rMT+T+\Delta t_1, & rMT+2T+\Delta t_2, & \ldots, rMT+kT+\Delta t_k, & \ldots, rMT+(M-1)T+\Delta t_{M-1}, \\
(r+1)MT+\Delta t_0, & (r+1)MT+T+\Delta t_1, & (r+1)MT+2T+\Delta t_2, & \ldots, (r+1)MT+kT+\Delta t_k, & \ldots, (r+1)MT+(M-1)T+\Delta t_{M-1}, \\
 & & & \cdots &
\end{array}
$$

etc.

The index r is simply a running index to indicate successive passes or cycles through the entire M-channel system. In this section and throughout this dissertation, the terms *nonuniform sampling, sampling with timing offsets,* and *systematic sampling skew* are used interchangeably. Note that the timing errors $\{\Delta t_k\}$ are fixed, i.e., the $\{\Delta t_k\}$ are deterministic, systematic timing offsets.

One way to view nonuniform sampling is to derive an "equivalent multiplicative sequence" — i.e., an impulse train on a uniform timebase, which, when it multiplies the input, yields the actual sampled values; this is illustrated in Fig. 6.26. Two points are apparent from the diagram: (i) the equivalent impulse train depends on the signal, and (ii) the equivalent impulse train depends on how quickly the signal is changing, i.e., the time derivative of the signal.

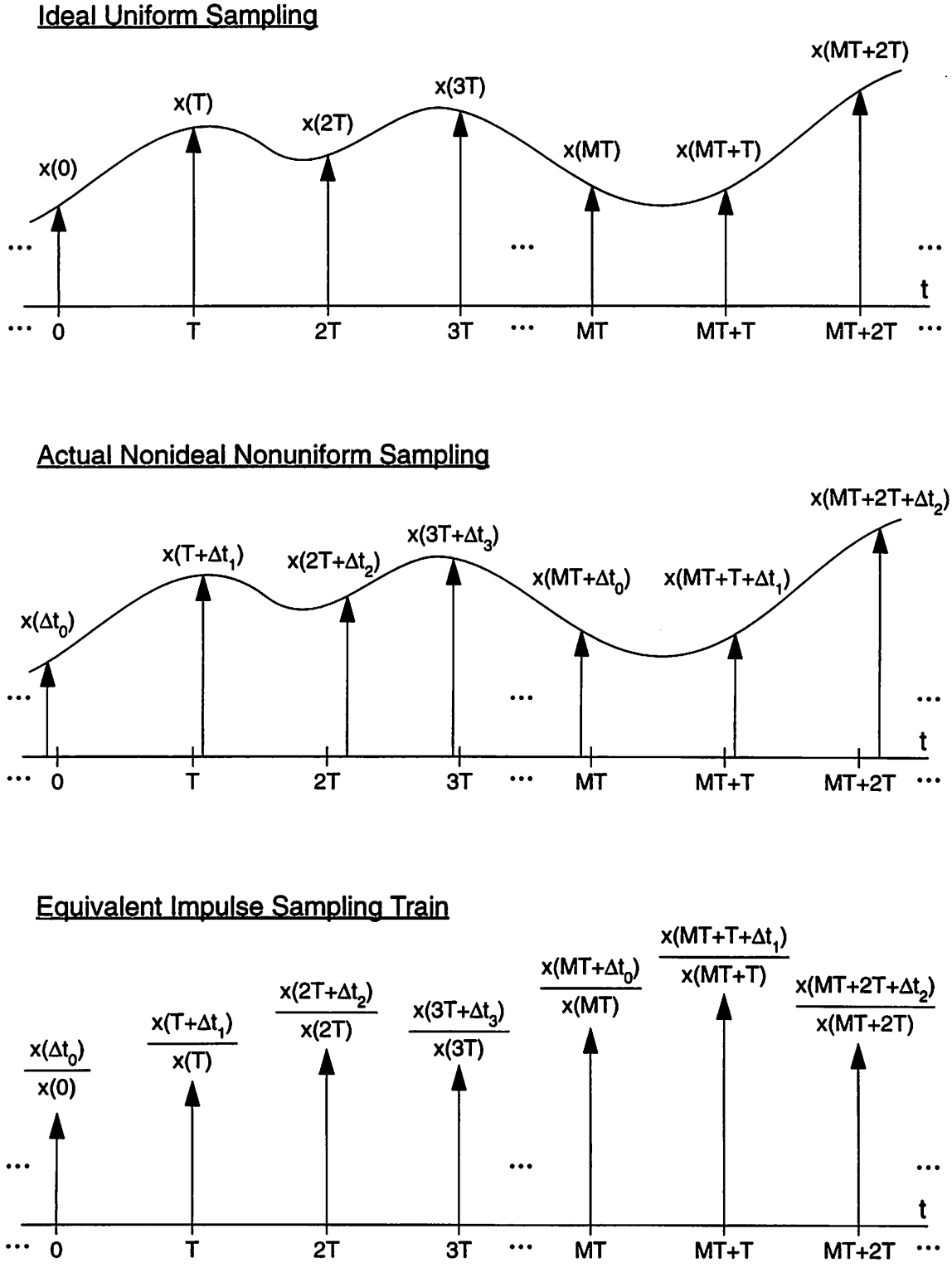


Fig. 6.26 Ideal and nonuniform sampling of a signal x(t) by an M-channel time-interleaved ADC system and the equivalent uniformly-spaced *signal-dependent* impulse sampling train.

6.3.1 General Analysis

This section presents two approaches for deriving the spectrum of a nonuniformly sampled continuous-time waveform x(t).

Nonuniform Sampling: Analysis Method 1

The analysis procedure here consists of the following three steps.

1. Calculate $P_k(j\omega)$, the Fourier Transform of $p_k(t)$, the sampling impulse train for each individual channel.
2. Convolve $P_k(j\omega)$ with $X(j\omega)$, the Fourier Transform of the analog input signal x(t), to obtain $Y_k(j\omega)$, the output spectrum of channel k.
3. Combine all the $\{Y_k(j\omega)\}$ to get the overall output $Y(j\omega)$.

Referring to Fig. 6.25, it is apparent [53] that, in channel k, the signal $x(t + \Delta t_k)$ is sampled at the rate $1/MT$ by the *delayed* impulse train given by

$$p_k(t) = \sum_{r=-\infty}^{\infty} \delta(t - rMT - kT) \tag{6.54}$$

The Fourier Transform of $p_k(t)$ is

$$P_k(j\omega) = \frac{2\pi}{MT} \sum_{n=-\infty}^{\infty} e^{-j(\frac{2\pi n}{MT})kT} \delta(\omega - \frac{2\pi n}{MT}) \tag{6.55}$$

and therefore the spectrum $Y_k(j\omega)$ at the output of channel k is given by

$$Y_k(j\omega) = \frac{1}{2\pi} [e^{j\omega\Delta t_k} X(j\omega)] \otimes P_k(j\omega) \tag{6.56}$$

$$= \frac{1}{MT} \sum_{n=-\infty}^{\infty} e^{-j(\frac{2\pi}{M})kn} e^{j(\omega - \frac{2\pi n}{MT})\Delta t_k} X\left[j(\omega - \frac{2\pi n}{MT})\right] \tag{6.57}$$

Combining the $\{Y_k(j\omega)\}$ terms from all channels, that is, for $k = 0, \ldots, M-1$, yields the final output spectrum:

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty}\left[\frac{1}{M}\sum_{k=0}^{M-1} e^{j(\omega-\frac{2\pi n}{MT})\Delta t_k} e^{-j(\frac{2\pi}{M})nk}\right] X\left[j(\omega-\frac{2\pi n}{MT})\right] \tag{6.58}$$

The expression on the right-hand side of the above equation can be written in a form more closely resembling the equations for gain mismatch, as follows:

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty}\Phi_{nn}(\omega)\, X\left[j(\omega-\frac{2\pi n}{MT})\right] \tag{6.59}$$

where

$$\Phi_{nn}(\omega) = \frac{1}{M}\sum_{k=0}^{M-1}\left[e^{j(\omega-\frac{2\pi n}{MT})\Delta t_k} e^{-j(\frac{2\pi}{M})nk}\right] \tag{6.60}$$

The following observations are pertinent.

- From (6.59), it is apparent that nonuniform timing results in a phenomenon similar to what occurs due to channel gain mismatches, namely, aliases of the input spectrum at the angular frequencies $\omega = 0, \pm 2\pi/MT, \pm 4\pi/MT, \ldots, \pm 2\pi n/MT, \ldots$, i.e., at *multiples of the channel sampling rate.*
- A key difference from the gain mismatch case is that the coefficient $\phi_{nn}(\omega)$ in (6.59) and (6.60) above *depends on frequency.*
- In general, $\phi_{nn}(\omega)$ is *not* a DFS coefficient; for example, it does not have period M.
- The general frequency-dependent coefficient $\phi_{nn}(\omega)$ is difficult to deal with analytically. However, in the important special case of a sinusoidal input signal, the analysis becomes relatively straightforward and is presented in Section 6.3.2.

Nonuniform Sampling: Analysis Method 2

An alternative approach for examination of the effect of nonuniform sampling is to consider an individual channel k. The basic approach of this analysis is similar to that used in the multirate analysis of gain mismatch: the transfer function for a single path in the parallel system is obtained, and then the terms for all channels are combined, taking appropriate account of phase. An individual channel k is shown in Fig. 6.27 below.

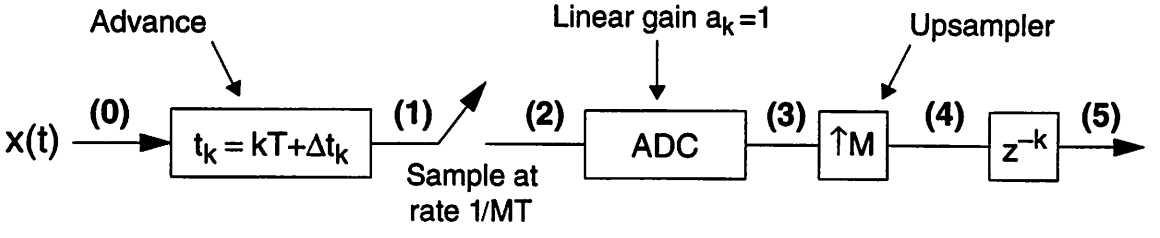


Fig. 6.27 Model for a single path in an M-channel time-interleaved ADC with nonuniform timing.

The frequency-domain representation of the signal at the various points indicated on the above block diagram is given below.

(0): $x(t) \leftrightarrow X(j\omega)$

(1): $x(t+t_k) = x(t+kT+\Delta t_k) \leftrightarrow e^{j\omega t_k} X(j\omega) = e^{j\omega(kT+\Delta t_k)} X(j\omega)$

(2),(3): $$\frac{1}{MT}\sum_{n=-\infty}^{\infty} e^{j(\omega-\frac{2\pi n}{MT})t_k} X\left[j(\omega-\frac{2\pi n}{MT})\right]$$

(5): $$e^{-j\omega kT}\frac{1}{MT}\sum_{n=-\infty}^{\infty} e^{j(\omega-\frac{2\pi n}{MT})t_k} X\left[j(\omega-\frac{2\pi n}{MT})\right]$$

$$= \frac{1}{MT}\sum_{n=-\infty}^{\infty} e^{j(\omega-\frac{2\pi n}{MT})\Delta t_k} e^{-j(\frac{2\pi}{M})kn} X\left[j(\omega-\frac{2\pi n}{MT})\right]$$

Note that as mentioned previously, the upsampler does not change the spectrum with respect to absolute frequency.

Finally, combining the expressions for the individual channel output — i.e., point **(5)** in Fig. 6.27 — for all channels yields

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty}\left[\frac{1}{M}\sum_{k=0}^{M-1} e^{j(\omega-\frac{2\pi n}{MT})\Delta t_k} e^{-j(\frac{2\pi}{M})nk}\right] X\left[j(\omega-\frac{2\pi n}{MT})\right] \tag{6.61}$$

which is identical to (6.58).

Aside

Note that in arriving at the final expression for the spectrum at point **(5)** in the above analysis, terms of the form

$$e^{-j\omega kT}\frac{1}{MT}\sum_{n=-\infty}^{\infty} e^{j(\omega-\frac{2\pi n}{MT})t_k} X\left[j(\omega-\frac{2\pi n}{MT})\right] \tag{6.62}$$

arise. The arguments of the exponentials simplify as follows.

$$-j\omega kT + j(\omega-\frac{2\pi n}{MT})\,t_k = -j\omega kT + j(\omega-\frac{2\pi n}{MT})\,(kT+\Delta t_k) \tag{6.63}$$

$$= -j\omega kT + j\omega kT - j(2\pi/M)\,kn + j(\omega-\frac{2\pi n}{MT})\,\Delta t_k \tag{6.64}$$

$$= -j(2\pi/M)\,kn + j(\omega-\frac{2\pi n}{MT})\,\Delta t_k \tag{6.65}$$

6.3.2 Special Case: Sinusoidal Input

This section discusses nonuniform sampling in a time-interleaved ADC for the important special case of a sinusoid input signal with angular frequency ω_0.

Basic Intuitive Analysis

A simplified intuitive analysis of sampling skew is now presented for the case of a complex exponential input. In general, ideally channel k samples $x(t+kT)$, but due to nonideal timing, it actually samples $x(t+kT+\Delta t_k)$. Let the input be given by

$$x(t) = e^{j\omega_0 t} \tag{6.66}$$

The actual value sampled $x(t+kT+\Delta t_k)$ can be written as

$$x(t+kT+\Delta t_k) = \frac{x(t+kT+\Delta t_k)}{x(t+kT)}\, x(t+kT) \tag{6.67}$$

$$= \frac{e^{j\omega_0(t+kT+\Delta t_k)}}{e^{j\omega_0(t+kT)}}\, x(t+kT) \tag{6.68}$$

$$= e^{j\omega_0 \Delta t_k}\, x(t+kT) \tag{6.69}$$

The term $e^{j\omega_0\Delta t_k}$ in the right-hand side of (6.69) may be interpreted as a *nonideal frequency-dependent gain.* If the $\{\Delta t_k\}$ are all equal, then the result is an inconsequential fixed delay. However, if the $\{\Delta t_k\}$ are not all equal, then the above suggests that behavior resembling what occurs due to gain mismatches is expected.

FM Analogy

If $\{\Delta t_i\}$ is the periodic extension of $\{\Delta t_k\}$, that is, $\Delta t_i = \Delta t_k$, for $k = 0, 1, ..., M-1$, and $\Delta t_{i+M} = \Delta t_i$, then the sampling instants are $\{t_i = iT + \Delta t_i\}$, and the individual samples can written as

$$x(t_i) = x(iT+\Delta t_i) = e^{j\omega_0 t}\Big|_{iT+\Delta t_i} = e^{j(\omega_0 iT + \omega_0 \Delta t_i)} \tag{6.70}$$

Since $\{\Delta t_i\}$ is a discrete-time periodic sequence with period M, therefore, the sequence of samples $\{e^{j(\omega_0 iT + \omega_0 \Delta t_i)}\}$ in (6.70) is seen to resemble *FM with tone modulation*. In this analogy, $f_0 = \omega_0/2\pi$ is the FM carrier frequency, and F_s/M is the (fundamental) modulating frequency.

Exact Analysis

The exact analysis is now given for a complex exponential input. From Section 6.3.1, the general expression for the output spectrum in the presence of timing mismatches is

$$Y(j\omega) = \frac{1}{T}\sum_{n=-\infty}^{\infty} \Phi_{nn}(\omega)\, X\left[j\left(\omega - \frac{2\pi n}{MT}\right)\right] \tag{6.71}$$

where

$$\Phi_{nn}(\omega) = \frac{1}{M}\sum_{k=0}^{M-1} e^{j(\omega - \frac{2\pi n}{MT})\Delta t_k} e^{-j(\frac{2\pi}{M})nk} \tag{6.72}$$

Here, the input is of the form

$$x(t) = e^{j\omega_0 t} \tag{6.73}$$

and its Fourier Transform is given by

$$X(j\omega) = 2\pi\delta(\omega - \omega_0) \tag{6.74}$$

From (6.71) the output is therefore given by

$$Y(j\omega) = \frac{2\pi}{T}\sum_{n=-\infty}^{\infty} \Phi_{nn}(\omega)\, \delta\left(\omega - \omega_0 - \frac{2\pi n}{MT}\right) \tag{6.75}$$

This can be simplified by noting that since the quantity $\Phi_{nn}(\omega)$ inside the summation in (6.75) is being multiplied by the delta function $\delta\left(\omega - \omega_0 - \frac{2\pi n}{MT}\right)$, $\Phi_{nn}(\omega)$ should be evaluated at $\omega = \omega_0 + \frac{2\pi n}{MT}$.

Thus, since the general expression for $\Phi_{nn}(\omega)$ is

$$\Phi_{nn}(\omega) = \frac{1}{M}\sum_{k=0}^{M-1} e^{j(\omega-\frac{2\pi n}{MT})\Delta t_k} e^{-j(\frac{2\pi}{M})nk} \tag{6.76}$$

substituting $\omega = \omega_0 + \frac{2\pi n}{MT}$ gives

$$\Phi_{n}(\omega_0) = \frac{1}{M}\sum_{k=0}^{M-1} e^{j\omega_0\Delta t_k} e^{-j(\frac{2\pi}{M})nk} \tag{6.77}$$

which is the DFS coefficient of the terms $\{e^{j\omega_0\Delta t_k}\}$. Therefore, the output spectrum due to nonuniform sampling of a complex exponential input is

$$Y(j\omega) = \frac{2\pi}{T}\sum_{n=-\infty}^{\infty} \Phi_n(\omega_0)\,\delta(\omega-\omega_0-\frac{2\pi n}{MT}) \tag{6.78}$$

Approximation

For small timing errors, the following approximation is valid:

$$e^{j\omega_0\Delta t_k} \approx 1 + j\omega_0\Delta t_k \tag{6.79}$$

and so the output spectrum becomes

$$Y(j\omega) = \frac{2\pi}{T}\sum_{n=-\infty}^{\infty} \Phi_n(\omega_0)\,\delta(\omega-\omega_0-\frac{2\pi n}{MT}) \tag{6.80}$$

where

$$\Phi_{n}(\omega_0) \approx \frac{1}{M}\sum_{k=0}^{M-1} (1 + j\omega_0\Delta t_k)\, e^{-j(2\pi/M)nk} \tag{6.81}$$

Comparing the right-hand side of the above equation with the expression derived earlier — for example, as given in (6.20) — for the coefficients $\{A_n\}$ that arise in the

analysis of *gain* mismatches in the parallel paths, it is apparent that the $\omega_0 \Delta t_k$ term can be regarded as an *input frequency dependent gain error.*

Numerical Example

Consider a 4-channel system with timing offsets in the channels given by

$$\Delta t_0 = 1.0 \text{ ns}, \quad \Delta t_1 = -0.4 \text{ ns}, \quad \Delta t_2 = -0.1 \text{ ns}, \quad \Delta t_3 = -0.5 \text{ ns}$$

Let the input be a sinusoid with frequency $f_0 = \omega_0/2\pi = 100\,\text{MHz}$.

It is necessary to consider what happens for both positive and negative frequencies, i.e., at $\pm\omega_0$. Evaluating the coefficients using (6.77) gives the following.

For positive frequency:

$$\Phi_0(\omega_0) = 0.931 - j0.008 \quad \Rightarrow \quad |\Phi_0(\omega_0)| = 0.932$$

$$\Phi_1(\omega_0) = -0.032 + j0.158 \quad \Rightarrow \quad |\Phi_1(\omega_0)| = 0.161$$

$$\Phi_2(\omega_0) = -0.028 + j0.271 \quad \Rightarrow \quad |\Phi_2(\omega_0)| = 0.272$$

$$\Phi_3(\omega_0) = -0.062 + j0.167 \quad \Rightarrow \quad |\Phi_3(\omega_0)| = 0.178$$

For negative frequency:

$$\Phi_0(-\omega_0) = 0.931 + j0.008 \quad \Rightarrow \quad |\Phi_0(\omega_0)| = 0.932$$

$$\Phi_1(-\omega_0) = -0.032 - j0.158 \quad \Rightarrow \quad |\Phi_1(\omega_0)| = 0.161$$

$$\Phi_2(-\omega_0) = -0.028 - j0.271 \quad \Rightarrow \quad |\Phi_2(\omega_0)| = 0.272$$

$$\Phi_3(-\omega_0) = -0.062 - j0.167 \quad \Rightarrow \quad |\Phi_3(\omega_0)| = 0.178$$

The corresponding spectrum, with relative magnitudes indicated, assuming a sampling frequency of approximately 20 times the input frequency is shown in Fig. 6.28.

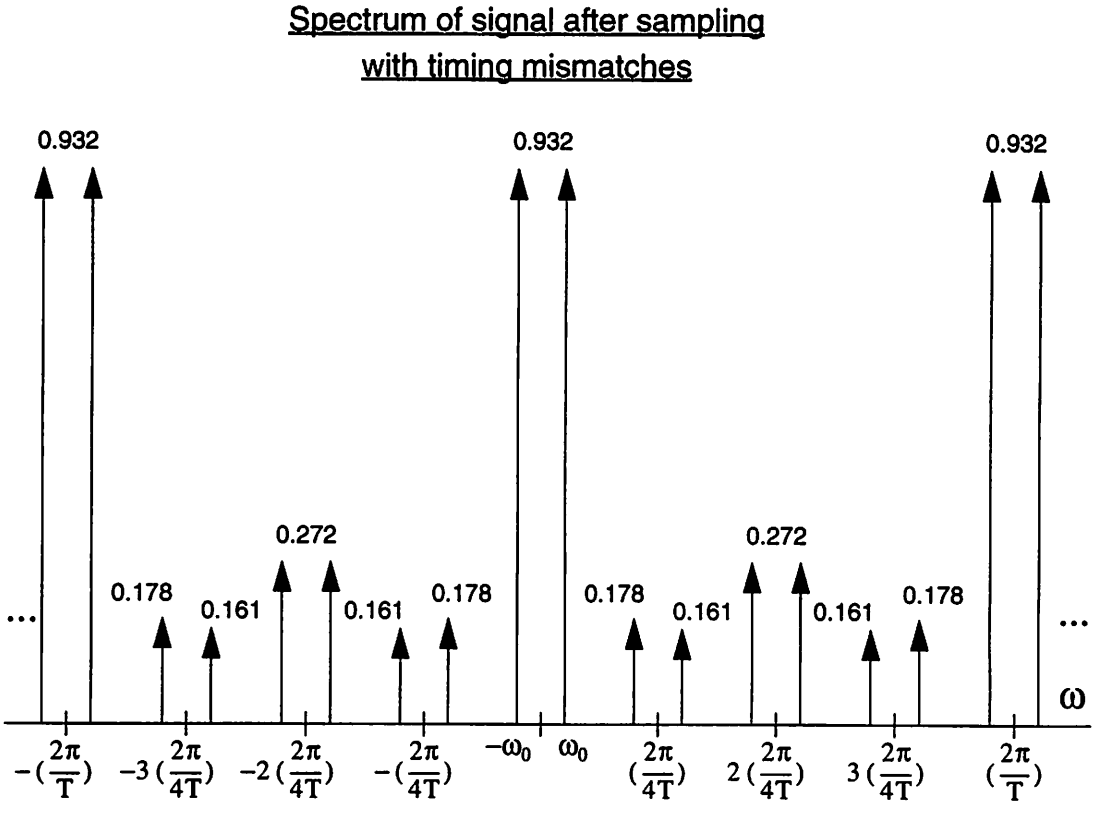


Fig. 6.28 Frequency-domain view when a sinusoid is sampled by a 4-channel time-interleaved ADC with timing mismatches: the output spectrum has sidebands around multiples of $F_s/4$. The sampling frequency is approximately 20 times larger than the input signal frequency.

It is apparent that timing mismatches result in sidebands around multiples of $2\pi/4T$. Note that the sidebands around multiples of the channel sampling frequency do not have equal magnitude, although for small timing skews they do. This is because since the terms $\{e^{j\omega_0\Delta t_k}\}$ are complex-valued, the DFS coefficients $\{\Phi_n(\omega_0)\}$ are not complex-conjugate symmetric. More generally, by considering the derivation of the output spectrum in the presence of timing mismatches — for example, the discussion leading to (6.58) — it is apparent that in a system in which *both* path gain mismatches and nonuniform timing are important, the overall equivalent frequency-dependent gain of channel k for a complex exponential input is $a_k e^{j\omega_0\Delta t_k}$, which clearly is complex-valued. Therefore, in such a case, the sidebands about multiples of the channel sampling rate do not have equal magnitudes.

6.4 STATISTICS: POWER AND SNR EXPRESSIONS

In this section, the results from the previous sections are extended to allow for random variations in the channel parameters, i.e., random variations over a batch or lot of circuits due to fabrication process variability, etc.

6.4.0 Basic SNR Analysis

Here, a basic, intuitive approach is presented. The input-output relationship for any A/D converter may be written as

$$y = (1+a)x + b + q \tag{6.82}$$

where x is the analog input, y is the amplitude-quantized output of the ADC, b is the offset, a is the gain *error* (different from previous notation), and q is the quantization error. The overall error (e) from input to output of the ADC may be defined as

$$e = y - x \tag{6.83}$$

Substituting from (6.82) gives

$$e = y - x = [(1+a)x + b + q] - x = ax + b + q \tag{6.84}$$

Assume that the gain error, offset, and quantization error are independent random variables with some statistical distributions, and have means zero and variances σ_a^2, σ_b^2, and σ_q^2 respectively.

Using (6.84), the mean square value (i.e., power) of the error e is then given by

$$E(e^2) = E[(ax + b + q)^2] \tag{6.85}$$

$$= E(a^2)E(x^2) + E(b^2) + E(q^2) \tag{6.86}$$

$$= \sigma_a^2\sigma_X^2 + \sigma_b^2 + \sigma_q^2 \tag{6.87}$$

where $\sigma_X^2 = E(x^2)$ is the energy of the signal (since the mean is assumed to be zero).

Therefore, the Signal-to-Noise-Ratio (SNR) — including all "noise" due to mismatch aliases and tones — is given by

$$\mathrm{SNR} = 10\log_{10}\left[\frac{E(x^2)}{E(e^2)}\right] \tag{6.88}$$

$$= 10\log_{10}\left[\frac{\sigma_X^2}{\sigma_a^2\sigma_X^2 + \sigma_b^2 + \sigma_q^2}\right] \tag{6.89}$$

$$= -10\log_{10}\left[\sigma_a^2 + \frac{\sigma_b^2}{\sigma_X^2} + \frac{\sigma_q^2}{\sigma_X^2}\right] \tag{6.90}$$

For the case of a sinusoidal input, systematic timing errors/sampling skew can be included in the above formulation, as follows. Previously in Section 6.3.2 — specifically, in equations (6.79)–(6.81) — it was shown that small timing errors may be considered equivalent to a frequency-dependent gain given by

$$a_k = e^{j\omega_0\Delta t_k} \approx 1 + j\omega_0\Delta t_k \tag{6.91}$$

Thus, the gain *error* is $\omega_0\Delta t_k$, and has variance $\omega_0^2\sigma_t^2$, where $\sigma_t^2 = \mathrm{Var}(\Delta t_k)$.

This variance may be incorporated into the expression on the right-hand side of (6.90) to give a frequency-dependent SNR that includes the effects of quantization noise (σ_q^2), channel gain mismatches (σ_a^2), channel offset mismatches (σ_b^2), and systematic timing mismatches (σ_t^2):

$$\mathrm{SNR}(\omega_0) = -10\log\left[\sigma_a^2 + \omega_0^2\sigma_t^2 + \frac{\sigma_b^2}{\sigma_X^2} + \frac{\sigma_q^2}{\sigma_X^2}\right] \tag{6.92}$$

6.4.1 Basic Distortion Analysis

Another way to view the cumulative effect of path mismatches in the frequency domain is to compute the total power in the "mismatch tones" generated when the input is a sinusoid — i.e., the total distortion power. Note, of course, that this is *not* harmonic distortion, since none of the spurious frequency components produced are *harmonically* related to the input frequency.

Suppose the input is a complex exponential, i.e., $x(t) = e^{j\omega_0 t}$. Then, the output spectrum from a parallel time-interleaved ADC with gain mismatches is given by

$$Y(j\omega) = \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} A_n \delta(\omega - \omega_0 - \frac{2\pi n}{MT}) \tag{6.93}$$

where

$$A_n = \frac{1}{M} \sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk} \tag{6.94}$$

and, in particular, the magnitude of the fundamental is the average gain through all the channels, i.e.,

$$A_0 = \frac{1}{M} \sum_{k=0}^{M-1} a_k \tag{6.95}$$

The total distortion power is the sum of the powers in the undesired tones, i.e.,

$$P_{DIS} = |A_1|^2 + |A_2|^2 + \ldots + |A_{M-1}|^2 \tag{6.96}$$

$$= \sum_{n=1}^{M-1} |A_n|^2 \tag{6.97}$$

In order to relate the power of the frequency-domain DFS coefficients $\{A_n\}$ to the time-domain quantities $\{a_k\}$, Parseval's Theorem for Discrete Fourier Series is employed,

as given below:

$$\frac{1}{M}\sum_{k=0}^{M-1}|a_k|^2 = \sum_{n=0}^{M-1}|A_n|^2 \tag{6.98}$$

Substituting relation (6.98) into equation (6.97) yields

$$P_{DIS} = \sum_{n=1}^{M-1}|A_n|^2 \tag{6.99}$$

$$= \left(\frac{1}{M}\sum_{k=0}^{M-1}a_k^2\right) - |A_0|^2 \tag{6.100}$$

If the channel gains $\{a_k\}$ are assumed to be independent identically distributed random variables with mean 1 and variance σ_a^2, then it is straightforward to show that the expected (i.e., average) signal power in the fundamental is

$$E\{P_0\} = E\{A_0^2\} = 1 + \frac{\sigma_a^2}{M} \approx 1 \tag{6.101}$$

and that the expected distortion power is

$$E\{P_{DIS}\} = \sigma_a^2\left(\frac{M-1}{M}\right) \tag{6.102}$$

Similar expressions arise for the expected power in the frequency components produced by offset and timing mismatches. Therefore, the expected Signal-to-Distortion Ratio due to mismatch tones, (SDR), is given by

$$SDR = \frac{E\{P_0\}}{E\{P_{DIS}\}} \approx \frac{1}{\left(\frac{M-1}{M}\right)(\sigma_b^2 + \sigma_a^2 + \omega_0^2\sigma_t^2)} \tag{6.103}$$

$$= -10\log_{10}\left[\left(\frac{M-1}{M}\right)(\sigma_b^2 + \sigma_a^2 + \omega_0^2\sigma_t^2)\right] \text{ dB} \tag{6.104}$$

6.4.2 Statistics of the Mismatch Coefficients

The previous two subsections focussed on *overall* SNR and SDR due to mismatches in a time-interleaved ADC. However, it is also important to know the detailed statistics of the individual mismatch coefficients [105], since these coefficients directly determine the magnitude of aliases and mismatch-related tones in the output spectrum. Accordingly, the focus of the following discussion is the distribution of the magnitudes $\{|A_n|\}$.

Expanding the expression for A_n into its real and imaginary components yields

$$A_n = \frac{1}{M}\sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk} \tag{6.105}$$

$$= \frac{1}{M}\sum_{k=0}^{M-1} a_k \cos\left[\left(\frac{2\pi}{M}\right)nk\right] - j\frac{1}{M}\sum_{k=0}^{M-1} a_k \sin\left[\left(\frac{2\pi}{M}\right)nk\right] \tag{6.106}$$

$$= \mathrm{Re}\{A_n\} + j\mathrm{Im}\{A_n\} \tag{6.107}$$

$$= R_n e^{j\Theta_n} \tag{6.108}$$

where the real and imaginary parts of A_n are, respectively,

$$\mathrm{Re}\{A_n\} = \frac{1}{M}\sum_{k=0}^{M-1} a_k \cos\left[\left(\frac{2\pi}{M}\right)nk\right] \tag{6.109}$$

$$\mathrm{Im}\{A_n\} = -\frac{1}{M}\sum_{k=0}^{M-1} a_k \sin\left[\left(\frac{2\pi}{M}\right)nk\right] \tag{6.110}$$

and the magnitude is

$$R_n = |A_n| = \sqrt{[\mathrm{Re}\{A_n\}]^2 + [\mathrm{Im}\{A_n\}]^2} \tag{6.111}$$

Assuming that the $\{a_k\}$ are independent and identically distributed Gaussian random variables with variance σ_a^2, then the real and imaginary parts $\mathrm{Re}\{A_n\}$ and $\mathrm{Im}\{A_n\}$ are zero-mean, normally-distributed random variables with variance $\sigma_a^2/2M$.

Furthermore, if two random variables x and y are normal, independent, and have zero mean and equal variance σ^2, then the random variable $r = \sqrt{x^2 + y^2}$ has a Rayleigh distribution with probability density function

$$f_r(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \tag{6.112}$$

for $r > 0$ and equal to zero otherwise, and cumulative distribution function

$$F_r(r) = 1 - e^{-r^2/2\sigma^2} \tag{6.113}$$

for $r > 0$ and equal to zero otherwise.

The expected value of a random variable that has a Rayleigh distribution is

$$E(r) = \sqrt{\frac{\pi}{2}}\,\sigma \tag{6.114}$$

Therefore, since the standard deviation of $\mathrm{Re}\{A_n\}$ and $\mathrm{Im}\{A_n\}$ is $\sigma_a/\sqrt{2M}$, the expectation of the magnitude of the mismatch coefficient is

$$E\{|A_n|\} = \frac{\sigma_a}{2}\sqrt{\frac{\pi}{M}} \tag{6.115}$$

For example [105], if $\sigma_a = 1\%$, $M = 4$, this gives $E\{|A_n|\} = 0.0044 = -47$ dB.

A useful question from the point of view of a practical design is the following: what is the maximum value of σ_a, σ_b, or σ_t permissible in order to guarantee that the magnitude of each mismatch tone is less than some value ε with probability $(1 - \alpha)$, that is, $100(1 - \alpha)\%$ of the time? (A typical value for α is 0.01.) The value of the appropriate σ obtained from this analysis may then be used in the circuit design process to specify, for example, the allowable offset mismatch, or capacitor ratio mismatch, or timing mismatch in a delay line. (These issues are discussed further in Chapter 7.) This question can be approached by employing the cumulative distribution of $|A_n|$ — i.e., the Rayleigh cumulative distribution function given in (6.113). Here, the general solution is derived first, and then it is applied to the specific cases of offset, gain, and timing mismatches.

From the preceding discussion, it is required that the probability that the value of the mismatch tone magnitude is less than or equal to ε is $1-\alpha$. Mathematically, this is

$$F_r(r)\big|_{r=\varepsilon} = 1 - e^{-\varepsilon^2/2\sigma^2} = 1-\alpha \tag{6.116}$$

which yields

$$\sigma = \frac{\varepsilon}{\sqrt{2}\sqrt{\ln(1/\alpha)}} \tag{6.117}$$

where ln ( ) is the natural log (base e). Note that because of the square root and the log function, $\sqrt{\ln(1/\alpha)}$ is quite a weak function of α: for typical values of α it is usually in the range 2.1–2.6. Equation (6.117) is now applied to the cases of gain, offset, and timing mismatch.

For gain mismatch, the σ in (6.116) corresponds to $\sigma_a/\sqrt{2M}$. Substituting this into (6.117) yields

$$\frac{\sigma_a}{\sqrt{2M}} = \frac{\varepsilon}{\sqrt{2}\sqrt{\ln(1/\alpha)}} \tag{6.118}$$

which gives

$$\sigma_a = \frac{\varepsilon\sqrt{M}}{\sqrt{\ln(1/\alpha)}} \tag{6.119}$$

For offset, the corresponding expression is

$$\sigma_b = \frac{\varepsilon\sqrt{M}}{\sqrt{\ln(1/\alpha)}} \tag{6.120}$$

For timing mismatch, the appropriate expression is

$$\omega_0\sigma_t = \frac{\varepsilon\sqrt{M}}{\sqrt{\ln(1/\alpha)}} \tag{6.121}$$

which gives

$$\sigma_t = \frac{1}{\omega_0}\cdot\frac{\varepsilon\sqrt{M}}{\sqrt{\ln(1/\alpha)}} \tag{6.122}$$

To summarize: (6.119), (6.120), and (6.122) give the values of the mismatch standard deviations σ_a, σ_b, and σ_t respectively required, so that, with probability $(1 - \alpha)$, the magnitude of the appropriate mismatch coefficient (equal to the amplitude of the mismatch tone) is less than ε relative to the fundamental. Note that the more channels there are (the higher M), the larger the mismatch standard deviation can be while ensuring that the tone has magnitude less than ε with the requisite degree of confidence. Intuitively, the *total* error energy is approximately the same, but it is spread over more aliased components. Note that in all these expressions it is assumed that the magnitude of the fundamental is 1.

6.5 SUMMARY

This chapter has presented DSP analysis of gain, offset, and timing mismatches in parallel time-interleaved ADC's. The key results discussed in this chapter are summarized qualitatively in the figure below and quantitatively in the table on the next page.

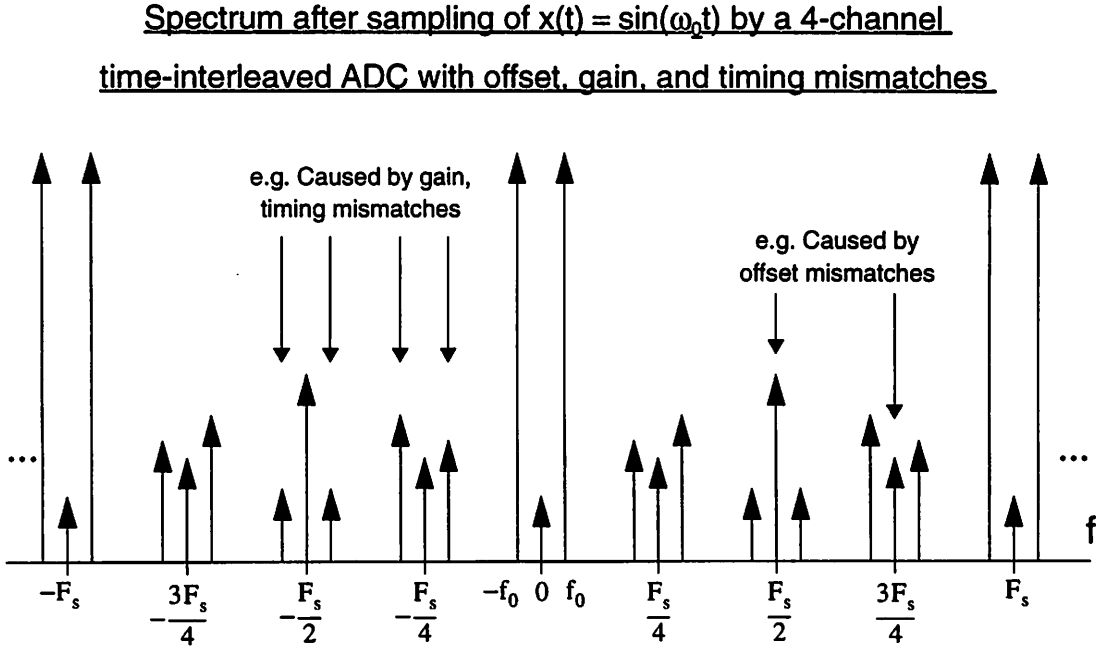


Fig. 6.29 Frequency-domain view when a sinusoid at frequency f_0 is sampled by a 4-channel time-interleaved ADC with offset, gain, and timing mismatches: the output spectrum has tones at multiples of $F_s/4$ and sidebands around multiples of $F_s/4$.

	Gain (a_k)	Offset (b_k)	Timing (Δt_k)
Input	$x(t)$	$x(t)$	$e^{j\omega_0 t}$
Spectrum at Output	$\frac{1}{T}\sum_{n=-\infty}^{\infty} A_n X\left[j(\omega-\frac{2\pi n}{MT})\right]$	$\frac{2\pi}{T}\sum_{n=-\infty}^{\infty} B_n \delta(\omega-\frac{2\pi n}{MT})$	$\frac{2\pi}{T}\sum_{n=-\infty}^{\infty} \Phi_n(\omega_0)\,\delta(\omega-\omega_0-\frac{2\pi n}{MT})$
Coefficient	$A_n = \frac{1}{M}\sum_{k=0}^{M-1} a_k e^{-j(2\pi/M)nk}$	$B_n = \frac{1}{M}\sum_{k=0}^{M-1} b_k e^{-j(2\pi/M)nk}$	$\Phi_n(\omega_0) = \frac{1}{M}\sum_{k=0}^{M-1} e^{j\omega_0\Delta t_k} e^{-j(2\pi/M)nk}$ $\approx \frac{1}{M}\sum_{k=0}^{M-1} (1+j\omega_0\Delta t_k)\, e^{-j(2\pi/M)nk}$
E{|Coeff.|}	$E\{\|A_n\|\} = \frac{\sigma_a}{2}\sqrt{\frac{\pi}{M}}$	$E\{\|B_n\|\} = \frac{\sigma_b}{2}\sqrt{\frac{\pi}{M}}$	$E\{\|\Phi_n(\omega_0)\|\} = \frac{\omega_0\sigma_t}{2}\sqrt{\frac{\pi}{M}}$
Condition for $\|Coeff.\| \le \varepsilon$ (99.9% confidence)	$\sigma_a \le \frac{\varepsilon\sqrt{M}}{2.6}$	$\sigma_b \le \frac{\varepsilon\sqrt{M}}{2.6}$	$\sigma_t \le \frac{\varepsilon\sqrt{M}}{2.6}\cdot\frac{1}{\omega_0}$

Chapter 7 Parallel Pipeline Architecture

7.0 INTRODUCTION

Previous chapters in this dissertation have discussed issues relating to dc transfer characteristics and transient response of pipelined multistage A/D converters *in which there is a single analog signal path — i.e., one "channel"*. Chapter 6 has considered the effects in the signal processing domain of using multiple parallel analog signal paths. This chapter describes an A/D converter architecture incorporating both pipelining and parallelism, and thus, in some sense, constitutes a logical fusion of the aforementioned chapters. In Section 7.1, some previous approaches to employing parallelism and time-interleaving in ADC's are reviewed. Section 7.2 introduces the parallel pipeline array A/D converter concept and describes the architecture. The main performance limitations are (i) mismatch-related issues associated with the use of parallel analog signal paths and (ii) settling time of the input sample-and-hold (S/H) and interstage residue amplifiers; these problems are discussed in Section 7.3 and appropriate solutions presented.

7.1 OVERVIEW OF PARALLELISM IN A/D CONVERTERS

The aim of this section is to examine various ADC topologies, paying particular attention to the use of parallelism — other aspects of the architectures are not emphasized. First, in Section 7.1.0, the concept of a time-interleaved A/D converter is reviewed; some implementations are described in Section 7.1.1. In Section 7.1.2, specific examples are presented of various other ADC's reported in the literature, which — although not usually considered parallel, time-interleaved converters — *do* incorporate parallelism in some interesting way. Finally, in Section 7.1.3, some architectural features underlying the examples presented are summarized, and some general observations are made.

7.1.0 Classical Time-Interleaved A/D Converter Topology: Overview

The idea of a time-interleaved array of ADC's is a well-known technique to achieve very high converter throughput; a 4-channel case is shown schematically below.

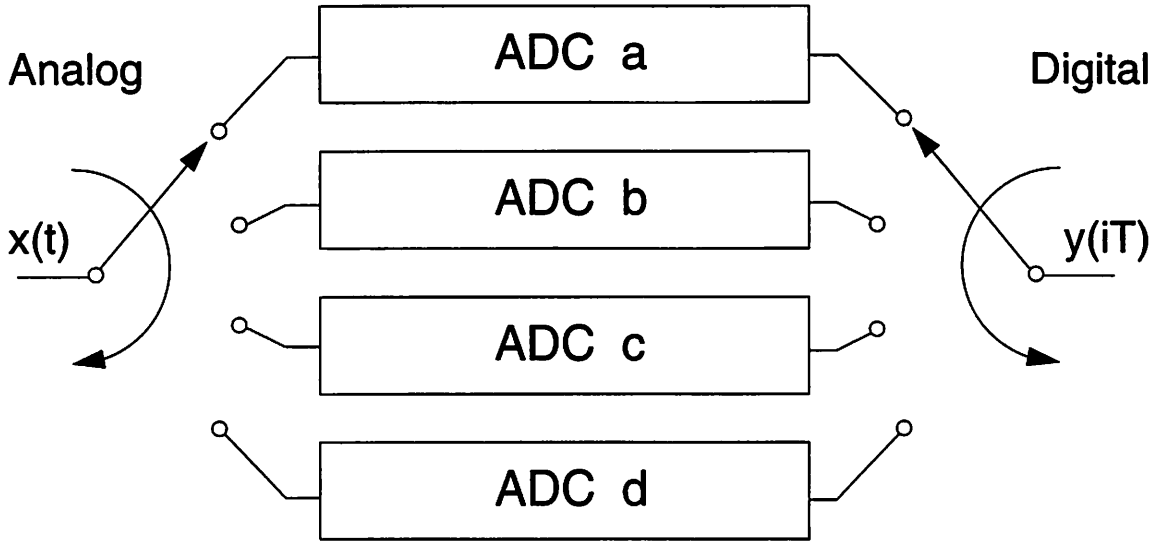


Fig. 7.0 Basic time-interleaved architecture showing front-end analog demultiplexer, multiple ADC's in parallel, and final digital multiplexer. The four channels are identified as a, b, c, and d.

A somewhat subtle point is that there are a number of possible ways to implement the system timing. In the DSP analysis in Chapter 6, this issue did not arise since ideal impulse sampling was assumed throughout. However, in a real implementation in which sampling of a time-varying input waveform is performed, a non-zero *acquisition* or *tracking* period is necessary, followed by a *hold* time. The quantizer part of the A/D conversion function (or at least one step of it, in the case of a multistep ADC) can then occur during the hold time of the front-end sample-and-hold, i.e., when the input is stable. Clock waveforms for ideal impulse sampling and for two possible simple track-and-hold schemes are illustrated schematically in Fig. 7.1 for a four-channel system — i.e., corresponding to the architecture of Fig. 7.0. It is assumed that each individual ADC has its own front-end S/H that can be connected directly to the analog input. In all cases, the sampling instants are indicated by an arrow.

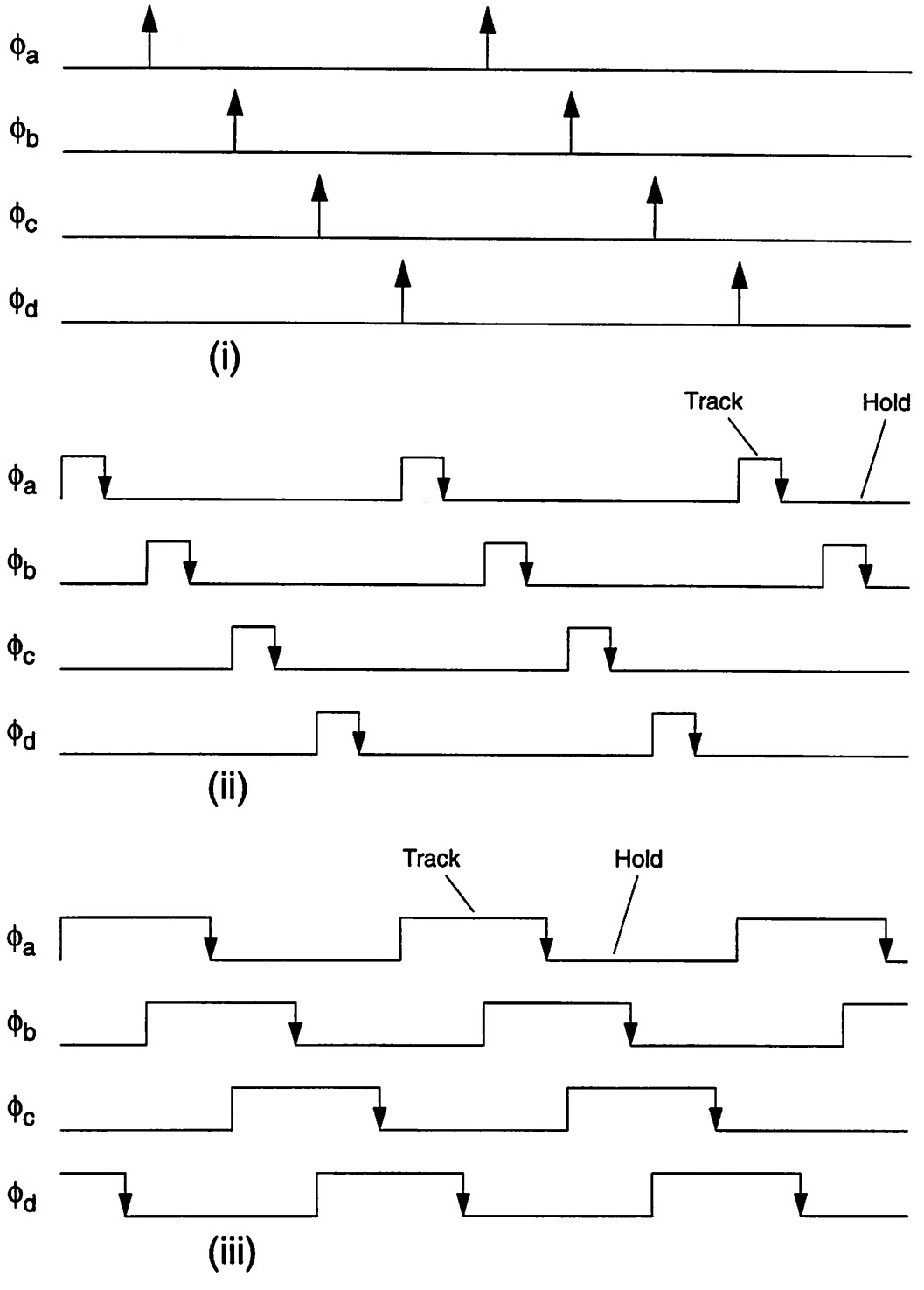


Fig. 7.1 Clock waveforms for an A/D converter system consisting of four parallel time-interleaved channels (labelled a, b, c, d). Three sampling schemes are shown: (i) ideal impulse sampling, (ii) implementation with track/acquisition time much smaller than hold time, (iii) track and hold times approximately equal. The arrow indicates the sampling instant.

Scheme (i) in Fig. 7.1 is the ideal (unrealizable) case, namely, impulse sampling, with an acquisition time of zero. For A/D converters in which acquisition occurs in a much shorter time than the time required to process the held input sample, a scheme resembling (ii) is appropriate. Examples of this are serial ADC architectures such as successive-approximation or recycling topologies (e.g., algorithmic), or multistage ripple-through (non-pipelined) architectures. For cases in which both clock phases are required to be of approximately equal duration, scheme (iii) is appropriate. An example of this category is a switched-capacitor (SC) pipelined multistage ADC architecture, since although the acquisition time of the front-end S/H is determined by a fast passive R-C network, the acquisition phase of a particular *intermediate* stage is coincident with the hold time of the adjacent stages, *during which the interstage gain amplifiers of those stages must settle.*

Note that nonoverlapping clocks are required to maintain the integrity of the critical sampling edges. In particular, at the instant that a particular channel samples the analog input — i.e., at the transition from acquisition/tracking mode to hold mode — no other clock transitions should be occurring that would affect the analog input signal or affect any switches connected to the input. Another issue related to the clock duty-cycle is as follows: in scheme (ii) of Fig. 7.1, the analog input is connected to only one channel at a given instant, whereas in case (iii) it is connected to two channels. In the latter case the input tracking network has to be fast enough to recover from any transient due to the sampling edge of the previous channel. The issue of input bandwidth is considered further in Section 7.3.4. From a practical viewpoint, one reason to choose a scheme resembling (iii) is that generation of clocks with approximately 50% duty cycle is often significantly less complex than realizing small duty cycles, especially if operation over a wide range of sampling frequencies is required. Note that for the case of four channels shown in Fig. 7.1, there are four critical clock edges — i.e., the high-to-low transitions of ϕ_a, ϕ_b, ϕ_c, ϕ_d — that have to be precisely matched in the time domain to avoid the problems due to systematic timing mismatches described in Chapter 6.

More generally, it is clear that the operation of sampling is fundamental to the A/D conversion process, and, consequently, the associated clocking and timing requirements merit careful consideration when comparing architectures and when evaluating feasibility of implementation. This is especially true for parallel time-interleaved topologies. The issues of timing accuracy and input bandwidth are discussed further in Section 7.3.3 and Section 7.3.4 respectively, in the specific context of the parallel pipeline architecture.

7.1.1 Classical Time-Interleaved A/D Converter Topology: Examples

Two time-interleaved A/D converter architectures are described briefly in the following two examples.

Parallelism Example 1: 1-GS/s, four-channel ADC system in Si bipolar and GaAs

A class of applications characterized by extremely high sample rate requirements and in which cost is often not the primary consideration is instrumentation — for example, high-speed digitizers for waveform acquisition, digital sampling oscilloscopes, etc. A block diagram is illustrated in Fig. 7.2 of a high-performance four-channel ADC system intended for such applications [109].

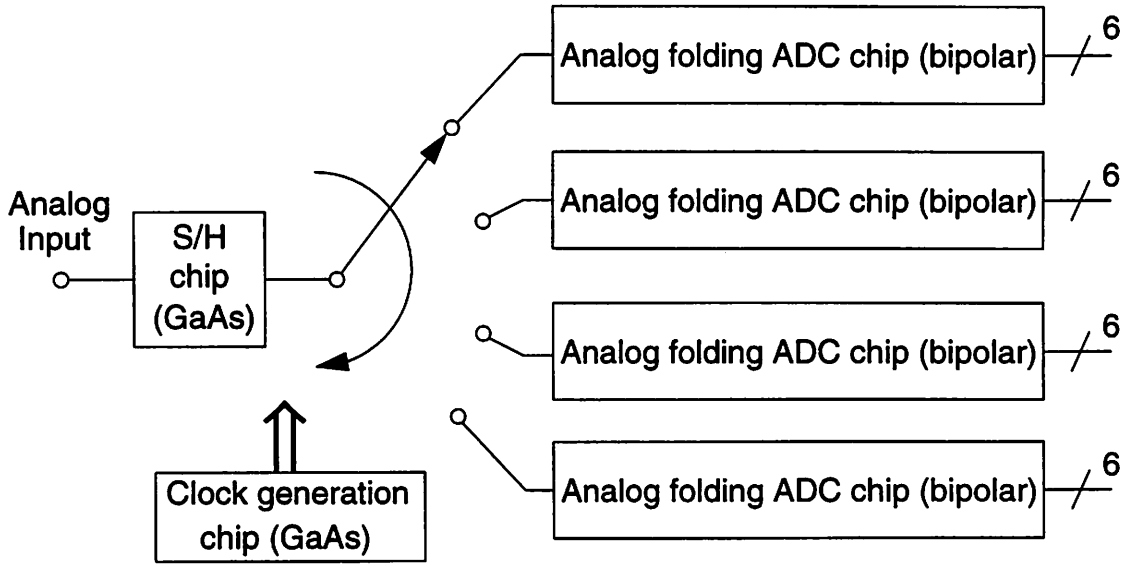


Fig. 7.2 Basic architecture of a 6-bit 1-GS/s ADC system [109].

As mentioned in Section 7.1.0 above, a primary issue in time-interleaved ADC architectures is how to implement the front-end analog demultiplexer function. The key feature of the topology of Fig. 7.2, in contrast to the configurations discussed previously, is the use of a *single high-speed S/H at the front end, clocked by a full sample-rate 1-GHz clock.* This in turn drives four other S/H's (one for each channel) in a two-rank configuration. *However, only the single front-end sampler sees the full-bandwidth analog input signal* — the second-rank S/H's acquire and sample the held output from the first stage, and so their timing is relatively non-critical [109], thereby dealing with the problem of timing mismatch between the multiple channels. Of course in many ways this is the ideal solution since automatically the problem of systematic timing mismatches is greatly alleviated; however, the cost is that this approach requires an extremely fast technology in order to be able to realize a S/H with both fast tracking *and* fast hold. The entire system described in [109] consists of two GaAs chips and four Si bipolar chips, consumes 16 W, and achieves an ac resolution of 5.2 effective bits for a 1-GHz input frequency. The voltage references for the four ADC's are calibrated in order to match the gains and offsets. Furthermore, a technique known as *voltage interleaving* [109], [93] is employed to enhance the resolution.

Parallelism Example 2: Array of 4 CMOS binary-weighted capacitor array ADC's

A simplified schematic of a time-interleaved converter array consisting of four 7-bit successive-approximation SC ADC's is shown below [8].

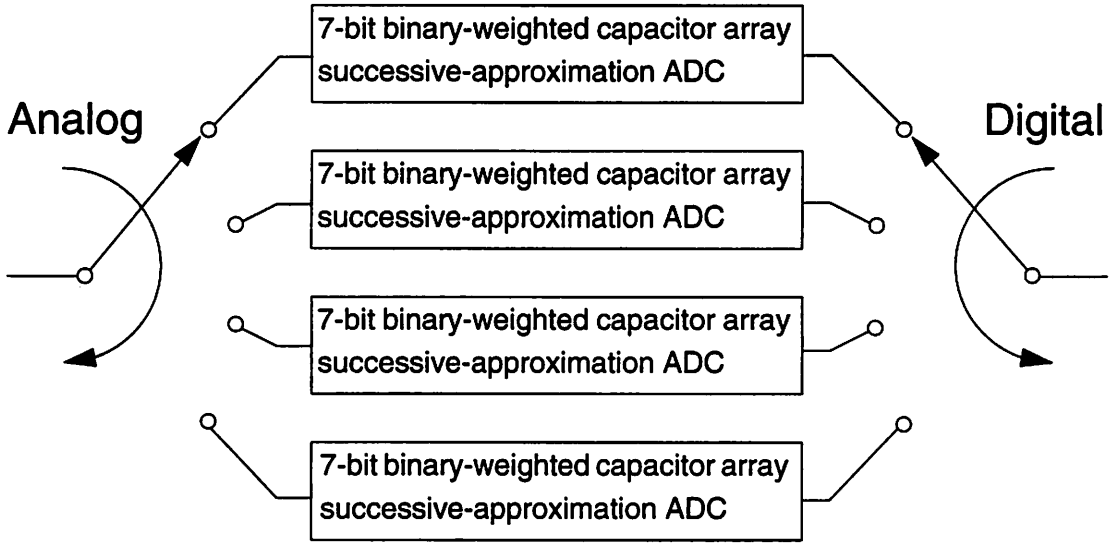


Fig. 7.3 Time-interleaved converter array containing four 7-bit switched-capacitor (SC) successive-approximation ADC's in parallel [8].

Each channel requires 8 clock cycles to perform a complete conversion — i.e., input acquisition followed by 7 bit decisions. Therefore, since there are four in parallel, the throughput is one conversion every two clock cycles. A single clock is supplied externally, and is used to generate the clocks required by each ADC and to implement the input demultiplexer function; the timing is similar to scheme (ii) of Fig. 7.1. During the acquisition phase each comparator undergoes an autozero/reset cycle; however, no gain calibration is used — the capacitors are sized reasonably large. The results reported achieve 39 dB SNR for a sample rate of 2 MS/s and with an input frequency of 100 kHz. Chip area was 208 mil × 278 mil (5.28 mm × 7.06 mm) in 10-μm CMOS.

7.1.2 Other Embodiments of Parallelism in ADC Topologies

In addition to the classical time-interleaved ADC configuration, parallelism has been incorporated in various ways in certain other A/D architectures in order to increase throughput significantly, while maintaining reasonable power and area. The next set of examples are in the class of subranging architectures, which are widely used for consumer and video applications typically with 8 bits of resolution. Before examining the specific use of parallelism in these cases, it is useful first as a preliminary note to review the basic two-step subranging topology represented schematically for an 8-bit case in Fig. 7.4.

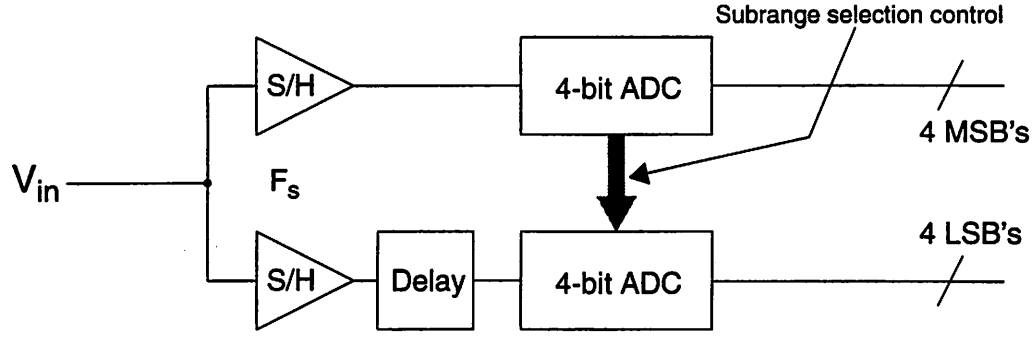


Fig. 7.4 Basic schematic of a subranging two-step flash ADC. The resistor string is not shown.

Subranging may be viewed as a multistage topology with an interstage gain of one. Consistent with the terminology used in Chapter 3, the two 4-bit ADC's comprising the Most Significant Bit (MSB) and Least Significant Bit (LSB) stages are sometimes referred to as *sub-ADC's*. Each sub-ADC is a 4-bit flash converter and consists of an array or *bank* of comparators. A single resistor string is used to generate all reference levels required for the sub-ADC thresholds — this constitutes a basic resource sharing. (For clarity, the resistor string is not shown explicitly in Fig. 7.4.) Assuming the comparators do not load the resistor string there is no static interstage gain error. It has become standard in actual implementations for decision level offsets in the first (MSB) stage to be handled by overrange detection — i.e., using extra threshold levels in the second (LSB) stage, as described in Section 3.2.1.

Consider the example of an 8-bit ADC consisting of two 4-bit sub-ADC's. When the MSB comparators are making their decisions — sometimes called the *coarse* decisions — the threshold levels used are resistor string reference levels spaced $2^4 = 16$ taps apart. Based on the result of the coarse decision, effectively, the ADC zooms in to the *subrange* (16 taps wide) selected by the coarse decision, and so during the LSB or *fine* decision phase, the levels switched to the LSB comparators are one tap apart but within the appropriate subrange selected by the coarse decision. Clearly, depending on the MSB decision, a different subrange is selected, and so different reference levels are switched to the LSB comparators — hence the term *subrange selection control* in Fig. 7.4. Note that the overall input capacitance of the converter is due to the contribution from the two comparator banks.

When implemented using a CMOS SC approach, a sequence of three basic operations occurs: (i) input acquisition and sampling, (ii) MSB comparison and decision, and (iii) LSB comparison and decision. Both the MSB and LSB banks can acquire the input signal onto their respective capacitors simultaneously, *but the LSB ADC must wait until the MSB bank has made its decision* — and the appropriate decision levels within the subrange that the MSB bank has selected are available — *before it can perform its comparisons.* It is apparent that the throughput of the basic two-step subranging structure is limited by the fact that the MSB comparison and decision and the LSB comparison and decision must be performed in a serial fashion. Thus, assuming a two-phase clocking scheme with each of the above-mentioned operations (i), (ii), and (iii) requiring at least half a clock period, the net throughput of the ADC is less than one conversion per clock cycle. Clearly, this motivates the use of parallelism and pipelining. In particular, the following three architecture examples use multiple banks of LSB (and sometimes also MSB) comparators, *thereby allowing one new sample to be taken each clock cycle.* These approaches exemplify an important concept: replication of hardware in order to ensure that resources are being used fully all the time, leading to power and area efficiency.

Parallelism Example 3: Two-step subranging architecture using two LSB ADC's

A block diagram of a modified two-step subranging ADC is shown below in Fig. 7.5. The key innovation is the introduction of a second 4-bit ADC for the LSB decisions. Each LSB ADC operates at rate $F_s/2$ on alternate input samples, thus allowing one new sample to be taken every clock cycle. The figure is drawn so as to emphasize the inherent parallelism in the architecture.

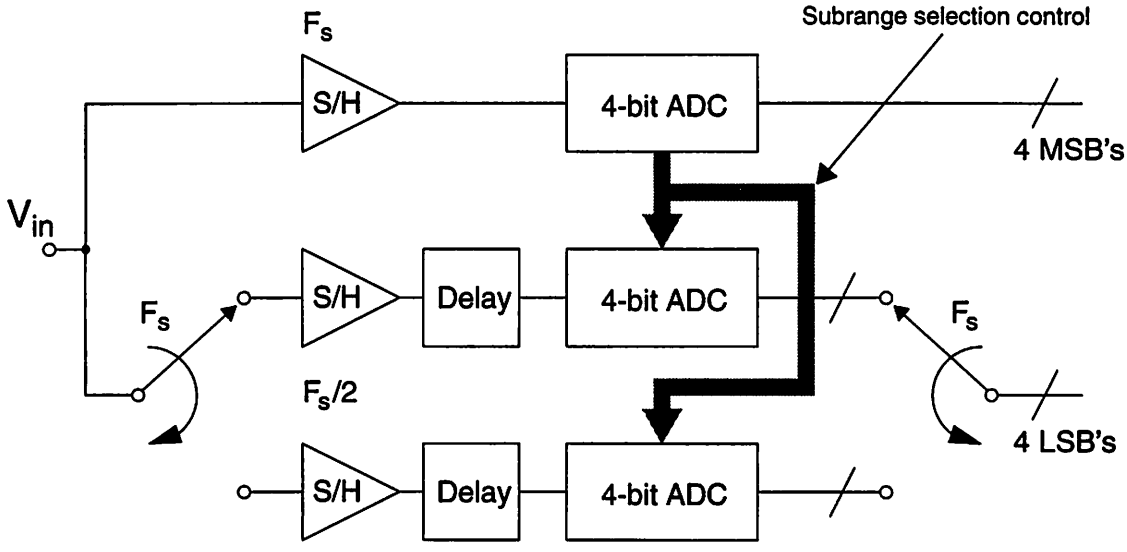


Fig. 7.5 Simplified functional schematic of an 8-bit subranging two-step flash ADC using two LSB ADC's — i.e., two banks of lower comparators. Each LSB ADC samples at rate $F_s/2$.

This basic architecture was first presented by Sony in [30], [31] and is available commercially in a product [130]. A similar topology was presented by Toshiba [145], [146], and a 9-bit implementation has been described by Hitachi [58]. At any particular instant, the input capacitance is from the MSB bank *and* one LSB bank. The descriptions in the cited references do not contain any discussion of mismatches, although clearly offset mismatch between the two LSB banks is a potential problem — especially at 9-bit resolution. It is possible that the offset cancellation is good enough to eliminate this problem.

Parallelism Example 4: Subranging ADC with two MSB and two LSB ADC's

A variant on the previous topology is to use two banks of LSB comparators *and* two banks of MSB comparators. Essentially, therefore, this architecture uses two distinct two-step ADC's in a parallel time-interleaved configuration.

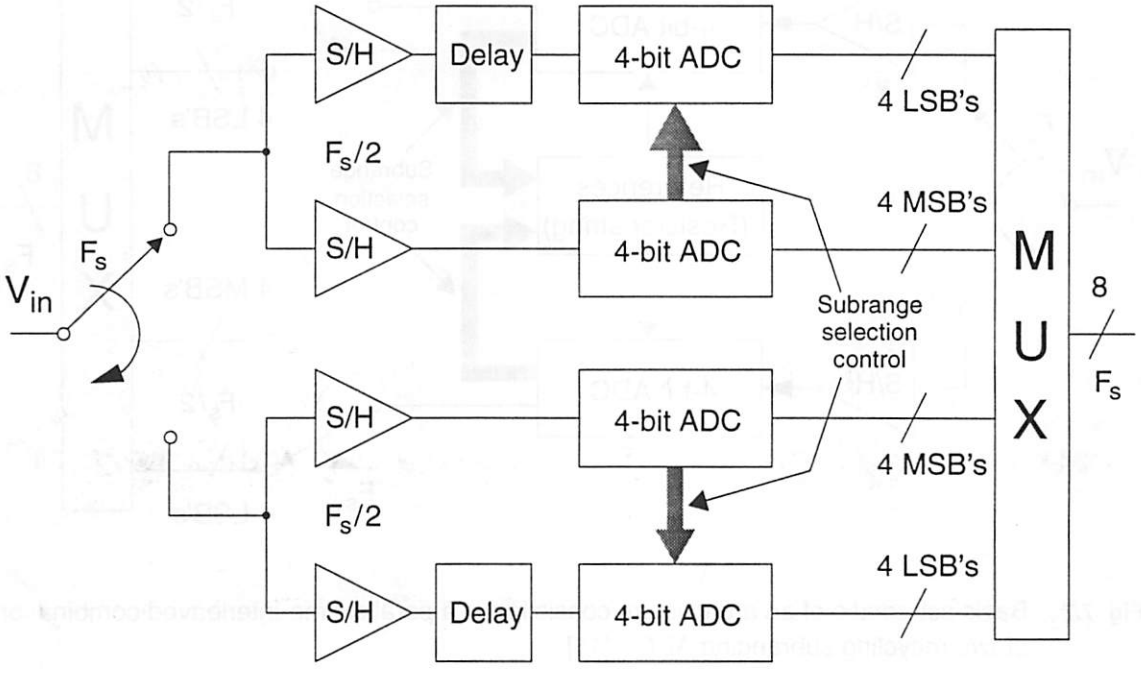


Fig. 7.6 Simplified schematic of a subranging two-step flash ADC using two banks of upper comparators and two banks of lower comparators. A single resistor string (not shown) is used to generate all required reference levels.

As may be seen from Fig. 7.6, four 4-bit sub-ADC's are present, i.e., an MSB comparator bank and an LSB comparator bank within each of the two parallel paths. A single resistor string is used to generate all reference levels. Implementations of this architecture have been reported by Hitachi in [86] and [87].

Parallelism Example 5: Interleaved combination of 2 recycling subranging ADC's

A configuration consisting of a parallel combination of two *recycling* subranging ADC's is illustrated in Fig. 7.7 below [45].

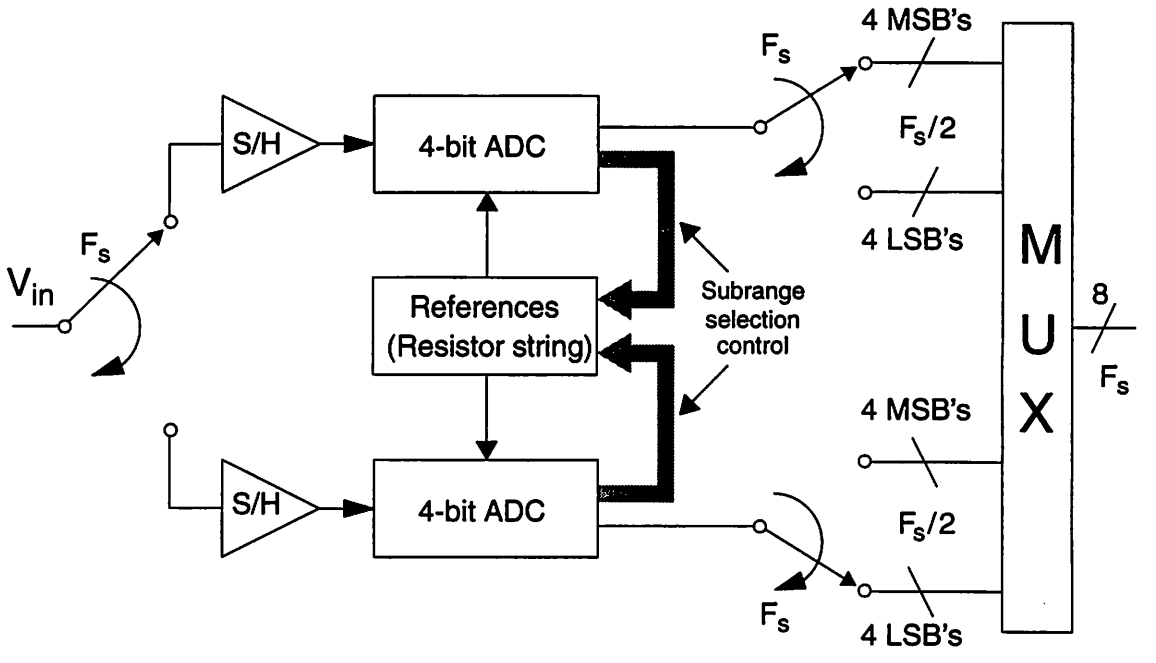


Fig. 7.7 Basic schematic of an architecture consisting of a parallel time-interleaved combination of two recycling subranging ADC's [45].

In this case, each channel uses a single 4-bit sub-ADC for *both* the MSB and the LSB decision processes. Again, as in the previous examples, there is a single resistor string, which is used by the two parallel paths.

Parallelism Example 6: Algorithmic ADC using two S/H's for higher throughput

Shown in Fig. 7.8 below is a modified SC algorithmic ADC architecture that uses *two* S/H blocks within the stage — rather than one, as illustrated for example in Fig. 4.3 — in order to increase the sample rate [56]. This architecture is an example of a case in which a key speed-limiting block was identified and replicated, resulting in twofold improvement in throughput.

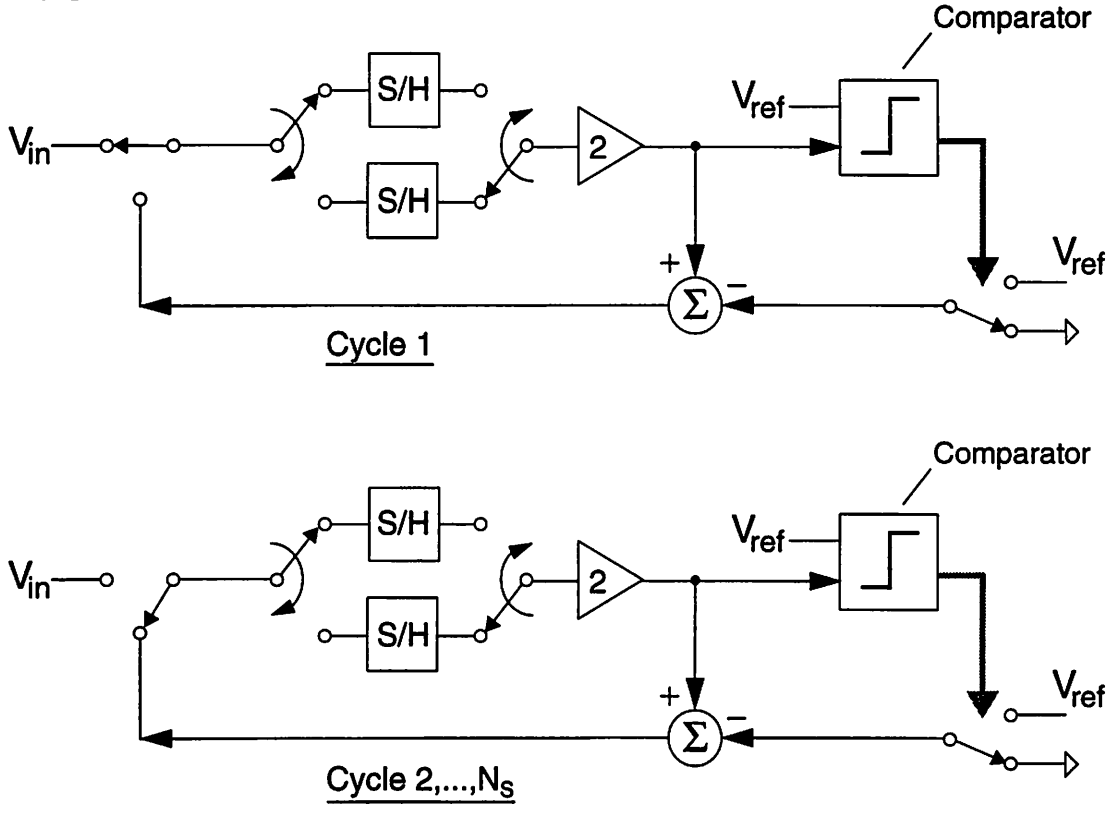


Fig. 7.8 Basic schematic of the modified algorithmic ADC structure used in [56]. The gain-of-two stage and subtractor is reset only once per conversion, and so is essentially operating as a continuous-time block.

Although a different S/H is used on alternate bit decisions, each input sample passes through the S/H's in the same sequence, however, and "sees" the same electrical path; thus, the issue of mismatch between parallel *signal paths* does not arise. Essentially, the use of the two S/H's in this ADC could be viewed as using "multiple function units" — as in digital VLSI. (Note that the prototype described in [56] employs two comparators and incorporates digital correction — this fact is not relevant to the present discussion.)

Parallelism Example 7: Two pipeline ADC's in parallel with shared sub-ADC's

An example of a pipelined multistage architecture using two time-interleaved paths with the sub-ADC blocks shared across the two channels is shown in Fig. 7.9 [161], [162].

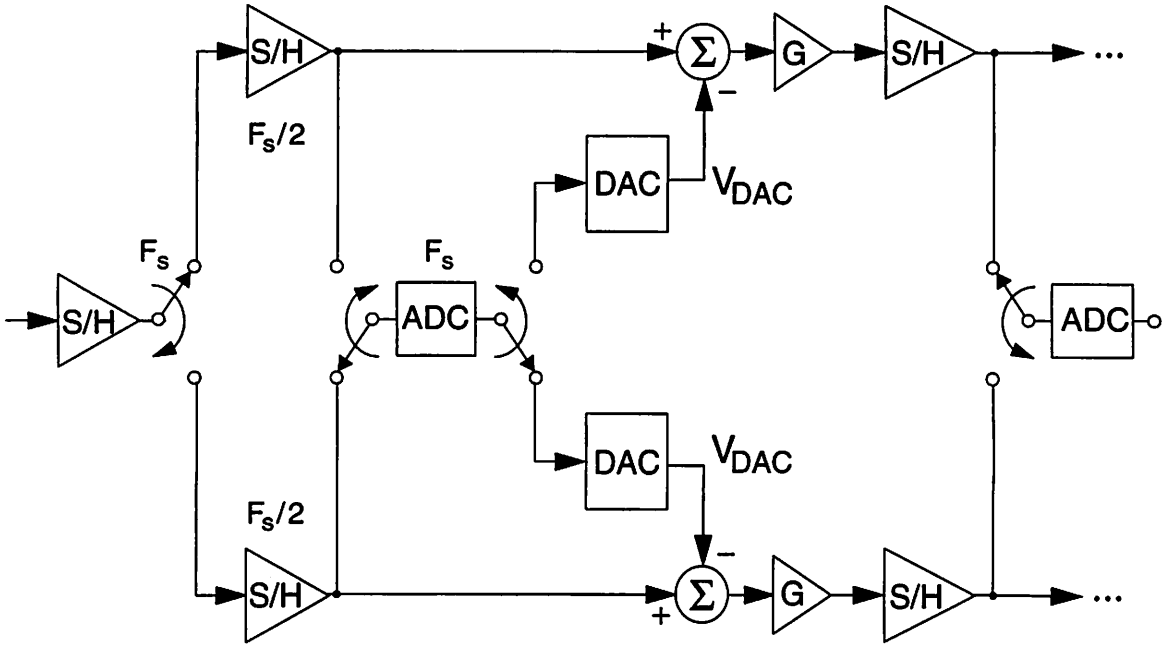


Fig. 7.9 Partial schematic of a pipeline ADC using two interleaved paths and sharing sub-ADC's.

This architecture incorporates a single high-speed front-end S/H in a similar way to Parallelism Example 1 — i.e., the two second-rank S/H's sample the held output from the first-rank S/H. There is no mismatch cancellation circuitry to deal with offset mismatch or gain mismatch between the DAC's or residue amplifiers in the two parallel paths. Offset mismatch is removed by manual adjustment of the reference voltages [162].

7.1.3 Remarks

Based on the preceding examples, it is possible to make some general observations regarding the use of parallelism in A/D converter architectures, as follows.

Note 1: The primary motivation for parallelism is to achieve the highest possible speed. In addition to the examples highlighted above, other instances of ADC subsystems that achieve samples rates of 1 GS/s or higher are described in [90], [114], and [141].

Note 2: Another general motivation for parallelism — or more precisely, for hardware replication — is resource utilization. In particular, if the system timing is such that a certain signal processing block (BLK_1) — e.g., a S/H or comparator bank — is idle and waiting for the outcome of an operation that is being performed by another block (BLK_2), then the inactivity of BLK_1 constitutes wasted silicon area and power. In such a scenario, *the addition of extra hardware often results in a proportionally greater increase in throughput*, and thus exemplifies the following broader strategy: replication and time-interleaving of hardware so as to keep all resources busy is a good way to obtain an efficient implementation.

Note 3: Many of the ADC's described above are examples of the use of *two* parallel analog signal paths: the bottleneck to the throughput in a single-channel architecture was identified, and the hardware for that block was doubled and time-interleaved, resulting in a modest increase in overall hardware — power, area, etc. — but roughly a factor of two improvement in throughput. In particular, the introduction of parallelism has been key in enabling the CMOS two-step subranging architectures to achieve 8-bit resolution at sample rates greater than or equal to 20 MS/s, which represented a significant improvement in performance as compared to the classical single-channel architecture. (For example, the single-channel two-step subranging implementation described in [28] achieved a sample rate of 5 MS/s.)

Note 4: Although, as mentioned above, parallelism may lead to greater efficiency and better resource utilization, it always increases hardware cost. Therefore, it is pertinent to ask: *what (if anything) has been done in each of the above-mentioned architectures to minimize hardware cost, in particular by sharing of resources?* Clearly, single-chip implementations have the greatest potential in this regard. In all the two-step subranging architectures a single resistor string is used. Furthermore, in general, the cost of clock generation is amortized over all parallel channels, although for the case of a chip driven by a full F_s clock and with only two channels, clock generation is relatively straightforward.

Note 5: With the exception of the architectures that use a fast front-end S/H, parallelism tends to lead to lower input impedance (depending on the way the analog demultiplexer is implemented) — i.e., higher input capacitance in the case of switched-capacitor ADC's, and lower input resistance in the case of ADC's with resistive input impedance. Furthermore, note that in the SC case, the input capacitance is *switched* at the full sample rate. These factors tend to increase the performance requirements of the driving stage.

Note 6: A complication associated with resource sharing across parallel analog signal paths is the following: any analog resource that is shared over M parallel signal paths must be fast enough to operate (i.e., settle) at a speed commensurate with the overall composite speed of the M-channel system. That tends to mean that shared bias levels and shared reference levels are required to have lower impedance than would be necessary in the case of a single channel. The important implication is that generation of these lower-impedance levels requires increased dc current, due to for example a lower-resistance resistor string, or reduced ($1/g_m$) in bias circuits. Furthermore, anything that is shared gives rise to the possibility of crosstalk and interference: in fact this represents a strong argument for maximum decoupling of parallel channels and minimum sharing of resources! In addition, interconnection and routing associated with shared resources tends to be complicated. This is obviously true for any resource that is shared between multiple chips in a hybrid implementation, but also tends to hold on a single chip.

Note 7: Although it has not been covered in detail in the examples described in this section, it is apparent that there may be significant complexity associated with clock generation for parallel time-interleaved A/D systems. In particular, whenever a time-varying analog signal is being sampled, the accuracy of the corresponding clock waveform sampling edge is critical, as is discussed later in Section 7.3.3. A description of a mixed-signal IC using parallel analog signal paths and with complex clock generation requirements may be found in [147]. A useful measure of the cost of clock generation is the number of low-jitter edges required [148].

Note 8: There is an interrelationship between pipelining and parallelism and, as in the case of digital VLSI systems, in analog systems also these tend to be present together.

Note 9: The issue of mismatches between the two parallel channels has not been addressed in any of the two-step subranging papers in the literature, possibly because the comparator autozeroing and offset cancellation are sufficient to eliminate this problem. In general, however, for any system employing parallel analog signal paths, offset, gain, and timing mismatches are of paramount importance, and must be considered at both at the system/algorithm level *and* circuit implementation level. Note that in some cases, mismatch considerations may dictate that a certain resource is shared — the parallel pipeline architecture that is described later in this chapter exemplifies this.

Summary: High-speed, highly parallel signal processing in the analog domain and on a single IC is still relatively immature and ad-hoc. Parallelism in the analog domain exhibits all of the same attributes as parallelism in digital systems, such as increased system throughput and increased hardware, which is accompanied by the potential for resource sharing. However, in the analog domain, there is an additional level of complexity (i) due to the path mismatch issue and (ii) due to the increased performance requirements on any shared resource in order to control unwanted coupling or interference. Note that sometimes these latter two considerations may be in opposition to each other, *since in*

some cases, sharing of resources may eliminate mismatches, thus leading to a key design tradeoff. Alternatively stated: given the problems associated with resource sharing, it should always be justified carefully; the most compelling reason to share an analog resource is that a signal-processing-level or system-level advantage results, such as elimination of a source of inter-channel mismatch.

In the context of time-interleaved A/D converter systems, the above examples and discussion motivate the following basic questions.

- Given that parallelism results in higher throughput, what ADC architecture should be used for the individual signal path?
- What are the properties of an ADC architecture and circuit implementation that make it particularly suitable for incorporation in an M-channel parallel interleaved array?

The following three points are relevant.

1. Use an architecture that is already hardware-efficient in the case of a single-channel implementation: pipelined multistage ADC's are characterized by good throughput per unit hardware at resolutions of about 8 bits or higher.
2. Use an architecture that facilitates sharing of resources, especially if in doing so inter-channel mismatches are alleviated.
3. Use an architecture that has intrinsically high input bandwidth and so is capable of handling the higher input slew rates and higher input signal frequencies associated with a sampling rate increased by a factor of M.

7.2 PARALLEL PIPELINE ARCHITECTURE

In Section 7.1, various examples of parallel A/D converter architectures have been described, including parallel arrays of flash ADC's, and subranging topologies employing parallelism. This section introduces a new parallel ADC structure: the *parallel pipeline array.* The key concepts of this architecture are (i) to combine a number of SC pipelined multistage ADC's in a parallel, time-interleaved configuration in order to increase throughput and (ii) to share resources across the multiple channels to keep area and power reasonable. In this way, hardware cost does not increase in direct proportion to the number of channels. A high-level block diagram view of the system is shown in Fig. 7.10.

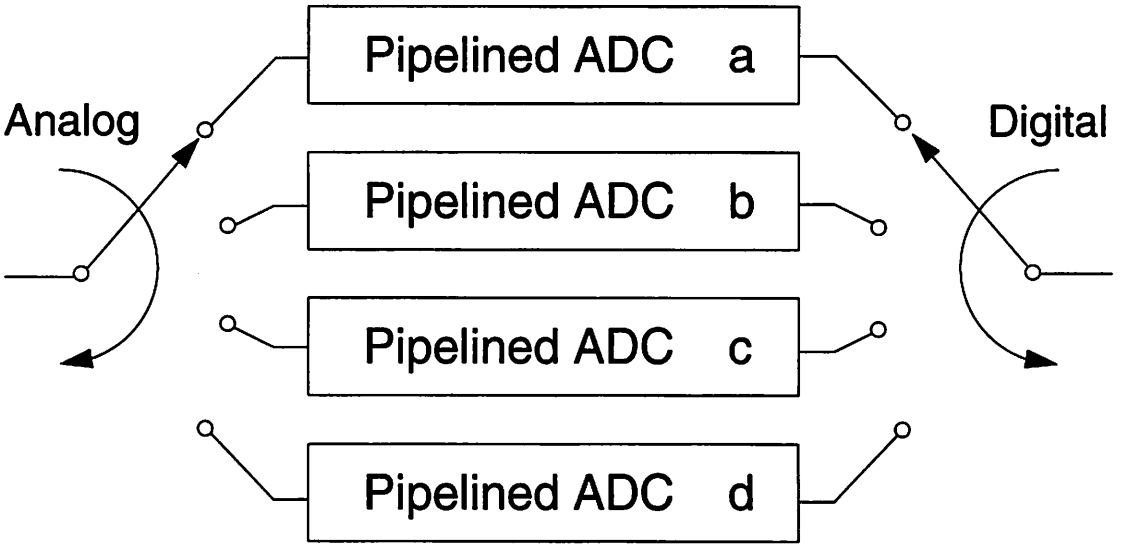


Fig. 7.10 Basic schematic of four-channel parallel pipeline array A/D converter. The individual channels/signal paths are identified as a, b, c, and d.

Clearly, any of the implementation examples described in Section 3.4.1 are possible for the pipeline ADC within each channel. However, the focus here is on Implementation Example 1 in Section 3.4.1. Thus, each ADC consists of a cascade of SC pipeline stages, with each stage comprising a k-bit resistor string ADC, a k-bit resistor string DAC, and a SC interstage residue amplifier [73]. In addition, each channel contains a front-end S/H circuit that is similar to the interstage amplifiers except that it has a gain of 1. A more detailed view of the architecture indicating the multiphase clocks is shown in Fig. 7.11.

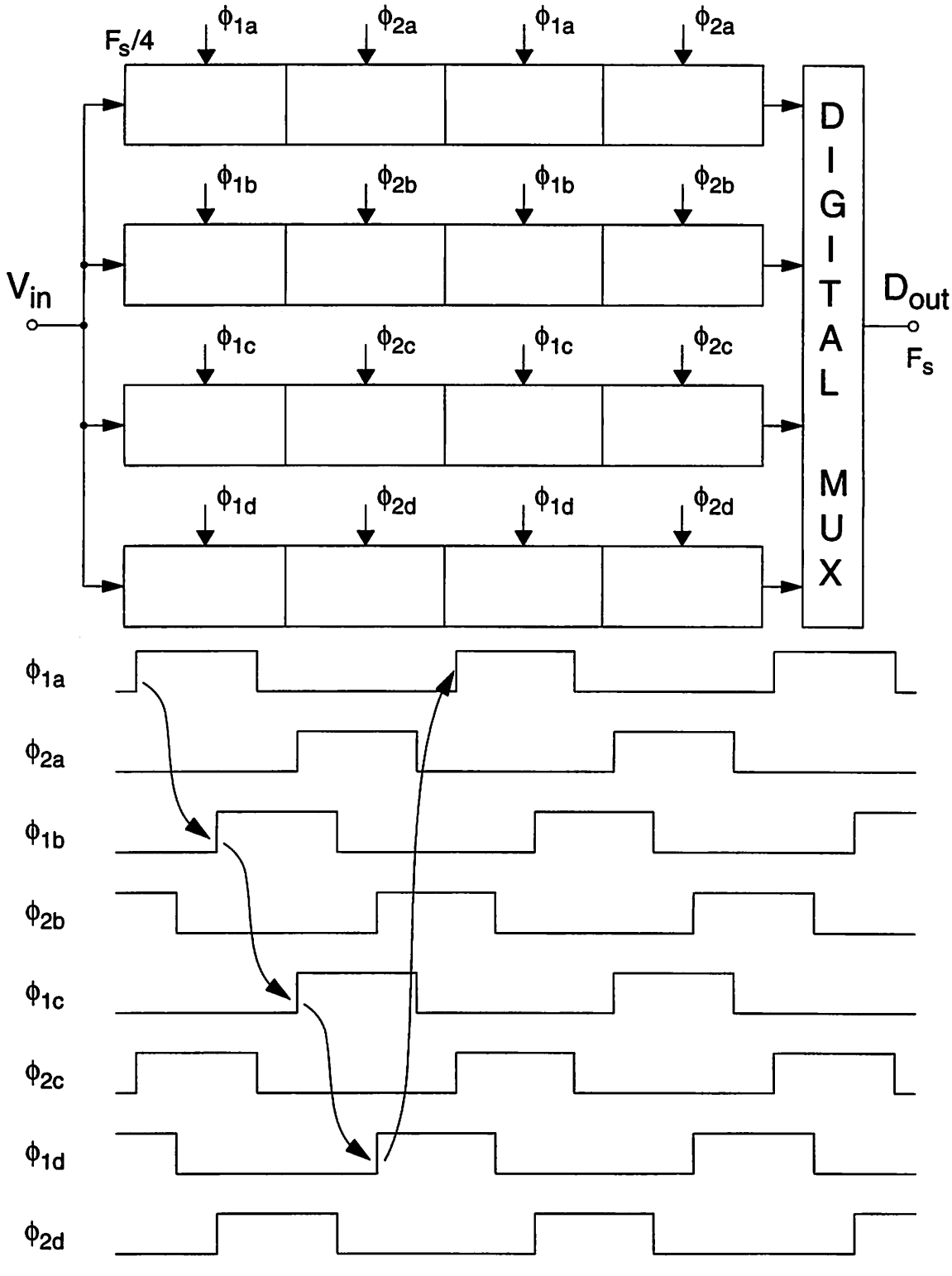


Fig. 7.11 Clocking of parallel pipeline ADC showing multiphase clocks and two-phase non-overlapping clocks for each channel. The arrows indicate the start of input acquisition — i.e., at the rising edge of ϕ_1. Each box represents one pipeline stage.

Each individual pipeline ADC channel uses a two-phase nonoverlapping clock, for example, ϕ_{1a} and ϕ_{2a} for channel "a". No explicit analog demultiplexer hardware is present — the demultiplexer function is achieved by the multiphase clocks that drive the S/H in front of each stage. The figure conveys some sense of the clocking complexity associated with the architecture. The topology results in a two-dimensional array of pipeline stages, displaying a loose resemblance to a digital systolic array. Thus, an alternative name for this ADC architecture is a *systolic A/D converter.*

Another view of the timing organization is shown in Fig. 7.12. Each horizontal bold solid arrow represents the cycle time of a particular pipeline stage within each channel, that is, the time taken for a stage to perform the following sequence of operations: acquisition and sampling of a new input or residue, a low-resolution A/D conversion, a low-resolution D/A conversion, and finally amplification and transfer of the resulting residue to the next stage. The concurrency *within* each channel is clear: adjacent stages operate 180° out of phase — e.g., the residue amplification phase of one stage occurs simultaneously with the signal acquisition phase of the next stage. This is stated explicitly on Fig. 7.12 for the stages within ADC "a". The concurrency *across the array* is indicated by the shaded arrow, which shows the channel sampling sequence: a→b→c→d→a, etc. In this four-channel example, each pipeline ADC operates at a rate of 1/(4T) and so the allowed time for op amp settling is 2T. In a single-channel pipeline ADC, with the same throughput $F_s = 1/T$ and employing 50% duty-cycle clocks, the time allowed for op amp settling is T/2. Therefore, for an ADC with sample rate F_s, implementation using a time-interleaved combination of M pipeline converters allows the S/H and interstage gain amplifier op amps M times longer to settle, which clearly makes circuit realization *significantly* easier.

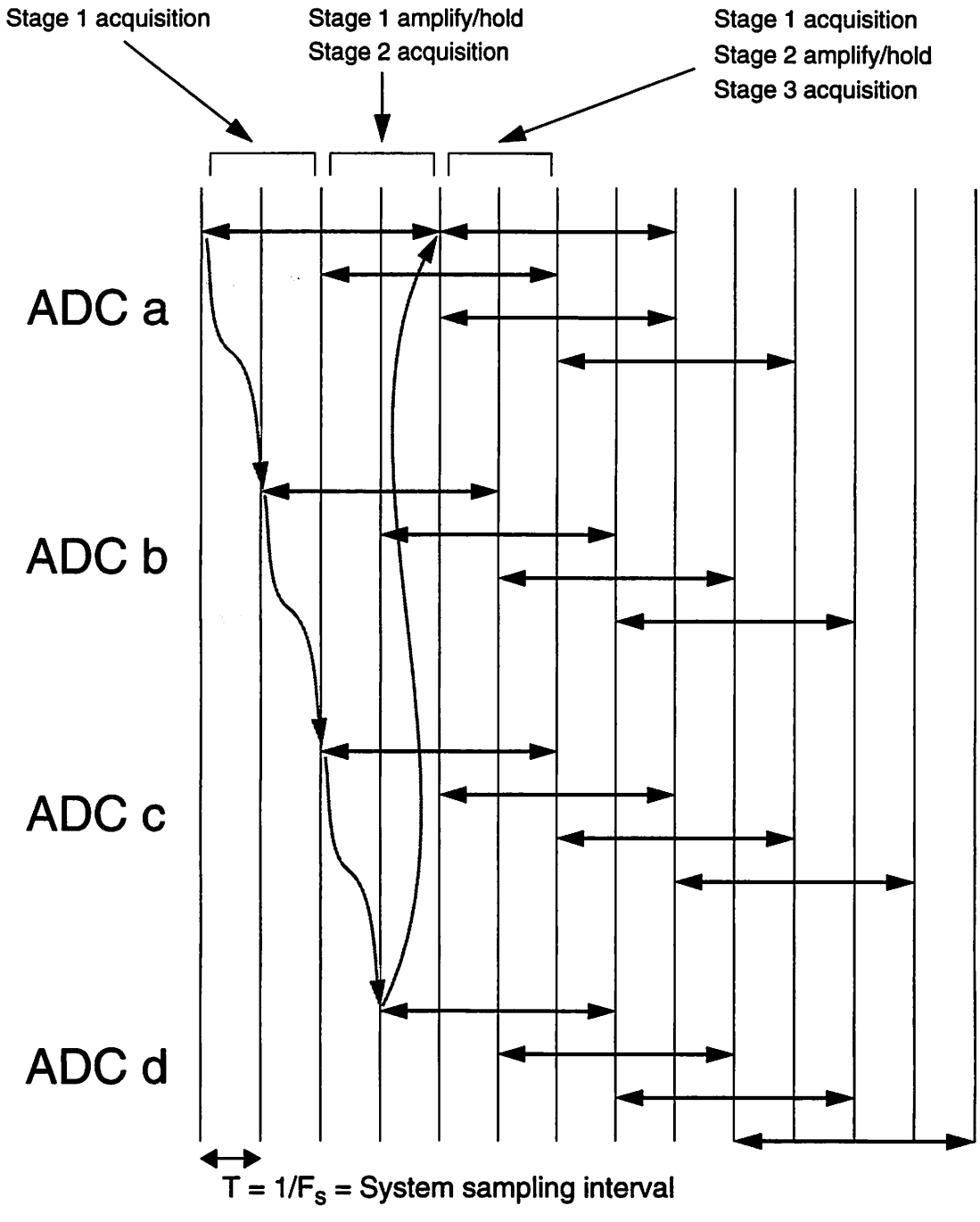


Fig. 7.12 Diagram of timing organization for parallel pipeline A/D converter.

The comparator clocking is a simple two-phase amplify-latch arrangement: during phase 1, the inputs are connected (i.e., reset) to the appropriate reference levels, and the latch latches the decision from the previous decision; during phase 2, the inputs are connected to the analog signal, and the comparator preamp amplifies the input difference, thereby precharging the latch. Note that since the comparators in the individual ADC channels are being clocked at F_s/M rather than at F_s, problems due to latch metastability are alleviated [163], [81], [82]. A simplified schematic of the comparator is illustrated in Fig. 7.13 below.

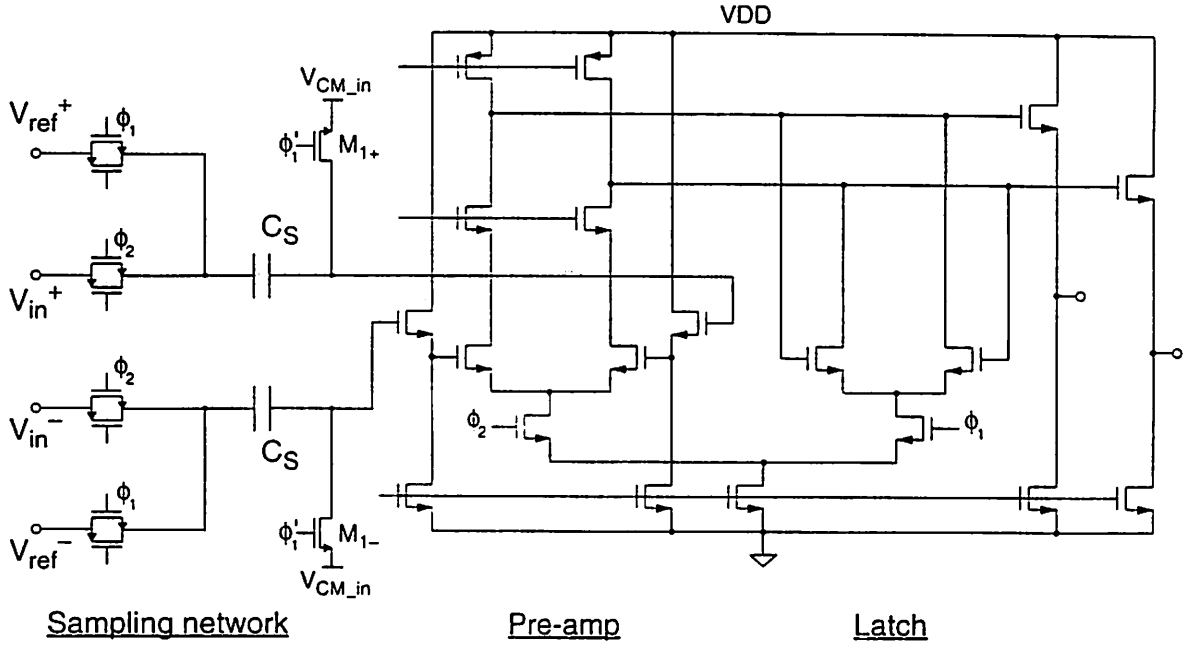


Fig. 7.13 Schematic of 2-phase switched-capacitor comparator: the implementation is a clocked preamplifier and regenerative latch configuration similar to that described in [73].

Note that because a front-end S/H is used, it is convenient to consider one "stage" of the pipeline as consisting of the SC gain stage *followed* by the sub-ADC and sub-DAC. This differs from the description in Chapters 3 and 4: compare, for example, the schematic of Fig. 3.31; however, this view is more appropriate when discussing issues relating to implementation and layout. As mentioned above, adjacent stages within each pipeline ADC channel operate 180° out of phase: the exact operations being performed by an individual stage during each of those two phases are now examined in more detail.

The operation of the first stage in acquisition mode is indicated in Fig. 7.14 below.

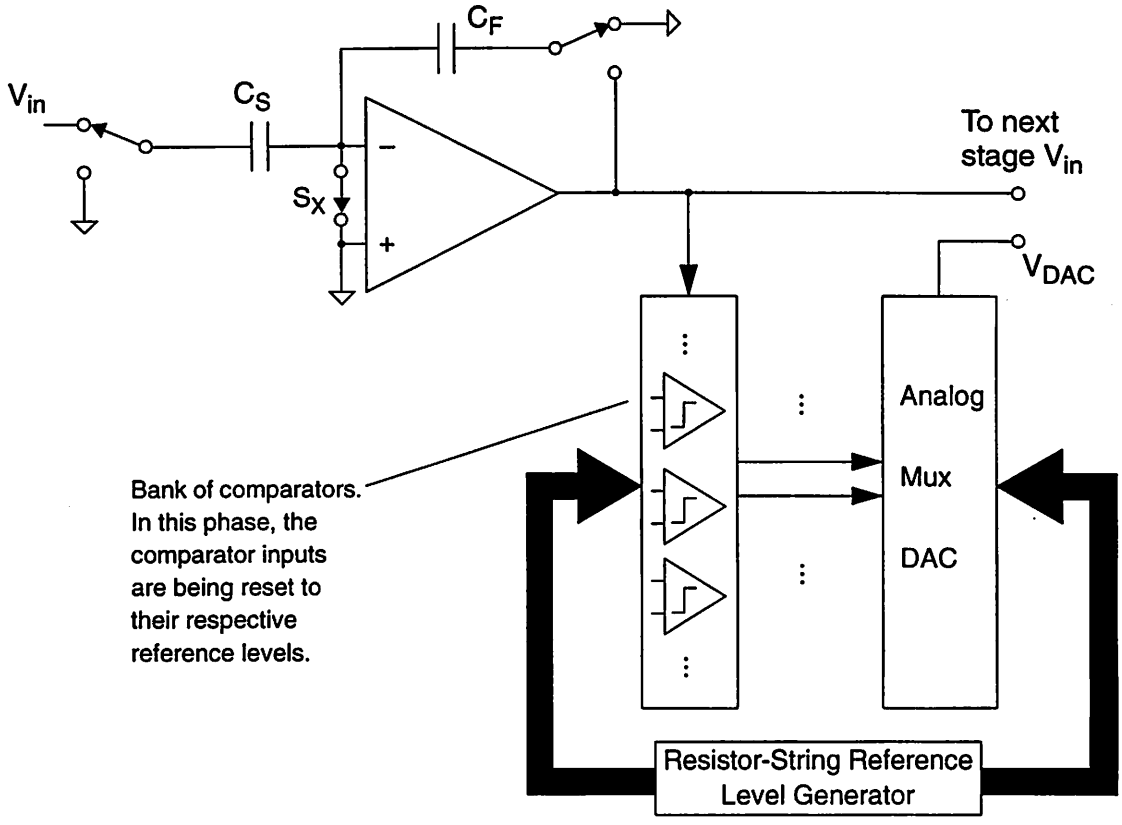


Fig. 7.14 Simplified single-ended circuit schematic of the first stage of a SC pipeline ADC. The switch configuration shown corresponds to acquisition/tracking/rest phase, i.e., phase 1.

In this phase, which may be thought of as the first half of the cycle, the following operations are occurring simultaneously: the input V_{in} is being acquired on sampling capacitor C_S of the S/H or interstage gain block; the input common-mode level of the op amp is being reset; the common-mode feedback (CMFB) circuitry is being reset and the op amp outputs are being reset to the output common-mode level; the inputs of the switched-capacitor comparators are being reset to the appropriate reference levels, i.e., their threshold levels; the latches are latching the decision from the *previous* sample and — via some digital logic — causing a new DAC output to be selected from the resistor string; this new DAC output is connected to the next stage, and is being used in that stage's amplification phase. Note that the only difference between the first stage and the internal stages is that first stage has a gain of 1 and has no DAC connection at its input.

The operation of an internal stage in amplification mode is shown in Fig. 7.15 below.

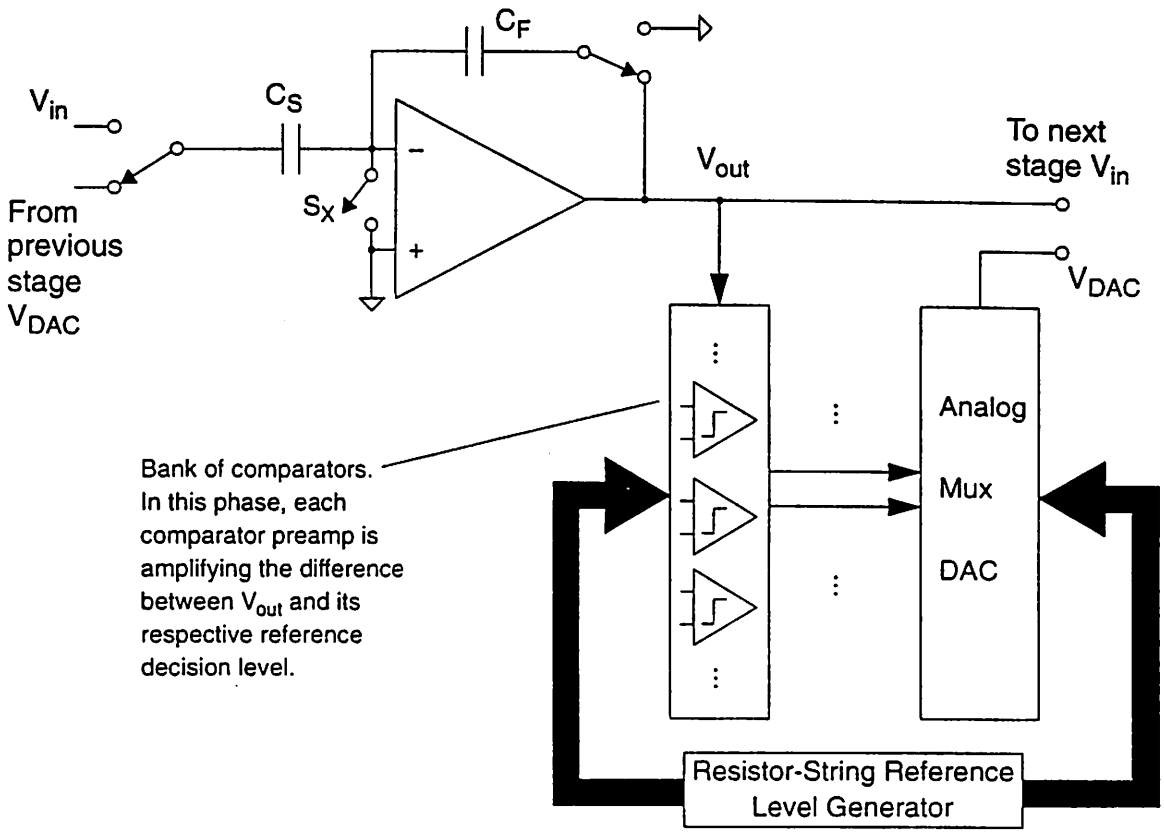


Fig. 7.15 Simplified single-ended schematic of one internal stage of a SC pipeline ADC; the switch configuration shown corresponds to the amplification/charge redistribution phase.

In this phase, which may be thought of as the second half of the cycle, the following operations are occurring simultaneously: the SC gain stage is connected to the DAC output from the previous stage; the amplifier is in its closed loop configuration and is amplifying the difference between the input voltage sampled on the previous phase and the DAC output from the previous stage; thus, the op amp output is settling to is final value given by $V_{out} = (C_S/C_F)(V_{in} - V_{DAC})$; this output voltage is connected to the inputs of the switched-capacitor comparators and the comparator preamps are amplifying the differences $(V_{out} - V_{threshold})$, and thereby precharging the latches; in addition, the output voltage from the gain stage is being acquired on the sampling capacitor of the next stage, which is operating 180° out of phase.

A simplified schematic/floorplan is shown in Fig. 7.16 illustrating four parallel channels and, within each channel, the input S/H and a portion of the first pipeline stage, including the bank of comparators. With reference to Fig. 7.16, the block marked "CAPS" contains the sampling capacitors for the comparators. The resistor string for all four pipeline ADC first stages is located in the center of the array (not shown in Fig. 7.16), and its reference levels are distributed vertically to all four channels, thereby eliminating inter-channel DAC mismatches. Bias levels for the comparators and op amps are also generated centrally and distributed vertically. (For simplicity, the op amp bias lines and the DAC connections to the interstage residue amplifier are not shown in Fig. 7.16.) The two connections for the interstage gain amplifier may be seen — i.e., one input is the analog output of the previous stage, the other input is the DAC output. The gain through each channel is set primarily by the matching of the input S/H capacitor ratio across the array, which is expected to be better than 0.1%. The differential input is shown on the left-hand side and is connected to the S/H of each channel. Signal flow and clock distribution are horizontal.

Clock routing is simplified by placing channels that are 180° out of phase *adjacent* to each other in the layout and routing the clock lines physically between them, so that a particular clock signal functions as ϕ_1 for one channel and ϕ_2 for the other channel, etc. With reference to the sampling sequence illustrated in Fig. 7.11 and Fig. 7.12, this implies that channels a and c are physically adjacent to each other in the layout, as are channels b and d. A similar strategy is possible in any multiphase time-interleaved system with an even number of phases.

Because accuracy and speed requirements are less stringent in the later stages of a multistage A/D converter, since any errors, when referred to the input, get divided by interstage gains, as explained in Notes 1 and 2 in Section 4.1, the capacitors and op amps may be made smaller in later stages, which results in a saving in chip power dissipation. The exact values used in the prototype implementation are given in Chapter 8.

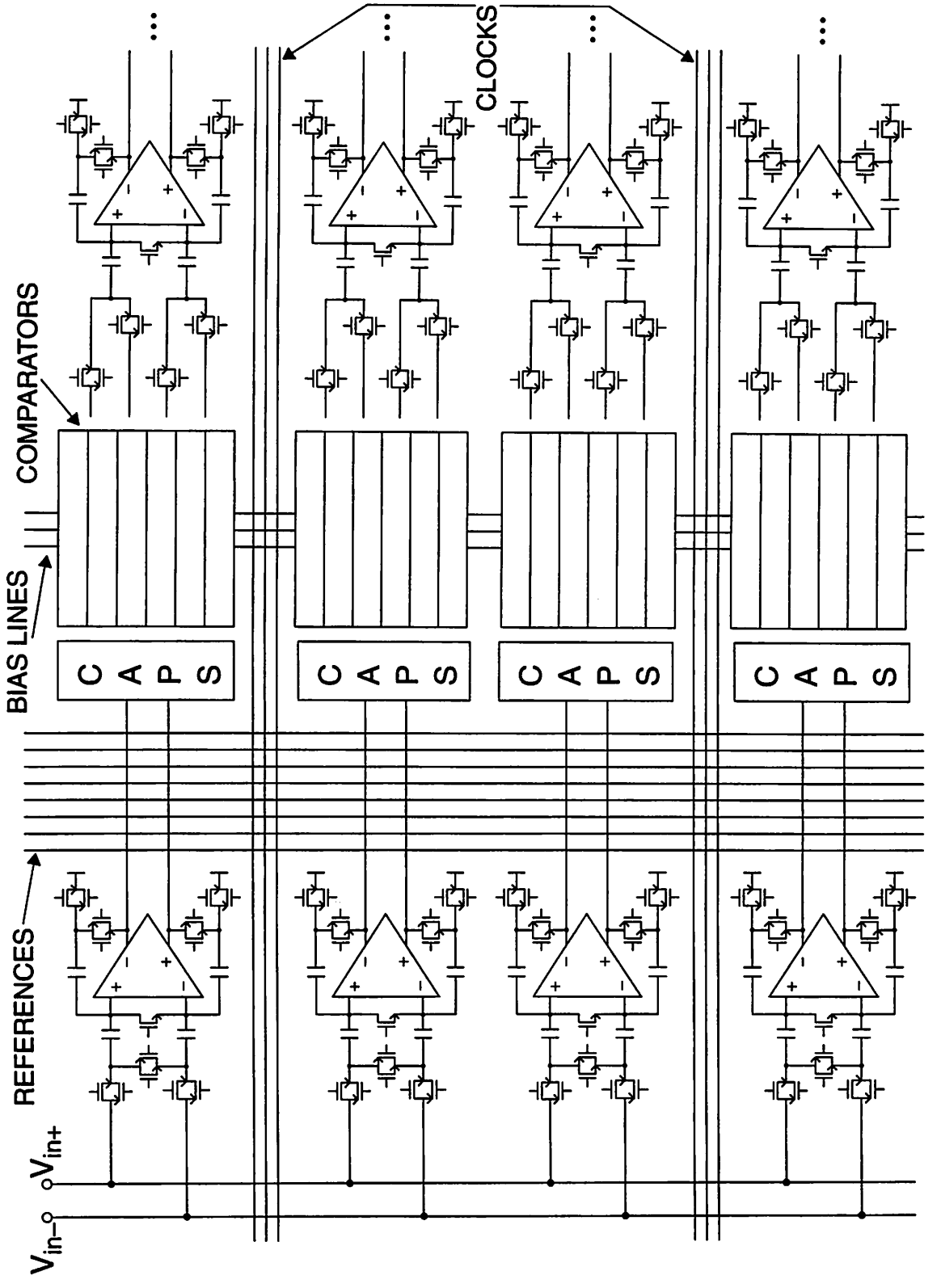


Fig. 7.16 Partial floorplan of parallel pipeline ADC with some layout information indicated.

7.3 PERFORMANCE LIMITATIONS

7.3.0 Overview

Performance limitations for parallel pipeline ADC's fall into two categories, as follows.

Performance Limitation 1: The use of parallelism in the *analog* domain introduces problems that do not arise in parallel digital systems, namely offset, gain, and timing mismatches [8], [53], [105], [109]. The topic of mismatches in parallel signal paths has been covered in Chapter 6 from a signal processing perspective. In Sections 7.3.1–7.3.3, the results from that chapter for offset, gain, and timing mismatches respectively are summarized, and then, for each one, pertinent implementation and circuit issues are considered. Note that in the following discussion, the notation of Chapter 6 is used — i.e., F_s is the overall system sampling frequency, M is the number of parallel channels, and so the individual channel sampling frequency is F_s/M.

Performance Limitation 2: Within each channel, the speed of each individual pipeline ADC is limited by the settling time of the S/H during "hold" mode, which is determined by an op amp settling time. Chapter 5 has discussed SC gain stage settling time analysis and optimization. Those results are reviewed briefly in Section 7.3.4 and then further implementation issues relating to the S/H and gain stages are discussed.

7.3.1 Channel Offset Mismatch

Channel offset mismatch across the array is due to op amp offset mismatches and charge injection mismatches from the switches in the S/H and interstage gain amplifiers in each channel. A schematic representation of the phenomenon along with its effect in the frequency domain for a four-channel case is illustrated in Fig. 7.17.

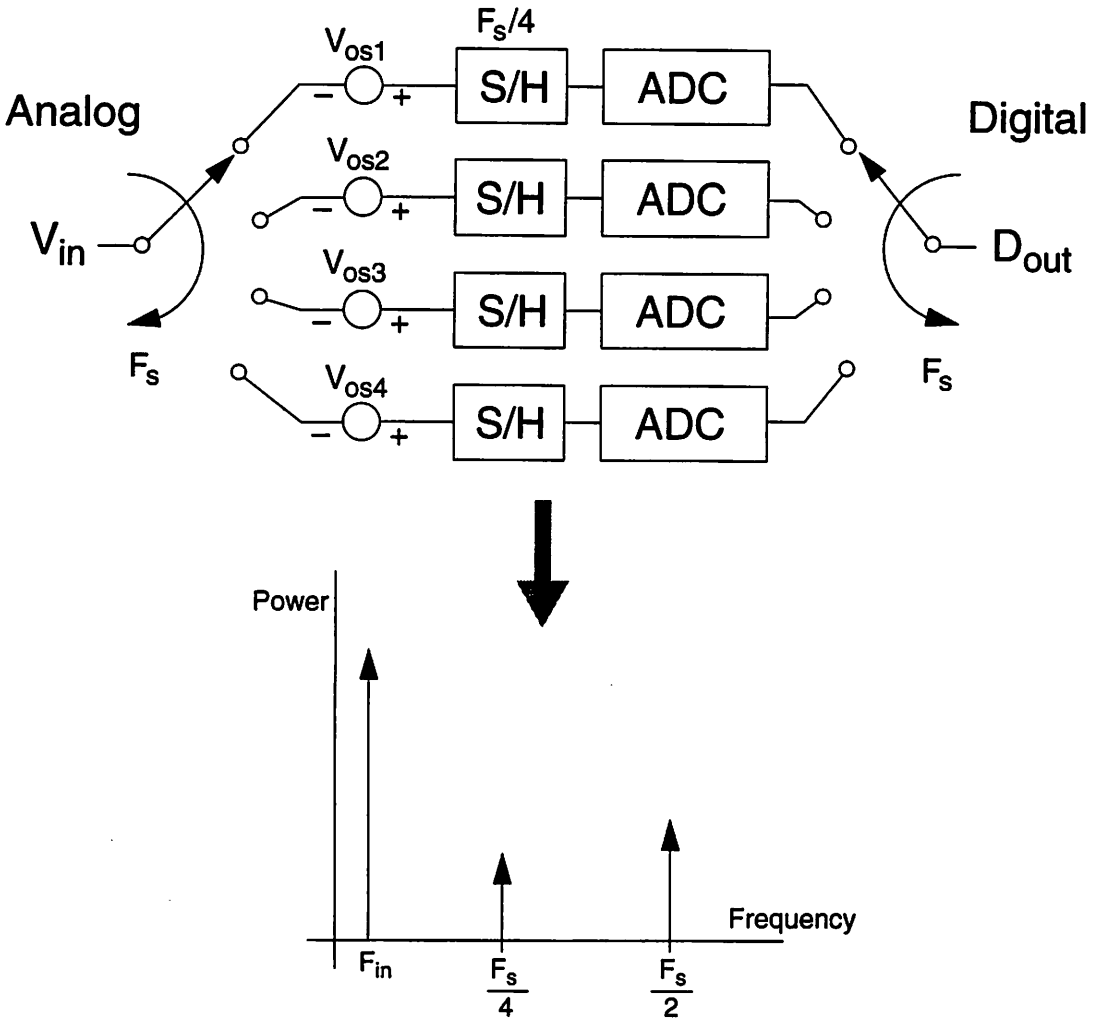


Fig. 7.17 Basic schematic of a 4-channel parallel time-interleaved ADC array showing different channel offsets and a typical effect in the frequency domain at the converter output.

Inter-channel offset mismatch gives rise to "fixed pattern noise": in an extreme case, for a DC input, each channel produces a different output code. In the frequency domain, as may be seen from the spectrum shown in Fig. 7.17, this is manifested as frequency components or tones at multiples of the channel rate F_s/M. In some applications — for example, video — such pattern noise is visible on a TV screen and not tolerable.

Another manifestation of offset mismatch [26] is the phenomenon of *pattern noise dependent on the input dc level*, as follows. Consider the behavior of a parallel interleaved ADC when the overall system sampling rate is F_s and the input V_{in} is ramped very slowly

from –Full Scale to +Full Scale: in some ranges of V_{in} the output codes from each of the M parallel ADC's are identical; in other ranges of V_{in} the different channels give different output codes resulting in the above-mentioned tones at multiples of F_s/M. Therefore, the existence or not of unwanted frequency components in the output spectrum, and their relative magnitudes, *depends on the value of the dc input* V_{in} — an unusual phenomenon. Furthermore, it is apparent that the smaller the mismatch in effective threshold levels across the array, the smaller the ranges of the analog input that produce mismatch tones.

Two possible solution approaches exist to dealing with offset mismatch.

Analog offset cancellation techniques may be employed such as the use of an auxiliary input stage to store a representation of the op amp offset and charge injection offset on capacitors [101]. Typically, the additional autozero circuitry is incorporated in the front-end S/H and in the next few stages and is not necessary in the later stages in the pipeline.

Digital offset cancellation is also possible: the offset for each channel is measured digitally and is subtracted from the "raw" output digital code of that channel. In order to ensure that the offset is less than 1/2 LSB of the desired overall system resolution, it is necessary to use one extra bit of resolution when digitizing the offset. For example, if the offset is measured using a 9-bit ADC, then it is known to 1/2 LSB accuracy at 8 bits. However, in multistage architectures with a small number (2–3) bits per stage, an extra bit in the overall system resolution necessitates only a modest increase in hardware — at most a single extra stage or perhaps modification to one stage. This is an important advantage of the pipeline architecture and is in contrast to a flash architecture, in which a doubling of the hardware would be required.

Note that for both the analog and the digital offset cancellation schemes, it is assumed that a suitable time exists (e.g., some system reset time) for autozeroing — it is not intended that measurement of the channel offsets is done "on the fly". More generally, analog autozero usually permits finer correction granularity, but suffers from the drawback

that it invariably involves additional analog circuit complexity and layout complexity, and adds extra parasitic capacitance, additional clocks, etc., and at high speeds tend to be difficult to realize. Purely digital calibration tends to be more flexible and permits programmability, etc.

7.3.2 Channel Gain Mismatch

Inter-channel gain mismatch is illustrated schematically in Fig. 7.18, and also shown is its effect in the frequency domain for a four-channel case.

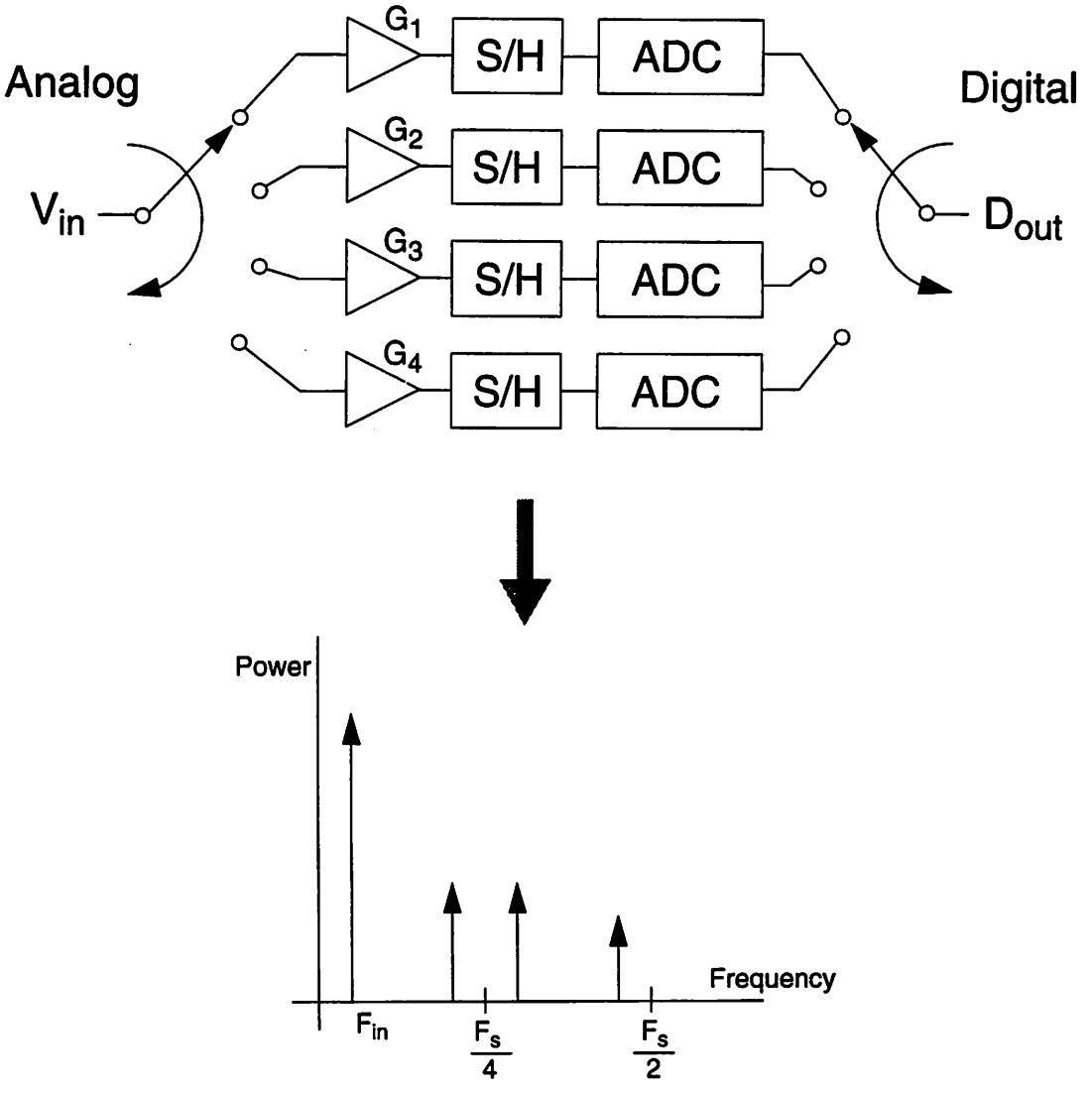


Fig. 7.18 Block diagram of a time-interleaved ADC array with four channels showing different channel gains and typical effect in the output spectrum.

In the frequency domain, inter-channel gain mismatch results in sidebands centered around multiples of F_s/M. Intuitively, the reason for this is as follows. Each individual channel samples the input signal at a rate of F_s/M causing the input spectrum to be repeated periodically at intervals of F_s/M. When the channels are perfectly matched these periodic repetitions cancel *except at integer multiples of* F_s. If some mismatch is present, however, these alias components do not cancel completely; thus, some residual alias components at multiples of F_s/M remain in the output spectrum. Alternatively, sampling with mismatched channel gains may also be regarded as multiplying (or modulating) the analog input by a periodic discrete-time sequence of period MT. The values of the sequence are (using the notation of Fig. 7.18)

$$\ldots, G_1, G_2, G_3, G_4, G_1, G_2, G_3, G_4, G_1, G_2, \ldots, \text{etc.}$$

For an input sinusoid at frequency F_{in}, this modulation gives rise to spurious mismatch products in the output spectrum at frequencies

$$F_s/M \pm F_{in}, 2F_s/M \pm F_{in}, 3F_s/M \pm F_{in}, \ldots, (M-1)F_s/M \pm F_{in}$$

There are two circuit causes of gain mismatch in the parallel pipeline ADC architecture. First, the gain of the input S/H and the interstage amplifiers in each channel is given by $G = C_S/C_F$, and thus set by a capacitor ratio. Usually with careful sizing and layout, capacitor matching good to about 0.1% is possible; this is satisfactory at the 8-bit level. Second, the DAC levels in each pipeline stage constitutes another error source, since DAC level mismatch across the channels causes both offset and gain mismatch. However, as discussed in the Section 7.2, in the parallel pipeline architecture the resistor string DAC level generator is shared over the array of parallel ADC's for hardware minimization reasons: therefore, perfect DAC matching is automatic.

7.3.3 Timing Mismatches and Clock Generation

Clock generation for a parallel pipeline ADC involves a number of considerations. In general, M multiphase clocks for the M time-interleaved channels are required, as depicted in Fig. 7.1. Furthermore, within each channel, each pipeline ADC requires two-phase nonoverlapping clocks, as illustrated in Fig. 7.11, and delayed sampling edges in order to implement "bottom-plate sampling" for minimization of signal-dependent charge injection effects in the S/H as described in [80]. In general, two types of sampling nonidealities exist for time-interleaved ADC's: random sampling jitter and systematic timing mismatches. Intuitively, as the frequency of the input signal increases, the effect of both these nonidealities becomes worse.

First, the phenomenon of random sampling jitter is considered. Sampling jitter is illustrated schematically in Fig. 7.19. Clearly, this issue arises for single-channel as well as multi-channel A/D systems.

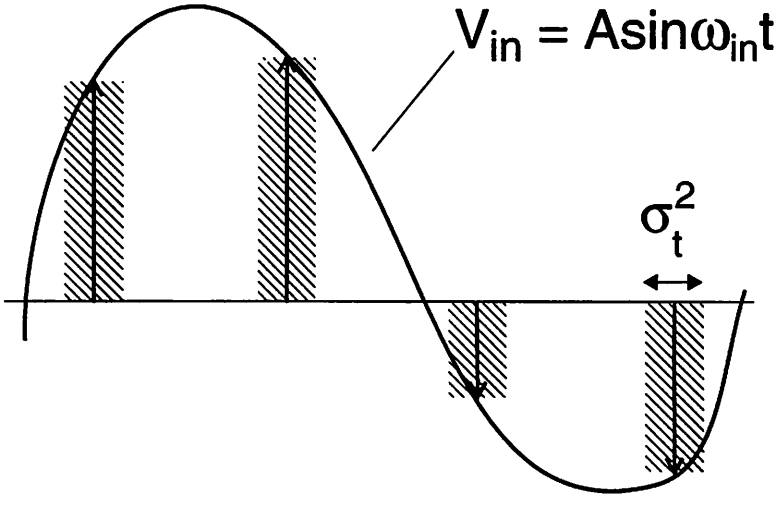


Fig. 7.19 Schematic representation of sampling jitter — the sampling instant has a random variation characterized by a variance σ_t^2, where σ_t is the RMS jitter — usually in ps.

Qualitatively, jitter results in a broadening of the spectral line and a raising of the noise floor. Quantitatively, for an input sinusoid of amplitude A and angular frequency ω_{in}, the presence of sampling jitter of variance σ_t^2 results in output noise power given by $1/2\, A^2\omega_{in}^2\sigma_t^2$, and thus degrades the SNR.

Second, a key timing-related problem is systematic skew or nonuniform sampling — i.e., *fixed* sample time mismatches between the channels. The resulting behavior in the frequency domain is qualitatively similar to gain mismatch errors: sidebands around F_s/M. However, the magnitude of the sidebands is frequency-dependent, as shown in Fig. 7.20.

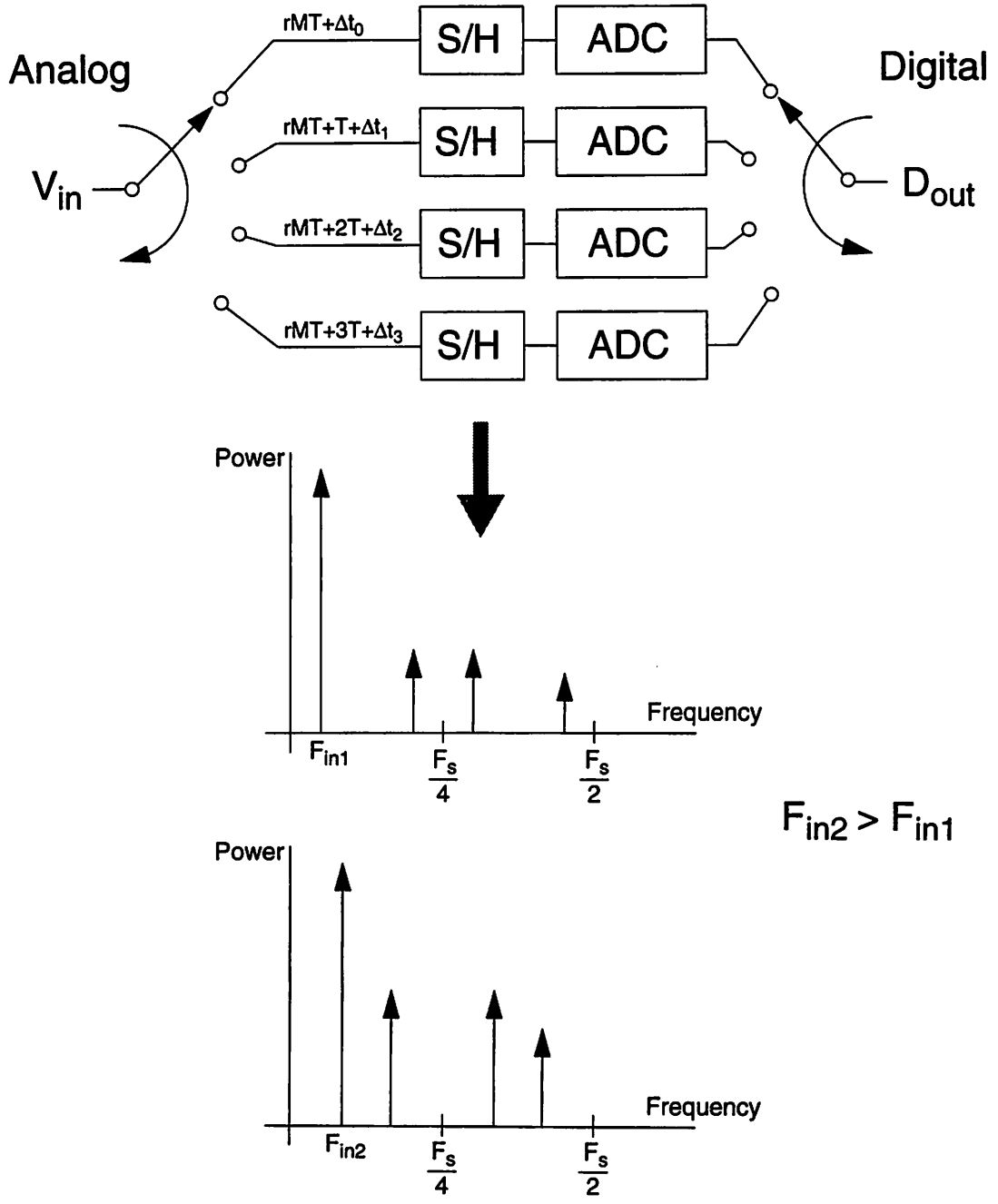


Fig. 7.20 Basic schematic of a parallel time-interleaved ADC indicating sample time mismatches and their effect in the frequency domain for two input frequencies.

Furthermore, as mentioned in Section 6.3.2, there are some other differences between gain and timing mismatches with respect to the resulting output spectra; specifically, the sidebands caused by nonuniform sampling do not, in general, have equal magnitude.

There are two radically different approaches for clock generation for time-interleaved ADC's or for any parallel analog system. The first possibility is to use a full sample rate clock (i.e., at frequency F_s) and use a purely digital approach using synchronous counters/dividers, etc. The second approach is to use some form of phase-lock technique such as a phase-locked loop (PLL) or delay-locked loop (DLL).

Here, the clock generation is implemented using a voltage-controlled delay line in a DLL configuration [44], [55], [62], [156], as illustrated schematically in Fig. 7.21.

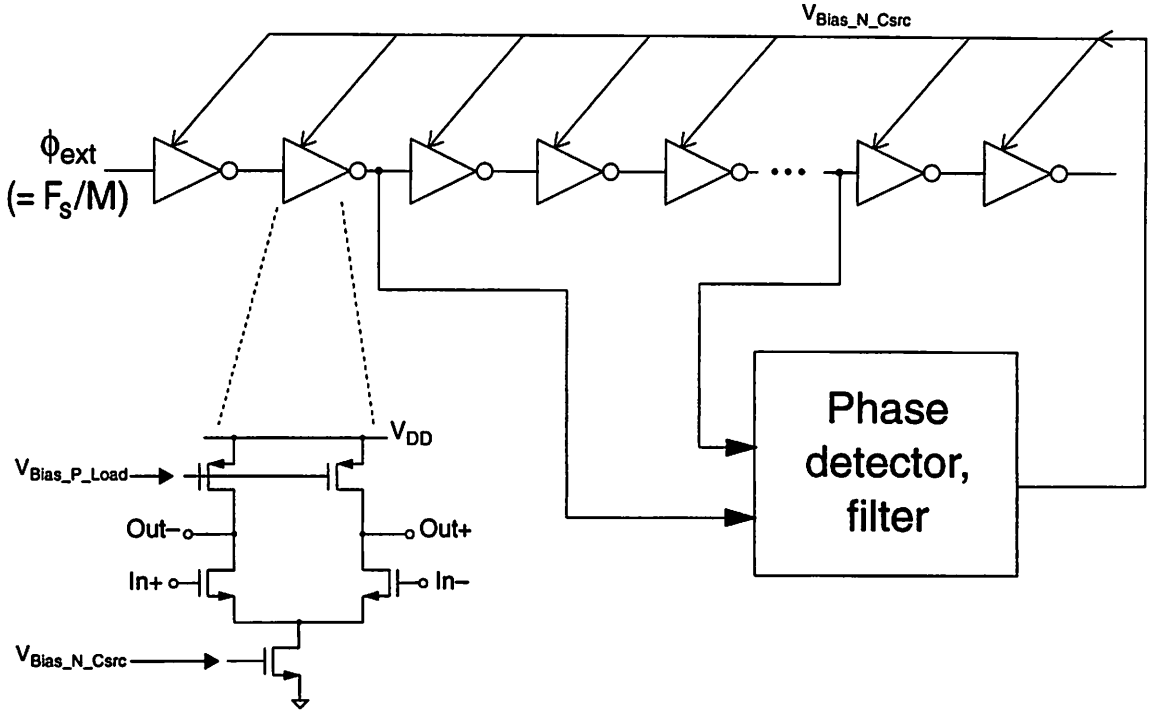


Fig. 7.21 Schematic of Delay Locked Loop based clock generation.

One advantage of a DLL over a phase-locked loop (PLL) is that a DLL topology has intrinsically lower jitter since it does not exhibit jitter accumulation — i.e., in the frequency domain, jitter peaking. The tapped delay line generates equally-spaced clock

edges that divide the period of the input reference clock ($\phi_{ext} = F_s/M$) into a number of equal intervals. Thus, this approach automatically provides the sampling phase granularity needed for the two-phase nonoverlapping clocks and delayed edges within each channel.

Furthermore, with this approach, it is not necessary to have an external clock at the full sampling rate. This is a key advantage since in a real system application, dealing with the input-output issues of a high-speed clock running at 100 MHz or higher can be very difficult. The delay cell itself is implemented as an NMOS differential pair with PMOS linear-region or triode loads [62], [41]. The differential implementation greatly reduces the sensitivity to power supply noise. The delay cell signals have approximately a swing of 1 V. These signals pass through source followers and then into a stage resembling a CMOS version of an ECL NOR gate in order to generate the larger-swing clock signals for the switch gates. The phase detector is a multiplying phase detector similar to that described in [156]. The DLL divides the input reference clock period into 64 intervals. However, because not all clock edges are needed by the ADC, and because the phase detector only requires two edges that are 90° out of phase, only 54 delay cells are necessary in the implementation.

7.3.4 Channel Input Sample/Hold and Op Amp

A key circuit issue in parallel pipeline ADC's is the design of the Sample/Hold (S/H) stage at the front end of each channel, since during signal acquisition each channel sees the full input signal bandwidth, as can be seen on the left-hand side of Fig. 7.16. At this point, it is appropriate to remark that in this dissertation, unless otherwise stated, it is assumed that for an A/D converter system with sampling rate F_s, the input signal may possess frequency components from dc to the Nyquist frequency $F_s/2$, and timing issues are considered on that basis. However, in some applications, when it is known from system considerations that the input signal spectral energy is limited to a lower band, the A/D architecture can take advantage of that fact. In particular, many ADC's are specifically intended for digital video applications where the peak input signal bandwidth is typically 4–5 MHz. When the input source is known to be from a CCD, which produces a discrete-time signal, the requirements on the S/H part of the ADC are greatly reduced [65], [84]. A simplified schematic representation of the input S/H is shown in Fig. 7.22.

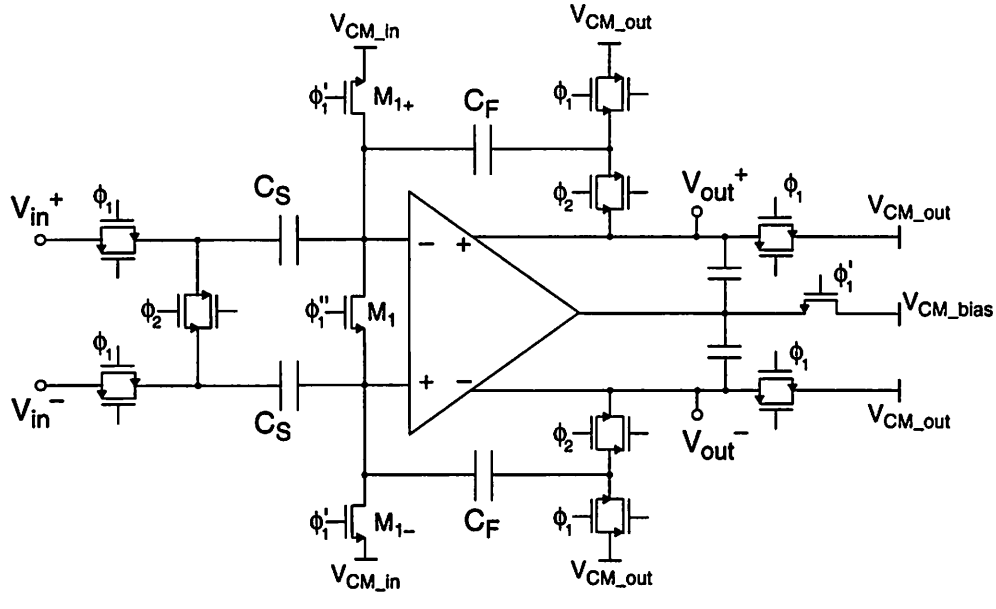


Fig. 7.22 Basic schematic of fully-differential switched-capacitor stage as used in the parallel pipeline ADC showing sampling capacitance (C_S) and feedback capacitance (C_F).

The transient response during acquisition/tracking mode and during amplification/hold mode are now examined.

Acquisition/Tracking

For the parallel pipeline architecture, signal acquisition time is limited by the RC delay of the input sampling network, which depends on the on-resistances of the input and center switches and the capacitances of the network. Since there are a number of resistances and capacitances, it is useful to examine the input network more carefully.

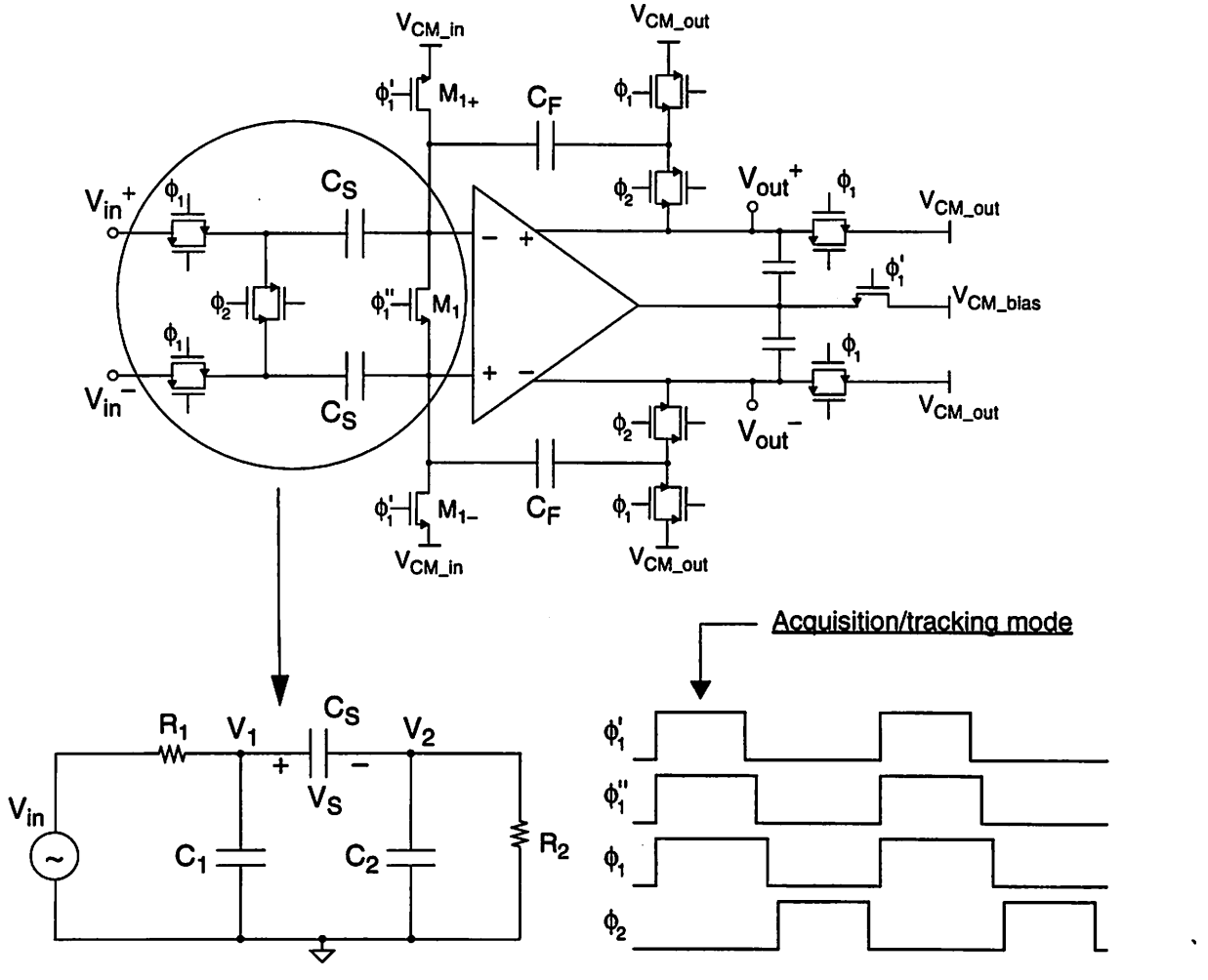


Fig. 7.23 Basic schematic of differential switched-capacitor stage showing detailed circuit model of input network.

Note that the input network is purely passive — this is sometimes referred to as *passive sampling* or *open-loop sampling*. The transfer function of interest is from the input

V_{in} to the voltage V_S across the sampling capacitor. Assuming that this network may be approximated by a first-order response, and using the technique of "Zero-Value Time Constants" [37], the effective overall time constant of this network is approximately given by

$$\tau = C_1R_1 + C_2R_2 + C_S(R_1 + R_2) \tag{7.0}$$

which may also be written as

$$\tau = (R_1 + R_2)\,(C_S + C_1 \cdot \frac{R_1}{R_1 + R_2} + C_2 \cdot \frac{R_2}{R_1 + R_2}) \tag{7.1}$$

In the input S/H stage in the 1-μm prototype described in Chapter 8, C_S and C_F are each 1 pF; the transistors in the input switch have sizes $(W/L)_N = 160/1$ and $(W/L)_P = 400/1$ (in microns); the NMOS center switch has size (W/L) = 150/1; and the bandwidth of the sampling network is estimated to be about 700 MHz.

Closed-loop Amplification/Hold

During the "hold" mode, speed is limited by the settling time of the op amp in its closed-loop configuration as shown below.

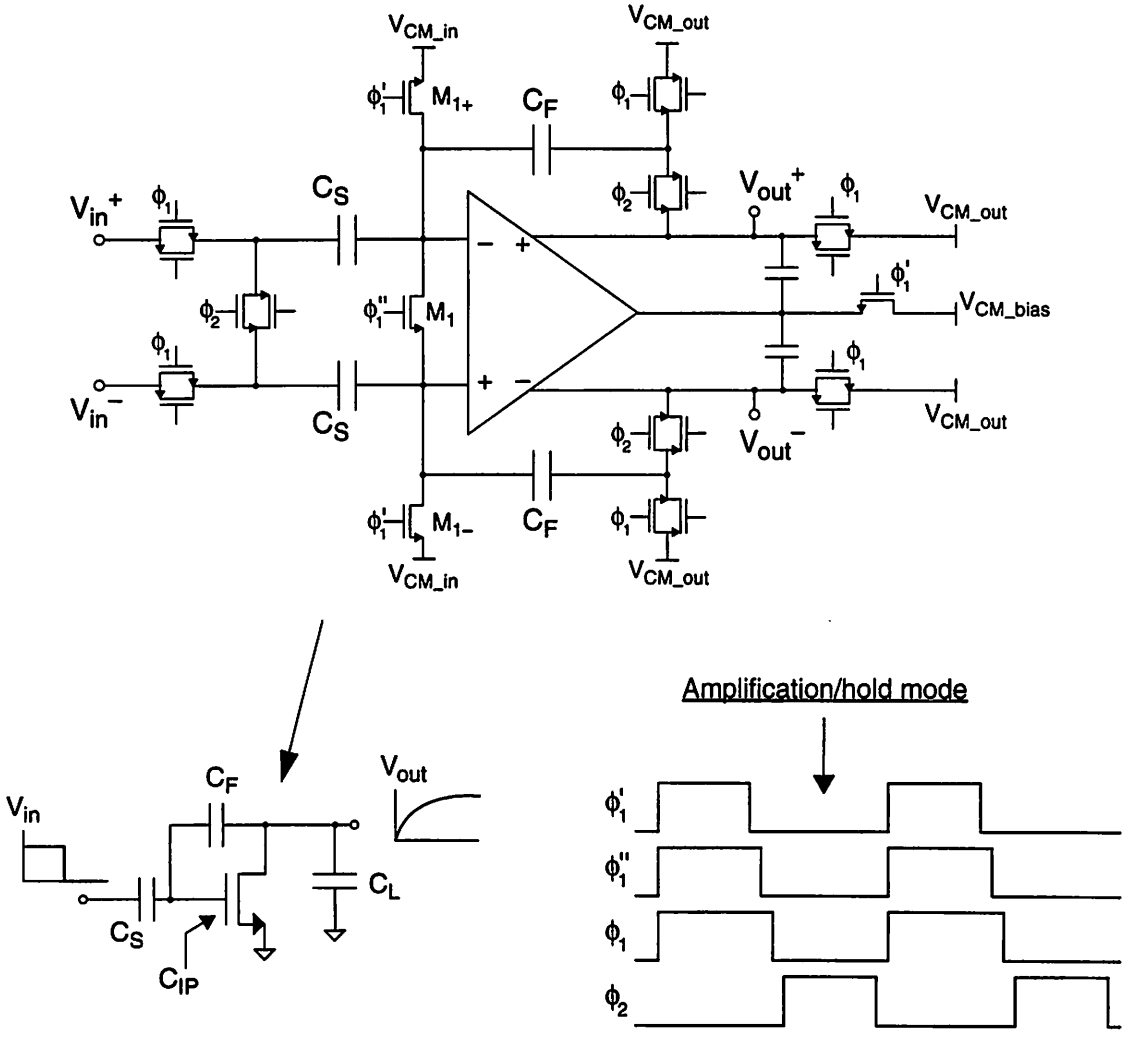


Fig. 7.24 Basic schematic of differential switched-capacitor stage showing equivalent circuit model for closed-loop amplification phase of operation.

From the above discussion, it is desirable to optimize the op amp settling time in the S/H "hold" phase. In order to achieve this objective, a strategy involving two key aspects is adopted: (i) a simple and hence intrinsically fast op amp topology is chosen, and (ii) systematic analysis and optimization is performed.

A fully-differential, mostly-NMOS, non-folded (or unfolded) cascode topology is used as shown in Fig. 7.25. (This is sometimes known as a *telescopic* amplifier [99].)

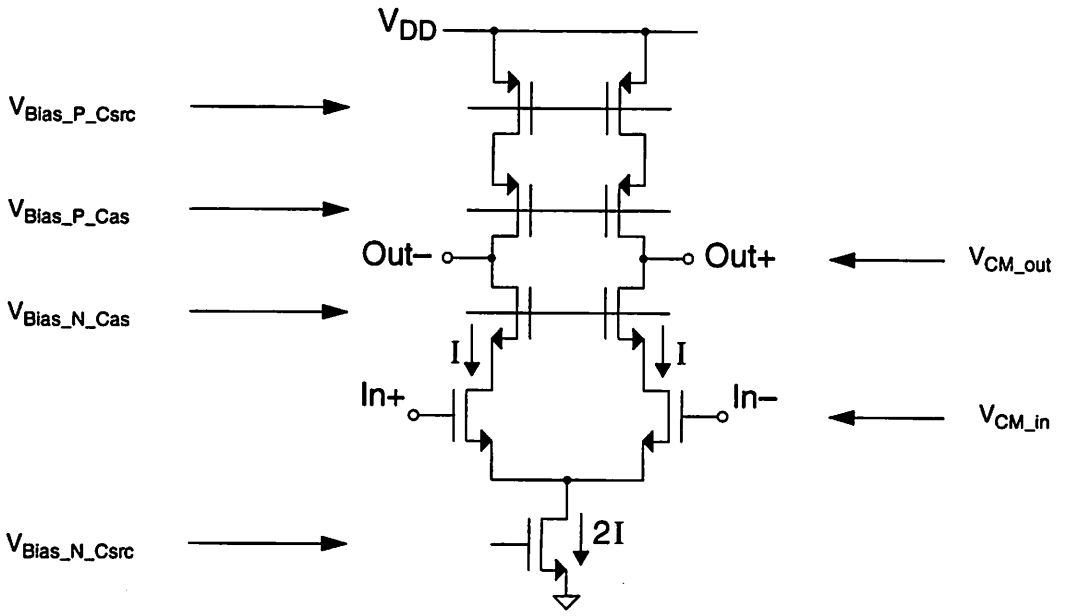


Fig. 7.25 Basic schematic of a fully-differential non-folded cascode op amp with bias voltages indicated.

The speed and simplicity advantages of having only NMOS devices in the signal path outweigh the loss of output swing relative to a folded cascode architecture. Note that there is no requirement for the common-mode levels at the input and output of the op amp to be equal. In practice, they are not equal and are chosen and optimized independently as follows: the common-mode level at the input V_{CM_in} is set up by a circuit that replicates the V_{DS} of the NMOS current source and the V_{GS} of the input devices; the common-mode level at the output V_{CM_out} is chosen to be in the center of the output swing of the amplifier so as to maximize the output swing. This is another advantage of the particular configuration used for the SC gain stage, which never requires the op amp input and output to be shorted together.

A number of bias level are required for this op amp topology as indicated on Fig. 7.25; these are defined below.

$V_{Bias\_P\_Csrc}$	PMOS current source bias level
$V_{Bias\_P\_Cas}$	PMOS cascode bias level
$V_{Bias\_N\_Cas}$	NMOS cascode bias level
$V_{Bias\_N\_Csrc}$	NMOS current source bias level
V_{CM_in}	Input common-mode level
V_{CM_out}	Output common-mode level

A biasing scheme similar to that reported in [66], [101] is used to maximize the amplifier swing. The open-loop dc gain obtainable from a single-stage op amp topology is limited. For n-bit accuracy, op amp gain greater than approximately 2^n is required. In the prototype chip described in Chapter 8, in order to ensure that the 1-μm channel-length devices operate in a sufficiently high-gain region of operation, the transistors are biased at a V_{DS} of 400–500 mV more than the classical "V_{DSAT}". DC Miller effect cancellation transistors are used in stage 2 (i.e., the first gain of 4 stage) [127]. In the prototype chip, voltage swing at the amplifier output is approximately 1.6 V single-ended with a single 5-V power supply. Clearly, this type of amplifier using many stacked devices becomes impractical at reduced power supplies. The S/H timing and input common-mode level reset are similar to the circuit described in [80]. The dynamic (i.e., capacitive) common-mode feedback (CMFB) loop at the amplifier output is through the NMOS tail current source. The bias level V_{CM_bias} in Fig. 7.23 is equal to the nominal V_{GS} of the current source devices, i.e., $V_{Bias\_N\_Csrc}$.

In order to optimize the op amp speed, the analysis and results of Chapter 5 are used; these are summarized briefly here as follows. The basic relation for the closed-loop –3 dB bandwidth ω_{CL} and corresponding time constant τ of the system shown in Fig. 7.24

incorporating the amplifier of Fig. 7.25 is given by

$$\omega_{CL} = \frac{1}{\tau} = f \cdot \frac{g_m}{C_{Ltot}} \tag{7.2}$$

where g_m is the transconductance of the op amp input devices, C_{Ltot} is the total capacitive loading at the output node including any loading due to the capacitive feedback network, g_m/C_{Ltot} is the unity-gain frequency of the closed-loop system, and the feedback factor (f) is given by

$$f = \frac{C_F}{C_F + C_S + C_{IP}} \tag{7.3}$$

where C_{IP} is the op amp input capacitance.

By fixing the sampling and feedback capacitances, and the current per unit width of the op amp transistors, the value of C_{IP} for minimum τ may be found. The analysis also includes the key fact that the output parasitic capacitance of the amplifier (C_{OP}) due to drain-bulk junction capacitances and overlap capacitance, etc., scales with the size of the op amp device sizes and hence with the input parasitic capacitance. The maximum bandwidth ω_{CLopt} is directly proportional to ω_T and depends only on *ratios* of the sampling, feedback, and extrinsic load capacitances — which are assumed to fixed. This approach was used as a basis for the choice of device sizes in the op amp.

The second pole frequency of the amplifier is approximately at ω_T of the NMOS cascode device. This frequency is much higher than the –3 dB frequency of the closed-loop system shown in Fig. 7.24 because of (a) the capacitive loading at the output node and (b) the effect of the feedback factor on bandwidth, and thus does not limit the speed, particularly in the gain of 4 stages.

It is useful to compare the equivalent circuits for the two modes of operation, that is, Fig. 7.23 and Fig. 7.24 and also to compare the corresponding equations, (7.1) and (7.2).

Note that the closed-loop time constant for the closed-loop amplifier is typically about 4–8 times longer than the RC time constant of the input sampling network: for swing and gain reasons, MOS transistors in op amps are usually biased at a $V_{GS} - V_T$ in the range 200–400 mV, whereas in MOS switches $V_{GS} - V_T$ in the range 1.0–1.5 V is common. This fact is one of the key motivations for the parallel pipeline architecture in CMOS: the throughput of a *single-channel* CMOS pipeline ADC is *always* limited by the op amp settling time in the S/H "hold" phase (assuming 50% duty-cycle clocks); by combining a number of pipeline ADC's in a parallel time-interleaved configuration it is possible to take advantage of the much higher input bandwidth of the passive input sampling network. Intuitively, in order to make maximum use of both the parallelism *and* the higher bandwidth of the input network, the number of parallel channels should approximately equal the ratio of the two time constants — i.e., the ratio of the amplifier closed-loop time constant to the RC time constant of the input sampling network.

7.3.5 Summary of Performance Limitations and Solutions

The performance limitations in parallel pipeline array ADC's along with the solutions proposed in this work are summarized in the table below.

Problem	Effects	Solutions
Channel offset mismatch	Tones at F_s/M, fixed pattern noise	Offset correction (digital, analog)
Channel gain mismatch	Modulation sidebands around F_s/M	Capacitor matching, shared DAC levels
Multiphase clock generation, jitter, skew	Spurious noise and tones	Delay locked loop
Op amp settling time	Limits throughput per channel	Optimized non-folded cascode op amp

Based on the above, the next chapter describes an experimental prototype that implements the parallel pipeline concept.

Chapter 8 Prototype Chip

8.0 CHIP IMPLEMENTATION DETAILS

A prototype 8-bit parallel pipeline array A/D converter was fabricated in a 1-μm double-metal CMOS technology with high-precision polycide capacitors [60]. The IC consists of a time-interleaved array of four switched-capacitor pipeline ADC's. Die photos are shown in Fig. 8.0 and Fig. 8.1. (Two photos are included since some features are clearer in one than the other.)

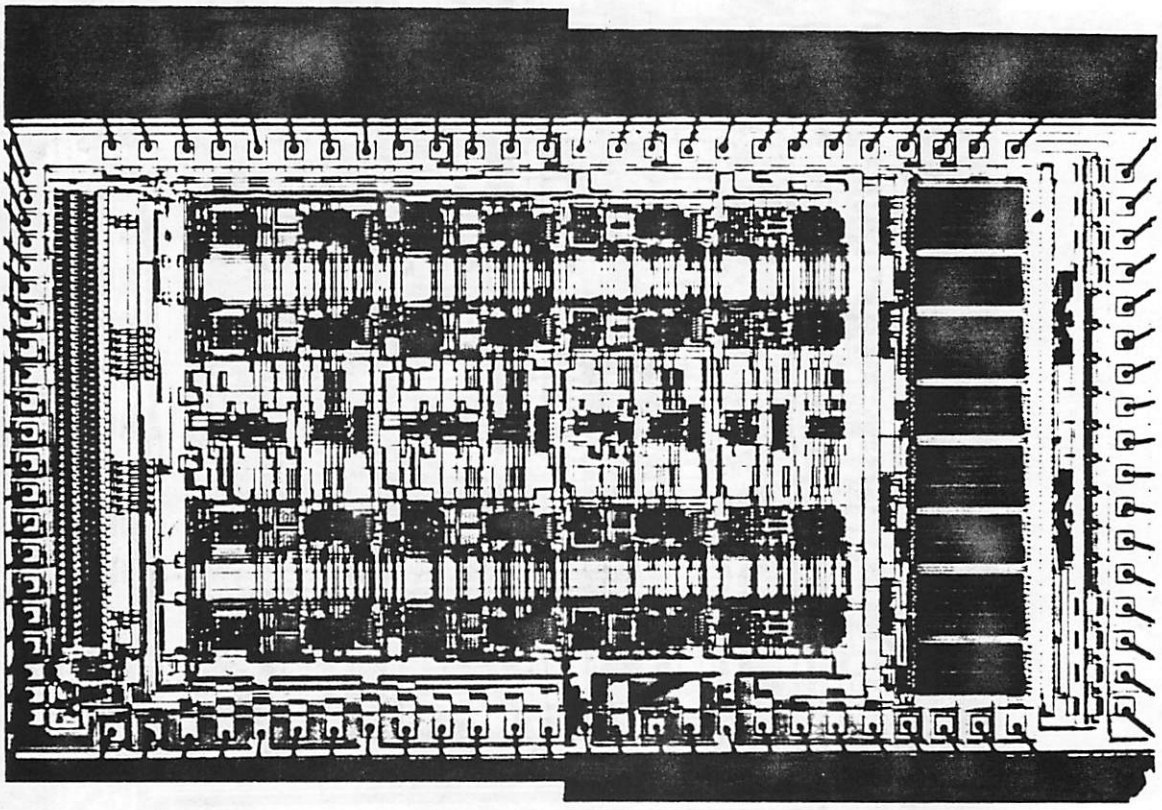


Fig. 8.0 Die photo of parallel pipeline array A/D converter chip.

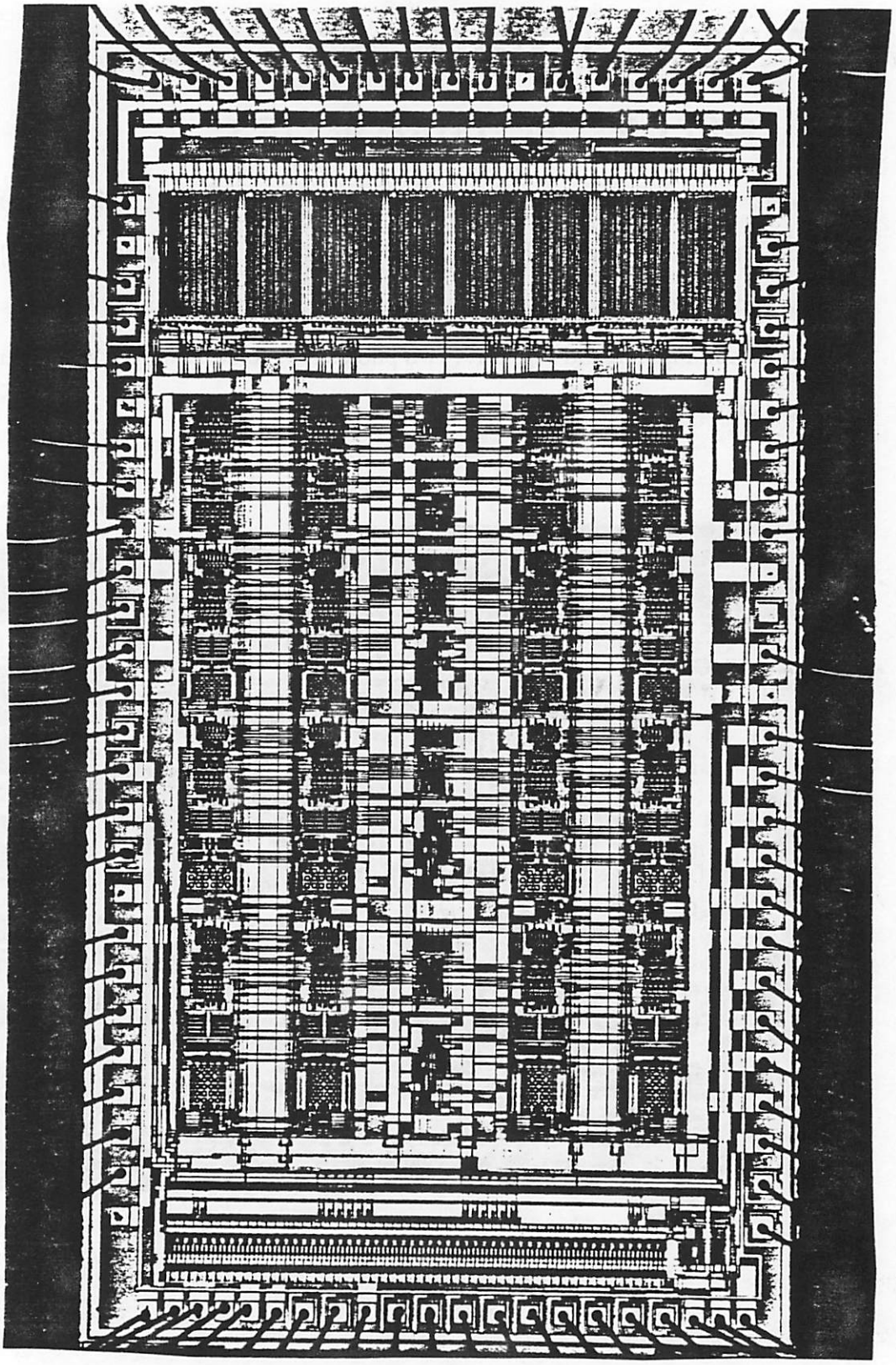


Fig. 8.1 Die photo of parallel pipeline A/D converter.

It is thought that the low frequency integral nonlinearity (INL) is due primarily to a layout error resulting in a large IR drop along the reference lines *within each channel* — and so not fundamental to the architecture.

A plot of SNDR versus input frequency is shown in Fig. 8.2 below.

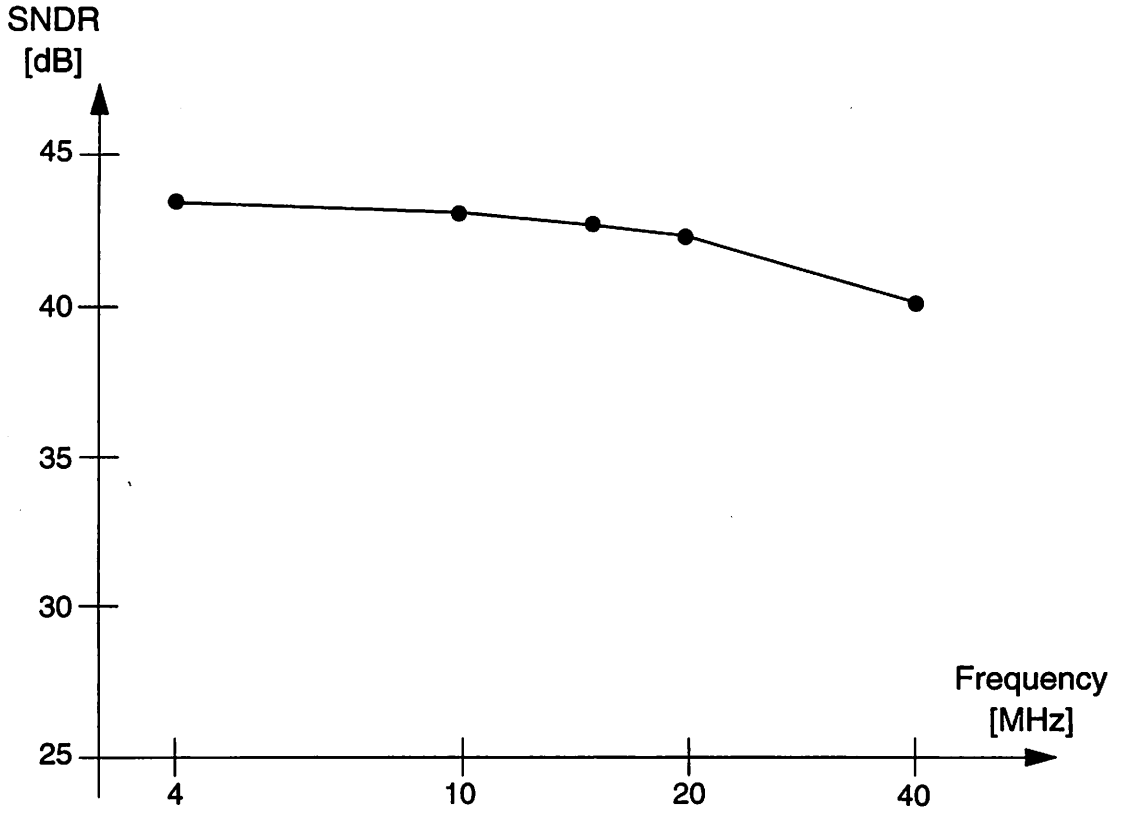


Fig. 8.2 Plot of SNDR versus analog input signal frequency.

There is a modest rolloff in SNDR as F_{in} varies from 4 MHz to 40 MHz — a few dB. An FFT of the output for a sampling rate of 85 MS/s and with a sinewave input at frequency 40 MHz is shown in Fig. 8.3 — this is somewhat like a beat frequency test.

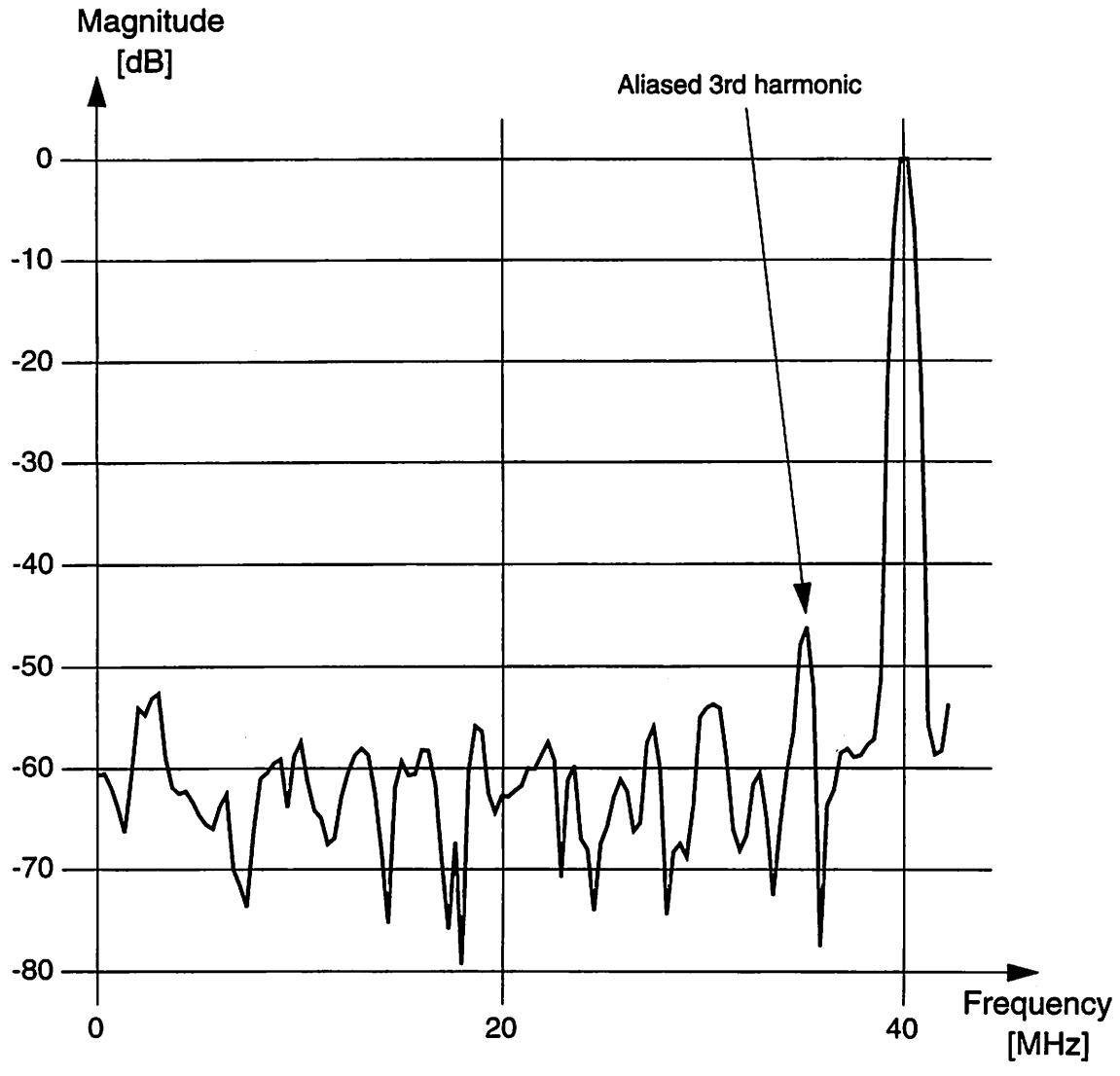


Fig. 8.3 Output spectrum (256-point FFT, Blackman-Harris window): F_s = 85 MS/s, F_{in} = 40 MHz.

The spectral line corresponding to the aliased 3rd harmonic is clearly visible and it is the dominant distortion component rather than mismatch-related tones. At high input frequency, SNDR is limited by nonuniform sampling due to systematic timing offsets from (a) the delay line used to generate the multiphase clocks and/or (b) the phase detector. The measured SNDR rolloff is consistent with about 25 ps of timing mismatch.

The total power dissipation of the resistor strings, op amps, comparators and associated bias circuits is 550 mW. The resistor strings consumed approximately 50 mW,

The following discussion is with reference to the die orientation of Fig. 8.0; in Fig. 8.1 the die appears rotated counter-clockwise. Along the left side of the die, extending almost along the entire side, is the voltage-controlled tapped delay line, which forms part of the delay-locked loop (DLL), and the associated clock generation circuitry. The clock lines are then routed horizontally. The four pipeline ADC's are arranged horizontally with the bias and reference block in the center between the second and third channels. Within each pipeline ADC, the four stages are visible. Progressing from left to right, each stage consists of a capacitor array, followed by an op amp, then a comparator bank and digital encoder, and finally the DAC switches. The symmetry of the two halves of the amplifiers can be clearly seen. It can also be seen that the amplifiers and capacitor arrays decrease in size going from the most significant stage to the least significant stage, i.e., horizontally from left to right on the die. The bias and reference blocks in the center contain bias circuitry and a resistor string for all parallel stages, i.e., there is one resistor string for all four first stages, one resistor string for all four second stages, etc. This is easier than using a single extremely low impedance resistor string, which would entail significant routing complexity. The key point is that sub-DAC mismatches *across the channels* are eliminated. Wide power and ground busses can be seen between the pipeline channels.

To aid circuit testability for this prototype, there is a 256-word on-chip shift-register buffer memory (on the right side of the die). This avoids the problem of high-speed digital outputs switching during conversion. During normal conversion, the ADC outputs go directly into the memory and the output buffers are disabled and so do not generate any Ldi/dt noise on the power supplies while sampling of the analog input is occurring. Then, when the memory is filled up, the data can be shifted out at a much slower frequency. The limited record length causes some loss in frequency resolution in the Discrete Fourier Transform computation, but is adequate. On the extreme right of the die the first-in-first-out shift-register array is visible. The package used for the prototype chip is a 100-pin ceramic Leadless Chip Carrier (LCC), of which 90 pins are actually used. The

total number of power supply connections, including various analog and digital VDD and GND connections, is 49.

The analog signal path is differential throughout the chip. As mentioned already, each channel contains four stages, with 3 bits per stage; one bit is used for digital error correction in each of the first three stages; interstage gain is 4. The nominal resolution of each pipeline ADC is therefore: $2 + 2 + 2 + 3 = 9$ bits. However, one bit is used for digital offset cancellation thus giving a net 8-bit overall system resolution. Digital offset correction is implemented by simply measuring the output code from each channel when the differential input is zero, and then subtracting this from each channel's output during normal operation. It is assumed that a suitable reset/autozero time is available for this offset measurement to be performed. As mentioned earlier, offset measurement *is not performed* "on-the-fly". The prototype did include the circuitry for analog autozero (AZ) on the first two stages of each channel — however, due to an implementation error this was not functional.

The front-end S/H in each channel has a gain of 1. Note that there is no relationship implied between the number of parallel channels and the number of stages — the fact that there are four ADC's in parallel and four stages within each pipeline is purely coincidental. There are four resistor strings (one for each pipeline stage), and each one is shared across all four channels. In order to ensure that the DAC settling does not limit the converter speed, the total resistance of each string is 320 Ω. This gives a worst case time constant, assuming a load of about 2 pF, of $(CR_{total})/2 = 0.3\,\text{ns}$. The reference voltages applied to the resistor strings are $V_{ref}^{+} = 3.4$ V and $V_{ref}^{-} = 1.8$ V. The resistor string DAC's are implemented differentially as in [73] allowing for a differential signal swing of ±1.6 V. The resistor strings and bias circuits for the op amps and comparators are located between the 2nd and 3rd channels on the die, and the reference and bias lines are routed vertically.

The input capacitance of the converter is approximately 2 pF single-ended, excluding pad capacitance. This corresponds to the input being connected to two channels at any one

time — the input capacitance of a single channel is the sampling capacitance of the front-end S/H, about 1 pF. Note, of course, that this is a switched capacitance.

As mentioned in Chapter 7, since accuracy requirements decrease going down the pipeline, it is possible to reduce sampling capacitor sizes and op amp device currents and widths. This was done in the prototype chip and the values used are tabulated below. (The transistor sizes are given in the form width/length, as drawn in microns.)

Stage	C_S	C_F	Current in each leg (I)	NMOS input	NMOS cascode	PMOS cascode	PMOS current source
1	1.0 pF	1.0 pF	1.0 mA	1024/1	512/1	512/1	2304/3
2	1.5 pF	0.375 pF	1.0 mA	1024/1	512/1	512/1	2304/3
3	0.52 pF	0.13 pF	1.0 mA	384/1	192/1	384/1	1536/3
4	0.52 pF	0.13 pF	0.3 mA	256/1	64/1	128/1	512/3

Note that the term "stage" here is used as defined in Section 7.2, i.e., each stage consists of a SC gain stage followed by a sub-ADC and sub-DAC. Thus, stage 1 in the above table is the input S/H, which has a gain of 1; the other stages have a gain of 4.

8.1 EXPERIMENTAL RESULTS

For generation of the analog input, a low phase noise, RF synthesized signal generator was used (HP8656B). The output from this signal generator was passed through a bandpass filter to eliminate harmonic distortion components. Passive L-C filters were used; specifically, 4-pole Chebyshev with a Q of 33 (manufacturer: TTE). The filter output was then passed through a wideband transformer — also known as a "2 way-180° power splitter" (manufacturer: Mini-Circuits). All of the above was in a 50-Ω impedance environment using standard coaxial cables, BNC connectors, etc. Note that for each distinct frequency of the analog input at which a Signal-to-Noise-plus-Distortion Ratio (SNDR) measurement was made, a different bandpass filter was used. The alternative of performing a wideband single-ended-to-differential conversion using op amps was

considered to be impractical from the point of view of generating low-distortion signals at the frequencies of interest (1–100 MHz). For clock generation, a HP8130A pulse generator was used *triggered externally* by another low phase noise RF sinewave signal source. This external triggering was found to have a noticeable effect on SNR especially at input frequencies of 40 MHz and 80 MHz. All the following measurements were taken from a test set-up consisting of a custom PCB incorporating analog and digital power and ground planes and separate routing layers for analog signals and digital signals. However, note that in these experiments, the IC was not fully surface-mounted — a socket was used. Digital data from the analog board was captured by a special-purpose digital data acquisition board [38] and transferred to a workstation for further processing. For calculation of the SNDR, software based on the technique of the Minimum Sinusoidal Error method was used [85]. Essentially this is a more sophisticated type of "best-fit sinewave" algorithm [10]. A histogram test [29] was used for obtaining INL and DNL.

Key chip parameters and results obtained at 5 V and 25°C are summarized below.

Sampling rate	85 MS/s
Active ADC circuit area	6.3 mm × 3.9 mm (= 25 mm^2) 250 mil × 150 mil (= 38 $kmil^2$)
Chip area including pads and on-chip shift-register/memory	8.6 mm × 4.7 mm (= mm^2) 340 mil × 180 mil (= $kmil^2$)
Power dissipation	1100 mW
Input signal range	3.2 V p-p differential
Input capacitance	2 pF
Dynamic range	49 dB
INL	1.0 LSB at 8 bits (typical)
DNL	0.8 LSB at 8 bits (typical)
Input-referred RMS noise	< 1/2 LSB at 8 bits
Magnitude of $F_S/4$ tone with zero input	< 1/2 LSB at 8 bits
Inter-channel gain mismatch	< 1/2 LSB at 8 bits

and the remaining power was divided roughly equally between the op amps and comparators. The static power dissipation in the comparators could be significantly reduced by fixing a level shifting problem. The delay line, associated level-shifters, and associated clock generation circuitry consumed approximately 550 mW, i.e., half the total power; thus, the intrinsic 8-bit 85-MS/s A/D converter consumed approximately 550 mW. This DLL power could be reduced by half by careful resizing of the transistors in the delay cell. It is important to note however, that the relatively high chip power in the implementation presented here is a function of the particular way the multiphase clocks were generated and not a fundamental aspect of the architecture.

In the prototype chip, one of the main reasons a DLL was used was to avoid the necessity of having an external clock at the full sampling frequency (F_s) in order to ease testing problems and the I/O problems of dealing with a very high-speed clock, etc. In general, for a parallel pipeline ADC with many channels in parallel and implemented in submicron technologies, this approach of using the channel sampling frequency as an external reference clock and using a DLL to perform clock multiplication on chip in order to generate the sampling edges is potentially an advantage. Moreover, this holds for any system doing oversampling, and is particularly appropriate for parallel analog systems. For many applications, however, a full F_s sampling clock is readily available, and in that case, it makes sense to implement the clock generation digitally with negligible static power.

Chapter 9 Summary and Conclusions

9.0 SUMMARY

This dissertation has described a new A/D converter architecture in CMOS — *the parallel pipeline array.* As background and context along the way, many associated topics have been examined in detail. These include the following:

- applications of and trends in high-speed, medium-resolution monolithic ADC's;
- many aspects of pipelined multistage ADC's, such as static transfer characteristics, digital error correction strategies, and error sources and nonidealities in switched-capacitor (SC) implementations;
- settling time of SC gain stages;
- effects in the frequency domain of mismatches in offset, gain, and timing in parallel time-interleaved analog systems, and the statistics of these effects;
- uses of parallelism in A/D converter architectures.

The prototype chip described in this dissertation is the fastest 8-bit CMOS ADC reported to date. The performance has been achieved using a new architecture consisting of a parallel time-interleaved array of SC pipeline ADC's. Key resources such as resistor strings, bias circuitry, and clock generation circuitry have been shared over the array in order to minimize hardware. Techniques have been employed to minimize the effects of offset, gain, and timing mismatches between the parallel channels. Within each channel, throughput has been maximized by careful sizing of op amp transistor widths in order to optimize S/H and interstage residue amplifier settling time.

9.1 CONCLUSIONS

The following lists the conclusions and insights obtained from this work. These conclusions are grouped under the chapter to which they most closely correspond, although some involve multiple considerations from a number of chapters.

Chapter 2:

- From an examination of recently-reported ADC's, sample rates of $f_T/100$ at 8-bit resolution are commonly attained. To achieve speeds above $f_T/50$ is extremely hard.

Chapter 3:

- Examination and comparison of a number of ADC architectures shows that many variations and special cases of a generalized SC gain stage occur: this is a key signal processing block.
- Sharing the feedback capacitor is highly advantageous in low closed-loop gain SC gain stages — both for speed and accuracy reasons.

Chapter 4:

- A generalized approach for analyzing multistep A/D conversion allows many algorithms and architectures to be viewed within a unified framework, and allows issues such as analysis of certain component errors and operation of non-binary radix algorithms to be considered in a systematic way.
- This generalized approach also leads naturally to consideration of A/D conversion as a decomposition into "weights" that can be digitally calibrated.

Chapter 5:

- Analysis of the SC gain stage yields a design approach for obtaining optimum — i.e., technology-limited — settling time performance. The approach considers the entire closed-loop system including the feedback factor, and also explicitly includes the scaling of the output parasitic capacitance with amplifier size.

Chapter 6:

- Analytical approaches for analyzing the effects of offset, gain, and timing mismatches in time-interleaved parallel analog channels from various digital signal processing viewpoints provide qualitative and quantitative insight into (a) the effect of these mismatches on the output spectrum and (b) the circuit precision requirements on the analog components. Expressions have been derived describing the behavior in the frequency domain caused by inter-channel mismatches, both for the deterministic case of known channel parameters, and also for the case of parameters with some statistical distributions.

Chapters 7 and 8:

- The parallel structure inherent in many two-step subranging architectures was identified and discussed and shown to be fundamental to those architectures.
- The parallel pipeline A/D converter architecture in CMOS allows very high throughput to be obtained. Measured results presented in this dissertation have demonstrated the fastest 8-bit CMOS A/D converter reported to date.
- The performance obtained corresponds approximately to a sample rate of $f_T/70$ (assuming an f_T of 6 GHz for a 1-μm CMOS device).
- One of the key aspects of the architecture is that it takes advantage of the high sample-mode bandwidth and consequent fast acquisition possible with a CMOS switched-capacitor S/H network. In general, for a SC gain stage, speed of *acquisition* is limited by a passive R-C delay, with a time constant of the form

$$\tau \sim RC$$

whereas speed of *amplification* — i.e., during the "hold" mode — is determined by the closed-loop settling of an op amp in a closed-loop feedback configuration, with a time constant of the form

$$\tau \sim \frac{1}{f} \cdot \frac{C}{g_m}$$

Thus, the key fact that allows such a high sample rate to be achieved is that the input bandwidth of a MOS switched-capacitor S/H circuit in acquisition mode is much higher than the closed-loop bandwidth of the S/H in "hold" mode. The use of parallelism means that the op amp closed-loop settling is no longer necessarily a bottleneck to throughput.

- This architecture is very suitable for moderate resolution levels (~7–8 bits) where mismatches can be dealt with. For higher resolutions, more calibration is necessary, which of course entails increased complexity.
- Clock generation in parallel A/D converter systems — and also, more generally, in any signal processing system that requires the generation of multiphase clocks with low-jitter sampling edges — is difficult, involves significant complexity, and often requires substantial power and area. More parallelism in the analog domain implies an increased number of critical sampling edges and increased clock generation complexity — this tends to have higher overhead than for purely digital systems.
- In some cases, there is a fundamental tradeoff between the system-level advantages of sharing resources for the reasons of minimizing effects due to mismatches, and the more circuit-level disadvantages associated with sharing since it tends to lead to interference and coupling.
- To zeroth order, sharing resources saves hardware; however, a complication associated with resource sharing across parallel analog signal paths is that any analog resource that is shared over M parallel signal paths is required to be fast enough to operate (i.e., settle) at a speed commensurate with the overall composite speed of the M-channel system, thus requiring correspondingly lower impedances and higher dc power. Furthermore, there is always some overhead and complexity associated with sharing.

- However, the converse to the above also holds: the most appropriate circuit block or resource to share is a block that is already faster than it is required to be, i.e., whose intrinsic speed is not being utilized or exploited to the maximum. Specifically, this is true for the sample-mode bandwidth of the input network. Alternatively stated: in parallel analog signal processing, if a certain resource or block is shared across a number of channels, the performance requirements of that block usually increase, with a consequent increase in power and area.

- An appropriate resource to be shared is one that affects the inter-channel mismatch — otherwise it may not be worth the apparent power or area saving. Therefore, for the parallel pipeline A/D converter architecture, it makes sense to share DAC levels rather than ADC levels — since sub-ADC levels are allowed to have relatively large error assuming digital correction is employed.

- Regarding an upper limit on the achievable speed: the question arises, how many channels are possible? A reasonable number is in the range 4–8; above that, the input bandwidth will limit performance. This number depends on the fact that the channel speed is determined by an op amp setting time, and the ratio of the closed-loop time constant of the S/H or interstage amplifier to the R-C time constant of the input network is typically in the range 4–8.

Overall Themes and Contributions:

In summary, the two main themes and overall contributions of this dissertation are (i) ways to look at and think about pipelined multistage A/D converters, especially CMOS integrated circuit implementations, and (ii) ways to look at and think about parallelism and time-interleaving in integrated circuit A/D converter architectures.

References

[Comments — noting, for example, interesting aspects about the content, the affiliation of the authors, etc., — are appended to some of the following references.]

[1] M. Azuma et al., "An FPN-free 2/3" 1.3M-Pixel CCD image sensor for HDTV camera system," in *ISSCC Dig. Tech. Papers*, Feb. 1991, pp. 212–213, 316. [Matsushita; FPN ⇒ fixed-pattern noise; interesting discussion of HDTV camera specifications — 2/3" means 9.6 mm horizontal X 5.4 mm vertical, a standard optical format, 70 dB dynamic range — 48.6 MHz transfer frequency.]

[2] H. Akimoto et al., "A 1/3-in 410000-Pixel CCD image sensor for HDTV camera system," *IEEE J. Solid-State Circuits*, vol. 26, no. 12, Dec. 1991, pp. 1907–1914. [Hitachi; 70 dB dynamic range; high-resolution video camera applications — operates at around 50 MHz.]

[3] Analog Devices, Data Sheets for AD872 (12 bits, 10 MS/s, ±5 V, 1.15 W), AD875 (10 bits, 15 MS/s, +5 V, 185 mW, no S/H — optimized for CCD input), AD9040 (10 bits, 32 MS/s, ±5 V, 940 mW), 1993; Preliminary Data for AD9022/9023 (12 bits, 20 MS/s, +5 V/–5.2V, 1.3 W).

[4] AT&T, Advance Data Sheet for CSPA1220, June 1993. [12-bit version of [75]].

[5] C. Basile, A. P. Cavallerano, and D. Teichner, "Baseband video processing and the transmission of HDTV signals," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 1, Mar. 1991, pp. 113–124. [Video applications; Philips Briarcliff Manor; discussion of basic HDTV parameters.]

[6] R. A. Bedford, "Industrial applications of HDTV," *IEEE Commun. Mag.*, vol. 29, no. 8, Aug. 1991, pp. 25–27. [HDTV applications in imaging including printing, cinema, computer applications, medicine, art museums; HDTV CCD cameras.]

[7] S. Bhattacharyya et al., *The Almagest: Shortened Manual for Ptolemy (0.4)*, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, 1992.

[8] W. C. Black Jr. and D. A. Hodges, "Time interleaved converter arrays," *IEEE J. Solid-State Circuits*, vol. SC-15, no. 6, Dec. 1980, pp. 1022–1029. [Time-interleaved A/D systems; mismatch analysis.]

[9] R. A. Blauschild, "An 8b 50ns monolithic A/D converter with internal S/H," in *ISSCC Dig. Tech. Papers*, Feb. 1983, pp. 178–179. [Ripple architecture; bipolar.]

[10] B. E. Boser, K.-P. Karmann, H. Martin, and B. A. Wooley, "Simulating and testing oversampled analog-to-digital converters," *IEEE Trans. Computer-Aided Design*, vol. 7, no. 6, June 1988, pp. 668–674.

[11] A. B. Carlson, *Communications Systems: An Introduction to Signals and Noise in Electrical Communication,* 3rd ed.,New York: McGraw-Hill, 1986, pp. 295–307.

[12] W. J. Chimiak, "The digital radiology environment," *IEEE J. Select. Areas Commun.*, vol. 10, no. 7, Sep. 1992, pp. 1133–1144.

[13] T. B. Cho, private communication.

[14] T. B. Cho, unpublished work.

[15] P. S. Chow, J. C. Tu, and J. M. Cioffi, "A multichannel transceiver system for asymmetric digital subscriber line service," in *IEEE Global Telecommunications Conference, Globecom '91,* Dec. 1991, pp. 56.5.1–56.5.5.

[16] Cirrus Logic, Advance Product Bulletin for CL-SH3300, Feb. 1993. [72 MHz.]

[17] D. W. Cline, private communication.

[18] D. W. Cline, unpublished work.

[19] J. D. Coker, R. L. Galbraith, G. J. Kerwin, J. W. Rae, and P. A. Ziperovich, "Implementation of PRML in a rigid disk drive," *IEEE Trans. Magnetics*, vol. 27, no. 6, Nov. 1991, pp. 4538–4543. [More general discussion of disk drive channel using chip presented at ISSCC in 1991.]

[20] W. T. Colleran, T. H. Phan, and A. A. Abidi, "A 10b, 100Ms/s pipelined A/D converter," in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 68–69, 265.

[21] C. S. G. Conroy, D. W. Cline, and P. R. Gray, "A high-speed parallel pipelined ADC technique in CMOS," in *1992 Symp. on VLSI Circuits Dig. Tech. Papers*, June 1992, pp. 96–97.

[22] C. S. G. Conroy, D. W. Cline, and P. R. Gray, "An 8-bit 85-MS/s parallel pipeline A/D converter in 1-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, April 1993, pp. 447–454.

[23] J. J. Corcoran, K. L. Knudsen, D. R. Hiller, and P. W. Clark, "A 400MHz 6b ADC," in *ISSCC Dig. Tech. Papers*, Feb. 1984, pp. 294–295, 357. [HP; uses analog encoding (folding); extensions of this architecture used in individual channels of the time-interleaved systems described in [109] and [114].]

[24] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Englewood Cliffs, New Jersey: Prentice-Hall, 1983, Section 2.3.

[25] Crystal Semiconductor Corporation, Data Sheets for CS5016, CS5014, CS5012A, Mar. 1992. [16, 14, and 12-bit, self-calibrating CMOS ADC's.]

[26] M. J. Demler, *High-Speed Analog-to-Digital Conversion*, San Diego: Academic Press, 1991, pp. 150–152.

[27] R. Dewell, "Electronics in medical imaging," *Electronics World + Wireless World*, vol. 96, no. 1651, May 1990, pp. 440–444.

[28] A. G. F. Dingwall and V. Zazzu, "An 8-MHz CMOS subranging 8-bit A/D converter," *IEEE J. Solid-State Circuits*, vol. SC-20, no. 6, Dec. 1985, pp. 1138–1143. [Gives typical measured performance at 5 MS/s.]

[29] J. Doernberg, H.-S. Lee, and D. A. Hodges, "Full-speed testing of A/D converters," *IEEE J. Solid-State Circuits*, vol. SC-19, no. 6, Dec. 1984, pp. 820–827.

[30] N. Fukushima et al., "A CMOS 8-bit 20MHz two-step parallel A/D converter with 95mW power consumption," in *International Conf. on Consumer Electron. Dig. Tech. Papers*, June 1988, pp. 238–239. [Sony.]

[31] N. Fukushima et al., "A CMOS 40MHz 8b 105mW two-step ADC," in *ISSCC Dig. Tech. Papers*, Feb. 1989, pp. 14–15, 273. [Sony; important paper.]

[32] J. Gallant, "ICs and modules for digital wireless telecommunications," *EDN*, vol. 38, no. 17, Aug. 19, 1993, pp. 77–90. [Review of standards and IC products.]

[33] B. Ginetti, A. Vandemeulebroecke, and P. Jespers, "RSD cyclic analog-to-digital converter," in *1988 Symp. on VLSI Circuits Dig. Tech. Papers*, Aug. 1988, pp. 125–126. [Redundant Signed Digit scheme.]

[34] B. Ginetti, P. Jespers, and A. Vandemeulebroecke, "A CMOS 13 bits cyclic RSD A/D converter," in *Proc. ESSCIRC'91, Seventeenth European Solid State Circuits Conf.*, Sep. 1991, pp. 245–248.

[35] B. Ginetti, P. G. A. Jespers, and A. Vandemeulebroecke, "A CMOS 13-b cyclic RSD A/D converter," *IEEE J. Solid-State Circuits*, vol. 27, no. 7, July 1992, pp. 957–965.

[36] L. Gomes, "Government gears up for 'data superhighway'," *San Jose Mercury News*, March 7, 1993, p. 24A.

[37] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed., New York: Wiley, 1984, pp. 434–440.

[38] J. D. Greco, "A performance evaluation system for analog to digital converters: research project," MS Report, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Jan. 1991.

[39] R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, New York: Wiley, 1986, pp. 505–510.

[40] M. Haas, D. Draxelmayr, F. Kuttner, and B. Zojer, "A monolithic triple 8-bit CMOS video coder," *IEEE Trans. Consumer Electron.*, vol. 36, no. 3, Aug. 1990, pp. 722–729. [Siemens; three 8-bit, 27-MS/s ADC's with some digital processing.]

[41] D. N. Helman, "A Multi-phase clock generator/parallel-phase sampler in 1 micron CMOS: research project," MS Report, Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, May 1990.

[42] R. M. Herman, C. H. Mason, H. P. Warren, and R. A. Meier, "A GPS receiver with synthesized local oscillator," in *ISSCC Dig. Tech. Papers*, Feb. 1989, pp. 194–195, 339. [Radio applications; GPS L-band — output IF near 60 MHz.]

[43] O. A. Horna, "A 150 Mbps A/D and D/A conversion system," *Comsat Technical Review*, vol. 2, no. 1, Spring 1972, pp. 39–72.

[44] M. Horowitz, "Clocking strategies in high performance processors," in *1992 Symp. on VLSI Circuits Dig. Tech. Papers*, June 1992, pp. 50–53. [Clock generation.]

[45] S. Hosotani, T. Miki, A. Maeda, and N. Yazawa, "An 8-bit 20 MS/s CMOS A/D converter with 50 mW power consumption," in *1989 Symp. on VLSI Circuits Dig. Tech. Papers*, May 1989, pp. 55–56. [Mitsubishi.]

[46] S. Hosotani, T. Miki, A. Maeda, and N. Yazawa, "An 8-bit 20-MS/s CMOS A/D converter with 50-mW power consumption," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, Feb. 1990, pp. 167–172. [Mitsubishi.]

[47] M. Hotta, T. Shimizu, K. Maio, K. Nakazato, and S. Ueda, "A 12mW 6b video frequency ADC," in *ISSCC Dig. Tech. Papers*, Feb. 1987, pp. 100–101, 357. [Hitachi; full flash; bipolar process with 8-GHz npn's and 860-MHz pnp's.]

[48] M. Hotta, T. Shimizu, K. Maio, K. Nakazato, and S. Ueda, "A 12-mW 6-bit video-frequency A/D converter," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 939–943. [Hitachi.]

[49] K.-C. Hsieh, "Noise limitations in switched-capacitor filters," Memorandum No. UCB/ERL M82/39, Electronics Research Laboratory, U. C. Berkeley, May 1982.

[50] P. J. Hurst, "Shifting the frequency response of switched-capacitor filters by nonuniform sampling," *IEEE Trans. Circuits Syst.*, vol. 38, no. 1, Jan. 1991, pp. 12–19. [Nonuniform sampling application.]

[51] T. L. Hylton, K. R. Coffey, M. A. Parker, and J. K. Howard, "Giant magnetoresistance at low fields in discontinuous NiFe-Ag multilayer thin films," *Science*, vol. 261, no. 5124, 20 Aug. 1993, pp. 1021–1024. [IBM ADSTAR research results with potential for densities of 10 Gbits/in^2.]

[52] M. Ishikawa and T. Tsukahara, "An 8-bit 50-MHz CMOS subranging A/D converter with pipelined wide-band S/H," *IEEE J. Solid-State Circuits*, vol. 24, no. 6, Dec. 1989, pp. 1485–1491. [NTT; 2-stage subranging; 1-μm CMOS.]

[53] Y.-C. Jenq, "Digital spectra of nonuniformly sampled signals: fundamentals and high-speed waveform digitizers," *IEEE Trans. Instrum. Meas.*, vol. 37, no. 2, June 1988, pp. 245–251. [Mismatch analysis; nonuniform sampling analysis.]

[54] R. Jewett, J. Corcoran, and G. Steinbach, "A 12b, 20MS/s ripple-through ADC", in *ISSCC Dig. Tech. Papers*, Feb. 1992, pp. 34–35, 235. [HP; interstage gain = 2; 3.5 W from +5 V, –5.2 V supplies.]

[55] M. Johnson and E. L. Hudson, "A variable delay line PLL for CPU-coprocessor synchronization," *IEEE J. Solid-State Circuits*, vol. 23, no. 5, Oct. 1988, pp. 1218–1223. [Clock generation.]

[56] G. Jusuf and P. R. Gray, "A 1-bit/cycle algorithmic analog-to-digital converter without high-precision comparators," Memorandum No. UCB/ERL M90/69, Electronics Research Laboratory, U. C. Berkeley, Aug. 1990.

[57] A. N. Karanicolas, H.-S. Lee, and K. L. Bacrania, "A 15b 1Ms/s digitally self-calibrated pipeline ADC", in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 60–61, 263.

[58] M. Kasahara et al., "A CMOS 9 bit 25MHz 100mW ADC for mixed analog/digital LSIs," in *Proc. Custom Integrated Circuits Conf.*, May 1991, pp. 26.7.1–26.7.4. [Hitachi; 2 sets of comparators.]

[59] S. Kauffman, "Multistage error correcting A/D converters", in *Wescon/82 Conference Record*, Sep. 1982, pp. 30-1/1–8. [Discusses some issues relating to the digitally-corrected subranging architectures used in the Analog Devices Computer Labs Division 10-bit, 20-MS/s and 12-bit, 5-MS/s products.]

[60] C. Kaya et al., "Polycide/metal capacitors for high precision A/D converters," in *IEDM Tech. Dig.*, 1988, pp. 782–785. [Discussion/comparison of different capacitor technologies for analog applications.]

[61] D. A. Kerth, N. S. Sooch, and E. J. Swanson, "A 667ns, 12-bit, two-step flash ADC," in *Proc. Custom Integrated Circuits Conf.*, May 1988, pp. 18.5.1–18.5.4. [Crystal; two-step recycling architecture with self-calibration; 3-μm CMOS; supplies: ±5 V; power: 700mW; full product details in Data Sheet for CS5412.]

[62] B. Kim, D. N. Helman, and P. R. Gray, "A 30-MHz hybrid analog/digital clock recovery circuit in 2-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1385–1394. [Clock generation.]

[63] H. Kimura, A. Matsuzawa, T. Nakamura, and S. Sawada, "A 10b 300MHz interpolated-parallel A/D converter," in *1992 Symp. on VLSI Circuits Dig. Tech. Papers*, June 1992, pp. 94–95. [Matsushita; 25-GHz Si bipolar process; die area: 9.0 mm X 4.2 mm; 4 W from –5.2 V.]

[64] H. Kimura, A. Matsuzawa, T. Nakamura, and S. Sawada, "A 10-b 300-MHz interpolated-parallel A/D converter," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, Apr. 1993, pp. 438–446. [Expanded version of above paper.]

[65] K. Kusumoto et al. "A 10b 20MHz 30mW pipelined interpolating CMOS ADC", in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 62–63, 264. [Matsushita.]

[66] C. A. Laber et al., "Design considerations for a high performance 3-μm CMOS analog standard cell library," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 2, Apr. 1987, pp. 181–189. [CMOS analog design; A/D integration.]

[67] C. Lane, "A 10-bit 60 MSPS flash ADC," in *Proc. 1989 Bipolar Circuits and Technology Meeting*, Sep. 1989, pp. 44–47. [AD; uses interpolation.]

[68] E. A. Lee and D. G. Messerschmitt, *Digital Communication*, Boston: Kluwer Academic Publishers, 1988, Chapter 6.

[69] S.-H. Lee and B.-S. Song, "A code-error calibrated two-step A/D converter," in *ISSCC Dig. Tech. Papers*, Feb. 1992, pp. 38–39, 237.

[70] S.-H. Lee and B.-S. Song, "Digital-domain calibration of multistep analog-to-digital converters," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, Dec. 1992, pp. 1679–1688.

[71] S. H. Lewis and P. R. Gray, "A pipelined 5MHz 9b ADC," in *ISSCC Dig. Tech. Papers*, Feb. 1987, pp. 210–211.

[72] S. H. Lewis, "Video-rate analog-to-digital conversion using pipelined architectures," Memorandum No. UCB/ERL M87/90, Electronics Research Laboratory, U. C. Berkeley, Nov. 1987.

[73] S. H. Lewis and P. R. Gray, "A pipelined 5-Msamples/s 9-bit analog-to-digital converter," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 954–961. [CMOS pipeline ADC's; digital error correction.]

[74] S. H. Lewis, H. S. Fetterman, G. F. Gross, Jr., R. Ramachandran, and T. R. Viswanathan, "A pipelined 9-stage video-rate analog-to-digital converter," in *Proc. Custom Integrated Circuits Conf.*, May 1991, pp. 26.4.1–26.4.4

[75] S. H. Lewis, H. S. Fetterman, G. F. Gross, Jr., R. Ramachandran, and T. R. Viswanathan, "A 10-b 20-Msamples/s analog-to-digital converter," *IEEE J. Solid-State Circuits*, vol. 27, no. 3, Mar. 1992, pp. 351–358.

[76] S. H. Lewis, "Optimizing the stage resolution in pipelined, multistage, analog-to-digital converters for video-rate applications," *IEEE Trans. Circuits Syst. II, Analog Digital Sig. Proc.*, vol. 39, no. 8, Aug. 1992, pp. 516–523.

[77] P. W. Li, M. J. Chin, P. R. Gray, and R. Castello, "A ratio-independent algorithmic analog-to-digital conversion technique," *IEEE J. Solid-State Circuits*, vol. SC-19, no. 6, Dec. 1984, pp. 828–836.

[78] Y.-M. Lin, "Performance limitations on high-resolution video-rate analog-digital interfaces," Memorandum No. UCB/ERL M90/55, Electronics Research Laboratory, U. C. Berkeley, June 1990, pp. 74–75.

[79] Y.-M. Lin, B. Kim, and P. R. Gray, "A 13 bit 2.5MHz self-calibrated pipelined A/D converter in 3 μm CMOS," in *1990 Symp. on VLSI Circuits Dig. Tech. Papers*, June 1990, pp. 33–34. [Single 5-V supply; 3 bits/stage, with 1 bit of redundancy.]

[80] Y.-M. Lin, B. Kim, and P. R. Gray, "A 13-b 2.5-MHz self-calibrated pipelined A/D converter in 3-μm CMOS," *IEEE J. Solid-State Circuits*, vol. 26, no. 4, Apr. 1991, pp. 628–636.

[81] C. W. Mangelsdorf, "Improving bandwidth and error rate in flash converters," in *1989 Symp. on VLSI Circuits Dig. Tech. Papers*, May 1989, pp. 53–54.

[82] C. W. Mangelsdorf, "A 400-MHz input flash converter with error correction," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, Feb. 1990, pp. 184–191.

[83] C. Mangelsdorf, H. Malik, S.-H. Lee, S. Hisano, and M. Martin, "A two-residue architecture for multistage ADCs", in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 64–65.

[84] C. Mangelsdorf, S.-H. Lee, M. Martin, H. Malik, T. Fukuda, and H. Matsumoto, "Design for testability in digitally-corrected ADCs", in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 70–71, 165. [Chip performance details in Data Sheet for AD875.]

[85] M. Mar, *mse*, minimum sinusoidal error method program for calculating signal to noise ratios for sinusoidal signals, Dec. 1991.

[86] T. Matsuura et al., "An 8b 20MHz CMOS half-flash A/D Converter," in *ISSCC Dig. Tech. Papers*, Feb. 1988, pp. 220–221, 376. [Hitachi.]

[87] T. Matsuura, H. Kojima, E. Imaizumi, K. Usui, and S. Ueda, "An 8-b 50-MHz 225-mW submicron CMOS ADC using saturation eliminated comparators," in *Proc. Custom Integrated Circuits Conf.*, May 1990, pp. 6.4.1–6.4.4. [Hitachi.]

[88] T. Matsuura et al., "A 95-mW, 10-b 15-MHz low-power CMOS ADC using analog double-sampled pipelining scheme," in *1992 Symp. on VLSI Circuits Dig. Tech. Papers*, June 1992, pp. 98–99. [Hitachi.]

[89] A. Matsuzawa et al., "A 10b 30MHz two-step parallel BiCMOS ADC with internal S/H," in *ISSCC Dig. Tech. Papers*, Feb. 1990, pp. 162–163, 290. [Matsushita; 1.2-μm BiCMOS with 7.5GHz npn; supplies: ±5 V]

[90] A. Matsuzawa et al., "A 6b 1GHz dual-parallel A/D converter," in *ISSCC Dig. Tech. Papers*, Feb. 1991, pp. 174–175, 311. [Bipolar ADC's; time-interleaved A/D systems; Matsushita; 2 time-interleaved ADC's; 13-GHz Si bipolar process; applications in digital 'scopes and logic analyzers; no discussion of mismatches.]

[91] K. J. McCall, M. J. Demler, and M. W. Plante, "A 6-bit 125 MHz CMOS A/D converter," in *Proc. Custom Integrated Circuits Conf.*, May 1992, pp. 16.8.1–16.8.4. [Micro Networks; full flash with occasional autozero, based on external asynchronous pulse.]

[92] R. H. McCharles, V. A. Saletore, W. C. Black, Jr., and D. A. Hodges, "An algorithmic analog-to-digital converter," in *ISSCC Dig. Tech. Papers*, Feb. 1977, pp. 96–97.

[93] M. T. McTigue and P. J. Byrne, "An 8-gigasample-per-second, 8-bit data acquisition system for a sampling digital oscilloscope," *Hewlett-Packard Journal*, vol. 44, no. 5, Oct. 1993, pp. 11–23.

[94] D. G. Messerschmitt, "Effects of bit-robbing in tandem digital switches," *IEEE Trans. Commun.*, vol. COM-27, no. 7, July 1979, pp. 1025–1027. [Mismatch analysis; gives expression for output power spectral density in case of gain mismatches.]

[95] Micro Linear Corporation, Data Sheet for ML2200, ML2208, Oct. 1990. [12-bit + sign self-calibrated ADC with S/H, input multiplexer, programmable gain, etc.]

[96] Y. Nakamura, S. Aikawa, and H. Takanashi, "Robust modem technology for outage-free digital microwave radio systems," *NTT Review*, vol. 3, no. 1, Jan. 1991, pp. 28–37. [Includes some discussion of trellis-coded modulation schemes.]

[97] National Semiconductor Corporation, Data Sheet for LM12458, Oct. 1991. [12-bit + sign self-calibrated ADC with S/H, 8-channel analog multiplexer, on-board FIFO, instruction RAM.]

[98] K. J. Negus, R. A. Koupal, D. Millicker, and C. P. Snapp, "Silicon bipolar mixed-signal parameterized-cell array for wireless applications to 4GHz," in *ISSCC Dig. Tech. Papers*, Feb. 1992, pp. 230–231, 293. [Radio applications; for wireless communication applications in 800–2500 MHz range, including: cellular systems at 800–950 MHz, L-band satellite transceivers at 1500–1700 MHz for applications such as GPS, the proposed Iridium system, etc.; example of GPS satellite receiver with final IF (IF2) at 25 MHz going into gain/ADC block and to CMOS DSP.]

[99] G. Nicollini, P. Confalonieri, and D. Senderowicz, "A fully-differential sample-and-hold circuit for high-speed applications," *IEEE J. Solid-State Circuits*, vol. 24, no. 5, Oct. 1989, pp. 1461–1465.

[100] K. A. Nishimura, "Optimum partitioning of analog and digital circuitry in mixed-signal circuits for signal processing," Memorandum No. UCB/ERL M93/67, Electronics Research Laboratory, U. C. Berkeley, July 1993.

[101] H. Ohara, H. X. Ngo, M. J. Armstrong, C. F. Rahim, and P. R. Gray, "A CMOS programmable self-calibrating 13-bit eight-channel data acquisition peripheral," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 930–938. [CMOS analog design issues; description of offset correction using auxiliary input stage.]

[102] S. Ohr, "Analog vs DSP: balancing speed and precision against cost," *Computer Design*, vol. 31, no. 5, May 1992, pp. 83–100. [Video applications; very interesting discussion of A/D signal processing issues.]

[103] M. Ohta et al., "A single-chip CMOS analog/digital mixed NTSC decoder," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1464–1469. [Video applications; Fujitsu; discusses some A/D integration issues; integrates 6-bit ADC, DAC's for RGB signals, peripheral and DSP circuits for a digital NTSC decoder; operates at 14.3 MHz ($4F_{sc}$) ; 11 mm X 11 mm; 1.2-μm CMOS; 1 W from 5 V; for low-end applications — e.g., compact LCD TV's, picture phones and video conference systems, multimedia systems and personal computers; discussion of some techniques to deal with the 6-bit quantization.]

[104] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 2nd ed., New York: McGraw-Hill, 1984, pp. 269–270, 279–281.

[105] A. Petraglia and S. K. Mitra, "Analysis of mismatch effects among A/D converters in a time-interleaved waveform digitizer," *IEEE Trans. Instrum. Meas.*, vol. 40, no. 5, Oct. 1991, pp. 831–835. [Gain and offset mismatch analysis.]

[106] R. Petschacher, B. Zojer, B. Astegher, H. Jessner, and A. Lechner, "A 10-b 75MSPS subranging A/D converter with integrated sample and hold," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1339–1346. [Bipolar ADC's; long version of 1990 ISSCC paper.]

[107] Philips Semiconductors, "High-performance 8-bit video data converters," in *Desktop Video Data Handbook*, 1992, pp. 2-3–2-11. [Overview of Philips TDA87xx 8-bit bipolar ADC's using the folding and interpolation architecture — e.g., TDA8703 (8 bits, 40 MS/s, +5 V, 290 mW, 14 pF input capacitance, 7.1 effective bits at 4.43 MHz input frequency).]

[108] R. Philpott, R. Kertis, R. Richetta, T. Schmerbeck, and D. Schulte, "A 7MB/sec (65MHz), mixed-signal, magnetic recording channel DSP using partial response signaling with maximum likelihood detection," in *Proc. Custom Integrated Circuits Conf.*, May 1993, pp. 10.4.1–10.4.4.

[109] K. Poulton, J. J. Corcoran, and T. Hornak, "A 1-GHz 6-bit ADC system," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 962–970. [Time-interleaved A/D systems; mismatch analysis — some qualitative discussion; got best paper.]

[110] P. Real, D. H. Robertson, C. W. Mangelsdorf, and T. L. Tewksbury, "A wide-band 10-b 20-Ms/s pipelined ADC using current-mode signals," *IEEE J. Solid-State Circuits*, vol. 26, no. 8, Aug. 1991, pp. 1103–1109. [2-μm BiCMOS with 2-GHz npn's; area: 6.35 mm X 7.61 mm; supplies: ±5 V; power: ≈ 1 W.]

[111] D. Robertson, P. Real, and C. Mangelsdorf, "A wideband 10-bit, 20Msps pipelined ADC using current-mode signals," in *ISSCC Dig. Tech. Papers*, Feb. 1990, pp. 160–161, 289.

[112] A. Rothermel, "BiCMOS circuits for DPCM coders in HDTV systems," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1470–1475. [BiCMOS; Some interesting discussion of ways to compare CMOS and BiCMOS (for digital).]

[113] P. A. Ruetz, "Architectures and design techniques for real-time image processing ICs," Memorandum No. UCB/ERL M86/37, Electronics Research Laboratory, U. C. Berkeley, May 1986.

[114] K. Rush and P. Byrne, "A 4GHz data acquisition system," in *ISSCC Dig. Tech. Papers*, Feb. 1991, pp. 176–177, 312. [Time-interleaved A/D systems; 16-W system including 13-GHz bipolar, 1-μm CMOS and 50-μm thick film technologies; two interleaved hybrid circuits — each hybrid contains one sampler chip, two ADC chips (each containing two interleaved 500-MS/s ADC's), and two memory chips.]

[115] T. S. Rzeszewski, "A technical assessment of advanced television," *Proc. IEEE*, vol. 78, no. 5. May 1990, pp. 789–804.

[116] H. Samueli and B. C. Wong, "A VLSI architecture for a high-speed all-digital quadrature modulator and demodulator for digital radio applications," *IEEE J. Select. Areas Commun.*, vol. 8, no. 8, Oct. 1990, pp. 1512–1519. [Radio applications; UCLA; some discussion of all-digital architectures and ADC requirements — e.g., typically the ADC runs at 100–200 MS/s.]

[117] G. J. Saulnier et al., "A VLSI demodulator for digital RF network applications: theory and results," *IEEE J. Select. Areas Commun.*, vol. 8, no. 8, Oct. 1990, pp. 1500–1511. [Radio applications; RPI and GE; system uses a single ADC to sample the IF signal and produces baseband I and Q signals — ADC samples at 4 times IF center frequency.]

[118] T. J. Schmerbeck, R. A. Richetta, and L. D. Smith, "A 27MHz mixed analog/digital magnetic recording channel DSP using partial response signalling with maximum likelihood detection," in *ISSCC Dig. Tech. Papers*, Feb. 1991, pp. 136–137, 304. [BiCMOS; mixed A/D integration issues; lots of care to minimize coupling from digital to analog; details of process and general design approach given in [123].]

[119] B. Schweber, "IF stages are going digital for both analog *and* digital signals," *Analog Dialogue*, vol. 26, no. 1, 1992, pp. 3–6. [Radio applications; discussion of direct digitization of IF signal.]

[120] S. Sheng, A. Chandrakasan, and R. W. Brodersen, "A portable multimedia terminal," *IEEE Commun. Mag.*, vol. 30, no. 12, Dec. 1992, pp. 64–75. [Overview of the U. C. Berkeley INFOPAD project at system and circuit level.]

[121] S. Sheng, private communication.

[122] N. Shiwaku et al., "A rail-to-rail video-band full Nyquist 8-bit A/D converter," in *Proc. Custom Integrated Circuits Conf.*, May 1991, pp. 26.2.1–26.2.4. [TI.]

[123] L. D. Smith et al., "A CMOS-based analog standard cell product family," *IEEE J. Solid-State Circuits*, vol. 24, no. 2, Apr. 1989, pp. 370–379. [A/D integration issues; BiCMOS; "Bipolar-CMOS" process: 1-μm CMOS; with 6-GHz npn's, low-threshold NMOS devices, MOS capacitors, etc.; 5-V supply.]

[124] C. G. Sodini, P.-K. Ko, and J. L. Moll, "The effect of high fields on MOS device and circuit performance," *IEEE Trans. Electron Devices,* vol. ED-31, no. 10, Oct. 1984, pp. 1386–1393.

[125] K. Sone, N. Nakadai, Y. Nishida, M. Ishida, Y. Sekine, and M. Yotsuyanagi, "A 10b 100MS/s pipelined subranging BiCMOS ADC," in *ISSCC Dig. Tech. Papers*, Feb. 1993, pp. 66–67, 264. [NEC; two-stage pipeline (6 bits + 5 bits); interstage gain is 1 — i.e., a subtractor.]

[126] B.-S. Song and M. F. Tompsett, "A 12b 1MHz capacitor error averaging pipelined A/D converter," in *ISSCC Dig. Tech. Papers*, Feb. 1988, pp. 226–225, 380.

[127] B.-S. Song, M. F. Tompsett, and K.R. Lakshmikumar, "A 12-bit 1-Msample/s capacitor error-averaging pipelined A/D converter," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, Dec. 1988, pp. 1324–1333.

[128] B.-S. Song and M. F. Tompsett, "A 10b 15MHz recycling two-step A/D converter," in *ISSCC Dig. Tech. Papers*, Feb. 1990, pp. 158–159, 289.

[129] B.-S. Song, S.-H. Lee, and M. F. Tompsett, "A 10-b 15-MHz CMOS recycling two-step A/D converter," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1328–1338.

[130] Sony Corporation, Data Sheets for CXD1176Q and CXD1179Q, 1993.

[131] D. K. Su, M. J. Loinaz, S. Masui, and B. W. Wooley, "Experimental results and modeling techniques for substrate noise in mixed-signal integrated circuits," *IEEE J. Solid-State Circuits*, vol. 28, no. 4, Apr. 1993, pp. 420–430.

[132] R. Subramanian, "Synchronization systems for spread-spectrum receivers," Ph.D. Dissertation, University of California, Berkeley, Nov. 1991.

[133] Y. Sugimoto and S. Mizoguchi, "An experimental BI-CMOS video 10bit ADC," in *1988 Symp. on VLSI Circuits Dig. Tech. Papers*, Aug. 1988, pp. 129–130. [Toshiba.]

[134] Y. Sugimoto and S. Mizoguchi, "An experimental BiCMOS video 10-bit ADC," *IEEE J. Solid-State Circuits*, vol. 24, no. 4, Aug. 1989, pp. 997–999. [BiCMOS ADC's; Toshiba; 3-μm BiCMOS with 4-GHz npn; no S/H; 0.5 W at 5 V.]

[135] S. Sutarja, "High-resolution pipelined analog-to-digital conversion," Memorandum No. UCB/ERL M88/27, Electronics Research Laboratory, U. C. Berkeley, May 1988, pp. 41–42.

[136] S. Sutarja and P. R. Gray, "A pipelined 13-bit, 250-ks/s, 5-V analog-to-digital converter," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, Dec. 1988, pp. 1316–1323. [Uses reference-feedforward correction to reduce effects of interstage gain errors in a multistage pipeline ADC. Achieves DNL < 0.5 LSB.]

[137] E. J. Swanson, N. S. Sooch, and D. J. Knapp, "Method for reducing effects of electrical noise in an analog-to-digital converter," U.S. Patent 4 746 899, May 1988.

[138] S. Takeuchi et al., "A 30-MHz mixed analog/digital signal processor," *IEEE J. Solid-State Circuits*, vol. 25, no. 6, Dec. 1990, pp. 1458–1463. [A/D integration; Mitsubishi; implements 30-MHz, 28-tap programmable FIR filter; die area is 9.71mm X 8.14mm; 1-μm CMOS.]

[139] I. Tamitani, H. Harasaki, and T. Nishitani, "A real-time HDTV signal processor: HD-VSP," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 1, no. 1, Mar. 1991, pp. 35–41. [Video applications; NEC; 8-bit system; mentions 74.25 MHz, 64.8 MHz, and 48.6 MHz as clock rates for ADC's and DAC's used in HDTV.]

[140] S. S. Taylor, "High speed analog-to-digital conversion in integrated circuits," Ph.D. Dissertation, University of California, Berkeley, Dec. 1978.

[141] Tektronix, Inc., "First in a new generation of lab instruments — the DSA 600 Series Digitizing Signal Analyzers," *Handshake*, Summer 1989, pp. 4–8. [The DSA 602 can achieve 8-bit resolution at 2 GS/s using four time-interleaved data acquisition channels each running at 500 MS/s; each channel contains two bipolar, 8-bit, 250-MS/s ADC's.]

[142] L. J. Thorpe, "HDTV and film — digitization and extended dynamic range," *SMPTE Journal*, vol. 102, no. 6, June 1993, pp. 486–497.

[143] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor,* New York: McGraw-Hill, 1987.

[144] K. Tsugaru et al., "A 10bit 40MHz ADC using 0.8μm BI-CMOS technology," in *Proc. 1989 Bipolar Circuits and Technology Meeting,* Sep. 1989, pp. 48–51.

[145] K. Tsuji, S. Saitoh, S. Morinaga, and T. Iida, "A new pipelined-subranging 8 bit 20 MHz CMOS AD converter," in *1988 Symp. on VLSI Circuits Dig. Tech. Papers*, Aug. 1988, pp. 127–128. [Toshiba; power: 700 mW.]

[146] K. Tsuji, H. Sugiyama, and N. Sugawa, "A CMOS 20MHz 8bit 50mW ADC for mixed analog/digital ASIC's," in *Proc. Custom Integrated Circuits Conf.*, May 1991, pp. 26.3.1–26.3.4. [Toshiba; core size is 2.9 mm^2 in 0.8-μm CMOS.]

[147] G. T. Uehara, "Circuit techniques and considerations for implementation of high speed CMOS analog-to-digital interfaces in DSP-based PRML magnetic disk read channels," Ph.D. Dissertation, University of California, Berkeley, Aug. 1993.

[148] G. T. Uehara, private communication.

[149] M. Umemoto, "Professional HDTV digital recorder," *IEEE J. Select. Areas Commun.*, vol. 10, no. 1, Jan. 1992, pp. 80–85. [Video applications; Hitachi; interesting discussion of HDTV digital recorder specifications — e.g., luminance sampling is 74.25 MHz at 8 bits.]

[150] R. E. J. van de Grift, I. W. J. M. Rutten, and M. van der Veen, "An 8-bit video ADC incorporating folding and interpolation techniques," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 6, Dec. 1987, pp. 944–953. [Philips; bipolar process with 7.5-GHz npn; achieves 7.6 effective bits at F_s = 50 MHz and F_{in} = 4.43 MHz; power is less than 300 mW; includes detailed discussion of folding and interpolation circuitry.]

[151] R. J. van de Plassche and P. Baltus, "An 8-bit 100-MHz full Nyquist ADC," *IEEE J. Solid-State Circuits*, vol. 23, no. 6, Dec. 1988, pp. 1334–1344. [Signetics HS3 bipolar process with 9-GHz npn; 800 mW from a single –5.2-V supply.]

[152] J. van Valburg and R. van de Plassche, "An 8b 650MHz folding ADC," in *ISSCC Dig. Tech. Papers*, Feb. 1992, pp. 30–31, 234. [Bipolar ADC's; Philips; bipolar process with 13-GHz npn; 850 mW from a single –4.5-V supply.]

[153] J. van Valburg and R. J. van de Plassche, "An 8-b 650-MHz folding ADC," *IEEE J. Solid-State Circuits*, vol. 27, no. 12, Dec. 1992, pp. 1662–1666. [Full length version of ISSCC paper.]

[154] P. Vorenkamp and J. P. M. Verdaasdonk, "A 10b 50MS/s pipelined ADC," in *ISSCC Dig. Tech. Papers*, Feb. 1992, pp. 32–33, 235. [Philips; bipolar production process with 3-GHz npn.]

[155] T. Wakimoto, Y, Akazawa, and S. Konaka, "Si bipolar 2GS/s 6b flash A/D conversion LSI," in *ISSCC Dig. Tech. Papers*, Feb. 1988, pp. 232–233, 383. [NTT.]

[156] K. M. Ware, H.-S. Lee, and C. G. Sodini, "A 200-MHz CMOS phase-locked loop with dual phase detectors," *IEEE J. Solid-State Circuits*, vol. 24, no. 6, Dec. 1989, pp. 1560–1568. [Clock generation.]

[157] B. C. Wong and H. Samueli, "A 200-MHz all-digital QAM modulator and demodulator in 1.2-μm CMOS for digital radio applications," *IEEE J. Solid-State Circuits*, vol. 26, no. 12, Dec. 1991, pp. 1970–1980. [Radio applications; UCLA; discussion of all-digital quadrature modulator and demodulator architectures.]

[158] H. C. Yang and D. J. Allstot, "Considerations for fast settling operational amplifiers," *IEEE Trans. Circuits Syst.*, vol. 37, no. 3, Jan. 1990, pp. 326–334.

[159] J. Yoshida, "Set-top box could be budget buster," *Electronic Engineering Times*, no. 774, Nov. 29, 1993, pp. 1,8.

[160] J. Yoshida and G. Leopold, "U.S. HDTV spec unveiled; grand alliance hands proposal to FCC for testing," *Electronic Engineering Times*, no. 769, Oct. 25, 1993, pp. 1,8.

[161] M. Yotsuyanagi, T. Etoh, and K. Hirata, "A 10bit 50Msample/sec pipelined CMOS A/D converter with S/H," in *Proc. Custom Integrated Circuits Conf.*, May 1992, pp. 16.7.1–16.7.4. [NEC.]

[162] M. Yotsuyanagi, T. Etoh, and K. Hirata, "A 10-b 50-MHz pipelined CMOS A/D converter with S/H," *IEEE J. Solid-State Circuits*, vol. 28, no. 3, Mar. 1993, pp. 292–300.

[163] B. Zojer, R. Petschacher, and W. A. Luschnig, "A 6-bit/200-MHz full Nyquist A/D converter," *IEEE J. Solid-State Circuits*, vol. SC-20, no. 3, June 1985, pp. 780–786.

[164] B. Zojer, B. Astegher, H. Jessner, and R. Petschacher, "A 10b 75MHz subranging A/D converter," in *ISSCC Dig. Tech. Papers*, Feb. 1990, pp. 164–165, 290. [Siemens; 7-GHz bipolar process; supplies: ±5 V; power: 2 W.]