MODELS FOR A COMPREHENSIVE METHODOLOGY FOR PARAMETER SELECTION AND DISCRETE DECISION-MAKING IN THE DESIGN OF INTEGRATED CIRCUITS

by

David C. Riley

Memorandum No. UCB/ERL M88/77

22 April 1988

MODELS FOR A COMPREHENSIVE METHODOLOGY FOR PARAMETER SELECTION AND DISCRETE DECISION-MAKING IN THE DESIGN OF INTEGRATED CIRCUITS

by

David C. Riley

Memorandum No. UCB/ERL M88/77

22 April 1988

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

MODELS FOR A COMPREHENSIVE METHODOLOGY FOR PARAMETER SELECTION AND DISCRETE DECISION-MAKING IN THE DESIGN OF INTEGRATED CIRCUITS

by

David C. Riley

Memorandum No. UCB/ERL M88/77

22 April 1988

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Models for a Comprehensive Methodology for Parameter Selection and Discrete Decision-Making in the Design of Integrated Circuits

David C. Riley

Ph.D.

Department of Electrical Engineering and Computer Sciences

ABSTRACT

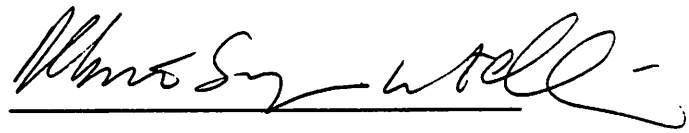
A new, optimization-based design methodology for integrated circuits is proposed that represents a multi-faceted generalization of conventional problem formulations for the statistical design of integrated circuits. It is applicable to the selection of parameter values, and to the selection of the most desirable design among limited numbers of given discrete design options. In both applications the criterion for the selection is a form of profit expected to be realized from the circuit, that is particularly relevant to the design of integrated circuits to be sold on the open market. This profit quantity is modeled as a function of the parameters to be selected and of random variables describing all statistical phenomena that affect the profit, including in particular processing disturbances and localized wafer defect phenomena. The number of performance categories of finished circuits allowed in the model is arbitrary, rather than just the two (meeting specifications, violating specifications) allowed in conventional statistical design methods.

In the parametric design application, the parameters to be selected are all those that have a first-order effect on the profit criterion. These include the rate of production of the

circuit, fabrication process test limits for wafer rejection, device dimensions, performance specifications, fractions of circuits to receive various packaging, testing, and reliability treatments, prices, and fractions of orders to be filled.

In the discrete decision-making application, designers are able to objectively choose between product design options which may differ in qualitative characteristics such as fabrication technology, fabrication process, circuit topology, and package and testing options made available.

In addition to its application to the two types of design selections described, the model also provides insights into ideal product development practice for integrated circuits.

Signature: 

Committee Chairman

Acknowledgement

I would like to express my gratitude to Professor Alberto Sangiovanni-Vincentelli, my research advisor, for his support and guidance throughout my doctoral program. I believe he has imparted to me skills that will be of value throughout my career, not only in the generation of useful technical knowledge, but also in its communication.

I would also like to express my thanks to Professors Charles Desoer, Richard Muller, Paul Gray, and Finbarr O'Sullivan (Department of Statistics) for their interest in my work and my success, and Professors Paul Gray and David Aldous (Department of Statistics) for their reading of this dissertation.

Since its inception in 1982, this research project has been supported by a number of organizations. Foremost among these is Harris Semiconductor. The project has benefited from having the opportunity to study real technical and economic considerations of an IC producer, by working in the Harris plant in Melbourne, Florida. There, the forward-looking interests and the efforts of James Spoto, Nick English, and Terry Coston, made it possible to gather information for and receive feedback about the model, both from them and from employees working in process engineering, circuit design, computer-aided design, test and evaluation, product engineering, marketing, sales, and management. Without this industrial contact, development of the new methodology described in this dissertation would not have been possible.

I would also like to acknowledge the financial support of the Joint Services Electronics Program at the Electronics Research Laboratory, (contract number F 49620-84-C-0057), the

Defense Advanced Research Electronic Systems Agency (Naval Electronic Systems Command contract number N00039-87-C-0182), the Micro Program of the State of California, Hughes Aircraft, and Texas Instruments.

Finally, I would like to express my appreciation to Bill Nye, now with Epsilon Active in Berkeley, and fellow students Jyuo-Min Shyu and Hormoz Yaghutiel, for their technical discussions on subjects peripheral to statistical circuit design, for their help with numerous software and administrative problems, and for their good humor.

Table of Contents

Chapter 1

INTRODUCTION

1.1. Scope of Application

In the design of integrated circuits, the types of information that can be considered given vary from one case to another. Given in almost every case is a required electrical functionality, which encompasses specifications such as numbers of input, output, and power terminals, and qualitative electrical input-output behavior characteristics. (The starting point of a definition of an electrical functionality might be a circuit class identifier such as "operational amplifier", or "2K × 8 synchronous random access memory".) Other characteristics such as quantitative performance levels may or may not be specified in advance. But whatever the starting point, the design process ends with a set of *manufacturing specifications* used in producing the circuit. These manufacturing specifications can be considered to be of two types.

One type consists of real numbers that are parameters values to be used in producing the circuit. In this dissertation, the parameters dctermined through the application of a particular design methodology are called its *design parameters*, and a choice of their values is called a *parametric design*. In conventional methods for integrated circuit design the design parameters are primarily dimensions of passive and active devices.

The other type of manufacturing specifications consists of all the remaining (non-parametric) specifications determined by the particular methodology. These are called *qualitative characteristics*, selections of which comprise a *qualitative design*. Clearly included among the qualitative characteristics of a circuit is its topology, but as defined here, qualitative characteristics can also be considered to encompass such design specifications as the choice of the fabrication process in which the circuit will be produced, and the choice of a set of package types for the circuit.

The design process itself can be considered to consist of three types of activities:

(1) identification of a usually limited number of qualitative design candidates that have the desired functionality.

(2) selection of the most desirable qualitative design

(3) selection of the most desirable parametric design, for a given qualitative design

Note that these activity types do not necessarily represents steps to be carried out once each, in the sequence listed. In fact selection among a number of qualitative designs ideally is preceded by a parametric design for each of them.

The initial aim in this research was the development of a methodology for parametric design activities. However, the work has also yielded a methodology for selection among qualitative design alternatives that has a most of the features of the parameter selection

methodology.

1.2. Characteristics of Comprehensive Methodologies for IC Parameter Selection

In this dissertation, the term *comprehensive*, when applied to methodologies for parametric design, is intended to have a specific objective meaning. Defining this meaning in detail is the next objective. It will enable easy understanding of the relationship between the parametric and qualitative design methodologies. It will also enable the new formulation for parametric design to be introduced and contrasted with formulations in previous research work in parameter selection. Before proceeding, however, it is necessary to define two terms.

By a *goodness criterion* for the design of IC's is meant a scalar quantity that (given a qualitative design) depends on the parametric design and summarizes the desirability of the design. The aim in developing a goodness criterion is to construct it so that the IC producer, in selecting among alternative designs would always choose the design having the largest (or smallest, as the case may be) goodness criterion.

By a *constraint quantity* is meant a scalar quantity that depends on the parametric design, that is used to specify a subset of the design parameter space representing unacceptable designs, and that does this through a convention that if the constraint quantity is, say, positive, the design is unacceptable.

Proceeding now with the main definition, a methodology for IC parameter selection is called a comprehensive methodology if it consists of the solution of a problem formulation that has the four fundamental characteristics named, defined, and discussed next.

I. Optimization Framework Utilization

For a methodology to have this characteristic, both of the following must be the case.

A. The formulation has a design model consisting of the following four elements in unambiguously specified form: a set of design parameters; a goodness criterion; optionally a set of one or more constraint quantities; a specification of the functional dependence of the goodness criterion and any constraint quantities on the design parameters.

B. The formulation allows solution methods to be applied that provide a comparison of their realized goodness criterion values with some estimate of the goodness criterion value at a local optimum point.

Note that, for example, nonlinear programming formulations provide such a comparison, using the tail of the truncated sequence of goodness criterion values, in convergence, to estimate a locally optimum goodness value.

II. Design Parameter Appropriateness

The requirement in A above that the formulation have a set of well-defined design parameters is not a distinguishing characteristic of comprehensive methodologies, since most design methods have this. Yet it is important that the design parameters of a parametric design methodology be appropriate to the design problem that is truly at hand.

Before this can be discussed, however, it is necessary to make some preliminary comments and to define some additional terminology.

First, note that some degree of familiarity with modern technology for the manufacture of integrated circuits is presumed in this dissertation. While the degree of expertise assumed is well below that possessed by fabrication process engineers, some common fabrication terminology will be used herein without definition.

The term *circuit* has been used above, in the abstract sense of a design consisting of an integrated interconnection of passive and active devices intended to realize some prescribed electrical functionality. In the sequel, "circuit" will also be used to connote an individual physical unit of production, whether in the form of a die or a finished packaged device, unless this alternative interpretation is not adequately distinguished from the other interpretation by the context of the discussion, in which case the term *unit* will be used.

After the design specifications for a circuit (abstract sense) are complete, there are a large number of additional attributes and objects that become associated with the circuit, such as a layout, a set of photolithographic masks, a set of package types in which the dice are to be encapsulated, a test program, one or more prices, and a marketplace niche, to name a few. In the context of more general attributes such as these, the term *product* is used in place of "circuit". However, without further clarification, there can be ambiguity in the interpretation of "product". The mask set is a useful entity in clarifying the intended meaning.

The mask sets of a producer can be considered to be in one-to-one correspondence with its circuits (abstract sense), but different mask sets of a producer might be used in manufacturing units having the same electrical functionality, and different units manufactured using the same mask set might be treated as different products

by marketing personnel and users, if only on the basis of different packaging or testing. In this dissertation, products are not meant to be one-to-one with functionalities or with marketing variations, but with mask sets (and hence with circuits).

In view of the fact that IC producers almost universally manufacture more than one IC product, the term *design product* is used to refer to the product to be designed.

The discussion of the selection of design parameters can now be continued, beginning with the introduction of some directly relevant terms.

First, a parameter in the design of a particular IC product is said to have *side effects* if changing the value of the parameter affects the desirability of the design of one or more other products, in a manner that is not accounted for in the goodness criterion of the methodology. An example of a parameter that has side effects is a process parameter such as an oven temperature in a process used to fabricate products other than the design product. A parameter in the design of a particular IC product is called *specific to the design product optimization* if it has no side effects.

Second, a design parameter is called *producer-controlled* if the producer is free to set its value as he wishes, at least over some open interval. An example of a parameter that is not producer-controlled is the cost of a piece of test equipment needed exclusively for the testing of the design product. Finally, a design parameter is called *directly implementable* if no further processing of its information content need be carried out by design engineers in order that it be implemented in the manufacturing of the product. An example of a parameter that is not directly implementable is the threshold voltage of an MOS transistor, since achieving a specified

threshold voltage requires the determination of a number of fabrication-related parameters.

With these definitions, the Design Parameter Appropriateness characteristic of formulations for comprehensive methodologies requires that both the following be the case.

A. The design parameters of the formulation form an independent set of parameters that are specific to the design product optimization, producer-controlled, directly implementable, and that in general have a first-order effect on the goodness criterion.

B. The set of design parameters of the formulation is the only set of parameters with the properties in A.

Note that B requires that every parameter that is specific to the design product optimization, producer-controlled, directly implementable, and that in general has a first-order effect on the goodness criterion, must be included in the set of design parameters, if doing so does not destroy their independence.

III. Goodness Criterion Appropriateness.

This characteristic is defined as follows.

The goodness criterion of the formulation explicitly reflects the producer's fundamental interests in manufacturing the design product.

In the design of almost all IC's, the producer's fundamental interest is the economic benefit expected to result from the implementation of the design. There are design cases in which the producer has no interest in economic considerations in

designing a particular IC. However, these are cases in which the producer has been given a "blank check" in developing the IC, and are very rare, even when the IC is being developed under government contract. Therefore, this exception is ignored without appreciable loss of generality. The implication, then, is that in a comprehensive methodology, the goodness criterion must be the economic benefit, or gain, expected to result from the implementation of the design (with designs of higher gain preferred to those of lower gain).

In reality, very few IC design methodologies use economic gain as a measure of the desirability of a design. In fact, very few use any analyses involving economic variables pertaining to the design product, such as its cost of production, sales revenue, or rate of production. In this respect IC design methodologies are like engineering design methodologies in general. This situation can be represented more clearly using some symbolic notation.

Let the qualitative and parametric design specifications for an engineering design be represented symbolically by the vector **d**, and let the expected economic gain from the product be denoted g. It has been asserted that a comprehensive design methodology must use a single measure of goodness, and that goodness criterion must be g. If this were commonly the case, the central problem in engineering design would be the development and exploration of the function g in

$$g = g(\mathbf{d}). \tag{1.1}$$

Instead the engineer typically uses as design criteria a number of performance attributes **a**, say, of a technical rather than economic nature. He generally considers the central problem in engineering design to be the development and exploration of the

function $\mathbf{r}$ in

$$\mathbf{a} = \mathbf{r}(\mathbf{d}).$$

However, if for a particular design problem, an explicit economic model h were known, i.e.,

$$g = h(\mathbf{d}, \mathbf{a}), \tag{1.2}$$

then the function g is known, since g can be expressed as

$$g = h(\mathbf{d}, \mathbf{r}(\mathbf{d})).$$

The reason that expected economic gain is not usually the goodness criterion for engineering design selection is that the explicit economic model of (1.1) is generally not readily available. Some effort must be expended to develop a model of the form of (1.2), and usually an implicit decision is made not to expend the effort. Nevertheless inclusion of explicit economic models in engineering design is the ideal, and a comprehensive design methodology must incorporate such models.

IV. Model Realism and Accuracy

The definition of this characteristic has the following two components.

A. Both the constraint quantities incorporated in the model, and the functional dependences in the model reflect realistic consideration of all first-order effects in actual IC production, including as appropriate, non-idealities in manufacturing processes, and non-deterministic phenomena.

B. The design model is numerically accurate.

The importance of non-deterministic phenomena in IC design and production has been recognized in the IC industry since its beginnings. The need for consideration of non-deterministic phenomena in comprehensive methodologies for IC design

exemplifies the same need in engineering design in general. This condition arises because it is frequently possible to exploit fully all reasonably manageable deterministic relationships. When this is done, design choices are made which tend to promote random effects to a dominant role in limiting further advances in performance.

In the integrated circuit case, it was evident in the earliest days of the industry that economic rewards generally increased as device dimensions decreased. Design practice immediately moved into a world of design with "as small as possible" devices. But what has limited the rewards of shrinking device dimensions is not any set of deterministic relationships. It is the variations in realized device dimensions from unit to unit which results in variations in performance, which in turn significantly degrades the economic value of the totality of units produced.

In summary, an IC parametric design methodology is called comprehensive if it consists of the solution of a problem formulation that has the characteristics of optimization framework utilization, design parameter appropriateness, goodness criterion appropriateness, and model realism and accuracy.

1.3. Scope of Application, Revisited

It is now easy to understand the relationship between the application of this research work to the selection of parametric designs and to the selection of qualitative designs. The basis for the selection of parameter values in the new methodology is the maximization of expected economic gain. If for each of the qualitative designs under consideration, coefficients needed for the models of the goodness criterion and the constraints are

correctly set for the particular qualitative design, a parametric design is carried out, the maximum goodness criterion value is determined, and the qualitative design with the highest goodness criterion is selected for implementation, then this selection clearly benefits from the features with which the goodness criterion has been endowed. The procedure just described is the qualitative design methodology proposed in this work.

Because this qualitative design selection is carried out simply by repeating parametric designs, it has not been necessary to treat qualitative design selection as an explicit goal in developing the models for the new formulation. Instead, all characteristics of the qualitative design, including the circuit topology, are assumed given, and the objective in the arguments that follow is the presentation of models for a comprehensive methodology for parameter selection in IC design.

1.4. Previous Work

Considerable research in methods for IC parameter selection has been carried out over the last two decades, and numerous methodologies proposed. The history of this work can be briefly sketched in terms of the above four fundamental characteristics of comprehensive methodologies, as follows. A decade ago no methodologies applicable to a wide range of circuit functionalities had any of the four characteristics. Since then, some methodologies have been implemented having the optimization framework characteristic [ZEI80], [NYE88], [SHY88]. Considerable work has been done on methods that have that characteristic, as well as a model realism and accuracy that is improved by consideration of non-deterministic phenomena occurring in circuit fabrication. Examples prior to 1980 are discussed and referenced in [BRA81]. Examples published since 1980 include [STY81],

[SIN81], [ANT81], [COO82], [VID82], [HOC83], [STY83], [HOC84], [STE86], [STY86], and [YU87]. This research area is referred to as *statistical design of integrated circuits* . This area, together with work aimed at Characteristic I alone, encompass almost all previous work that is at all relevant to comprehensive design techniques.

What remains is the limited amount of work that has been aimed at developing a comprehensive parametric methodology in the full sense that has been defined in Section 1.2. In order to discuss this work, it is first necessary to describe at a more detailed level some important assumptions commonly made in research in IC parameter selection. This will also serve to further introduce and place in perspective the formulation of the new methodology.

Most of these commonly made assumptions restrict, to some degree, the full incorporation of the four fundamental characteristics described above. First assumptions that are in common with past work and the new formulation are discussed, then assumptions that differ between past work and this work.

Before beginning this discussion of assumptions, a term must be defined. Part of the qualitative design of an IC product is a specification of a set of tests to be performed on each unit produced. A unit is called *operational* if when excited by the full range of input signals in the test program established for the product, any deviation of any of its required output signals from the ideal response of circuits in its class is of a quantitative rather than qualitative nature. For example, if an analog switch has a resistance across a particular closed contact of 4 ohms, this represents a quantitative deviation from the ideal. If on the other hand the resistance of the contact is constant independent of the state of the inputs, this represents a qualitative deviation. Specific detailed definitions of operational units

would of course be particular to the product class in question.

1.4.1. Assumptions in Common Between Existing Formulations and the New Formulation

The assumptions that are shared by past work and this work relate to comprehensive methodology Characteristic IV, Model Realism and Accuracy.

(1) Nature of Performance Measures

> The electrical performance of each operational unit can be adequately summarized by the values of a set of real-valued performance measures (e.g. offset voltage, power consumption) that are circuit responses, or (usually trivial) functions of circuit responses.

This assumes in particular that, in order to characterize the performance capabilities of the circuits produced, it is not necessary to use performance measures that are given as continuous functions of an independent variable, such as frequency domain characteristics and transient response waveforms. One circumstance in which such a performance measure would be necessary is if it were required that a frequency domain gain or phase response lie within some two prescribed functions of frequency (an envelope), for all frequencies within some given range. However, assumption (1) is not significantly restrictive in practice, because testing costs associated with measuring such function-valued entities and comparing with reference functions are prohibitive compared to the costs associated with measuring real-valued performance measures such as frequency domain responses at specific frequencies, and transient responses at specific time points. As a consequence, measuring function-valued performance measures is usually not done.

(2) Presence of Statistical Effects

In circuit fabrication there are parameters subject to statistical (non-deterministic) variation from unit to unit, that are not compensated for by any subsequent trimming process on individual units, and that are of sufficient magnitude that the resultant variations in performance measures among the units causes in turn significant variation in their desirability.

Some examples of non-deterministic parameters that pertain to wafer fabrication and vary from unit to unit are effective device dimensions, dopant concentrations, polysilicon thicknesses, and ion implantation profile parameters.

This assumption is slightly restrictive in that there are a minority of products that have device dimensions laser-trimmed, with the trimming process on each chip controlled by one or more measured performance characteristic of the chip.

If there are significant variations in the desirability of the units produced under a particular design, then the performance of the least desirable units is significantly inferior to the potential performance of the topology used, with obvious implications for the contribution of those units to the economic gain from the product. Intuition, mathematical analyses, and industrial examples all support the notion that favorable selection of design parameter values can minimize the undesirable consequences of statistical fluctuations. It is this objective that has provided the prime impetus for work in statistical circuit design, and a major impetus in this work.

(3) Model Availability

There is available a model expressing the functional dependency of every performance measure established for the design product on every parameter that is either a

design parameter, or a parameter that varies non-deterministically from unit to unit and causes significant variations in unit desirability. Furthermore, there is available the joint statistical distribution of all parameters of the latter type.

Developing models and statistical distributions such as just described, and developing an IC design methodology are research problems that are rather distant, in terms of their knowledge prerequisites and the various characteristics that influence their attractiveness to individual researchers. The theoretical significance of the above assumption is that it limits the scope of the latter research problem to exclude the former.

In practice, the assumption is not restrictive, in that examples of the required models have been in existence for years. The modeling consists of two steps. In one step the functional dependence of the performance measures needed on device and parasitic parameters such as dimensions and saturation currents (bipolar transistors only) are obtained through circuit simulation. In the second step, the modeling of the device and parasitic parameters needed for the other step is generally approached in one of two ways.

In one approach, the needed parameters are divided into two groups - deterministic and non-deterministic. The deterministic parameters consist of design parameters and the non-deterministic parameters are modeled statistically, frequently based on laboratory measurements of device and parasitic parameter values. For examples of this approach, see [DIV78A], [DIV78B], [INO82], and [DIV84]. Considerable statistical modeling of device parameters based on laboratory measurements has been carried out in industry without extensive reporting of results.

In the other approach, some of the needed parameters are design parameters, and the dependence of the remaining parameters on parameters of the wafer fabrication process are

determined using simulation or direct function evaluation, and these process parameters are either design parameters or are modeled statistically, again based on laboratory data. For examples, see [MAL81], [MAL82], and [STY84].

The former approach has significant disadvantages in that distributions of the device parameters of interest must be assumed, and large amounts of data must be collected. This is because it is necessary to infer values not only of device parameter means and variances, but also of a large number of parameter correlations. Furthermore, this must be repeated for all sets of device dimensions of interest. It is not likely that methods for reducing the severity of these difficulties will ever be developed. The main disadvantage of the latter approach, on the other hand, which is the need for sophisticated and accurate models of device parameters in terms of process parameters, can be expected to diminish in severity as the state-of-the art in this type of modeling advances.

(4) Performance Characterization Practices

When an IC is needed for a particular application, and before it is decided that a particular product of a particular producer will be used, the user demands a characterization of the performance of the units of the product. Producers adequately satisfy this need of users through the following actions. They select a set of performance measures for the circuit in question (a qualitative design decision) and a set of one or more real-valued threshold values called *specs levels* (e.g. power consumption of 120 mW), for each performance measure (a parametric design decision). Then they subdivide the totality of operational units produced into two or more categories (a qualitative design decision) based on inequality comparisons of their performance measure values with the specs levels. This is done in such a way that all but one of the categories

are defined by inequalities requiring performance measures of a certain minimal desirability, and the remaining *least desirable* category has no performance constraints, and thus consists of units with performance not qualifying them for membership in any of the other categories. The performance of units to be provided to users is characterized by specifying their performance category membership.

Note that since the units in the least desirable category lack any performance characterization, they generally can make no positive contribution to the economic gain associated with the product. In statistical circuit design work such circuits are referred to as *failed* circuits.

In theory the assumption that the procedure by which producers satisfy the performance characterization needs of users is restrictive. For example, the means of characterizing circuits described above would not be adequate if users were interested in being provided with collections of units with performances conforming to user-prescribed statistical distributions. In practice, neither this particular counterexample nor any other occurs, so the assumption is not at all restrictive.

Note furthermore that this circumstance provides incentive for the producer to select spec levels for the remaining categories that are loose, so the number of failed units that make no contribution to the economic gain is small. On the other hand, making spec levels loose means the desirability of the circuits characterized by these specs levels, as judged by users, decreases, which affects the economic gain adversely. In simplest terms, this is the nature of the tradeoff that is entailed in the selection of spec levels.

1.4.2. Assumptions That Differ Between Existing Formulations and the New Formulation

Listed below are five pairs of assumptions. The first assumption of each pair applies to almost all previous work in the statistical design of IC's, and is identified with a P for "previous". The second assumption of each pair is the assumption of the new formulation that corresponds to the first in the pair, and is identified with an N for "new". Except in cases noted otherwise, each assumption of the new formulation represents a relaxation of the corresponding assumption of previous work.

As each assumption pair is named, the characteristic of comprehensive parametric design methodologies to which it corresponds is noted in parentheses.

(5) Choice of Design Parameters (Characteristic II, Design Parameter Appropriateness")

(P) Each design parameter is a parameter of a distribution used in the statistical modeling (discussed in assumption (3) above) (e.g. [STY81], [ANT81], [HOC83]).

Most frequently the parameters are means of distributions of device dimensions.

Note that with the above assumption, the design parameters cannot include parameters associated with effects that are best modeled deterministically. One example of such parameters is performance specs. Hence these must be prescribed and fixed.

(N) The design parameters must form an independent set of directly implementable parameters that are producer-controlled, specific to the design product optimization, and that in general have a first-order effect on the goodness criterion.

Clearly with this assumption, the new formulation has characteristic IIA of comprehensive design methodologies, since the two are paraphrases of each other. In the

interest of endowing the new formulation with Characteristic IIB as well, it has been an objective in this work to identify and include *all* parameters meeting the description above.

(6) Number of Performance Categories (Characteristic IV, Model Realism and Accuracy)

(P) the number of performance categories the design product has is two.

This assumption does not hold for most integrated circuit designs, and hence is significantly restrictive.

Units that do not fall into the failing category are called *passing* circuits. Also, the fraction of the operational units produced that are passing is called the *parametric yield.* In spite of the restrictiveness of assumption (6P), parametric yield is either the goodness criterion, to be maximized, or, occasionally, a constraint quantity, to be bounded from below, in most methodologies proposed in IC statistical design (e.g. [STY81], [ANT81], [COO82], and [HOC83]).

(N) the number of performance categories the design product has is arbitrary.

(7) Occurrence of Non-operational Units (Characteristic IV, Model Realism and Accuracy)

(P) all units resulting from production of the circuit are operational.

Whether or not this assumption is significantly restrictive depends on the die area of the circuit. For any circuit, regardless of die area, wafer defect phenomena, primarily of the point defect type, may cause individual units to not be operational. While the incidence of non-operational units due to these phenomena is low for circuits of small die area, it can constitute a majority of the units produced for sufficiently large die area.

(N) The incidence of non-operational circuits may cause significant degrading of the goodness criterion

(8) Choice of Goodness Criterion (Characteristic III, Goodness Criterion Appropriateness)

(P) the goodness criterion most appropriate for parametric designs in which assumption 6P holds, is the parametric yield of the design. (The larger the parametric yield, the better.)

In view of the claim that goodness criteria for comprehensive parametric design should involve economic variables, the implications of assumption 8P itself can be understood best in terms of such variables. A basic tenet in the new formulation is that it is possible to associate a *cost* and an *economic return* with a particular IC product. For the purposes of this discussion, these are assumed well-defined, and denoted by c and r, respectively. If n^t denotes the total number of units fabricated, say in a given time interval, n^p denotes the number of units with passing performance (non-operational units are ignored here, as in conventional design methodologies), and y denotes the parametric yield, then,

$$y = \frac{n^p}{n^t}$$

But in a first-order economic model, n^t is proportional to to the cost, and n^p to the economic return. Hence there is a constant, say β, such that

$$y = \beta \frac{r}{c} \tag{1.3}$$

Hence in approaches to parametric design based on assumption 8P, there is an economic rationale underlying the choice of criterion - that maximizing the yield maximizes the ratio of economic return to cost.

The following is a preliminary form of the assumption corresponding to assumption 8P, for the new formulation.

(N′) It is possible to associate a cost and an economic return with a given IC product. Furthermore, the goodness criterion most appropriate for parametric design of IC's is the economic return minus the cost. (The larger the difference, the better.)

This difference is considered to be the economic gain associated with the product. Thus,

$$g = r - c \tag{1.4}$$

Further perspective on assumption 8P can be gained by comparing parametric yield as a criterion with the economic gain criterion just proposed as the ideal. It is intuitive that maximizing y and g as given by the above equations are not equivalent. For example, if in a certain time period one design had an expected economic return of \$110 000 and an expected cost of \$100 000, and a second design had an expected economic return of \$1 100 000 and an expected cost of \$1 000 000, the second would have ten times the expected economic gain than the first. Yet methodologies based on maximizing the ratio of economic return to cost would fail to distinguish between the two designs. More generally, a simple analysis of the functions of two variables in (1.3) and (1.4) shows that the cost and return of any (c, r) pair in which c and r are unequal can be incremented so that the resultant changes in y and g are of opposite signs. Hence parametric yield as a goodness criterion fails to reflect the producer's fundamental interests in manufacturing the design product. Note that this deficiency has its roots in the independence of parametric yield from the rate of production of the design (units per month, say).

Assumption (8N′) implies that it is feasible to model the cost and economic return associated with the design product. The basic structures of IC production cost pertinent to parametric design are the same throughout the industry. However, appropriate models for

economic return depend on the economic environment of the producer. Organizations producing IC's are of two basic types: firms that sell their products on the open market, called *IC houses* here; divisions within firms, called *IC-supplying divisions* here, that turn over their finished IC's to other divisions of the firm for use in the production of electronic systems.

For an IC house, the economic return from a particular IC product is the total *revenue* derived from sales of the product in the open market, and is well-defined and easily computed.

For the case of IC's designed and manufactured by an IC-supplying division, the producer is the firm to which the division belongs. The economic return the firm realizes is revenue from the sale of the electronic systems that use the IC product. There is no natural and obvious way to associate an economic return with the IC in question. One approach is to divide the revenue among the divisions participating in the development of the electronic system product according to some artificial, possibly cost-based rule. Another approach, in which economic return associated with the IC in question can more readily be tied to the performance of the IC, is to use as a simulated economic return an estimate of the price the firm would have had to pay had it purchased IC's of like performance from another firm. In any case, the economic return of an IC produced by an IC-supplying division is less well defined than for one produced by an IC house.

Because of this, and, more significantly, the fact that the IC houses comprise a larger part of the industry than do IC-supplying divisions (both in terms of designs carried out and numbers of units produced), explicit modeling of economic return has been limited to the case of the IC house.

The actual assumption related to assumption 8P above that applies to the new formulation is as follows.

(N) Producers of IC's are IC houses that sell their goods on the open market. It is possible to associate a cost and a revenue with a given IC product. Furthermore, the goodness criterion most appropriate for parametric design of IC's is the revenue minus the cost. (The larger the difference, the better.)

This clearly represents a restriction in generality of the formulation explicitly presented in this dissertation. Unlike the case for assumption pairs 5, 6, and 7, assumption 8N is not in all respects less restrictive than assumption 8P. However, assumption 8N allows application of the new methodology to many more design problems in the real world than does assumption 8P and its prerequisite assumption 6P. Also note that, as already suggested, the cost models presented in this dissertation for the IC house are directly applicable for the *IC-supplying division*. Also, some basic principles of the revenue modeling presented herein apply to that case. Modeling of economic return for the IC-supplying division is an objective deserving of high priority among possible extensions of the formulation.

Under the above assumption, the economic return associated with a product is its revenue, and the economic gain of the product is its profit. Let the symbol r be interpreted from this point on as revenue instead of economic return, and let π denote profit. Then from (1.4) the goodness criterion in the new formulation is written as,

$$\pi = r - c \,. \tag{1.5}$$

Now, finally can be discussed, quite conveniently, previous work intended to have not only Characteristic I, Optimization Framework Utilization, and consideration of statistical

phenomena in fabrication, but also some of the remaining characteristics of a comprehensive methodology given in Section 1.2.

The first known reference to work relevant to comprehensive design techniques, in 1984, was a brief research summary pertaining to the work on which this dissertation is based [RIL84]. Since then, six writings on this subject are known to have appeared [RIL85], [MAL85], [OPA85] [RIL86A], [RIL86B] and [OPA86]. [RIL85], [RIL86A], and [RIL86B] are co-authored by this writer and Alberto Sangiovanni-Vincentelli, are based on the work which is also the subject of this dissertation, and, except for the restriction to the IC house application, present forms of the new methodology that achieve essentially all the characteristics of a comprehensive methodology. [RIL85] is a preliminary report on an early version of the profit model. [RIL86A] is a description that is as detailed in parts as is this dissertation, but that is based on a profit model lacking a number of the features that are included herein. [RIL86B] is a highly condensed presentation of the new methodology based on the full profit model discussed here.

Regarding the work of other researchers [MAL85] essentially makes assumptions 6P, 7P, and 8P, (defined above), and restricts its set of design parameters to what are essentially performance specifications. The closely related papers [OPA85] and [OPA86] in effect assume 7P and 8P. While all three of these papers give attention to assigning economic value to finished circuits, the essential purpose in using a profit criterion rather than a cost criterion is that it uses a meaningful determination of the economic value (revenue here) of finished circuits as a function of their performances. This requires the elimination of the role of designer's subjective evaluations of circuit performances. Achieving this purpose of the use of the profit concept is not attempted in these two papers.

1.5. Remarks on the Focus of The Formulation

The previous section has made it clear that the problem formulation proposed in this dissertation is a multi-faceted generalization of previous formulations. The essence of the proposed formulation is not to be found in any one or two dominant distinguishing assumptions, but in its balanced approach toward achieving the four characteristics that have been described as comprising a formulation for a comprehensive parameter-setting methodology. For example, as has been suggested, most work relevant to achieving comprehensive methodologies has been aimed at achieving Characteristics I, Optimization Framework Utilization, and an aspect of Characteristic IV, Model Realism and Accuracy, namely the inclusion of statistical models for fabrication fluctuations.

Lack of attention to the other characteristics of comprehensive methodologies would be justified if it could be argued that the two characteristics just named solely limit the state of the art of parametric design. But more realistically the emphasis on Characteristics I and IV is attributable to the high visibility to the design engineer of current limitations in the state of the art in these areas: without Characteristic I, parameter selection is a mere art; without realistic consideration of statistical effects under Characteristic IV, the designer is almost powerless to limit the production of useless units of the design. Penalties associated with the failure to achieve Characteristic II, Design Parameter Appropriateness, and Characteristic III, Goodness Criterion Appropriateness, are considerably less tangible.

The treatment of performance spec parameters provides a somewhat more detailed example of imbalanced emphasis in problem formulation.

Any application of the existing formulations has associated with it some choice of design parameters (that does not include performance specifications). Let these design parameters be represented by the vector **u**. This would almost certainly include nominal device dimensions, and possibly means and dispersion parameters needed to characterize the consequences of statistical fluctuations occurring in fabrication. Let **e** be a vector of electrical performance specifications all of which are formal specifications for the product being designed, and all of which are salient to customers, i.e., sufficiently important to influence their selection of which IC product to buy. As has been stated, in existing formulations the parametric yield generally plays a role. It depends on both u and **e**, that is,

$$y = y(\mathbf{u}, \mathbf{e})$$

If no other types of degradation of the economic value of the manufacturing output are considered, the cost c can be written as

$$c = c(\mathbf{u}, y(\mathbf{u}, \mathbf{e})) \tag{1.6}$$

The revenue realized by an IC house depends on **e**, and also on the price at which it sells its circuits, say p. That is,

$$r = r(\mathbf{e}, p) \tag{1.7}$$

Hence, combining (1.5), (1.6), and (1.7),

$$\pi = r(\mathbf{e}, p) - c(\mathbf{u}, y(\mathbf{u}, \mathbf{e})).$$

In terms of this expression, conventional formulations can be characterized as adjusting u in order to either maximize y or, occasionally, to minimize c. But the incentive for using such formulations is that, if after the nominal parametric design of a circuit is complete, a fine-tuning of design parameters according to the formulation is carried out, the resultant decrease in cost will be worth the resources expended. Yet underlying this incentive is that of maximizing profit. This mandates instead a balanced approach in which π is

maximized by adjusting not only u, but also e and p. Among the deficiencies of conventional formulations implicit in this characterization is that they would presumably involve considerable effort and computer time to precisely adjust parameters based on a definition of what a "good" set of specifications is, which has been, to a degree, pulled out of the proverbial air.

One final remark is in order regarding the level of detail to which the profit model of this dissertation is described. The assumptions of the new methodology that have been discussed to this point can be thought of as affecting the highest levels of a hierarchical profit model. The arguments of Chapters 2 through 13 serve to fill out lower levels of the hierarchy. The final expression for expected profit, presented in Chapter 13, contains some functions (with the quantities on which they depend enclosed in parentheses, in standard fashion). For these functions, all the arguments are defined, and considerations in developing their explicit form are discussed, at appropriate places in the dissertation, but for some of these functions no explicit form is ever presented. This omission occurs only for functions belonging to the lowest level of the model hierarchy, and occurs primarily because their most accurate explicit specification depends on properties of the manufacturing facilities used to produce the IC's, and so would differ from one IC house to another.

1.6. Some Notational Conventions

The notation used in the detailed modeling will be introduced as needed throughout the dissertation. However, there is a great deal of it, and it is desirable to discuss some notational conventions at this point. Also, some specific notations are briefly introduced here in concentrated form, because together with the conventions, it allows the pattern in the notations to be seen.

As has already been evident, vectors are in boldface.

Following conventional practice in statistics, random variables are denoted by upper-case letters, and their realizations by the corresponding lower-case letters.

Many variables in the model must be indexed by natural-number indices. All subscripts will be such indices. All letters that serve to qualify the meaning of variables with which they are associated, and which would normally appear as subscripts, appear here as superscripts. There are no variables raised to powers.

Note that C and c always denote cost (with the units of dollars), while c used as a subscript indexes potential chips.

The symbol "1" when subscripted with one or more indices represents an indicator random variable, taking on the values one and zero only.

In conventional mathematical notational practice, if "*expr*" stands for some expression, then an expression of the form $h(expr)$ could represent the function h evaluated at the value *expr*, or it could represent multiplication of h and *expr*. In this dissertation, if the latter interpretation is desired, some type of brackets other than ( ) are used, so that the use of ( ) always stands for the evaluation of a function.

Next some natural-number variables representing total quantities of various objects are defined.

The detailed cost modeling is done for a *batch* of wafers numbering n^w. Formally, then,

n^w number of wafers in a batch.

It is assumed the photolithographic mask set for the product defines two types of patterns on the wafers: one for the *chips* to be packaged and sold, and another for a set of *test patterns* (variously termed "drop-ins", or "process control modules"), for monitoring the effects of fabrication processes on the wafer. Let

N^{cl} number of chip (die) locations on each wafer, as defined by the mask set for the design product.

Note that N^{cl} is almost always in the range from 20 to 2000. Similarly, let

n^t number of repetitions of the test pattern on the wafer.

Note that n^t is typically in the range from 1 to 5.

Now let

n^{dc} number of devices in the chip being designed.

n^{dt} number of devices in the test pattern.

Corresponding to the variables n^w, N^{cl}, n^t, and to either n^{dc} or n^{dt} are defined the natural-number indices w, c, t, and d, as follows. Let,

w index of the wafers in a batch.

c index of the chips on a wafer.

t index of the test patterns on a wafer.

d index of the devices in a chip or test pattern.

These indices are used in combinations. For example, if there is some quantity, say z, which has different values for different devices, its value for device d of chip c on wafer w is denoted z_{wcd}.

1.7. Organization of the Dissertation

This dissertation is organized as follows. In Chapter 2 some fundamental perspectives on the determination of an appropriate form of profit for statistical design of IC's are presented. In Chapter 3 preliminary descriptions of the design parameters of the new formulation are presented. Chapter 4 describes how the methodology would be most appropriately utilized, from a product development perpective. The detailed modeling of the profit explicitly associated with the product to be designed begins in Chapter 5, with a description of the statistical modeling of the fabrication process disturbances. In Chapter 6 is presented the high-level structure of the profit model. Several elements of the model pertaining to the die area of the circuit under design are presented in Chapter 7. Chapter 8 presents the model for the costs of manufacturing operations performed before die slice. In Chapter 9 is modeled the flow of individual circuits that takes place after die slice. The objective is to model the numbers of circuits that emerge from the post-die-slice operations in various categories of interest. The costs of these operations is modeled quite simply in terms of these numbers, in Chapter 10. Chapter 11 presents a detailed discussion of the revenue model of the formulation. In Chapter 12 a class of constraints which tie together the cost and revenue models is derived. In Chapter 13, the models of Chapters 5 through 12 are coalesced into a single statement of the problem formulation of the new methodology. Finally, in Chapter 14, the profit model is summarized, both in words and with block diagrams. Also, the main features of the methodology that is based on the profit model are discussed, including its application to discrete decision-making, and future work is outlined.

Chapter 2

PERSPECTIVES ON PROFIT

In most applications of parametric optimization in engineering, discussions of the modeling aspects of the work may consist of a relatively brief but precise definition of the quantity which is to be the optimization criterion, an identification of the design parameters, and a possibly lengthy derivation of the functional dependence of the former on the latter.

In this application, the optimization criterion has been defined briefly, but not precisely. Equation (1.5) defines profit precisely in terms of revenue and cost. It represents the root of a tree with two branches representing the two additive terms on the right-hand side of the equation. However, both these terms are ambiguous in their application. Like profit, revenue and cost each consist of a summation of terms representing different types of revenues, and costs, respectively. But which types of contributions should be included varies from one economic context to another. So, just as for profit, no brief but precise definition of revenue and cost can be given. Instead, as for profit, revenue and cost can each be thought of as nodes of a tree with multiple branches that represent additive

contributions to them. These branches represent a second level in the tree. Additional levels are in fact needed, the process ending only when terms representing contributions can be defined sufficiently precisely to be treated as leafs of the profit tree. These terms might be explicit functions of design parameters, for example.

The profit model in this dissertation is presented essentially by arguing the existence of the leafs at each node, and moving downward to generate the tree. There is, however, one fundamental issue that cannot be presented under this format - the treatment of the time axis in the profit model. It underlies the definitions of economic variables at all levels of the tree.

The term "profit" is frequently used, incorrectly, to refer to quantities which have the dimensions of monetary value per unit time (and profit data is labeled with the units of dollars, rather than dollars per year, for example). Such quantities are more appropriately referred to as "rates of profit". The rate of profit of a firm could be computed on a daily basis, for example, and reported in units of dollars per day. No rate of profit quantity can reasonably serve as an optimization criterion, because rates of profit fluctuate with time. What may sensibly serve as an optimization criterion, however, is the time integral of an appropriately defined rate of profit quantity, over a specified time interval. The widely worshiped quarterly profit is such a criterion, although of course summation substitutes for integration in its computation.

It remains to select a time interval on which the accumulation of profit is based in the new methodology. The time period which has been selected begins when the newly designed or newly redesigned product is introduced in the market, and, in the ideal application of the methodology, ends when it is expected that at least some of the data used in the

models of the methodology will be updated and the product at least partially redesigned. This time period is called the *design lifetime*. The profit model of the new methodology is an essentially static model based on this lifetime. There is no variable, either continuous or discrete, representing time, and running from beginning to end of the design lifetime. (Of course there is in general a time axis implicit in time-domain circuit simulations required by the methodology.) In static models, economic variables such as costs, revenues, and profit are defined for the design lifetime as a whole. So in particular there is no integration required to compute accumulated profit. For the modeling here, a certain duration is chosen for the design lifetime, based on expected accuracy stability of the models. It is not even required to define a symbol for this duration, for the model equations.

With the treatment of the time axis established, the development of the profit model can begin. The above discussion has implied without so stating that the arguments begin at the root of the single-product profit model tree, which represents the expected profit associated with the design product, (accumulated over the design lifetime). But each firm has the more fundamental interest of maximizing the total expected profit of their entire product line. If changes in the design parameters of one of a firm's products can significantly affect the profit associated with other products of the firm, it is in the interest of the producer to carry out parametric design using a more ambitious methodology than has heretofore been described. Specifically, this ideal methodology would choose design parameters for the firm's entire product line in a single integrated design process that would maximize total expected product line profit.

In practice, simultaneous design of an entire product line would entail an unworkable manipulation of engineering manpower and cannot be seriously considered. Yet it is very

desirable that a comprehensive methodology have some mechanism for treating interactions among the products of the firm.

For this reason, surprising as it may seem, the presentation begins by looking at models for total product line profit. This is just the sum of the profit contributions of the products that comprise the product line. Symbolically, the single-product profit tree that has been described is a sub-tree of a larger tree with root representing product line profit and first-level nodes representing profits of individual profits. The presentation begins at the root of this larger tree. The end objective is not to derive a complete profit model for the entire product line. Instead it will be demonstrated that if a model for the profit explicitly associated with the design product is modified by the addition of revenues associated with other products of the firm that are in the same circuit class, and in some cases as well by the subtraction of a term proportional to the rate of production of the design product, maximizing the resultant optimization criterion will tend to maximize the profit of the entire product line. The somewhat lengthy argument required to demonstrate this begins with the introduction of concepts and relationships that focus on the key role of rates of production in determining product line profit.

2.1. Rates of Production, Fixed and Variable Costs, and Contribution Profit

Before focusing attention on product line profit, it is convenient to introduce one of the design parameters of the new formulation, which of course is primarily concerned with design product profit. The design parameter to be introduced is one that summarizes the rate of production of the design product. It is of interest to examine whether such a parameter might have all the characteristics individual design parameters should have, as listed earlier in Characteristic II of comprehensive formulations.

Clearly a parameter that summarizes the rate of production of the design product would have a first-order effect on its profit. And certainly such a parameter would be specific to the design product optimization, and producer-controlled. However, whether a rate-of-production parameter is directly implementable depends on its precise definition.

The quantity that might seem to be the most natural measure of rate of production, the rate of production of non-failing finished units, is not directly implementable. This is because this rate depends on the rate at which circuits are initiated in production, and on the survival rate, roughly speaking, of the circuits, through production. Attempts to maximize the latter are the subject of many papers in statistical circuit design. Hence the number of finished units is not implementable as a design parameter. However the rate at which circuits are initiated in production is directly implementable in the form of instructions to personnel involved the earliest stages of the fabrication process. Hence rate of production defined in this manner is a directly implementable parameter, has all the characteristics individual design parameters are required to have as described in Characteristic II of advanced formulations, and is a design parameter in the new methodology.

The explicit definition below is based on the design lifetime of the design product and uses the fact that the fundamental indivisible unit of production in the first steps of fabrication is the wafer. n^{wl} is defined as follows, with *wl* standing for wafer, lifetime.

n^{wl} number of wafers of the design product started in production during its design lifetime.

Note that n^{wl} is a dimensionless quantity (number of wafers, not number of wafers per unit time). The quantity is loosely referred to as a rate of production variable because production personnel, in implementing a particular value of n^{wl}, would, at least in principle,

divide it by the known design lifetime (which might be many months long) measured in weeks or days, to compute a weekly or daily rate of production.

Just as in a methodology that designs products one at a time, the rate of production of the product of interest is an appropriate design parameter, in a methodology that designs all products of an IC house simultaneously, rates of production of each of the products are appropriate design parameters. These are defined next. First, let n^P denote the number of products of the firm, and let the products be indexed by $p \in \{1, 2, \ldots, n^P\}$. (The design product is considered among these.) Considerable complexity is avoided with little loss of insight if it is assumed all products share the same design lifetime. Accordingly, n_p^{wl} is defined as follows.

n_p^{wl} number of wafers of product p started in production during the design lifetime of the product line.

In a methodology that designs all products simultaneously, each product would have associated with it additional design parameters of various types. It is not feasible to introduce all of these at this point, but, for example, they include device dimensions. Let

$\boldsymbol{\delta}_p$ a vector of design parameters pertinent to product p.

As already stated, a goodness criterion appropriate for the design of a product line is total expected product line profit. Let

$E\pi^T$ total expected profit from the entire product line, during its design lifetime

The total profit depends on all the product production rates and other design parameter vectors. That is, symbolically,

$$E\pi^T = E\pi^T(n_1^{wl}, n_2^{wl}, \ldots, n_{n^P}^{wl}, \boldsymbol{\delta}_1, \boldsymbol{\delta}_2, \ldots, \boldsymbol{\delta}_{n^P})$$

Now let,

$\mathbf{n}^{wl}$ a vector of all the product rates of production n_p^{wl}

Note that the use of boldface **n** here differentiates this vector from n^{wl}, the unsubscripted variable defined above for use in the design product profit modeling in the sequel.

Also, let,

$\boldsymbol{\delta}$ a vector all the product design parameter vectors δ_p,

With these notations, the problem formulation suggested here is,

$$\max_{\mathbf{n}^{wl}, \boldsymbol{\delta}} E\pi^T(\mathbf{n}^{wl}, \boldsymbol{\delta}) . \tag{2.1}$$

As in the case of expected design product profit, expected product line profit has revenue and cost contributions. Let

c^T total expected cost of production for the entire product line, during its design lifetime

r^T total expected revenue from the entire product line, during its design lifetime

Then,

$$E\pi^T(\mathbf{n}^{wl}, \boldsymbol{\delta}) = r^T(\mathbf{n}^{wl}, \boldsymbol{\delta}) - c^T(\mathbf{n}^{wl}, \boldsymbol{\delta}) . \tag{2.2}$$

Now it is necessary to make the contributions of individual products to the total profit explicit. Note that the arguments used to do this, unless otherwise noted, are applications of microeconomic models applying to a wide range of industries beyond IC production.

Revenue is realized in the form of payments to the producer for circuits delivered, and determining the revenue associated with each product is a mere accounting problem. Let

r_p expected revenue arising from the sale of finished units of product p during the design lifetime of the product line

Then,

$$r^T(\mathbf{n}^{wl}, \boldsymbol{\delta}) = \sum_{p=1}^{n^P} r_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) \; . \tag{2.3}$$

Unlike in the case of revenue, there are contributions to total cost that are not associated with particular products, and which are incurred independent of their rates of production. In economic theory these costs are conventionally called *fixed costs*. The prototypical example of fixed cost is depreciation of capital equipment. For the analysis to follow, let,

c^F total expected fixed cost of production in the design lifetime

The contributions to total cost that are associated with particular products result from performing manufacturing operations on units of production, i.e., either wafers or individual circuits. These contributions have a first-order dependence on rates of production, and hence are called *variable costs*. These include actual costs outlays associated with the various manufacturing operations, such as those for raw materials consumed, and for a large part of the firm's labor cost. As in the case of revenue, it is again a mere accounting problem to associate these actual variable cost contributions with particular products. Let

c_p^A total expected variable cost associated with product p (summed over all manufacturing operations needed to manufacture the product)

The total product line variable cost can of course be formed as the sum over product variable costs. Combining all of these arguments regarding cost suggests the following model.

$$c^T(\mathbf{n}^{wl}, \boldsymbol{\delta}) = c^F(\boldsymbol{\delta}) + \sum_{p=1}^{n^p} c_p^A(\mathbf{n}^{wl}, \boldsymbol{\delta}) .$$

The fixed cost is so named because, as implied by this equation, it is independent of the rates of production n_p^{wl} of the various products. It is a sum of component costs of being in business. A priori, as also implied by the equation, there is no reason to expect it to be independent of all other design parameter values in the methodology, δ_p. In fact, although it will not be argued at length here, close examination of the constituents of c^F has lead to the conclusion that it is also independent of the types of parameters that are treated as design parameters in the new formulation. A typical example of the ingredients of this conclusion is the assumption that the cost of depreciation of IC fabrication equipment is, for a given technology, independent of choice of device dimensions of the circuits it is used to produce.

The independence of the fixed cost term from $\boldsymbol{\delta}$ can be represented simply by dropping it as an argument of c^F in the above equation, which then becomes,

$$c^T(\mathbf{n}^{wl}, \boldsymbol{\delta}) = c^F + \sum_{p=1}^{n^p} c_p^A(\mathbf{n}^{wl}, \boldsymbol{\delta}).$$

This can now be combined with (2.2) and (2.3) to yield an expression for total expected profit,

$$E\pi^T = -c^F + \sum_{p=1}^{n^p} r_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) - \sum_{p=1}^{n^p} c_p^A(\mathbf{n}^{wl}, \boldsymbol{\delta})$$

Now it becomes convenient to define the quantity $E\pi_p$ as,

$$E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) = r_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) - c_p^A(\mathbf{n}^{wl}, \boldsymbol{\delta}) . \qquad (2.4)$$

In microeconomic theory this quantity would be referred to as the (expected) *contribution profit* associated with product p, so named because it is the contribution of product p to the "covering" of fixed costs. With this, the total expected profit can be written as,

$$E\pi^T = -c^F + \sum_{p=1}^{n^P} E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) .$$

The problem formulation of (2.1) has thus been transformed to,

$$\max_{\mathbf{n}^{wl}, \boldsymbol{\delta}} \{-c^F + \sum_{p=1}^{n^P} E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta})\}. \tag{2.5}$$

In the formulation of nonlinear programming problems, whenever a criterion function contains an additive term that is independent of the unknowns of the problem, the additive term has no effect on the solution set, and may be set to zero. This fact will be used repeatedly in this dissertation, and for convenience will be called the *principle of the irrelevance of constants* when its application is clear.

An immediate consequence of this principle is that the problem formulation of (2.5) is equivalent to,

$$\max_{\mathbf{n}^{wl}, \boldsymbol{\delta}} \{\sum_{p=1}^{n^P} E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta})\}. \tag{2.6}$$

Fixed costs comprise a larger portion of total manufacturing costs in the IC industry than in most industries, frequently dominating variable costs. Yet as already implied, they do not explicitly enter in the profit model of the new methodology, a fact that permits considerable model complexity, and data gathering for the actual computations, to be avoided.

The sum in (2.6) is appropriately referred to as the total expected contribution profit of the product line. For later convenience, this is denoted π^{cT}. That is,

$$E\pi^{cT} = \sum_{p=1}^{n^P} E\pi_p. \tag{2.7}$$

The problem formulation of (2.6) presents an opportunity to make the purpose of examining product line profit more explicit. Each of the product contributions in (2.6)

depends on the design parameters of the entire product line. Clearly if the expected profit of each product depended only on the design parameters explicitly associated with the product, the problem could be decomposed into a collection of non-simultaneous, product-by-product designs. Formulation (2.6) greatly overstates the degree of interaction among product profits, but as suggested at the outset, to assume complete decoupling is to ignore interaction effects which may approach first-order significance. One of the terms in (2.6) represents profit associated with the design product. The purpose of this discussion is to derive modifications that can be made to expressions for explicit design product profit to account for interactions between design product profit and that of other products, in order to improve the capability of the new methodology to design the design product in a way that tends to maximize $E\pi^T$.

The product profit interactions needing to be treated consist of two forms of competition among the products of the firm - for its manufacturing resources, and for its customers. The former is treated first, in the next section.

2.2. Maximization of Product Line Profit with Capacity Constraints

It is well known that, in the IC industry, there occur periods of low sales volume when firms operate their fabrication lines below their capacities, in order to reduce variable costs. However, most of the time IC producers operate their fabrication lines at full capacity, and wafers cannot be produced at a higher rate no matter how economically attractive it might be to do so. In this case producers are faced with a short-term capacity constraint (a constraint that could be alleviated, but not quickly, and not without significant capital investment). It is important to modify the formulation of (2.6) to account for this phenomenon and determine the implications for the design of individual products.

The model for production capacity used in this work is now introduced. It is assumed the producer has a number of wafer fabrication lines each of which is capable of performing all fabrication steps needed to manufacture at least one product. Let,

n^f number of wafer fabrication lines the producer has

The fabrication lines of the producer are indexed by $f \in \{1,2,\ldots,n^f\}$. It is essentially universal in the industry that all the processing steps needed to produce each product are performed in only one fabrication line. Hence each product has associated with it a unique fabrication line in which its wafers are manufactured. For brevity, this fabrication line will be referred to as the fabrication line in which the product is manufactured, or in which it is made, even though much of the manufacture of IC's is not wafer processing.

Now let,

n_f^p number of products having their wafers manufactured in product line f

and for $p \in \{1,2,\ldots,n_f^p\}$, let,

n_{fp}^{wl} number of wafers of product p of fabrication line f started in production during the design lifetime

Furthermore it is assumed that, for each fabrication line, there is, in the short term, some natural-number upper bound on the total number of wafers, summed over all products, that the line can produce during the design lifetime. Let,

n_f^{wT} maximum number of wafers that can be produced in fabrication line f during the design lifetime

The competition among products for manufacturing capacity, then, can be expressed as,

$$\sum_{p=1}^{n_f^p} n_{fp}^{wl} \leq n_f^{wT}, f = 1,2,\ldots,n^f . \tag{2.8}$$

It is also assumed known whether or not each fabrication lines is expected to operate at full capacity during the design lifetime. This information is represented for fabrication line f by the 1,0 indicator 1_f^{fc}. That is,

$$1_f^{fc} = \begin{cases} 1 & \text{if fabrication line } f \text{ at full capacity} \\ 0 & \text{otherwise} \end{cases}$$

(The superscript fc here is intended to stand for *full capacity*.) Then the optimization problem to be solved can be represented as,

$$\max_{\boldsymbol{\delta}, \mathbf{n}^{wl}} \{ \sum_{p=1}^{n^p} E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) \mid 1_f^{fc}[\sum_{p=1}^{n_f^p} n_{fp}^{wl} - n_f^{wT}] = 0, f = 1,2,\ldots,n^f \}, \tag{2.9}$$

where it should be understood that the constraint $0 = 0$ simply represents a missing constraint.

In order to study the effect of the capacity constraints, it is desirable to focus attention on the optimization of contribution profit with respect to $\mathbf{n}^{wl}$ alone. First observe that the problem of (2.9) can in principle be solved in the following nested form. (Solving an optimization problem here means finding the set of all locally optimum points.)

$$\max_{\boldsymbol{\delta}} \max_{\mathbf{n}^{wl}} \{ \sum_{p=1}^{n^p} E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) \mid 1_f^{fc}[\sum_{p=1}^{n_f^p} n_{fp}^{wl} - n_f^{wT}] = 0, f = 1,2,\ldots,n^f \}.$$

The problem of interest, then, is the inner optimization,

$$\max_{\mathbf{n}^{wl}} \{ \sum_{p=1}^{n^p} E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) \mid 1_f^{fc}[\sum_{p=1}^{n_f^p} n_{fp}^{wl} - n_f^{wT}] = 0, f = 1,2,\ldots,n^f \}, \tag{2.10}$$

with δ held fixed.

Considerable insight into this problem can be obtained by examining in some detail the typical dependence of the contribution profit of a single arbitrarily chosen product on its rate of production, for the case that the production line in which it is produced is not operating at full capacity. $E\pi_p$ is given in terms of r_p and c_p^A in (2.4). Figure 1a contains a plot of the general form of r_p and c_p^A versus n_p^{wl} with δ_p held fixed. The cost curve passes through the origin by definition, and the linearity of the model of c_p^A is widely accepted. The form of the revenue can be explained as follows. Almost universally, as was stated earlier, the product is produced in several versions which sell at different prices. At a sufficiently low value of n_p^{wl}, all versions of the product "sell out", and revenue is proportional to n_p^{wl}. As n_p^{wl} increases, however, a point is reached where one of the versions of the product ceases to sell out, whereafter further increase in revenue comes only from those versions still selling out. Eventually a point is reached beyond which none of the versions sell out, and revenue is independent of n_p^{wl}. In Figure 1b is the implied plot of $E\pi_p$. It is a piecewise-linear unimodal function. Suppose for the sake of argument that instead it were a differentiable unimodal function. With no capacity constraint, the maximizing value of n_p^{wl} would be that corresponding to the unique maximum of the curve, at which

$$\frac{\partial E\pi_p}{\partial n_p^{wl}} = 0 .$$

Now if the maximum capacity constraint of the line in which the product in question is manufactured is considered hypothetically to decrease to the point where the capacity constraint becomes active, and beyond, this forces the maximizing values of n_p^{wl} to decrease for all products sharing the same fabrication line as the product in question. (For a

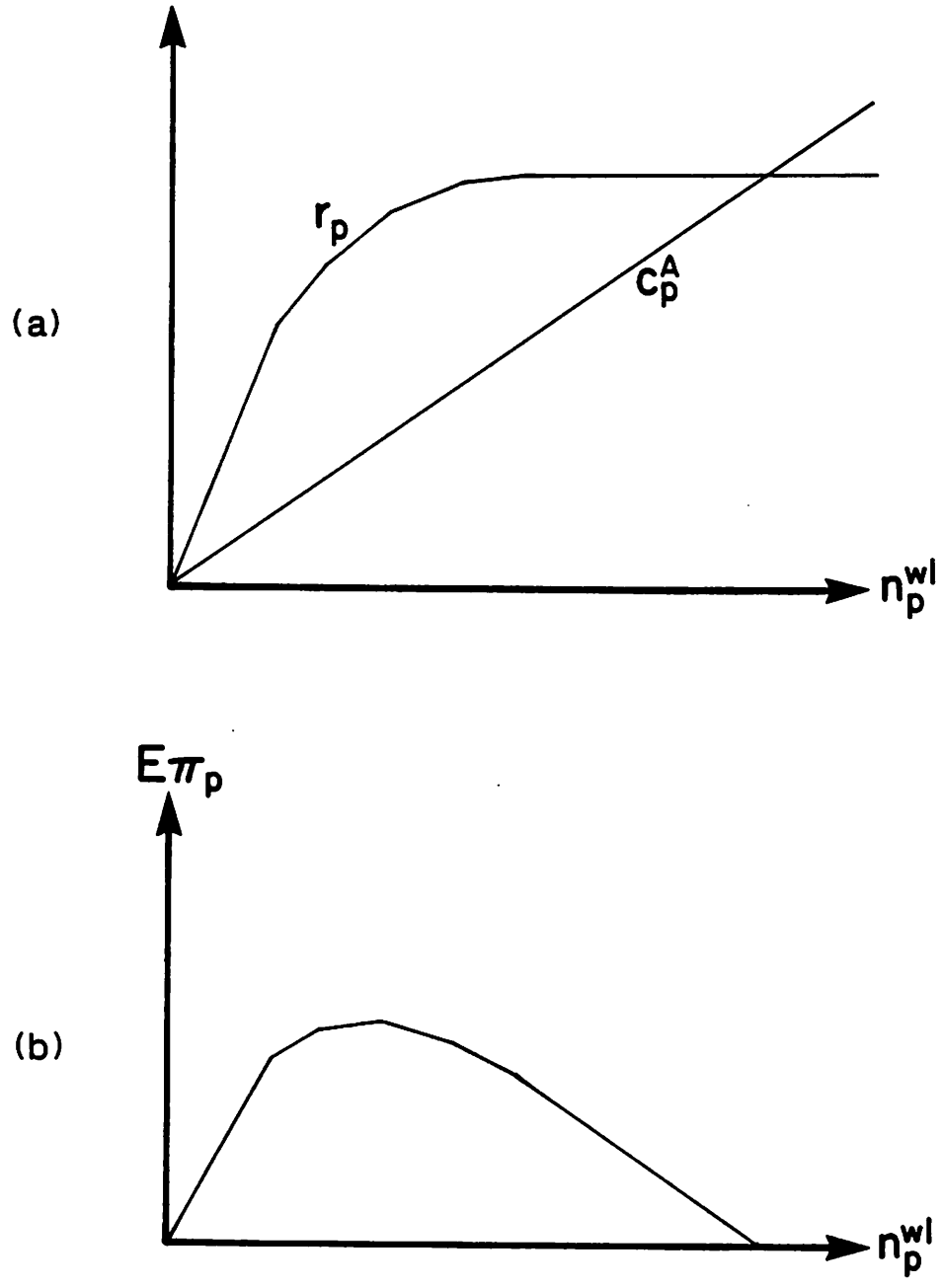


Fig. 1. Economic variables pertaining to IC product p, as a function of the number of wafers of the product started in fabrication during the design lifetime of the product line. (a) Variable cost and revenue. (b) Expected contribution profit.

production rate to increase would be counterintuitive since this would both decrease its profit and exacerbate the shortage of capacity.)

Necessary conditions for the profit-maximizing values of the production rates can be obtained using a Lagrange multiplier analysis. First it is necessary to introduce some additional notation. Let $\mathbf{n}_f^{wl}$ be defined by,

$$[\mathbf{n}_f^{wl}]' = [n_{f1}^{wl}\ n_{f2}^{wl} \cdots n_{fn^p}]'$$

where $'$ denotes matrix transpose. And let,

$\mathbf{u}^{n_f^p}$ a column vector of length n_f^p with each of its elements unity

For problems in which the theory of Lagrange multipliers applies, it asserts, in words, that there exists a linear combination of the gradients of the constraints which when added to the gradient of the criterion function, produces the zero vector. Here, it is convenient to view gradient vectors, which consist of partial derivatives with respect to elements of $\mathbf{n}^{wl}$, as being divided into partitions with one partition for each of the n^f fabrication lines. The resultant Lagrange multiplier equations are similarly partitioned. Explicitly, the assertion is that (for each $\boldsymbol{\delta}$) there exists a set of real numbers denoted for this discussion as $c_1^{O/w}, c_2^{O/w}, \ldots, c_{n_f}^{O/w}$ such that the following equations hold.

$$[\frac{\partial E\pi^{cT}(\mathbf{n}^{wl}, \boldsymbol{\delta})}{\partial n_f^{wl}}] - c_f^{O/w} 1_f^{fc} \mathbf{u}^{n_f^p} = 0\ ,\ f = 1,2,\ldots,n^f\ . \tag{2.11}$$

where use of the total contribution profit $E\pi^{cT}$ defined in (2.7) has been made for notational simplicity. Note that the real numbers $c_1^{O/w}, c_2^{O/w}, \ldots, c_{n_f}^{O/w}$ are in one-to-one correspondence with fabrication lines. For the problem at hand, they play the role of the Lagrange multipliers.

In words, (2.11) says that at the production rates that maximize profit subject to capacity constraints, the slopes of the profit versus production rate curves for all products sharing the same fabrication line must be the same. Note that since the maximizing production rates with active constraints are less than their unconstrained values, these slopes are positive, and hence by (2.11) so are the $c_f^{O/w}$.

The following step is key in deriving a modification to the profit model that accounts for fabrication line capacity constraints. The method of Lagrange multipliers allows conversion of an equality-constrained optimization problem to an equivalent unconstrained problem having the same necessary conditions. Here, the necessary conditions are those of (2.11) (together with the equality constraints) and the unconstrained problem equivalent to (2.10) is,

$$\max_{\mathbf{n}^{wl},\,\mathbf{c}_f^{O/w}} \{ \sum_{p=1}^{n^P} E\pi_p(\mathbf{n}^{wl},\delta) - \sum_{f=1}^{n^f} c_f^{O/w} 1_f^{fc} [\sum_{p=1}^{n_f^p} n_{fp}^{wl} - n_f^{wT}] \} , \tag{2.12}$$

where $\mathbf{c}^{O/w}$ is a vector of the Lagrange multipliers. Note that the modification to the total explicit product line profit is an additive one.

Considerable intuitive motivation for this result can be obtained by considering the optimization of this criterion function with respect to $\mathbf{n}^{wl}$ alone. That is, as in an earlier equivalence, the problem of (2.12) can be considered equivalent to the nested formulation,

$$\max_{\mathbf{c}_f^{O/w}} \max_{\mathbf{n}^{wl}} \{ \sum_{p=1}^{n^P} E\pi_p(\mathbf{n}^{wl},\delta) - \sum_{f=1}^{n^f} c_f^{O/w} 1_f^{fc} [\sum_{p=1}^{n_f^p} n_{fp}^{wl} - n_f^{wT}] \} .$$

Consider the inner optimization. In seeking maximizing $\mathbf{n}^{wl}$ values, by the principle of irrelevant constants, all terms of the form $c_f^{O/w} 1_f^{fc} n_f^{wT}$ can be omitted, so that the formulation,

$$\max_{\mathbf{n}^{wl}} \{ \sum_{p=1}^{n^p} E\pi_p(\mathbf{n}^{wl}, \delta) - \sum_{f=1}^{n^f} c_f^{O/w} 1_f^{fc} \sum_{p=1}^{n_f^p} n_{fp}^{wl} \}$$

will do. Comparison with (2.10) shows that the criterion function of this formulation has been decremented by a term for each product made in a full-capacity fabrication line, equal to the production rate of the product times a coefficient associated with the fabrication line in which it is manufactured. Since these terms detract from profit, they can be viewed as artificially introduced cost terms. By the necessary condition of (2.11), the coefficient $c_f^{O/w}$ can be thought of as the increase in the profit realized by each product manufactured in fabrication line f, per wafer increase in its production rate. Subtracting the artificial cost term corresponding to a particular product made in a full-capacity fabrication line, is a reflection that a decision to increase the production rate of the product forces a decrease in the production rates of other products in the line, and hence a loss of opportunity to make a profit through those products. Accordingly, such cost terms are called *opportunity costs*, in microeconomics. They are exclusive of the actual variable costs introduced earlier, and are introduced to optimally reduce production rates from their unconstrained values, to the point that the capacity constraints are satisfied. Note that the notation $c_f^{O/w}$ arises here from from its description as the *opportunity cost per wafer of fabrication line f*, or *opportunity cost coefficient of fabrication line f.*

Continuing with the thread of the argument that lead up to (2.12), it is now convenient to consider the effects of relaxing the assumption operative since (2.10), that δ is fixed. If the intervening discussion is reviewed, it is not difficult to see the following: the fundamental basis of the argument is the existence of the Lagrange multipliers now referred to as opportunity costs per wafer; certainly these parameters may vary with δ;

allowing the opportunity costs per wafer to be functions of δ is the only required modification of the discussion.

With this, and since (2.12) is an equivalent formulation to (2.10), it is possible to state the following as the conclusion of this section of the chapter. The maximization of total product line contribution profit subject to possibly active capacity constraints on fabrication lines, as formulated in (2.9), can be based on the formulation,

$$\max_{\delta,\, \mathbf{c}^{O/w},\, \mathbf{n}^{wl}} \{ \sum_{p=1}^{n^p} E\pi_p(\mathbf{n}^{wl}, \delta) - \sum_{f=1}^{n^f} c_f^{O/w}(\delta)\, 1_f^{fc} [\sum_{p=1}^{n_f^p} n_{fp}^{wl} - n_f^{wT}] \} \,. \tag{2.13}$$

2.3. Competition for Customers, and Model Reduction

The second form of competition between the design product and other products of the firm, that needs to be addressed, is competition for customers. The effect is simple.

Let d denote the value of the product index p representing the design product. Any change in the design parameter vector δ_d of the design product that increases, say, the desirability of product d, may cause customers of the product type in question to switch their purchase choice to this product. If the product from which a particular customer switches is that of some other firm, the only profit term affected is π_d. But the firm may have products other than product d with the same electrical functionality, called *brother products* of product d. If the product from which the customer switches is a brother of product d, the revenue and hence the profit of the brother product is decreased. This unpleasant outcome of engineering diligence is conventionally called *cannibalism.*

In order to obtain an appropriate treatment of the effect, it is convenient to write the criterion of (2.13) in a particular expanded form. Additional notation will need to be

introduced in the process.

First, it is convenient to affect a change of notation and rewriting of (2.13) in a different form. Since it has been assumed that each product has a unique fabrication line associated with it, it is possible to define a map from products to fabrication lines. This function is denoted $f(\,)$, that is, let,

$f(p)$ index of the fabrication line in which wafers of product p are manufactured.

Then by equality of their criterion functions, the formulation of (2.13) is equivalent to the following formulation,

$$\max_{\boldsymbol{\delta},\, \mathbf{c}^{O/w},\, \mathbf{n}^{wl}} \{ \sum_{p=1}^{n^p} [E\pi_p(\mathbf{n}^{wl}, \boldsymbol{\delta}) - c^{O/w}_{f(p)}(\boldsymbol{\delta})\, 1^{fc}_{f(p)}\, n^{wl}_p] + \sum_{f=1}^{n^f} c^{O/w}_f(\boldsymbol{\delta}) 1^{fc}_f n^{wT}_f \} \qquad (2.14)$$

where it should be clear that all of the $c^{O/w}_{f(p)}$ are in fact elements of $\mathbf{c}^{O/w}$. Note that the symbol f without arguments continues to be used as before as a free-running index.

Next, two subsets of $1, 2, \ldots, n^p$ are defined. Let

Ψ^b the set of indices all of the firm's products that are brothers of the design product

and,

$\Psi^{\bar{b}}$ the set of indices of all of the firm's products that are neither the design product nor its brothers.

With this notation, and applying the definition of (2.4) for the brother products, the formulation of (2.14) is equivalent by equality of criterion function to,

$$\max_{\delta,\, \mathbf{c}^{O/w},\, \mathbf{n}^{wl}} \{E\pi_d(\mathbf{n}^{wl}, \delta) - c^{O/w}_{f(d)}(\delta)\, 1^{fc}_{f(d)}\, n^{wl} \tag{2.15}$$

$$+ \sum_{p \in \Psi^b} [r_p(\mathbf{n}^{wl}, \delta) - c^A_p(\mathbf{n}^{wl}, \delta) - c^{O/w}_{f(p)}(\delta)\, 1^{fc}_{f(p)}\, n^{wl}_p]$$

$$+ \sum_{p \in \Psi^{\bar{b}}} [E\pi_p(\mathbf{n}^{wl}, \delta) - c^{O/w}_{f(p)}(\delta)\, 1^{fc}_{f(p)}\, n^{wl}_p]$$

$$+ \sum_{f=1}^{n^f} c^{O/w}_f(\delta)\, 1^{fc}_f\, n^{wT}_f \} \,.$$

Now it is necessary to define vectors with elements that are subsets of the elements of n^{wl}. In order to limit notational complexity in so doing, the superscript wl is temporarily dropped from all rate of production variables. In minor abuse of the convention that indices of vector elements are written as subscripts, let n^d denote the number of wafers of the design product started in production during the design lifetime. This has previously been denoted as n^{wl}_d. And let,

$\mathbf{n}^b$ a vector of the numbers of wafers started in production of the products in Ψ^b, during the design lifetime.

$\mathbf{n}^{\bar{b}}$ a vector of the numbers of wafers started in production of the products in $\Psi^{\bar{b}}$, during the design lifetime.

The design parameter vectors δ^d, δ^b, and $\delta^{\bar{b}}$, are analogously defined vectors with elements from δ. (Note that δ^d, unlike n^d, is a vector.)

With these notations, the formulation of (2.15) can by equality of criterion functions be expanded to,

$$\max_{\delta, \mathbf{c}^{O/w}, \mathbf{n}} \{ E\pi_d(n^d, \mathbf{n}^b, \mathbf{n}^{\bar{b}}, \delta^d, \delta^b, \delta^{\bar{b}}) - c^{O/w}_{f(d)}(\delta)\, 1^{fc}_{f(d)} n^d$$

$$+ \sum_{p \in \Psi^b} [r_p(n^d, \mathbf{n}^b, \mathbf{n}^{\bar{b}}, \delta^d, \delta^b, \delta^{\bar{b}}) - c^A_p(n^d, \mathbf{n}^b, \mathbf{n}^{\bar{b}}, \delta^d, \delta^b, \delta^{\bar{b}}) - c^{O/w}_{f(p)}(\delta)\, 1^{fc}_{f(p)}\, n_p]$$

$$+ \sum_{p \in \Psi^{\bar{b}}} [E\pi_p(n^d, \mathbf{n}^b, \mathbf{n}^{\bar{b}}, \delta^d, \delta^b, \delta^{\bar{b}}) - c^{O/w}_{f(p)}(\delta)\, 1^{fc}_{f(p)}\, n_p]$$

$$+ \sum_{f=1}^{n^f} c^{O/w}_f(\delta)\, 1^{fc}_f\, n^{wT}_f \, .$$

Now the process of reducing this formulation to one appropriate for single-product design begins. Neither the actual costs or revenues of the design product or its brothers depend on the rates of production or the other design parameters of the products that are not their brothers. Conversely, neither the actual costs or revenues of the products in $\Psi^{\bar{b}}$ depend on the production rates or other design parameters of the design product and its brothers. Furthermore, the actual costs of the brother products do not depend on the rate of production or the other design parameters of the design product. Therefore, the functional dependencies in the above formulation can be simplified to,

$$\max_{\delta, \mathbf{c}^{O/w}, \mathbf{n}} \{ E\pi_d(n^d, \mathbf{n}^b, \ , \delta^d, \delta^b, \ ) - c^{O/w}_{f(d)}(\delta)\, 1^{fc}_{f(d)} n^d \tag{2.16}$$

$$+ \sum_{p \in \Psi^b} [r_p(n^d, \mathbf{n}^b, \ , \delta^d, \delta^b, \ ) - c^A_p(\ , \mathbf{n}^b, \ , \ , \delta^b, \ ) - c^{O/w}_{f(p)}(\delta)\, 1^{fc}_{f(p)}\, n_p]$$

$$+ \sum_{p \in \Psi^{\bar{b}}} [E\pi_p(\ , \ , \mathbf{n}^{\bar{b}}, \ , \ , \delta^{\bar{b}}) - c^{O/w}_{f(p)}(\delta)\, 1^{fc}_{f(p)}\, n_p]$$

$$+ \sum_{f=1}^{n^f} c^{O/w}_f(\delta)\, 1^{fc}_f\, n^{wT}_f \, ,$$

where by minor abuse of notation, blank spaces are left in place of design parameters that do not affect the term in which they appear.

Almost the entire discussion in this chapter on perspectives in profit has pertained to the development of a formulation appropriate for the simultaneous integrated design of an entire product line. The above formulation, notwithstanding its awkwardness and asymmetry, is such a formulation. Now, however, the reality must be faced that the design lifetimes (specified by an introduction date and a cancellation date) assigned to different products in carrying out their design, do not coincide.

There are three major issues that arise in attempting to apply the above formulation to single-product design. Resolving them calls for less of concise mathematical arguments and more of the art of modeling.

First, it is not clear which products should be included in the summations of (2.16). Certainly all existing products, i.e. products the designs of which have been determined at the time the new methodology is applied to a new product, and that have not been cancelled, should included. And it seems clear that there is no way to anticipate the design parameter values and contributions to profit of products that have yet to be designed. Hence in the best possible treatment, the products included should consist of all the existing products of the firm.

The second, more substantial issue is that heretofore, economic variables for existing products other than the design product have been defined in terms of a common product-line design lifetime that is not defined in the context of single-product design. It is necessary to redefine a design lifetime of other products for purposes of accounting in the design of the product of interest. This redefined lifetime is temporarily referred to as the *accounting lifetime*.

Clearly there is no point in including contributions to profit from other products, that occur prior to the lifetime of the design product. Evidently the accounting lifetime of other products should start with the lifetime of the design product. There are two conceivable points at which the accounting lifetime might end - at the end of the design product lifetime, or at the end of the lifetime assigned to the other product when it was designed. If the former precedes the latter in time, clearly the accounting lifetime should end with that of the design product, to avoid overestimate of profit contributions. If the latter precedes the former, again the accounting lifetime may be taken to end with the end of the design product lifetime, even though sales may have been discontinued at an earlier point in time. In other words, suppose, for example, that for products other than the design product, n_p^{wl} is defined according to,

r_p expected revenue arising from the sale of finished units of product p during the design lifetime of the design product

Then the appropriate accumulated revenue is included regardless how it accumulates over time. Hence the accounting lifetime for products other than the design product is taken to be the design lifetime of the product of interest, and all economic quantities pertaining to other products, including π_p, c_p^A, and n_p^{wl}, are defined on this basis, as was r_p above.

The third issue arising in the application of (2.16) to single-product design is determining the appropriate role of $c^{O/w}$, since it cannot remain an optimization variable. It is intuitive that, if the design product is manufactured in a full-capacity line, its profit should be penalized for increases in its production rate, to account for lost profits from other products made in the same line. Furthermore, a penalty proportional to its production, as are those in the above formulations, seems quite adequate to first order. But it does not seem

clear what should be the value of the opportunity cost coefficient, say $c_d^{O/w}$ of the fabrication line in which the design product is made, and it may seem that the concept of opportunity cost must be a casualty of the infeasibility of simultaneous design.

On the other hand, the typical fabrication line is used to produce at least a moderately large number of products. In a simultaneous design methodology, it is unlikely that any one product would have a dominant effect on the optimizing value of $c_d^{O/w}$. Thus its value can be thought as primarily determined by the other products. This suggests the opportunity cost coefficient be treated as a constant the value of which is based on the incremental variation of the profit of existing products on their rates of production. One of the equations in (2.11) pertains to the fabrication line of the design product. It is proposed that the $c_{f(d)}^{O/w}$ be taken as the average of the $(n_f^p - 1)$ partial derivatives of the gradient vector in that equation, that correspond to existing products other than the design product. No attempt is made here to further detail the estimation of these partial derivatives, since information from both design simulations and actual historical data are potentially useful, and the availability and form of such information is expected to differ from firm to firm.

Note that this approach salvages the concept of opportunity cost, but there is no means to determine how the resultant $c_d^{O/w}$ varies with any of the elements of δ. On the other hand, this dependence was presumed in the previous section, merely for lack of reason to assert that $c_d^{O/w}$ is independent of δ; there is no known reason to believe there is a strong dependence.

Presuming that the other products were also designed using the new methodology, it seems likewise reasonable to treat their opportunity cost coefficients as constant. Hence all opportunity cost coefficients are treated as constant.

This completes the resolution of the three issues arising in the application of (2.16) to single-product design. The next and final step is based on the assumption that, in general, the design parameters of products other than the design product have been fixed at the time this product is designed. Hence these must be treated as constant. These consist of $\mathbf{n}^b$, δ^b, $\mathbf{n}^{\bar{b}}$, and $\delta^{\bar{b}}$. It has already been argued that all opportunity cost coefficients be treated as constant. Note also the n_f^{wT} are given constants (see (2.8)). With all these parameters constant, by the principle of irrelevant constants, the single-product formulation corresponding to (2.16) is,

$$\max_{n^d,\,\delta^d} \{\, E\pi_d(n^d,\delta^d) - c^{O/w}_{f(d)}(\delta)\, 1^{fc}_{f(d)}\, n^d + \sum_{p\in\Psi^b} r_p(n^d,\delta^d) \,\} \qquad (2.17)$$

Note that the severity of assumptions made in addressing non-simultaneity of design lifetimes, such as the choice of accounting lifetimes, is diminished by the fact that the only type of economic variable in (2.17) directly dependent on the accounting lifetimes, is brother product revenue.

Modulo the inclusion of terms to account for opportunity cost and brother product revenue, the remainder of the dissertation focuses on design product profit modeling. It is of interest to write (2.17) in a notation better suited to this subject. First it is desirable to revert to n^{wl} instead of n^d, for the design product rate of production. Second, the product identifier d standing for design product can be dropped from π and δ. And the product $c^{O/w}_{f(d)}\, 1_{f(d)}$ can be denoted as $c^{O/w}$, with the understanding that the opportunity cost coefficient used should be that of the fabrication line in which the design product is made, and that if that line is not operating at full capacity, $c^{O/w}$ should be zero. With these notational changes, (2.17) becomes,

$$\max_{n^{wl},\delta} \{ E\pi(n^{wl},\delta) - c^{O/w} n^{wl} + \sum_{p \in \Psi^b} r_p(n^{wl},\delta) \}$$

The important perspectives presented in this chapter on the interpretation of profit most appropriate for the new methodology have lead to the following conclusion. If the criterion on which the methodology is based is the expected value of the contribution profit associated with the starting in production of n^{wl} wafers of the design product, plus the revenues associated with its brother products, minus the rate of production of the design product times the opportunity cost coefficient of the design product's fabrication line, maximizing the resultant optimization criterion will tend to maximize the profit of the entire product line. In this description, it should be understood that if the design product is not manufactured in a full-capacity fabrication line, the opportunity cost coefficient is zero.

Chapter 3

IDENTIFICATION OF DESIGN PARAMETERS

One effective means of providing an overview of an optimization-based design methodology is to identify its goodness criterion and its design parameters. The meaning of the optimization criterion of the methodology has at this point been sufficiently refined for this purpose, but aside from the design product rate of production n^{wl}, already introduced, little has been said about the identity of the design parameters. It is appropriate to do so now, especially because it is needed for the discussion in the next chapter of the practical utilization of the methodology.

Paraphrasing the discussion at the beginning of Section 1.4.2, it has been an objective in the development of the new methodology to identify and include as design parameters all directly implementable parameters that are producer-controlled, specific to the design product optimization, and that in general have a first-order effect on the goodness criterion,

unless they destroy independence of the complete set of design parameters.

It has not proven particularly difficult to apply these criteria in selecting the design parameters for the methodology. Nevertheless, just as it is difficult to briefly and precisely define the goodness criterion of the methodology, as explained at the beginning of Chapter 2, it is also not feasible to briefly and precisely define the remainder of its design parameters, and justify their status, in advance of the detailed model development details presented later. The list below is presented merely to provide a preliminary familiarity with the design parameters. (n^{wl} is included for completeness.)

- n^{wl} the number of wafers of the design product started in production during its design lifetime.
- $\mathbf{l}$ a vector of process test limits.
- $\mathbf{x}$ a vector of device dimensions.
- $\mathbf{e}$ a vector of electrical specifications for the product.
- $\boldsymbol{\alpha}$ a vector of fractions used in specifying the proportions of circuits which are subjected to various processing operations which take place after the wafers are sliced into individual dice. Among these operations are the packaging and testing.

In addition to these, there are two types of design parameters that pertain to product revenue. Note that the model recognizes the existence of different versions of the design product, which are distinguished primarily by their electrical grade and their package type. The additional parameters are,

- $\mathbf{p}$ a vector with each element equal to the average selling price of one version of the product.

q a vector with each element equal to the quantity sold of one version of the product.

In the real world the dependence of the expected contribution profit on these parameter types is rather complex. In any reasonably accurate model of this profit, the different parameter types would interact with one another in their effects on the profit. That is, optimal values of the parameters of any given type depend on the values of the remaining types of parameters.

An important perspective on the new methodology can be gained by considering how these design parameters are set in conventional product design practice. Figure 2 is a representation of the portion of the organizational structure of a typical IC house which is pertinent to product design, together with a block representing potential customers of the firm. Adjacent each organizational unit is listed a design parameter type, except for Marketing, which has three such types listed. In conventional practice, each unit is given authority for the setting of their associated design parameter type(s). In setting their parameters, each unit seeks and receives quantitative information from other units, generally in the form either of verbal or printed information, or possibly of plots. Then it sets its parameters by "satisficing" some measure of its performance which has been defined by management.

There are three primary limitations of this product design methodology:

(1) The units are satisficing their performance measures, rather than maximizing them. This is primarily for lack of appropriate tools.

(2) The collective effect of the performance measures of the various units is not necessarily to maximize profit. One example pertains to the Wafer Fabrication

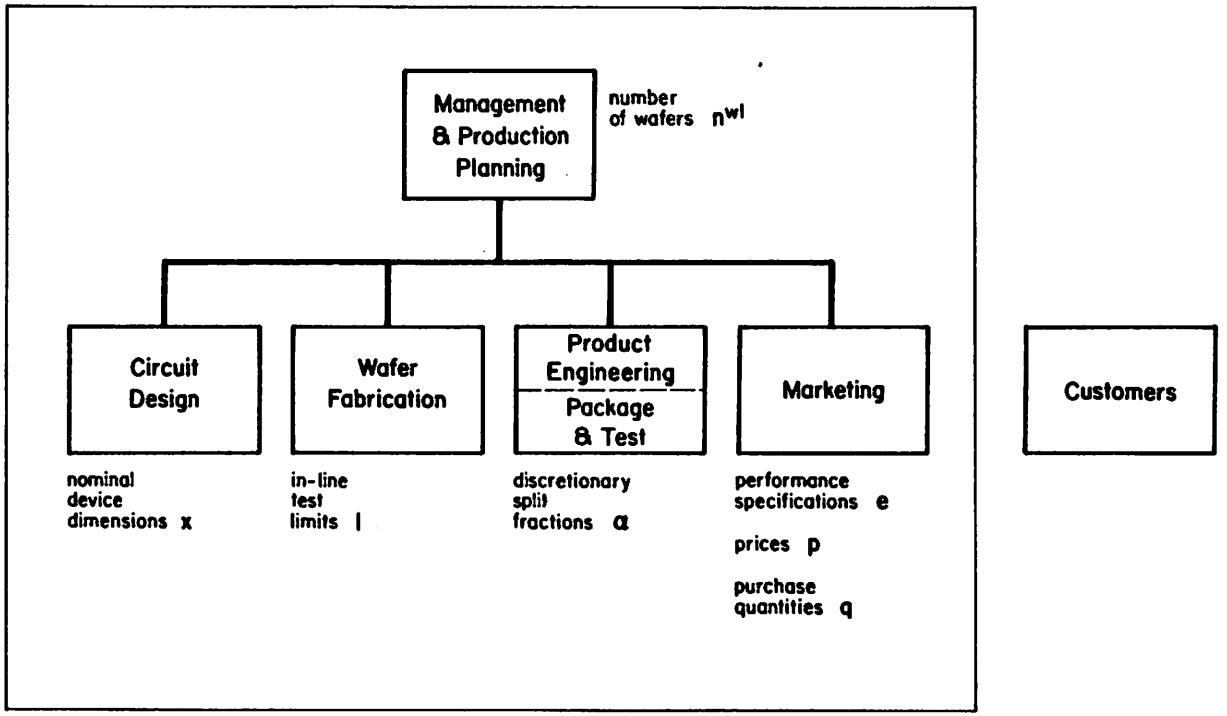


Fig. 2. Typical organizational structure of an IC house, as it pertains to the conventional determination of the design parameters of the new methodology. (Customers also depicted.)

Unit. Its performance measure is essentially the number of wafers it produces in a specified time interval, given a specified set of equipment and manpower resources. This means Wafer Fabrication has an incentive to set in-line test limits loosely. But if the limits are set too loosely, wafers which are destined to yield poorly when their individual circuits are tested later, are continued in production, which can represent a degrading of profit.

The attempt to bring elements of the parametric design of IC products typically performed in different organizational units of the firm into a unified optimization framework has been a important goal of this work.

Chapter 4

OVERVIEW OF THE METHODOLOGY UTILIZATION

The proposed design methodology consists of a collection of coordinated activities built around a central core consisting of an optimization iteration which, given one point in the design parameter space, produces another for which the expected contribution profit is greater. In order to describe this in more detail, it is necessary to first introduce three types of activity required in the routine utilization of the methodology.

(1) *market survey activities.* It is necessary to collect data to enable the quantitative estimation of the following quantities: for each available version of the design product, the number of units that customers will want to purchase during the design lifetime, as a function of the price and salient electrical specifications of the version. This is done using market survey techniques especially developed for the new methodology. The techniques are primarily based on carefully constructed

and conducted surveys of customers. The surveys are carried out by marketing personnel.

(2) *optimization setup and execution.* This is primarily a design engineering responsibility.

(3) *design implementation.* This is a product engineering responsibility.

(4) *redesign-time selection activities.* Each time the optimization procedure is applied to a product, the resultant design depends on the market data used. This data is essentially a snapshot of market conditions taken prior to the running of the optimization. After the snapshot is taken, market conditions can gradually shift until the point that it is in the interest of the firm to change the design of the product. Deciding when this is the case is what is meant by redesign-time selection. It must be done through a joint effort among marketing, design engineering, and product engineering.

The structure connecting these activity types is heavily influenced by economic, i.e. cost, considerations. No statistical design methodology is useful if it costs more to carry out than the profit increase it achieves. In conventional formulations this is expressed simply as a concern for cpu expense in running optimization iterations. There are two contrasts between the methodology utilization costs in conventional formulations and in the new formulation. First, although cpu costs will certainly be greater in the new, more ambitious formulation, market survey costs will probably dominate them. Second, since the optimization criterion is itself in units of dollars, there are natural and rational bases for utilization decisions (in contrast, for example, to the baseless criterion frequently applied that a computer design tool is too slow if the design engineer becomes bored while using

it).

A consequence of the appreciable cost of market survey activities which is very important for the methodology, is that, to the greatest degree feasible, individual surveys should be directed at the market modeling not of individual products, but of individual product classes. It is not possible to give a precise prescription for dividing IC's into classes. Basically, two IC products should be considered to be in different classes if their most potential customers of one of the products would rule out considering the purchase of the other on the basis that its nominal electrical functionality is qualitatively different. Hence, two analog switches differing only in their *on-resistance*, no matter by how much, are in the same product class. And, as suggested in the first paragraph of Chapter 1, "operational amplifiers" and "2K × 8 random access memories" are in different classes. But, examining the issue more closely, operational amplifiers designed to operate with different power supply configurations should not be considered as belonging to the same class. So, although it is usually not difficult to divide IC products into classes, it must be done with detailed consideration of the electrical functionality of the products in question.

All these considerations lead to a structure connecting the activities required in applying the methodology to a product class, which is represented by the flow chart in Figure 3. It basically consists of an outermost loop representing repetitions of a market survey for the product class, containing a loop over the products in the class, which in turn contains the optimization loop. Note that the "blocks" having a pointed shape on one end are unconventional flow diagram symbols to be interpreted as follows: control flows out the pointed side if the statement within the symbol is true. The circled numbers will be used to refer to the various blocks in the figure. Regarding block 1, n^{rsp} is some number of

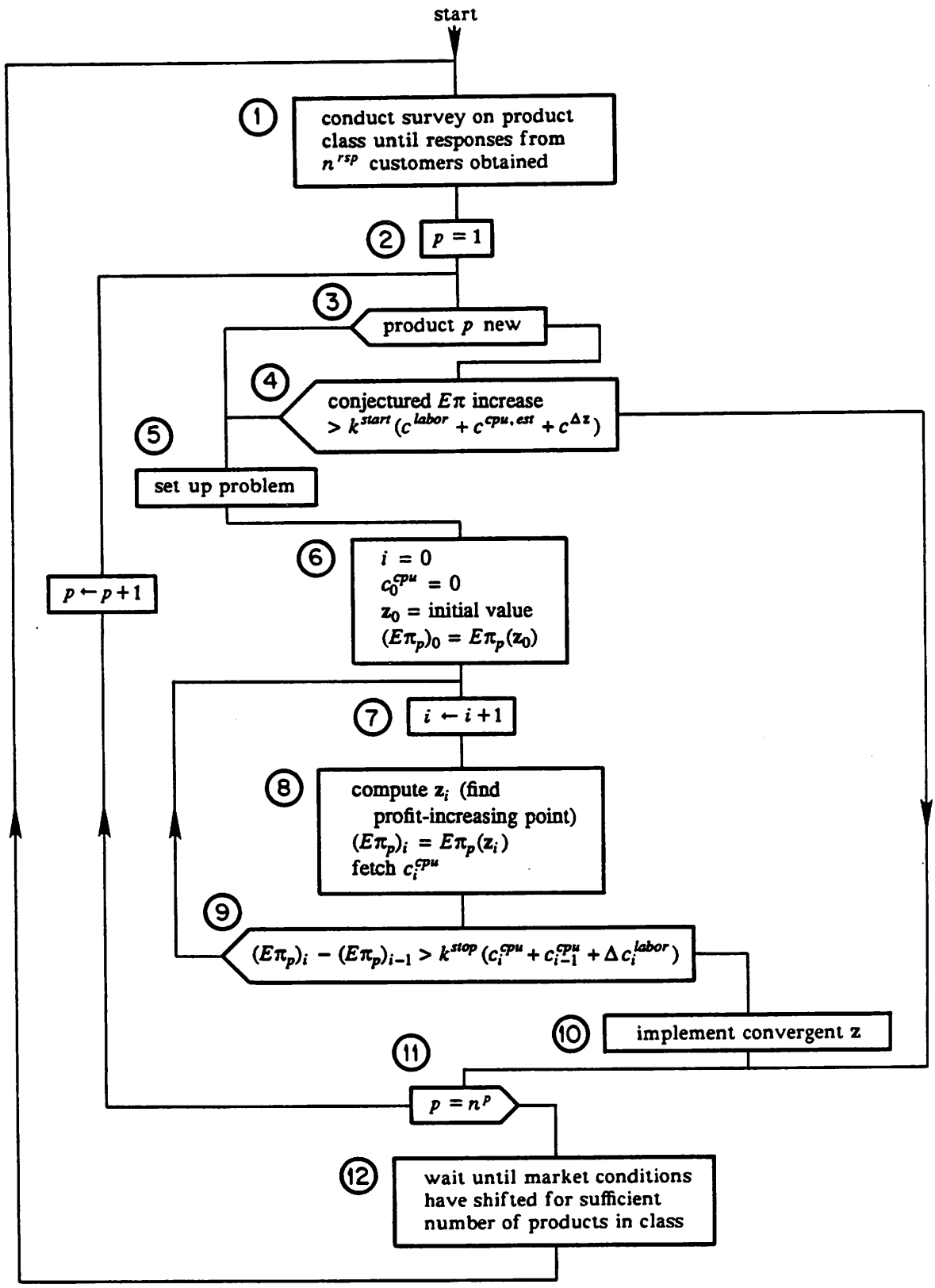


Fig. 3. Flowchart representation of the parametric utilization of the new methodology.

complete survey responses considered necessary for the statistical significance of the market modeling. In block 2 the product index p is initialized. In block 4, $c^{labor,s}$ is an estimated cost of design engineer labor required to set up the optimization software for product p, $c_1^{cpu,est}$ is an estimated cost to run the first optimization iteration for the product, $c^{\Delta z}$ is an estimated cost to implement a design change for the product, and k^{start} is a dimensionless constant greater than one. $E\pi_p$ represents the product p profit increase realized in executing optimization iterations. The decision represented symbolically in this block is a redesign-time selection activity. It is decided to redesign a product if the conjectured increase in $E\pi_p$ is exceeds the costs entailed. The conjectured increase in profit is based on the degree of change seen in the market for the product, and on experience with the performance of the methodology. k^{start} is greater than unity to account for uncertainty in this conjectured increase. Regarding the optimization blocks 5 through 9, i is the iterations index, c_i^{cpu} denotes the cost of the cpu time spent on the problem since the start of the first iteration, $\mathbf{z}_i$ denotes design parameter value after iteration i (except for z_0, which is an initial guess, for a new product, and the actual design value of the "old" design, for an existing product), $E\pi_p(\,)$ is the profit model presented later in the dissertation, and Δc_i^{labor} is the cost of design engineering time to carry out the i'th iteration. In block 9, k^{stop} is set to a value greater than unity to account for the fact that the profit increase on the iteration completed immediately prior to the test is presumably greater than would be achieved on the next iteration, and to account for uncertainty in the profit model. Block 12 represents a redesign time selection activity.

The primary feature of the methodology reflected in the diagram is that it attempts to guarantee a net economic gain through its utilization.

Note that because of the high cost of market surveys, n^{rsp} could be determined iteratively through a process of alternately surveying to get some fraction of n^{rsp} responses, and running the optimization for two or three products, and observing the resultant sequence of profit values to determine when enough total responses have been obtained.

Note that the diagram depicts explicitly what is considered to be the complete utilization of the methodology. There are numerous other potentially very useful ways the methodology or its profit model can be used. Among the most useful is a class of partial utilizations. These can be contrasted with the complete utilizations with the aid of the flowchart. Note that in block 4 a decision is made whether in response to new product class market data, product p should be completely parametrically redesigned. The decision depends on estimated costs to implement the new design, which can be appreciable. A decision to redesign the product means in general that all design parameters of the product will change. A decision not to redesign the product means that none of them will change. But there is a middle ground in which some subset of the design parameters are held fixed while profit is maximized with respect to the remaining parameters. The parameters held fixed would be those for which the costs of changes are relatively high. Specifically, the costliness of changes in the various design parameter types is typically in the following: descending order: $\mathbf{x}$, $\mathbf{e}$, $\mathbf{p}$, $\mathbf{q}$, n^{wl}, α, $\mathbf{l}$. The cost of changing the last three is negligible. Note that many products for which an application of the complete utilization cannot be justified, could be redesigned with respect to the last five, say, parameter types in the list.

Chapter 5

STATISTICAL MODELING OF PROCESS DISTURBANCES

The disturbances in the fabrication of IC's can be statistically modeled using various types of parameters. The two major types that can be used are device parameters and fundamental process parameters. The use of device parameters has the serious drawback that correlations among them must be adequately modeled, a difficult if not impossible task. Until recent years, the use of fundamental process parameters has had the serious drawback that the solution of partial differential equations was necessary to compute the device parameters corresponding to the process parameters needed for the evaluation of circuit performance. For statistical circuit optimization, the cost of this computation, which in general needs to be repeated many times, is prohibitive. However, in 1981 the FABRICS IC process simulator was introduced, in [MAL81]. FABRICS represents an attempt to eliminate the need for costly partial differential equation solutions through the use of pro-

gram parameters that are adjusted to the particular industrial fabrication process in which the product of interest is to be fabricated. It provides a means of calculating device parameters from fundamental process parameter with a cpu cost less than or comparable to the cost of computing circuit responses. As a consequence of these basic considerations, it was decided that the new methodology presented here should use fundamental process parameters to statistically model process disturbances, and should incorporate FABRICS code in its software implementation.

As a result of this decision, some properties of the new methodology are dictated by properties of FABRICS. The two most important of these are the identities of the fundamental process disturbances that are modeled as random variables, and their distributions.

Regarding the fundamental process disturbances, note that they are modeled statistically as a collection of independent normally distributed random variables, and the term "fundamental" is meant to imply that every other random variable occurring within FABRICS is some function of at least one of the fundamental random variables. For the detailed identities of the fundamental process disturbances, the reader is referred to the FABRICS documentation [NAS83, NAS84]. However this information is summarized here as follows. The approximately 40 fundamental process disturbances include: line widths; diffusivities and segregation coefficients of impurities; implantation profile spread quantities; oxidation growth coefficients; oxide charge and fast state densities; substrate impurity concentration; poly resistivity and thickness; contact resistivity model parameters.

Each of the n^w wafers that comprise the fundamental unit of analysis and simulation in this methodology have n^t test locations, each with n^{dt} devices, and N^{cl} chip locations, each with n^{dc} devices. This means that there are a total of $n^w n^t n^{dt}$ test location devices

and $n^w N^{cl} n^{dc}$ chip devices in the fundamental unit of analysis. FABRICS generates distinct realizations of all of the (approximately 40) fundamental process disturbances for every device in the fundamental unit of analysis. Let

$\mathbf{D}_{wtd}$ a vector for device d of test location t of wafer w, the components of which are the fundamental process disturbance random variables for that device.

$\mathbf{D}_{wcd}$ a vector for device d of chip c of wafer w, the components of which are the fundamental process disturbance random variables for that device.

The set of all disturbance random variables pertaining to the test locations, in the fundamental unit of analysis, will be denoted as $\{\mathbf{D}_{wtd}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}$. Similarly the set of all disturbance random variables pertaining to the chip locations, in the fundamental unit of analysis, will be denoted as $\{\mathbf{D}_{wcd}\}_{d=1\ c=1\ w=1}^{n^{dc}\ N^{cl}\ n^w}$.

Chapter 6

THE HIGH-LEVEL STRUCTURE OF THE MODEL

The profit model for the new methodology was described early in Chapter 2 as being essentially a static model. In a strictly static model, no attempt is made to define and model economic quantities associated with subintervals of the full time interval on which the model is claimed to apply. For the system of interest here, however, it is desirable to model economic variables relating to profit that are associated with intervals of production smaller than the design lifetime. There are two reasons for this - one theoretical and the other practical.

The theoretical reason can be described as follows. Recall from Section 2.1 that the contribution profit that is to serve as the criterion in the methodology is the one resulting from the starting in production of a total of n^{wl} wafers of the product during its design lifetime. The use of a strictly static profit model suggests that the producer accumulates all

the circuits from this number of wafers until the end of the design lifetime, then offers them for sale. However in reality producers manufacture a relatively much smaller group of wafers, then use them to fill orders. When their stock is approaching depletion, producers manufacture another small group, and use them to fill orders, and so forth. When wafers are manufactured in these smaller groups, the mix of electrical grades of the circuits for each group differs from that of the n^{wl} wafers taken as a whole, fluctuating from group to group due to the manufacturing disturbances present. It is desirable to account in some way for this effect. The number of wafers in the groups varies, but modeling this dynamic effect is beyond the scope of this dissertation. As an approximation, however, it is assumed that the number of wafers in these groups is the same for each group. These groups are the batches of n^{w} wafers introduced formally in Section 1.6.

The practical reason for focusing on subsets of the entire manufacturing output of the design lifetime is related to the typically very large number of circuits associated with a batch. In the application of the methodology, potentially cpu-costly circuit simulation will be required to estimate the electrical performance of the circuits in some way. It is anticipated that sufficient statistical significance in the computation of profit can be achieved by computing it on the basis of no more than few batches. This would represent considerable computational savings.

Note that focusing on a batch of wafers does not represent a reversion to a truly dynamic model; the subdivision of the design lifetime is done on the basis of production intervals, not time intervals.

The profit associated with the entire design lifetime results fundamentally from subtracting each of a number of cost contributions from a revenue contribution. It might be

inferred from the discussion to this point that each of these contributions can be computed by multiplying the corresponding contribution associated with a batch by $\frac{n^{wl}}{n^{w}}$. This is the case for revenue, and for the opportunity cost which has already been modeled. However, the actual production costs to manufacture a batch of wafers depend on the point during the production history of the product when the batch is manufactured. This is because during the design lifetime, the employees of the firm who carry out the various manufacturing operations accumulate experience in performing them efficiently. Consequently, the time required and hence cost for the operation generally decreases with the cumulative number of wafers manufactured since the beginning of the design lifetime. (The effect may be small after redesigns, since employees already have relevant experience at the beginning of a "redesign lifetime".) The function which yields the costs of an operation as a function of production experience is called the *learning curve* for the operation. Learning curves are typically adequately modeled with elementary functions, e.g. the sum of a constant plus a decaying exponential. In the IC manufacturing case of interest here, let $\hat{w}^h$ denote the number of wafers produced since the start of the design lifetime. Figure 4a plots the cost per wafer of a typical manufacturing operation as a function of $\hat{w}^h$ as it varies over the design lifetime of a product. The total expected cost of producing each wafer is therefore subject to a decrease similar in shape to that shown in the figure. This is plotted in Figure 4b. If all the manufacturing operations shared the same learning curve, the total expected cost plot would have identically the same shape as that of each ot its constituent operation costs.

Also shown in Figure 4b is a production interval of width n^w representing a typical batch, centered at $\hat{w}^h = w^h$. The goal of the first part of the modeling is to calculate the

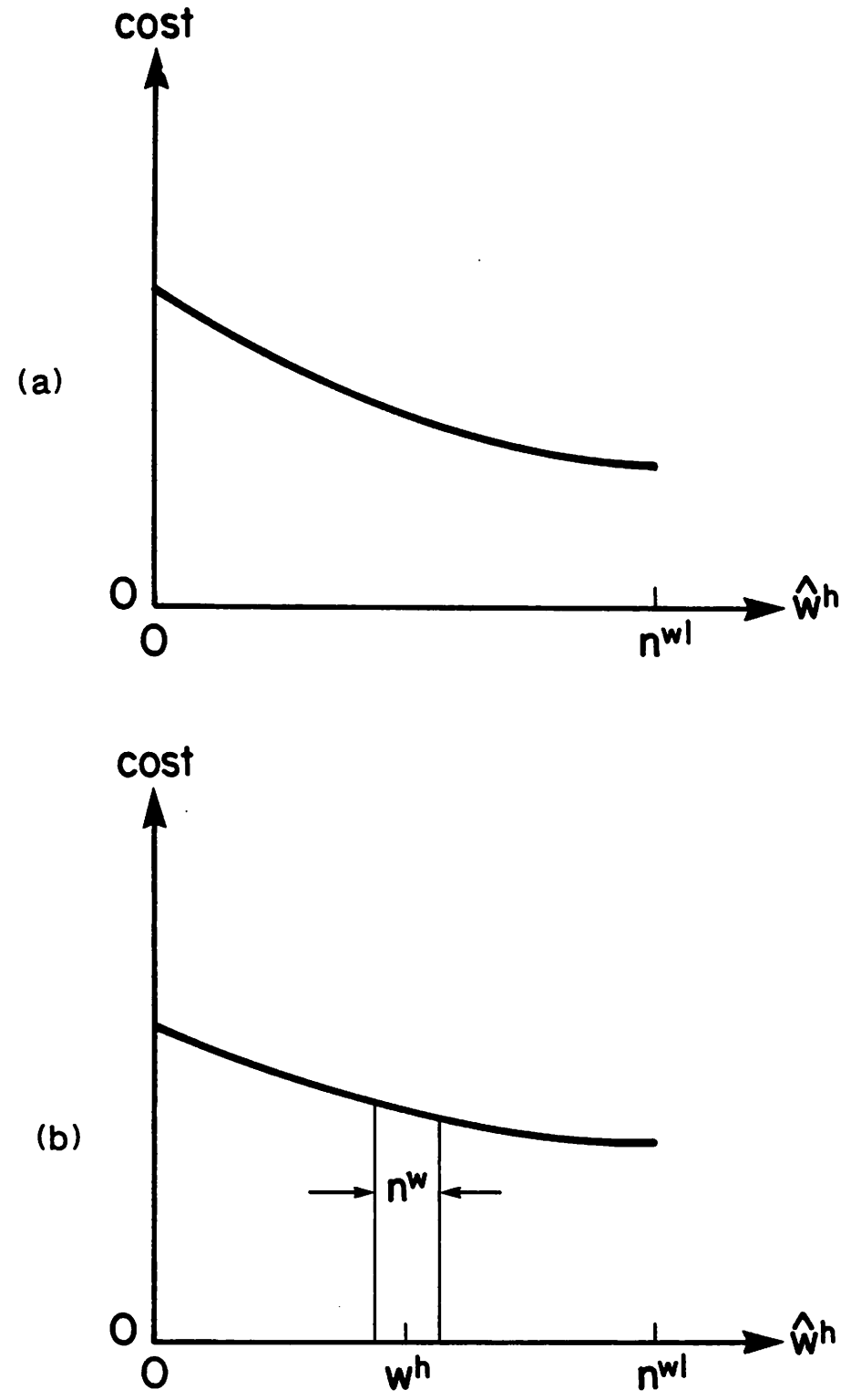


Fig. 4. Variable manufacturing costs per wafer of an IC product as a function of the number of wafers produced in its entire history. (a) for a typical individual manufacturing operation. (b) total over all operations.

total expected cost and profit associated with the starting in production of n^w wafers centered at the point w^h in the production history, assuming that for each manufacturing operation the cost to perform the operation on every wafer in the batch is the value of the learning curve cost for that operation, evaluated at $\hat{w}^h = w^h$.

This will yield two useful results. First, if in the total cost expressions n^w is set to one, they give the exact total cost for wafer w^h. This can then be used to give an exact result for the profit associated with all n^{wl} wafers, which provides insight into the modeling problem. However, computing this would be computationally prohibitive, as was described above. Without the single-wafer restriction, the cost expressions remain approximately correct, since generally $n^w << n^{wl}$, and the learning curve varies slowly with $\hat{w}^h$. They can then be used to derive an approximate expression for the profit associated with all n^{wl} wafers, which requires computations only for the circuits from n^w wafers, yet accounts for the learning curve effect.

Let

π^{n^w} contribution profit associated with the starting in production of a batch of n^w wafers, centered at position w^h in the design lifetime production history.

In order to study the modeling of π^{n^w}, it is helpful to imagine temporarily that it is meaningful to define the contribution profit of a wafer, say wafer w, neglecting opportunity costs. This is denoted π_w. Values of π_w would vary from wafer to wafer, if only due to processing fluctuations. With this assumption, which will be reexamined later, the contribution profit associated with n^w wafers can be computed as the sum of the wafer profits. That is,

$$\pi^{n^w} = \sum_{w=1}^{n^w} \pi_w$$

The high-level structure of the model for wafer profit π^{n^w} is closely related to the treatment of one of the six classes of design parameters in the model - in-line test measurement limits. It is assumed the values of some physical quantities which are to some degree indicative of the success of the processing, are measured at each test pattern on each wafer, at one or more points in the fabrication sequence.

These physical quantities might include resistivities, contact resistances, capacitances, threshold voltages for MOS test devices, and current gains and breakdown voltages for bipolar test devices. If the mask set defines more than one identically-designed test pattern on each wafer, as is typically the case, the measurement results from these repeated structures are assumed to be averaged to produce a single value for each measurement type made.

The test results for the various wafers in theory provide some indication of the electrical performance to be expected from the individual circuits originating on the wafers, once all the processing is complete. In particular, a wafer may have in-line measurement results sufficiently deviant that the expected economic benefit from completing the fabrication of the wafer and sale of the sellable circuits from it, is negative. When such in-line test results are observed, it is not in the interest of the producer to continue processing the wafer. The new methodology assumes the producer accepts the principle that some wafers should be discarded on this basis.

The points at which in-line tests are made, on the basis of which wafers may be rejected divide the fabrication sequence into *processing segments*. The number of such segments is the number of testing points plus one. Let

n^{σ} number of processing segments in the wafer fabrication sequence.

σ index of the processing segments in the wafer fabrication sequence.

Each processing segment may consist of any number of processing operations, or steps, and each segment has costs associated with it. In order to correctly calculate total costs in a fabrication process in which wafers are discarded, it is necessary to distinguish between two types of processing operations. As is well known, it is customary practice to produce wafers in groups called *runs* or *lots*. (These groupings are unrelated to the groupings introduced earlier.) This means that wafers belonging to the same run are subjected to the various processing steps (a step is smaller than segment) at nearly the same time, and records are kept on the fabrication history of only the run as a whole. In addition, there are many processing operations including diffusion and oxidation, in which the wafers comprising a run are also physically together in trays. These will be called run-based operations. It is crucial to distinguish such operations from those in which the wafers are processed "one at a time", called *wafer-based* operations. This is because if due to in-line test results a particular wafer is discarded at a certain testing point, no more wafer-based operations need be performed on that wafer, but the run to which the wafer belongs will nevertheless need to be subjected to all the remaining run-based operations in the fabrication recipe for the product. Therefore the cost savings realized from the discarding of wafers are associated only with the costs of wafer-based operations.

Let

C_{0w} sum of variable costs over all run-based operations in the fabrication of wafer w.

$C_{\sigma w}$ sum of variable costs to perform all wafer-based operations on wafer w, in processing segment σ, $\sigma = 1, 2, \ldots, n^{\sigma}$.

$1_{\sigma w}$ $\begin{cases} 1 & \text{if wafer } w \text{ survives the test point after processing segment } \sigma. \\ 0 & \text{if wafer } w \text{ is discarded at the test point after processing segment } \sigma. \end{cases}$

r_w revenue from the sales of all sellable circuits originating from wafer w.

Here it has been assumed that it is meaningful to define $C_{n^\sigma w}$ and r_w, which is not strictly the case. However this can be accounted for in the interpretation of the results of the derivation. Continuing, then, the profit associated with wafer w can be written as,

$$\pi_w = -\{C_{0w} + C_{1w} + 1_{1w}[C_{2w} + 1_{2w}[C_{3w} + \cdots + 1_{(n^\sigma-2)w}[C_{(n^\sigma-1)w} + 1_{(n^\sigma-1)w}[C_{n^\sigma w} - r_w]]\cdots]]\}.$$

Therefore,

$$\pi^{n^w} = -\sum_{w=1}^{n^w}\{C_{0w} + C_{1w} + 1_{1w}[C_{2w} + 1_{2w}[C_{3w} + \cdots + 1_{(n^\sigma-2)w}[C_{(n^\sigma-1)w} + 1_{(n^\sigma-1)w}[C_{n^\sigma w} - r_w]]\cdots]]\}. \tag{6.1}$$

Since the $1_{(n^\sigma-1)w}$ is the indicator for a testing point within wafer fabrication, $C_{n^\sigma w}$ includes the costs of all operations performed on the chips after the die separation operation. These operations are referred to as *back-end* operations, in essential agreement with industry jargon. It has been found advantageous to model the costs of the back-end operations separately from the remaining costs. More concisely, let

C_w^{BE} total variable back-end cost, summed over all back-end operations, for wafer w.

$C_w^{\#}$ total variable cost incurred between most-downstream testing point, and die separation, for wafer w.

Then the cost of the last fabrication "segment" is just,

$$C_{n^\sigma w} = C_w^{\#} + C_w^{BE}.$$

Combining this with (6.1) gives,

$$\pi^{n^w} = -\sum_{w=1}^{n^w} \{C_{0w} + C_{1w} + 1_{1w}[C_{2w} + 1_{2w}[C_{3w} + \cdots + 1_{(n^\sigma-2)w}[C_{(n^\sigma-1)w} + 1_{(n^\sigma-1)w}[C_w^{\#} + C_w^{BE} - r_w]]\cdots]]\}. \tag{6.2}$$

Now, the cost of *front-end* operations is defined quite naturally as,

$$C^{FE} = \sum_{w=1}^{n^w} \{C_{0w} + C_{1w} + 1_{1w}[C_{2w} + 1_{2w}[C_{3w} + \cdots + 1_{(n^\sigma-2)w}[C_{(n^\sigma-1)w} + 1_{(n^\sigma-1)w}[C_w^{\#}]]\cdots]]\}. \tag{6.3}$$

Let

$$1_w^{BE} = \prod_{\sigma=1}^{n^\sigma-1} 1_{\sigma w}. \tag{6.4}$$

This is just the indicator function which is unity if and only if wafer w has survived all the wafer fabrication testing points. The total back-end cost is defined as,

$$C^{BE} = \sum_{w=1}^{n^w} 1_w^{BE} C_w^{BE}, \tag{6.5}$$

and the design product revenue r^{dn^w} as,

$$r^{dn^w} = \sum_{w=1}^{n^w} 1_w^{BE} r_w. \tag{6.6}$$

Also, since only explicit contributions to profit of the design product have been considered in the analysis, it is necessary to reintroduce consideration of the brother product revenue and opportunity cost. For the purposes of this discussion, which do not include the precise modeling of brother product revenue, it suffices to define r^{bn^w} as the total brother product revenue corresponding to r^{dn^w}. Also, let,

$$r^{n^w} = r^{dn^w} + r^{bn^w} \tag{6.7}$$

Regarding opportunity cost, let,

$c^{O/w}$ opportunity cost coefficient (dollars per wafer) associated with the production of the design product.

Then the opportunity cost contribution to the profit associated with n^w wafers is just $-n^w c^{O/w}$, and combining (6.2) through (6.7) gives,

$$\pi^{n^w} = -n^w c^{O/w} - C^{FE} - C^{BE} + r^{n^w}. \tag{6.8}$$

All of the terms in the definition of C^{FE} are well-defined quantities, unaffected by the temporary assumption made above. In contrast, the above expressions for C^{BE} and r^{dn^w} are not strictly valid because of the assumption, and r^{bn^w} is not well defined. C^{BE} is redefined according to

C^{BE} cost of performing all back-end operations associated with the starting in production of a batch of n^w wafers, centered at position w^h in the design lifetime production history.

and modeled from scratch from this definition. Design and brother product revenue variables corresponding to n^{wl} wafers are also defined and modeled from scratch in the sequel. Nevertheless (6.5) and (6.6) have shown that the models for back-end cost and design product revenue must reflect the fact that if a particular wafer does not survive the front-end operations, it can make no contribution to these quantities.

A more important result of the analysis of this section, however, is that (6.8) decomposes to an extent the profit modeling problem into four smaller and well-defined modeling problems corresponding to the four terms on the right of the expression:

(1) $n^w c^{O/w}$: The modeling of opportunity cost is complete at this point, since both factors in this product are elementary quantities in the methodology.

(2) C^{FE}: The modeling of front-end cost, begun in this section, is completed in Chapter 8.

(3) C^{BE}: The modeling of back-end cost is essentially confined to Chapters 9 and 10.

(4) r^{n^w}: The modeling of revenue is essentially confined to Chapters 11 and 12.

Chapter 7

MODELING DIE AREA EFFECTS

An important component of the modeling of profit relates to effects that depend on circuit die area. This is reflected in (6.8) of the preceding chapter, in that all of the last three terms on the right in that equation depend on the number of chip locations on each wafer, N^{cl}, and the last two terms depend on the defect yield.

Note that with the exception of the dependence of front-end cost on N^{cl}, these dependencies are in general first-order effects, with major economic significance. For example, for some large-area digital products, defects may be responsible for forcing producers to consistently discard more than 90% of the chips produced.

The dependencies of the last three terms in (8.8) on N^{cl} and defect yield are developed in later sections, as outlined above. The dependencies of the latter two quantities on design parameters and random variables must also be modeled. It is most

convenient to discuss this modeling at this point.

Let

A^c total area of each chip, including associated streak (border) area.

Clearly N^{cl} depends on A^c. N^{cl} is of course a positive integer and a discontinuous function of A^c (see [GUP72]), however the magnitude of the steps in the N^{cl} versus A^c function are relatively small except for the largest VLSI circuits. Analysis of the dependency has indicated that there can be determined an explicit differentiable function of A^c that, when rounded to the nearest integer, would give values of N^{cl} that would cause an error in the computation of profit that is small relative to those from other components of the profit model. Let $\hat{N}^{cl}(A^c)$ denote the function that models the dependence of N^{cl} on A^c. Having this, it remains to discuss the dependence of A^c on design parameters and random variables. However, it is most convenient to do this in conjunction with the discussion of defect yield modeling which follows.

The economic implications of yield losses due to localized defects has fostered considerable research effort aimed at developing accurate predictive models. Such models have the potential to be applied in a number of ways in the industry. They can be used for short-term process monitoring, long-term process development monitoring, and production planning [STA83].

Unfortunately the problem of modeling defect losses is difficult. There are numerous physical effects which can cause catastrophic defects. The magnitude of the analysis problem is reduced by an appropriate categorization of these effects. They are usually considered to belong to one of the following categories: point defects, line defects, area defects, and defect clusters [GUP74]. Accurate analyses potentially can be obtained to

account for all the physical effects within each category. However the modeling of losses in the last three categories has had limited success. More to the point, since the most prevalent type of defects is generally point defects, is that there is no single widely accepted and utilized model for this category. Among the difficulties is that the statistical modeling of the distribution of number defects per chip has proceeded axiomatically, and there has been disagreement, albeit subsiding of late, over selection of the most appropriate distribution.

Nevertheless reasonable accuracy has been reported when sophisticated models are ambitiously applied to and adjusted for specific stable industrial processes (e.g. [HEM81],[STA81]). Such accuracy is assumed here on the basis that the new methodology provides both a new incentive and a new framework for IC producers to provide an accurate defect yield model for their own fabrication processes. The objective of the study of defect yield in this dissertation has been to determine a form of defect yield model sufficiently general to accomodate any such producer-supplied model.

Most of the literature utilizes a characteristic of a particular chip design called the *active area*, which is intended to refer to all the area of a chip design such that if a localized defect occurs in that area on a particular chip, that chip will fail catastrophically. The exact definitions of the active area for each device type and for the interconnect are usually not specified. These details, like the details of the defect yield model, are also not specified here, but are left to the user of the methodology. (Some recent work has introduced more sophisticated concepts of "critical area", in which such effects as finite size of "point" defects are considered. See for example, [FER85]. Such concepts can potentially be incorporated in the profit model, as long as any need for details of chip layout, which

cannot possibly be available at the time of application of the new methodology, is eliminated by appropriate statistical modeling of area effects.) In the literature utilizing the active area concept, there is agreement that the expectation value of the defect yield is a decreasing function of the "active area" of the chip design, a dependency which is first-order for larger chips [PRI70], [GUP74], [WAR74], [HU79], [WAR81]. The active area, denoted A herein, is in fact the primary independent variable in the models.

In view of this, the acceptable form for defect yield models for the profit model could be established as that of an arbitrary function of active area. However, many defect yield models provide not only an expression for the expectation value of the defect yield, but its distribution. With this, fluctuations in yield from wafer to wafer can be accounted for. Consequently, the model form for the profit model is enhanced by introducing a set of identically distributed random variables $\{Y_w\}_{w=1}^{n^w}$, one for each of the n^w wafers in the analysis. The distribution of these random variables is denoted f_Y. The defect yield model used in the profit model can be any model in which the yield of wafer w, denoted Y_w^{df} can be expressed as depending only on Y_w and the active area A, i.e.,

$$Y_w^{df} = \hat{Y}^{df}(Y_w, A), \; w = 1, 2, \ldots, n^w. \tag{7.1}$$

What remains, as for A^c, is to discuss the dependence of A on design parameters and random variables.

Device dimensions were among the parameters that in Chapter 3 were identified without justification as design parameters in the methodology. They are the only class of parameters listed there that could conceivably influence A^c or A. From device dimensions one can trivially calculate an additional area parameter a, defined by,

a total device active area

Note that some devices may, at the discretion of the user of the methodology, have none of their dimensions specified by elements of the x vector, in which case their active area is a constant. (Note also that in FABRICS, misalignment errors are superimposed on all nominal values of device dimensions in the fabrication simulation.) The model for total device area as a function of x is denoted $a(\mathbf{x})$.

A^c and A certainly depend on a, and hence on device dimensions, but the dependencies are not deterministic. This is because A^c and A are determined only after the physical layout of the design circuit is complete, and because solutions to the circuit layout problem are non-unique. Consequently, it is proposed that a regression model be used for the dependence of each of these areas on a, with the parameters of the models determined from data on existing circuits that are judged to have layout considerations (i.e. approximate chip size range, ratio of interconnect to device area, degree of regularity, and so forth) similar to those of the design product. It is expected that affine dependencies of the expectations of the desired areas on a should suffice. In view of the high correlation between A^c and A, their models are assumed to share a common underlying random variable Λ, with probability density function f_Λ, characterizing the effectiveness of the layout process. The models can be written more explicitly, then, as,

$$A^c = A^c(\Lambda, a)$$

and,

$$A = A(\Lambda, a).$$

The models for N^{cl} and defect yield can now be assembled. Beginning with N^{cl}, recalling that the dependence of N^{cl} on A^c is denoted $\hat{N}^{cl}(A^c)$, it can be written as,

$$N^{cl} = \hat{N}^{cl}(A^c(\Lambda, a(\mathbf{x}))$$

For later convenience, the composite function

$$N^{cl}(\Lambda, \mathbf{x}) = \hat{N}^{cl}(A^c(\Lambda, a(\mathbf{x}))$$

is defined, so that the final model for the number of chip locations on each wafer is symbolized as,

$$N^{cl} = N^{cl}(\Lambda, \mathbf{x}). \tag{7.2}$$

For the defect yield, from (7.1), Y_w^{df} can be written as,

$$Y_w^{df} = \hat{Y}^{df}(\mathrm{Y}_w, A(\Lambda, a(\mathbf{x}))), \; w = 1, 2, \ldots, n^w.$$

For later convenience, the composite function

$$Y^{df}(\mathrm{Y}_w, \Lambda, \mathbf{x}) = \hat{Y}^{df}(\mathrm{Y}_w, A(\Lambda, a(\mathbf{x}))).$$

is defined, so that the final model for the defect yield random variable for wafer w is symbolized as,

$$Y_w^{df} = Y^{df}(\mathrm{Y}_w, \Lambda, \mathbf{x}), \; w = 1, 2, \ldots, n^w. \tag{7.3}$$

The above dependencies of N^{cl} and the wafer defect yields Y_w^{df} on the device dimension vector x mean that these dimensions potentially have a first-order effect on expected contribution profit, even if the critical role they play in determining circuit performance is ignored. Since they also have the other required properties listed in Section 1.2, they are treated as design parameters in the methodology.

Chapter 8

MODELING OF FRONT-END COST

The discussion here is concerned with the dependence of the front-end cost C^{FE} on design parameters and random variables.

First, refering to (6.3), $C_{\sigma w}$, $\sigma = 1,2,\ldots,n^{\sigma}-1)$ and $C_w^{\#}$ are the same for all wafers, $w = 1,2,\ldots,n^{w}$. They were defined earlier as indexed only in deference to the back-end and revenue terms (see (6.2)).

Second, the segment costs for $\sigma = 1,2,\ldots,(n^{\sigma}-1)$ need not be modeled as random variables. The cost of wafer probe is modeled as a random variable, as will be discussed later, but this cost is not included in the first $(n^{\sigma}-1)$ processing segments, as will be argued next.

Clearly one candidate for a testing point is at "wafer probe", the point near the end of the processing when for the first time the individual circuits on the wafer are probed. That

is, if the collective performance of all the die on a particular wafer is sufficiently poor, the wafer could be discarded after wafer probe. This practice can be modeled starting from (6.3). However it is assumed here that wafers are never discarded after wafer probe. This is because the only incentive to do so is to save the cost of die separation. It would be a very rare event to find a "yield" at wafer probe so low that it is not economically worthwhile to separate the dice. Also, there is generally no basis for modeling a testing point as occurring between wafer probe and die separation. Therefore, the cost of wafer probe (and of die separation) are included in $C^{\#}$, and not in the first $(n^{\sigma}-1)$ processing segments.

With these two observations, the notation in (6.3) can be simplified. Let

c_0 sum of variable costs per wafer over all run-based operations in the front end.

c_σ sum of variable costs per wafer over all wafer-based operations in processing segment σ , $\sigma = 1,2,\ldots,(n^{\sigma}-1)$.

Then (6.3) can be replaced by,

$$C^{FE} = \sum_{w=1}^{n^w} \{c_0 + c_1 + 1_{1w}\,[c_2 + 1_{2w}\,[c_3 + \cdots \\ 1_{(n^{\sigma}-2)w}\,[c_{(n^{\sigma}-1)w} + 1_{(n^{\sigma}-1)w}\,[C^{\#}]]\cdots]]\}. \tag{8.1}$$

c_σ, $\sigma = 1,2,\ldots,(n^{\sigma}-1)$ are not constant, however. They are in general cost components which are subject to the learning curve effect. The model for the dependency of c_σ on w^h is denoted $c_\sigma(w^h)$.

As stated above, $C^{\#}$ includes the costs of wafer probe testing and die separation. Both these costs vary significantly with N^{cl}. On one hand, they are typically relatively small, numerically, and their variation will generally not have a major influence on the

outcome of the optimization. On the other hand, the only effort required on the part of the user of the methodology would be to supply a rough model of the dependence of these costs on N^{cl}, since N^{cl} is modeled for its other, first-order effects on profit. By curve-fitting techniques which need not be extremely accurate, any IC house should be able to generate a reasonable continuous-function model for the dependence of each of these costs on N^{cl} and w^h. A product of a function of N^{cl} and one of w^h should be adequate. In any case, let $\hat{C}^{\#}(N^{cl}, w^h)$ denote the resultant model for the final front-end segment cost. And let $C^{\#}$ be the composite function,

$$C^{\#}(\Lambda, w^h, \mathbf{x}) = \hat{C}^{\#}(N^{cl}(\Lambda, \mathbf{x}), w^h). \tag{8.2}$$

The front-end testing point survival indicators $1_{\sigma w}$, $\sigma = 1, 2, \ldots, (n^{\sigma}-1)$ depend on the entire set $\{\mathbf{D}_{wtd}\}_{d=1\,t=1}^{n^{dt}\;\;n^{t}}$ of disturbances associated with wafer w, and on the process test limits. The latter, as stated previously, comprise one of the classes of design parameters in the methodology, because they meet the criteria for the treating of parameters as design parameters described in Section 1.2. The models for the in-line test measurement results are embodied in the FABRICS program. If at a particular testing point there are n^z physical quantities measured, the indicator function for that testing point is unity if and only if the n^z-vector of measurements lies in some set in R^n parametrized by $2n^z$ in-line test limit parameters. The most effective form for this set remains to be determined. In any case, let

$\mathbf{l}$ vector of all in-line test limits for the product.

Then the model for the dependence of $1_{\sigma w}$ on $\{\mathbf{D}_{wtd}\}_{d=1\,t=1}^{n^{dt}\;\;n^{t}}$ and $\mathbf{l}$ is represented by

$$1_{\sigma w}(\{\mathbf{D}_{wtd}\}_{d=1\,t=1}^{n^{dt}\;\;n^{t}}, \mathbf{l}).$$

Combining this and the definitions of the c_σ functions with (8.1) and (8.2) gives,

$$C^{FE} = \sum_{w=1}^{n^w} \{c_0(w^h) + c_1(w^h)$$
$$+ 1_{1w}(\{\mathbf{D}_{wtd}\}_{d=1\,t=1}^{n^{dt}\;n^t}, 1)\,[c_2(w^h) + 1_{2w}(\{\mathbf{D}_{wtd}\}_{d=1\,t=1}^{n^{dt}\;n^t}, 1)\,[c_3(w^h) + \cdots \tag{8.3}$$
$$+ 1_{(n^\sigma-2)w}(\{\mathbf{D}_{wtd}\}_{d=1\,t=1}^{n^{dt}\;n^t}, 1)\,[c_{(n^\sigma-1)w}(w^h) + 1_{(n^\sigma-1)w}(\{\mathbf{D}_{wtd}\}_{d=1\,t=1}^{n^{dt}\;n^t}, 1)\,[C^{\#}(\Lambda, w^h, \mathbf{x})]] \cdots]] \}.$$

The dependence of C^{FE} on design parameters and random variables can be represented as,

$$C^{FE} = C^{FE}(w^h, \{\mathbf{D}_{wtd}\}_{d=1\,t=1\,w=1}^{n^{dt}\;n^t\;n^w}, 1, \Lambda, \mathbf{x}). \tag{8.4}$$

Chapter 9

A BACK-END FLOW MODEL

9.1. Motivation

The output of what has been labeled herein as front-end processing is the input to what has been labeled as back-end processing. It consists of a collection of dice, which are potentially the active components for marketable realizations of an IC product. The output of the back-end processing consists of a collection of "finished" circuits, ready to be shipped to customers, but in a number of categories, each with different potentialities to be sold as one of the (typically between two and ten) advertised variations of the product.

Since in general finished circuits which are in different categories on whatever basis cost differently to process through the back-end, and sell at different prices, it is necessary to characterize the various categories of finished product, and to model the numbers of cir-

cuits originating with a prescribed number of input wafers, which emerge from the back-end in the each of the categories. This is the objective of this Chapter. The notational system developed in this section will serve as the foundation for the modeling in later sections of both back-end costs and revenue.

Unlike the front-end processing, in which all the material being processed follows the same path, the material flowing through back-end processing undergoes a sequence of splits, interspersed among processing operations, which place it in successively more narrowly defined categories, until it is in the finished-product categories just mentioned. The splitting of circuits in the back-end processing flow is done on the basis of two basic types of conditions:

(1) which of the various electrical performance categories the circuits belong in, as determined by their performance in tests.

(2) which of a number of discrete *back-end treatments* the circuits are to be subjected to. Among the treatments to which the circuits may be subjected are packaging in one of a number of packaging options, and various treatments designed to increase customer confidence in the reliability of the product. (The prototypical example of the latter is the "burn-in" procedure, in which the circuits are operated under specified electrical and environmental conditions for some large specified number of hours, then performance tested.)

Note that the splitting-flow characteristic of the back-end processing is in strong contrast with the flow of material in most industrial manufacturing processes. The latter is a merging flow. One-piece parts are assembled into multi-piece parts, which in turn are assembled into sub-assemblies, and so on up a manufacuring tree until one final product is

produced.

Achieving the objective stated above requires the development of an internal model of the back-end processing which keeps quantitative track of the splitting which takes place.

Note that starting with this section, when the term *circuit* is used, it is used when there is no interest in distinguishing between a die and a die mounted in a package.

9.2. General Features of the Back-end Flow Tree

Before proceeding with the main line of argument, a clarification regarding the output of the front-end processing is needed. That output could be considered to include not only the dice, but for each die, the specification of in which of a number of wafer-probe performance categories established for the product, the die belongs. (Most commonly, there are but two such categories - one for circuits to be rejected without further processing or testing, and the other for all other dice.) So the cost of obtaining the wafer-probe test results contributes to the front-end cost. The results themselves do not affect the front-end cost, but the back-end cost. In fact the specification of many salient characteristics of the wafer-probe test procedures is part of the design of the back-end processing of a product. In none of this irony is there any loss of model accuracy.

Study of real-world back-end flows has shown that they can be modeled at the highest level as having a tree structure, as depicted in Figure 5. For this application, the tree is called a *back-end processing stage tree*. The total collection of circuits resulting from die separation are thought of as starting at the (level 0) root node of the tree, then being split into meaningful categories as they proceed from left to right. Eventually, each circuit arrives at a leaf of the tree, representing either that it is ready to be shipped to a

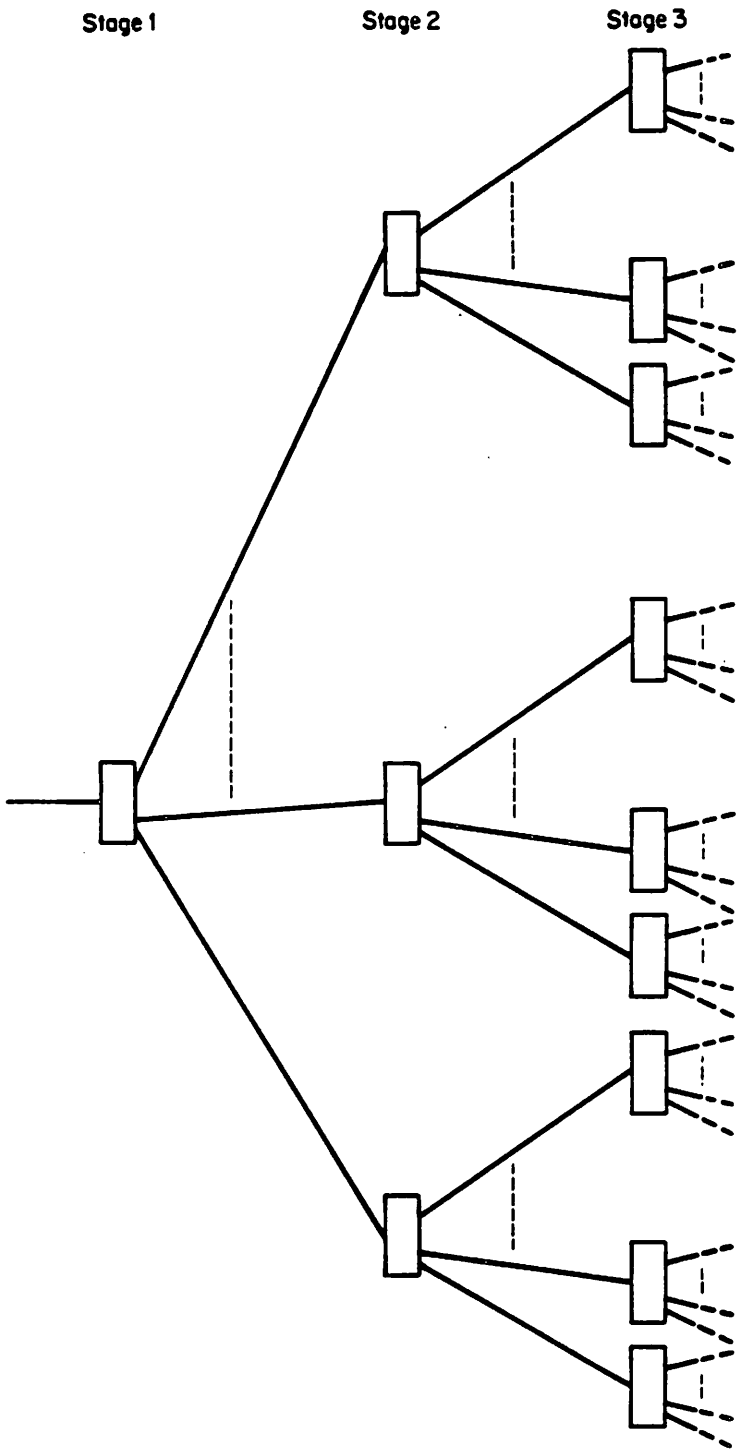


Fig. 5. General tree structure of the back-end processing stage tree, for modeling circuit flow in the back-end processing.

customer, or discarded as not worthy of sale. Although not depicted in the figure, the level at which this occurs in general varies among the circuits processed, from one, to the depth of the tree. The depth of the tree varies from product to product, from one to an arbitrary finite number, which in real cases seldom exceeds four. The levels of the tree are referred to as back-end processing stages.

The nodes of the tree, depicted as rectangles in the figure, have a special internal structure. Namely, they contain within them a depth-2 general tree, as depicted in Figure 6. The first (leftmost) node in the depth-2 tree, represents a split based on condition (1) of the previous section. Such a split is called a *test-outcome* split, and branches emanating to the right of such nodes are called *test-outcome branches*. Each of the nodes at the next level represents a split based on condition (2) of the previous section. Such a split is called a *discretionary* split, and branches emanating to the right of such nodes, which are shown with heavy lines, in most cases containing rectangular boxes, are called *discretionary branches*. The terms "test-outcome" and "discretionary" are suggestive of the detailed meanings of these two types of splits, which are summarized in the following preliminary descriptions.

At each test-outcome node, the producer has a collection of circuits impingent on the node all of which have been subjected to a certain set of tests since they passed through the father of the node. The circuits are placed in categories based on their performance in the tests. These categories are in one-to-one correspondence with the sons of the node. Which category a particular circuit is placed in is dictated by its test results. Thus the producer can systematically influence the assigning of circuits to categories only through some basic (and non-obvious) modification of the design of the product.

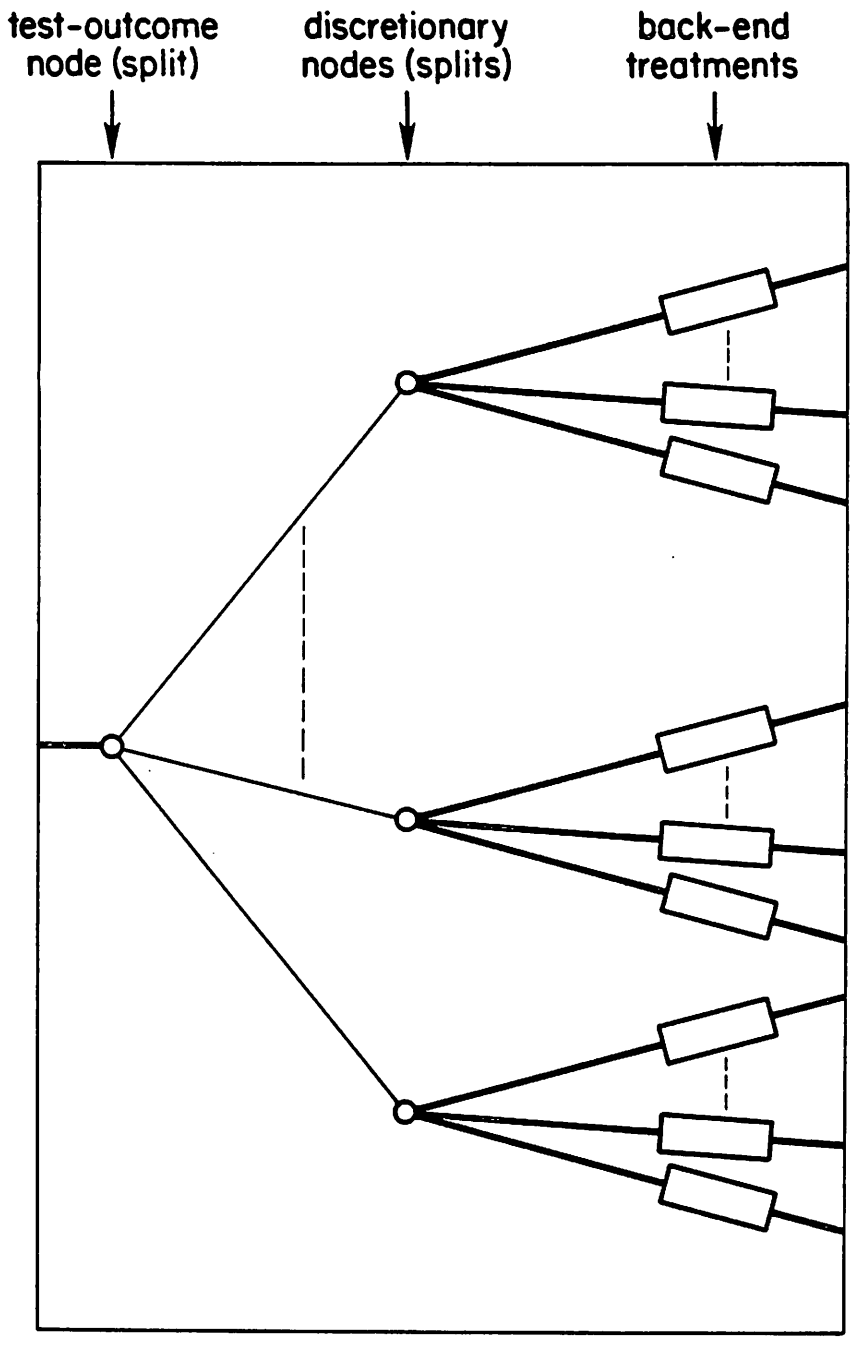


Fig. 6. Depth-2 tree structure of the test-outcome and discretionary substages comprising each back-end processing stage.

At the discretionary splits nodes, the producer has a collection of circuits impingent on the node and a set of back-end treatments to which the circuits might next be subjected. The assignment of circuits to treatments at the split is not dictated by any pre-existing constraints, but is entirely at the discretion of the producer (hence the term). (The assignment is in principle carried out with the maximization of profit as the underlying goal, and information on market demand for the various treatments an essential input.)

The majority of the discretionary branches have rectangular boxes depicted in them, each of which represents a back-end treatment, which includes all the processing operations required to be performed before the test-outcome split at the next level of the tree. These operations include the testing which is the basis for the next, test-outcome split.

The tree which results when the depth-2 tree is substituted for the nodes in the back-end processing stage tree is called the *back-end processing tree*. It consists of an alternation between test-outcome substages and a discretionary substages.

In Figure 7 is depicted a plausible example of such a tree. There are just two points about the figure which should be noted here. First, the underlying process stage tree of the tree in the figure has a depth of two, and since the depth of any back-end process tree is twice that of its underlying back-end process stage tree, the tree in the figure has a depth of four. Second, note that the discretionary branches having no operation blocks are "dummy" branches representing flow paths with no processing operations associated with them.

In order to most easily discuss in general terms the role of back-end operations in the flow structure, it is helpful to recognize that instead of associating each operation with a discretionary branch, it can be associated with the node attached at the right end of the

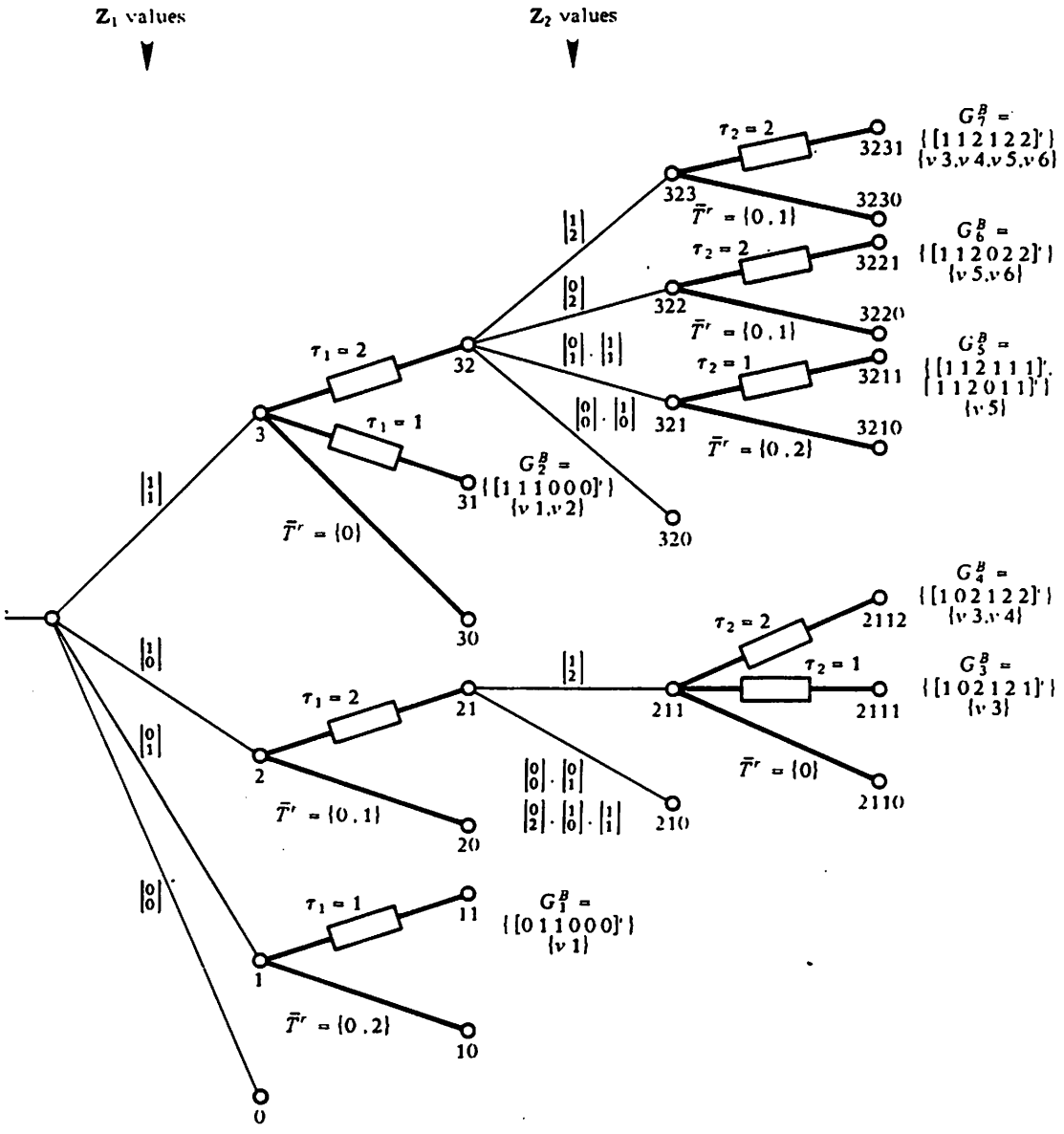


Fig. 7. Back-end processing tree for a hypothetical circuit.

branch. Doing so allows standard terminology for trees to be utilized, and the earlier characterization of the role of operations in the flow structure to be replaced by the following succinct description.

Most test-outcome nodes are treated as having a set of back-end processing operations associated with them, which are simply all the processing operations required to be performed after the test-outcome split at the next previous level. A universal exception is the root node, because the costs of operations to get circuits to the root node are the front-end costs, which certainly should not contribute to the back-end costs. Most leaf nodes are also treated as having a set of back-end processing operations associated with them, again those required to get the circuits to the node from the next previous level. A frequent exceptions occurs in the representation of circuits to be discarded. Such circuits are represented by a "discard" test-outcome branch and dummy leaf node emanating from the pertinent test-outcome split node, as exemplified by the nodes in the figure labeled "0","210", and "320".

Note that discretionary split nodes never have processing operations associated with them.

9.3. Overview of Circuit Characteristics

The previous section has described general properties of a flow structure capable of representing the back-end flow of almost any IC product. The way in which the detailed structure of the tree and the bases for the splits in it are determined, for a particular given IC product, remains to be discussed.

The information needed to determine these consists of a specification of the number of versions in which the product is to be produced and all the characteristics which each

version is to have, with the possible exception of some electrical performance parameters which are to be treated as design parameters in this methodology. All of this information is considered given, in the methodology. (Whether the methodology is being applied to the development of a new product or to the redesign of an existing one, the information is generated through a collaboration of a group of employees of the producer which may include design engineers, process engineers, product engineers, marketing personnel, and high-level management. The task involves considerable non-parametric and not necessarily optimal decision-making, and is not directly part of the utility of this methodology.) The term "grade" is used when distinctions between versions of a product are of interest, but usage varies. Let *product variant*, or simply *variant* refer to the discrete purchase options that customers of the design product have. For each variant, the producer has a set of characteristics defining the variant. This set consists of the two types of characteristics corresponding to conditions (1) and (2) of the previous section (namely electrical performance and back-end treatment), plus the price structure of the variant (the price for each of the several order quantity ranges). There is some subset of these characteristics that the producer chooses to communicate by whatever means, to prospective customers. These are referred to here as *published* characteristics. Naturally it is assumed that for any two variants there is at least one published characteristic other than price which the producer claims differs between them.

In Figure 8 is shown a chart of characteristics for a hypothetical linear amplifier having 6 variants, numbered from v1 to v6. Note that the first four rows pertain to electrical characteristics which are not precisely defined unless details of the pertinent test circuits and excitations are specified. No attempt is made to present plausible information of this kind, since it would not be germane.

Product variant number	v1	v2	v3	v4	v5	v6	units
Voltage offset, max., θ = 25° C	–	2	2	2	2	2	mv.
Output current, min., θ = 25° C	25	25	–	–	25	25	ma.
Slew rate, min., θ = 125° C	–	–	120	120	–	–	v./ μsec.
Bandwidth, min., θ = 125° C	–	–	10	10	5	10	MHz
Package type	none	none	TO–99	TO–99	TO–99	TO–99	–
Burn–in	no	no	no	yes	no	yes	–
Price, quantity [1–24]	4.85	9.22	14.22	20.58	11.87	16.80	$
Price, quantity [25–99]	3.93	7.47	11.52	16.67	9.61	13.61	$
Price, quantity [100–999]	3.30	6.27	9.67	13.99	8.07	11.42	$
Price, quantity [above999]	2.67	5.07	7.82	11.32	6.53	9.24	$

Fig. 8. Characteristics (guaranteed) for the six variants of a hypothetical linear IC product.

The flow tree of Figure 7 is the model of the back-end flow appropriate to the hypothetical circuit specified in the chart of Figure 8, under the totality of back-end modeling assumptions in the methodology. There are only two points about Figure 7 that need be made here. First, the format, although not necessarily coinciding with that appearing in typical industry product literature, is nevertheless suitable for representing the product characteristics of almost all real IC products. Second, the numbers of specifications and back-end treatments in Figure 8 are small compared to those of typical real products.

In order to clarify the correspondence between Figures 7 and 8, it is necessary to develop an understanding of how the tree structure and bases for the splits are determined for particular products, as well as the assumptions underlying this procedure. This is the purpose of the next four sections.

First, the modeling of electrical performance and of back-end treatments must be made more precise.

9.4. Modeling of Electrical Performance

The electrical performance of the variants of a product can be specified with a set of distinct performance specifications, with each specification imposing a bound on some circuit response. Each specification has a temperature value associated with it, specifying the ambient temperature at which the response is to be measured.

It is now argued that, under certain simplifying assumptions, each specification needed in defining the variants of a product can be associated with one and only one processing stage.

First, note that the testing of circuits in the back-end is performed in a number of distinct test facility types, the capabilities of which are generally not interchangeable. Specifically, such facility types include wafer-probe, and may include room-temperature, low-temperature, and high-temperature.

Without significant loss of generality, it is assumed no circuits are subjected to testing in a particular test facility type, then subjected to cost-incurring processing, then returned for testing in the same test facility type (unless the second testing plays a negligible role in affecting the flow of circuits, in which case it can be treated as a cost-incrementing operation rather than a test-outcome split).

Of course the wafer-probe facility type is for all circuits the first facility type encountered. For simplicity, it is also assumed that for any subsequent testing, all circuits share the same order of test facility types encountered (although some circuits are destined by

their performance not to make it through the entire sequence). Back-end flows in which this assumption cannot be made can be modeled by an extension of the model presented herein.

With the assumption, and the fact that various back-end treatments must be performed on the circuits between test facility types, the test-outcome substage of each processing stage has a unique facility type associated with it. In a typical four-stage flow, the test-outcome substages of the last three stages might consist of the room-temperature, low-temperature, and high-temperature test facilities, respectively.

It is generally clear in which facility type each specification should be measured. If for example "high-temperature" (meaning anything above room temperature) is part of the specification of a test-facility type used in the back-end testing, then that facility is uniquely associated with any product specification requiring measurement at an ambient temperature above room temperature. Tests which purport to measure high-speed circuit responses cannot be performed in the wafer-probe facility. Thus if there is only one other room-temperature facility type used, all high-speed responses must be measured there. Furthermore, whenever there is freedom of choice of facility type for a particular specification, the choice is dictated by the cost-minimization principle that specifications should be measured at the earliest possible point in the back-end test sequence.

Finally, then, since every specification has a unique test facility associated with it, and every test facility a unique back-end stage, then every specification has a unique back-end stage associated with it.

In order to proceed it is necessary to introduce some notation and terminology. First, the notion of specification is made precise. The total set of "specifications", as the term

has been used above, are considered be grouped by processing stage in which they are measured. Let

n^b number of stages in the processing stage tree.

b natural-number index of the stages.

Let the s'th *specification type* of the b'th processing stage, denoted (Ψ_{bs}, E_{bs}, e^*), consist of the following:

(1) a *specification description* Ψ_{bs}, which in turn consists of the following:

(a) a specification of the test circuit in which the circuit is to be tested, and the electrical excitations to be applied to it;

(b) a value of temperature θ_b to be imposed as a test ambient temperature.

(2) a *response function* E_{bs}, which is a function of measurable circuit responses, and, without loss of generality, positive by definition.

(3) a positive *specification level* e^*.

The following assumptions regarding the interpretation of E_{bs} are made without significant loss of generality, and, in the case of the third assumption, without any loss of generality:

(1) for each specification type in the problem, if all other characteristics of the design product are held fixed, the "desirability" of the product for each of its potential customers is a function of E_{bs} which is either non-decreasing or non-increasing.

(2) for each specification type in the problem, all potential customers agree as to whether the above directionality should be non-decreasing, or non-increasing.

(3) the E_{bs} can be so defined that the directionality is non-decreasing, for if a particular function of circuit responses lacks this property, its inverse does not.

These assumptions can be summarized as follows: every customer prefers large values of every response function E_{bs} to small ones.

In conventional statistical circuit optimization formulations, each specification type would have only one specification associated with it, and that specification would have only one specification level. In this formulation, each specification type may have a producer-determined number n_{bs}^{l} of specifications associated with it, which differ (only) in the value of their specification levels. Furthermore, these levels satisfy the criteria listed in Section 1.2 for eligibility as design parameters in the methodology. The capability to treat specification levels as design parameters is an important feature of the methodology. However, for a number of reasons, it is generally not possible and may not be desirable to treat all specification levels as design parameters.

Let

$\mathbf{e}_{bs}$ the n_{bs}^{e}-vector of all the specification levels pertaining to specification type s of processing stage b, which the producer wishes to treat as design parameters.

ε_{bs} the set of the elements of $\mathbf{e}_{bs}$.

ε_{bs}^{k} the set of all specification level constants that the producer wishes to apply to specification type s of processing stage b.

The vector $\hat{\mathbf{e}}_{bs}$ consists of the elements of $\varepsilon_{bs}^{k} \cup \varepsilon_{bs}$ ordered so that the smallest elements come first. Let the elements of $\hat{\mathbf{e}}_{bs}$ (which may be a mixture of constants and variables) be indexed by the natural-number variable l. Without loss of generality it can be assumed

that,

$$0 < \hat{e}_{bs1} \leq \hat{e}_{bs2} \leq \cdots \leq \hat{e}_{bsl} \leq \cdots \leq \hat{e}_{bsn_{bs}^{l}}$$

Note that for all b and s, there exists a $n_{bs}^{l} \times n_{bs}^{l}$ matrix with elements 0,1, or -1, say, $\mathbf{A}_{bs}^{e}$, and a n_{bs}^{l}-vector, say $\mathbf{k}_{bs}^{e}$, such that the above inequalities can be written as the linear constraint,

$$\mathbf{A}_{bs}^{e}\ \mathbf{e}_{bs} + \mathbf{k}_{bs}^{e} \leq 0. \tag{9.1}$$

Let

$$\mathbf{e}_{b} = [\mathbf{e}'_{b1}\ \mathbf{e}'_{b2}\ \cdots\ \mathbf{e}'_{bn_{b}^{s}}]',$$

and

$$\mathbf{e} = [\mathbf{e}'_{1}\ \mathbf{e}'_{2}\ \cdots\ \mathbf{e}'_{n^{b}}]'.$$

Specification $(\Psi_{bs}, E_{bs}, e_{bsl})$, $l \in \{1,2,\ldots,n_{bs}^{l}\}$ is said to be *satisfied* if and only if, when E_{bs} is measured according to specification description Ψ_{bs}, the inequality

$$E_{bs} \leq e_{bsl},$$

is satisfied.

The electrical performance of a particular circuit depends on device process disturbance random vectors having distributions differing not only from circuit to circuit, but from wafer to wafer. This has been reflected in the notation in Chapters 5 and 8, in which circuits are referenced as chip c of wafer w. The fact that in the back-end processing all the circuits from a number of wafers are pooled into a single collection does not eliminate the necessity for referencing quantities associated with individual circuits via the wafer index of its parent wafer. Hence, some quantities appropriately pertaining to the electrical performance of a circuit will be subscripted with w and c.

Let the *specification indicator random variable for circuit c of wafer w and specification type s of processing stage b,* denoted $Z_{bs,wc}$, be defined as follows. $Z_{bs,wc}$ is 0:

if after all back-end testing of the circuit is complete, it has not been subjected to testing for specification type s of processing stage b, or, it has been subjected to such testing but did not satisfy specification $(\Psi_{bs}, E_{bs}, e_{bsl})$, for any value of l,

and 1:

if after all back-end testing of the circuit is complete, it has been subjected to testing for specification type s of processing stage b, and specification index l^*, $l^* \in \{1,2,\ldots,n_{bs}^l\}$ is the smallest value of l such that $(\Psi_{bs}, E_{bs}, e_{bsl})$ is satisfied.

In words, then, $Z_{bs,wc}$ increases from zero over the integers to n_{bs}^l as the performance of circuit c of wafer w improves. Note that if two finished circuits named A and B, say, have received identical back-end treatments, and are in the same electrical performance category except that,

$$Z_{bs,(wc)_B} < Z_{bs,(wc)_A}$$

then the producer may if he so desires ship circuit A to any customer who is expecting a circuit with the characteristics of circuit B. This observation is trivial, yet this kind of substitution freedom will be seen to have a major impact on the modeling of profit.

Proceeding as in the definitions of $\mathbf{e}_b$ and $\mathbf{e}$, let the *stage specification indicator* (SSI) vector $\mathbf{Z}_{b,wc}$ be defined by,

$$\mathbf{Z}_{b,wc} = [Z_{b1,wc}\ Z_{b2,wc}\ \cdots\ Z_{bn_b^t,wc}]',$$

and the *specification indicator* vector $\mathbf{Z}_{wc}$ be defined by,

$$\mathbf{Z}_{wc} = [\mathbf{Z}'_{1,wc}\ \mathbf{Z}'_{2,wc}\ \cdots\ \mathbf{Z}'_{n^b,wc}]'.$$

$\mathbf{Z}_{wc}$, then, is a summary of the electrical performance of circuit c of wafer w.

The specification indicator notations established above to summarize the electrical characteristics of particular circuits also enables more general theoretical discussion of the characteristics of product variants, as well as possible outcomes of the back-end processing. In such discussions, the subscripting of variables by wc in order to reference particular circuits is omitted.

9.5. Modeling of Back-end Treatments

Let

n_b^τ number of back-end treatments which can be performed on the circuits in the discretionary substage of stage b.

And let the *stage treatment indicator* (STI) τ_b index the back-end treatments to which circuits may be subjected in stage b. τ_b is allowed to take on any value in $\{0, 1, \ldots, n_b^\tau\}$, with the value 0 reserved for the *null treatment*, that is for no treatment (and no incurred cost).

Certainly for at least one back-end stage for any product, if a customer orders a variant of the product which is claimed to have been subjected to a particular treatment in that stage, the customer will not accept circuits subjected to a different treatment. This is clear if only due to the first stage, which presumably includes at least one packaging treatment option. One package cannot be substituted for another. On the other hand, there are also back-end stages such that the producer can safely provide any of the customers of a particular variant of a product, a different treatment in that stage from that which the customers, in effect, ordered. As a trivial example, the producer can safely provide circuits which

have been subjected to burn-in operation and checking to customers who ordered ones that have not. A back-end stage in which substitution of alternative treatments is allowed, is called a *treatment-substitutable stage.*

For treatment-substitutable stages, the substitution relationship between any two treatments is almost always unidirectional. That is, one of the two can be substituted for the other but not vice versa. This being the case, unless n_b^τ is one, there may be a rather complex set of substitution relations between different treatments. What is assumed about this is that for all customers, treatment $(\tau_b)_A$ can be substituted for treatment $(\tau_b)_B$ if and only if $1 \leq (\tau_b)_B < (\tau_b)_A$. This is restrictive, but not significantly so, primarily due to the fact that for treatment-substitutable stages, n_b^τ seldom exceeds two.

The above implies a situation similar to one described in the previous section on electrical performance. It implies that if two finished circuits named A and B, say, are in the same electrical performance category, and have received identical back-end treatments, except that in treatment-substitutable stage b^*,

$$(\tau_{b^*})_B < (\tau_{b^*})_A$$

then the producer may if he so desires ship circuit A to any customer who is expecting a circuit with the characteristics of circuit B. The observation is similarly trivial but important.

Let the *treatment indicator* vector τ be defined as,

$$\tau = [\tau_1 \; \tau_2 \; \cdots \; \tau_{n_b^\tau}]'.$$

Then τ is a summary of all the treatments to which circuits can be subjected in the back-end processing.

9.6. Characteristic Combinations

Combining the results from the above two sections, the vector pair $(\mathbf{Z},\tau)$ can represent both the set of variants of a product and the characteristics that circuits exiting the back-end processing can have. So can the *characteristic combination* vector χ, defined as the following reordering of the elements of $\mathbf{Z}$ and τ:

$$\chi = [\mathbf{Z}'_1\ \tau_1\ \mathbf{Z}'_2\ \tau_2\ \cdots\ \mathbf{Z}'_{n^b}\ \tau_{n^b}]'.$$

As an example, it is next shown that the characteristics of the variants detailed in Figure 8 can be represented using the notation. The data and identifications which enable the representation are as follows.

Since only two test temperatures are called for, and packaging and burn-in can be accomplished in two stages, let $n^b = 2$. The voltage offset and output current specification types can be measured at wafer probe, and slew rate and bandwidth at final test, in the second stage. Therefore, let n_1^s and n_2^s both be 2. Let

E_{11} (1 mv.)/(voltage offset)

E_{12} output current

E_{21} slew rate

E_{22} bandwidth

$$Z_{11} = \begin{cases} 1 & \text{if } E_{11} \geq .5 \\ 0 & \text{otherwise} \end{cases}$$

$$Z_{12} = \begin{cases} 1 & \text{if } E_{12} \geq 25\,\text{ma.} \\ 0 & \text{otherwise} \end{cases}$$

$$Z_{21} = \begin{cases} 1 & \text{if } E_{21} \geq 120\text{v./}\mu\,sec. \\ 0 & \text{otherwise} \end{cases}$$

$$Z_{22} = \begin{cases} 2 & \text{if } E_{22} \geq 10 \text{ MHz} \\ 1 & \text{if } 5 \text{ MHz} \leq E_{22} < 10 \text{ MHz} \\ 0 & \text{otherwise} \end{cases}$$

Regarding the treatment substages, in stage 1, in which packaging must be done, let

$$\tau_1 = \begin{cases} 2 & \text{for circuits packaged in the TO-99 package} \\ 1 & \text{for raw dice (no package)} \\ 0 & \text{for circuits subjected to no treatment in stage 1} \end{cases}$$

Similarly, in stage 2, in which burn-in may be done, let

$$\tau_2 = \begin{cases} 2 & \text{for circuits subjected to normal treatment in stage 2, plus burn-in} \\ 1 & \text{for circuits subjected to normal treatment in stage 2, only} \\ 0 & \text{for circuits subjected to no treatment in stage 2} \end{cases}$$

Note that for both $\tau_1 = 1$ and $\tau_1 = 0$, the dice are not assembled into any package, but nevertheless the two cases must be distinguished, since for $\tau_1 = 1$, in general some incidental, cost-bearing operations must be performed on the dice prior to shipment, while $\tau_1 = 0$ represents the case in which dice are discarded.

Under these identifications and data, the product variants defined in Figure 8 can be summarized by their χ vectors as follows:

variant #	χ'
v1	[0 1 | 1 | 0 0 | 0]
v2	[1 1 | 1 | 0 0 | 0]
v3	[1 0 | 2 | 1 2 | 1]
v4	[1 0 | 2 | 1 2 | 2]
v5	[1 1 | 2 | 0 1 | 1]
v6	[1 1 | 2 | 0 2 | 2]

Not all the circuits resulting from the back-end processing have a χ vector which is among the definitions of the product variants. In the above example, Z_{11}, Z_{12}, and Z_{22} can take on two values, and τ_1, Z_{21}, and τ_2 three. The number of combinations of these values is 216. In the general case, the number of *characteristic combinations* of circuits is

$$\prod_{b=1}^{n^b} \{ [\prod_{s=1}^{n_b^s} (n_{bs}^l + 1)] [n_b^\tau + 1] \}$$

Let

Ω the set of χ vectors corresponding to all the characteristic combinations of the design product.

Ω^v the set of χ vectors representing variants of the design product.

In Figure 9 is a type of representation of the elements of Ω for the hypothetical circuit. The 54 (=216/4) lines represent all combinations of τ_1, Z_2, and τ_2. The complete set of 216 characteristic combinations can be generated by successively appending the 54 combinations listed "on the left" by the following allowed values of Z'_1: [0 0], [0 1], [1 0], and [1 1]. All information to be conveyed using the figure, pertaining to individual characteris-

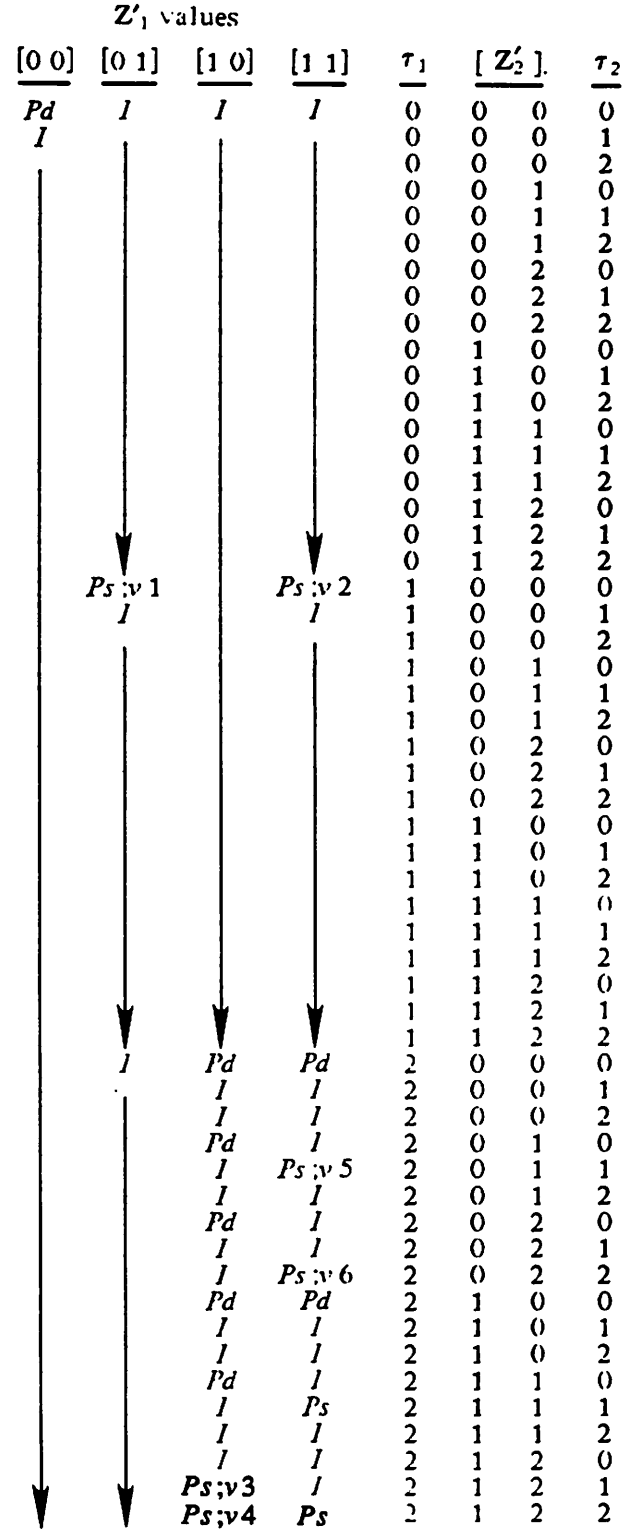


Fig. 9. Assignment of characteristic combinations to the three combination classes Ω^{I}, Ω^{Ps}, and Ω^{Pd}, and identification of characteristic combinations of the variants, for the hypothetical circuit.

tic combinations, is presented in columns labeled with the value of $\mathbf{Z'}_1$ of the combination of interest.

Note that the elements of Ω^v for the hypothetical circuit, listed above, are labeled in the figure with their variant number.

9.7. Generation and Properties of the Back-end Flow Tree

The generation of the back-end flow tree is a fundamental step in the profit modeling of individual IC products. With the notations which have been established to summarize the possible characteristics, electrical and otherwise, of the design product, the information needed to generate the tree is as follows:

(1) the structure of the $\mathbf{X}$ vector, as can be specified by the value of n^b and the number of elements in each vector $\mathbf{Z}_b$, $b = 1, 2, \ldots, n^b$.

(2) the set of all characteristic combinations Ω (or some smaller set of data which uniquely specifies it).

(3) the variant set Ω^v.

(4) which, if any, of the back-end stages are treatment-substitutable stages.

(5) the selection of back-end treatments which can be performed on circuits impingent on each discretionary node, from among all the treatments which are performed in the back-end stage to which the node belongs.

An explicit procedure for the generation of the back-end flow tree for individual products has been developed. Not surprisingly, generation begins with the root node and proceeds from "left to right", substage by substage, until the depth of the tree is reached. Each non-leaf node is "processed" to accomplish the following tasks:

(1) determine the number of its sons.

(2) determine the criteria that determine to which son individual circuits passing through the tree will move.

(3) determine which, if any, of the sons are leaf nodes.

(4) perform a categorization of some characteristic combinations that may be associated with the node.

In contrast, leaf nodes are "dummy" nodes - they serve as placeholders and require no processing.

The detailed generation procedure is more elaborate than one might expect. The goal here is only that of describing the salient results of the process.

In this description, it necessary to be able to refer to individual nodes of the back-end tree. They can be identified using a string of l^n natural numbers, where l^n is the level of the node. The root is identified by a null string. The sons of a particular node are indexed by the string of their father appended with a natural number to distinguish between the sons. Figure 7 illustrates the notation. Because no node in the tree for the hypothetical circuit has more than 10 sons, it is not necessary in this case to separate the numbers in the string by delimiters. The symbol η will be used generically to denote these strings. Variables that are associated with nodes use the η value of their node as a subscript.

The description uses the example of the hypothetical circuit. Its tree is shown in Figure 7. For ease of reference, the nodes in the tree have been labelled with figures having the number of their digits increasing according to the level of the node (the root label has no digits, i.e. is null).

One result of the generation is that, at each test-outcome node in stage b^*, the set of all possible stage specification indicator vectors $\mathbf{Z}_{b^*}$ is divided into groups, each of which is associated with a son (and hence a test-outcome branch) of the node. In Figure 7, every test-outcome branch is labeled with the members of its SSI vector group. Node 32 is particularly illustrative of this grouping. The test-outcome branch from node 32 to node 321 is the path taken by all the circuits impingent on node 32 that, in the tests represented by that node, demonstrate electrical performance represented either by $\mathbf{Z}_2 = [0\ 1]'$ or $\mathbf{Z}_2 = [1\ 1]'$.

Recall that each discretionary substage has a set of back-end treatments associated with it. However, individual nodes in a particular discretionary substage may not have all the treatments of their substage associated with them. It may not be necessary for any of the circuits impingent on such a node to be subjected to one or more of the treatments of the substage, in order that some finished circuits sellable as each variant be produced. In this connection, another of the results of the tree generation is that, for each discretionary node in stage b^*, say, the back-end treatments associated with stage b^*, indexed by STI τ_{b^*}, are divided into those that need to be performed on at least some of the circuits impingent on that node, and those that do not need to be performed on any circuits impingent on that node. The indices of the former set of treatments is said to belong to the *required treatment set* of the node, denoted T^r, and the indices of the treatments not needed are said to belong to the complement of T^r with respect to the set of all stage b^* treatments, denoted $\overline{T}^r$. Each discretionary node having its $\overline{T}^r$ non-empty has a dummy son and discretionary branch to represent this fact. And every treatment index belonging to the set T^r of a node has a discretionary branch and a son associated with it.

Note that in Figure 7, every discretionary branch is labeled either with a stage treatment indicator (τ value) or a set of τ values comprising the set $\overline{T}^r$ for the node to which the branch is connected on the left. The interpretation of this labeling can be effectively explained using an example. Those circuits impingent on node 211 in Figure 7 which are to be subjected to the treatments represented by $\tau_2 = 1$ and $\tau_2 = 1$, flow to nodes 2111 and 2112, respectively. These are the only treatments performed on circuits impingent on node 211.

Another of the results of the tree generation is that it assigns every element of Ω to one of three disjoint subsets of Ω, called *combination classes*:

(1) the class of *impossible combinations*, denoted Ω^I. This is the set of combinations such that it is impossible for any circuit exiting the back-end processing to be represented by that characteristic combination.

(2) the class of (possible) *sellable back-end outcomes*, denoted Ω^{Ps}. This is the set of combinations such that it is possible for circuits to exit the back-end processing with the combination, and the combination represents circuit characteristics that are equal to or better than those of at least one product variant (thus representing circuits which the producer intends to sell). The elements of this class are further divided into disjoint *sellable back-end outcome groups* (*sellable groups* for short), with the members of each group sharing the same leaf node by which they exit the flow.

(3) the class of (possible) *discard back-end outcomes*, denoted Ω^{Pd}. This is the set of combinations such that it is possible for circuits to exit the back-end processing with the combination, and the characteristics of the combination equal or exceed

those of no product variant (and thus the combination represents circuits which presumably must be discarded). The elements of this are further divided into disjoint *discard back-end outcome groups* (*discard groups* for short), with the members of each group sharing the same leaf node by which they exit the flow.

Membership of the elements of Ω in these classes is determined in the course of the generation of the tree. For example, in the generation of the tree for the hypothetical circuit, after node 21 is processed, all the elements of Ω which have $\mathbf{Z}_1 = [1\ 0]'$, $\tau_2 = 2$, and do not have $\mathbf{Z}_2 = [1\ 2]'$ are assigned to the set Ω^{Pd} of discard back-end outcomes. Figure 9 is used to record the assignment of elements of Ω to one of its three combination subsets, for the hypothetical circuit. If a particular χ vector is assigned to Ω^{Pd}, its subset membership is indicated in the figure with the superscript Pd, and similarly for the other combination subsets Ω^{Ps} and Ω^{I}.

As will be evident in the sequel, the back-end outcome groups play an important role in the profit model. The following notation will be used in identifying them. Let

n^{os} the number of sellable back-end outcome groups.

n^{o} the total number of back-end outcome groups (sellable plus discard).

Let the the back-end outcome groups be denoted by G_o^B, indexed by $o \in \{1, 2, \ldots, n^{os}\}$ for the sellable groups, and $o \in \{n^{os}+1, n^{os}+2, \ldots, n^{o}\}$ for the discard groups. Note that except for the requirement that the first n^{os} integers be assigned to sellable groups, the assignment of o values to back-end groups is arbitrary.

For the hypothetical circuit, $n^{os} = 7$. The seven sellable groups for the product are specified in Figure 7 adjacent the leaf nodes to which they correspond. The cardinality of G_5^B establishes that the use of the term "group" in this context is not a misnomer. The

discard groups are not labeled on the figure, due to their high cardinality.

The important role of the back-end outcome groups in the profit model is due to two properties they have, described next.

First, in general, two circuits share the same costs of back-end processing if and only if the circuits are represented by characteristic combinations in the same back-end outcome group. This is because they share the same costs if and only if they travel the same path through the tree. Since the path through the tree taken by a particular circuit is uniquely determined by its characteristic combination vector χ, and the back-end outcome groups are disjoint, two circuits travel the same path through the tree if and only if they are represented by characteristic combinations in the same back-end outcome group.

Second, two sellable circuits share the same potential to create revenue for the producer if and only if the circuits are represented by characteristic combinations in the same sellable back-end outcome group. This is because two sellable circuits share the same potential to create revenue for the producer if and only if the set of variants as which they can be sold is identical. This is the case if and only if the circuits are represented by characteristic combinations in the same sellable back-end outcome group. This latter equivalence is not obvious, but is a fundamental property resulting from the generation of the tree. Not having demonstrated how this property is achieved, it is important to at the least state the property concisely. This will also be needed in subsequent discussion of the relationship between the back-end outcome groups and the characteristic combinations of the variants. First some preliminary definitions are needed.

For $b \in \{1, 2, \ldots, n^b\}$, one SSI vector (see Section 9.4) $(\mathbf{Z}_b)_A$ is said to *cover* another $(\mathbf{Z}_b)_B$, denoted $(\mathbf{Z}_b)_B \leq (\mathbf{Z}_b)_A$, if and only if

$$(Z_{bs})_B \leq (Z_{bs})_A \, , \, s = 1,2,\ldots,n_b^s.$$

For $b \in \{1,2,\ldots,n^b\}$, one STI (see Section 9.5) $(\tau_b)_A$ is said to *cover* another $(\tau_b)_B$, denoted $(\tau_b)_B \precsim (\tau_b)_A$, if and only if

$$(\tau_b)_B \leq (\tau_b)_A,$$

for treatment-substitutable stages, and

$$(\tau_b)_B = (\tau_b)_A,$$

for stages which are not treatment-substitutable stages.

Also, given two characteristic combination vectors

$$(\chi)_A = [(Z'_1)_A \ (\tau_1)_A \ (Z'_2)_A \ (\tau_2)_A \ \cdots \ (Z'_{n^b})_A \ (\tau_{n^b})_A]' \in \Omega,$$

and

$$(\chi)_B = [(Z'_1)_B \ (\tau_1)_B \ (Z'_2)_B \ (\tau_2)_B \ \cdots \ (Z'_{n^b})_B \ (\tau_{n^b})_B]' \in \Omega,$$

$(\chi)_A$ is said to *cover* $(\chi)_B$, denoted $(\chi)_B \precsim (\chi)_A$, if and only if for $b = 1,2,\ldots,n^b$,

$$(Z_b)_B \precsim (Z_b)_A$$

and

$$(\tau_b)_B \precsim (\tau_b)_A.$$

For any $\chi^* \in \Omega$, define the set $V(\chi^*) \subset \Omega^v$ by

$$V(\chi^*) = \{ \chi \in \Omega^v \mid \chi \precsim \chi^* \}.$$

Then the second property that each back-end outcome group G_o^B has, $o = 1,2,\ldots,n^o$, is that for any χ_1 and $\chi_2 \in \Omega^{Ps} \cup \Omega^{Pd}$, $V(\chi_1)$ and $V(\chi_2)$ are the same if and only if χ_1 and χ_2 belong to the same back-end outcome group. As a consequence each back-end outcome group G_o^B has a unique set of variants associated with it.

For the hypothetical circuit, the names of the variants associated with the seven sellable back-end outcome groups are shown in Figure 7 beneath the specifications of the groups.

Note that the objective of endowing the back-end outcome groups with these important two properties is the fundamental principle shaping the tree generation procedure.

9.8. Modeling of Circuit Counts in the Back-end Processing.

Now that the categories of finished product which are important to the profit model have been identified as the back-end outcome groups, it remains to model the numbers of circuits originating with n^w input wafers, which emerge from the back-end in each of these groups. These are the *back-end outcome circuit counts* N_o^c, $o = 1,2,\ldots,n^o$, defined by,

N_o^c the number of circuits originating from n^w wafers, represented by χ vectors in back-end outcome group o.

Note that generically the symbol N^c is used as a random variable denoting circuit counts.

The modeling is accomplished by modeling N_o^c for arbitrary o. The modeling of N_o^c is decomposed at the highest level using the fact that the circuit count from n^w wafers is the sum of the counts associated with each of the wafers. Formally, let

N_{ow}^c the number of circuits originating from wafer w, represented by χ vectors in back-end outcome group o.

Then,

$$N_o^c = \sum_{w=1}^{n^w} N_{ow}^c. \tag{9.2}$$

The modeling of N_{ow}^c is accomplished by modeling it for arbitrary w.

This outcome count is in general diminished by the numbers of circuits on wafer w found at wafer probe to have been the victim of catastrophic defects. The basic approach in modeling this count is to first model it ignoring catastrophic defect phenomena. The

resultant model is then integrated with the general defect yield model discussed in Chapter 7 to give the desired result.

Quantities and models defined assuming no catastrophic defects occur, are referred to as *zero-defect* quantities and models. Circuit count variables appended with the symbol ~ denote the same counts under the zero-defect assumption. In particular, the quantity of immediate interest is the *zero-defect back-end outcome circuit count* $\tilde{N}^c_{ow}$ (for arbitrary o and w):

$\tilde{N}^c_{ow}$ the number of circuits originating from wafer w, represented by χ vectors in back-end outcome group o, assuming no circuits have been discarded due to catastrophic defects.

To this point, the focus of the modeling has been narrowed from $\{N^c_o, o = 1,2,\ldots,n^o\}$, to $\tilde{N}^c_{ow}$ for arbitrary o and w. Two more assumptions will be invoked before an expression for a circuit count is presented. First, note that given o, the path through the flow tree taken by circuits having outcome o is determined, as was implied near the end of the previous section. In general there are some discretionary nodes in that path, at each of which certain numbers of circuits are assigned to the various treatments associated with the node. It is temporarily assumed that at each discretionary node in the path in question, the proportion of circuits which are assigned to the son that is in the path of interest, is unity, i.e. 100%. Second, it is assumed that w does not denote a wafer that has been discarded in the front-end processing.

Now recall from Section 9.6 that corresponding to a given characteristic combination vector $\boldsymbol{\chi}$, there is a specification indicator vector $\mathbf{Z}$, obtained by deleting the treatment indicators from the χ vector. Similarly, a given set of characteristic combination vectors,

say Ω^*, has a corresponding set of $\mathbf{Z}$ vectors, denoted as $Z(\Omega^*)$. Recall also that for circuit c of wafer w, the specification indicator is denoted $\mathbf{Z}_{wc}$.

$\mathbf{Z}_{wc}$ is a function of quantities already introduced in the modeling. Specifically, the set of disturbances $\{\mathbf{D}_{wcd}\}_{d=1}^{n^d}$ together with the device dimension vector $\mathbf{x}$ determine through circuit simulation all the response functions of circuit c of wafer w, and when the response functions are compared with appropriate elements of the specification level vector $\mathbf{e}$, $\mathbf{Z}_{wc}$ can be determined. Hence,

$$\mathbf{Z}_{wc} = \mathbf{Z}_{wc}(\{\mathbf{D}_{wcd}\}_{d=1}^{n^{dc}}, \mathbf{x}, \mathbf{e}),\ c = 1,2,\ldots,N^{cl}(\Lambda,\mathbf{x}),\ w = 1,2,\ldots,n^w.$$

Using an indicator function, the event that circuit wc is represented by a characteristic combination vector in outcome group o is denoted as,

$$1_{\mathbf{Z}_{wc} \in Z(G_o^B)} = 1.$$

Therefore, under the prevailing assumptions,

$$\tilde{N}_{ow}^c = \sum_{c=1}^{N^{cl}(\Lambda,\mathbf{x})} 1_{\mathbf{Z}_{wc} \in Z(G_o^B)}.$$

It is not difficult to remove the front-end survival assumption. If a particular wafer is discarded in the front-end processing, the number of circuits it can contribute to outcome group o is certainly zero. Using the survival indicator 1_w^{BE} defined in Chapter 6, the extended result is just,

$$\tilde{N}_{ow}^c = 1_w^{BE} \sum_{c=1}^{N^{cl}(\Lambda,\mathbf{x})} 1_{\mathbf{Z}_{wc} \in Z(G_o^B)}. \tag{9.3}$$

Generalizing this expression to remove the assumption that at discretionary nodes all circuits flow to the son in the path corresponding to outcome o, is not nearly so trivial. Some digression regarding the quantification of the splits at discretionary nodes is required.

It is instructive to consider the most ambitious approach to controlling the splits that occur at discretionary nodes. This would be for the producer to decide circuit by circuit which of the treatments selectable at a particular discretionary node the circuits is to receive, utilizing all the information about the circuit which has been previously generated. This information would include all test-limit measurement data pertaining to the parent wafer of the circuit, and all previously generated electrical test results (by which is meant raw real-valued test data, not SSI vectors). In current practice in the industry none of this is done. Indeed, it is likely that no analysis or optimization tool designed to make such decisions, no matter how sophisticated, could bring about a profit increase, when the high cost of maintaining such data is considered. In particular, no information about individual circuits impingent on a particular discretionary node is used. The practice is to simply control the number of circuits that are assigned to the various treatments selectable at the node.

Clearly these numbers can effect the quantities of circuits ending up in the various back-end outcome groups. Hence it is necessary to model the numbers of circuits to be subjected to the various treatments associated with the node. More concisely, let j generically index the treatments that belong to the required treatment sets T^r of the various nodes. Let η represent an arbitrarily chosen discretionary split node. And let

n_η^j cardinality of the required treatment set at node η.

Then it is necessary to model the number of circuits resulting from the fabrication of n^w wafers which receive each of the n_η^j treatments required at the node.

Let

$\tilde{N}^c_\eta$ the number of circuits resulting from the n^w wafers started in production, which are impingent on node η, assuming no circuits have been discarded due to catastrophic defects.

$\tilde{N}^c_{\eta j}$ the number of circuits resulting from the n^w wafers started in production, which move to son j of node η,

Clearly, if the producer is given some fixed number $\tilde{N}^c_\eta$ of circuits impingent on a discretionary node, since the assignment of circuits to treatments takes place within his production facilities, the producer has freedom to set the values of the $\tilde{N}^c_{\eta j}$. Of course the freedom is hardly complete. First, clearly

$$0 \le \tilde{N}^c_{\eta j}, \; j = 1,2,\ldots,n^j_\eta \tag{9.4}$$

must be satisfied. Also, since circuits are not created in the midst of the back-end flow, there is the constraint that

$$\sum_{j=1}^{n^j_\eta} \tilde{N}^c_{\eta j} \le \tilde{N}^c_\eta.$$

In general, some small portion of circuits subjected to a back-end treatment may be fatally damaged by the treatment and need to be discarded. For simplicity, this phenomenon is ignored. Consequently, the conservation constraint

$$\sum_{j=1}^{n^j_\eta} \tilde{N}^c_{\eta j} = \tilde{N}^c_\eta \tag{9.5}$$

is assumed. From this it is clear that the producer has $n^j_\eta - 1$ degrees of freedom at node η, i.e., he is free to pick $\tilde{N}^c_{\eta j}$ for $j = 1,2,\ldots,n^j_\eta - 1)$ and $\tilde{N}^c_{\eta n^j_\eta}$ is then determined by,

$$\tilde{N}^c_{\eta n^j_\eta} = \tilde{N}^c_\eta - \sum_{j=1}^{n^j_\eta - 1} \tilde{N}^c_{\eta j} \tag{9.6}$$

Then the inequalities of (9.4) can be expressed in terms of $\tilde{N}^c_{\eta j}$ for $j = 1,2,\ldots,(n^j_\eta - 1)$,

as,

$$0 \le \tilde{N}^c_{\eta j} \ , \ j = 1,2,\ldots,(n^j_\eta - 1) \tag{9.7}$$

and, from (9.6),

$$\sum_{j=1}^{n^j_\eta - 1} \tilde{N}^c_{\eta j} \le \tilde{N}^c_\eta \tag{9.9}$$

There is another sense in which the producer is not free to set values of $\tilde{N}^c_{\eta j}$. $\tilde{N}^c_{\eta j}$ was taken above to be a given fixed number of circuits, yet absolute numbers of circuits associated with any of the nodes of the back-end tree are random variables depending in complex and important ways on phenomena that take place in the fabrication of the wafers. The control the producer has over the splitting at discretionary nodes is much more appropriately parametrized using ratios of numbers of circuits flowing out of these nodes to the number flowing in, providing the latter is not zero. (If it is zero, note that no modeling of the split at the node in question is needed.) Accordingly, let the *discretionary split fraction (DSF) of the j'th treatment of discretionary node* η, denoted $\alpha_{\eta j}$, be defined for $j = 1,2,\ldots,n^j_\eta$ as,

$$\alpha_{\eta j} = \tilde{N}^c_{\eta j} / \tilde{N}^c_\eta .$$

Dividing (9.5) by $\tilde{N}^c_\eta$ gives the corresponding conservation constraint,

$$\sum_{j=1}^{n^j_\eta} \alpha_{\eta j} = 1. \tag{9.10}$$

Similarly (9.6) becomes,

$$\alpha_{\eta n^j_\eta} = 1 - \sum_{j=1}^{n^j_\eta - 1} \alpha_{\eta j} .$$

reflecting that given $\alpha_{\eta j} \ , \ j = 1,2,\ldots,(n^j_\eta - 1)$, $\alpha_{\eta n^j_\eta}$ can be determined. Division of (9.7) and (9.9) by $\tilde{N}^c_\eta$ gives the constraints which apply to the first $(n^j_\eta - 1)$ $\alpha_{\eta j}$ values,

$$0 \leq \alpha_{\eta j}\ ,\ j\ =\ 1,2,\ldots,(n_{\eta}^{j}-1), \tag{9.11}$$

and

$$\sum_{j=1}^{n_{\eta}^{j}-1} \alpha_{\eta j} \leq 1. \tag{9.12}$$

Inasmuch as the numbers of circuits in the definitions of the discretionary split fractions are non-negative integers, they should be restricted to the set of positive rationals. However, typical values of numbers of circuits are sufficiently large that negligible error is introduced by allowing them to take on any real value satisfying (9.11) and (9.12) above.

The modeling of circuit counts at arbitrary discretionary split node η can be restated in terms of the new parametrization. The producer is free to select the values of the $n_{\eta}^{j}-1$ directly implementable real-valued parameters $\alpha_{\eta j}\ ,\ j\ =\ 1,2,\ldots,(n_{\eta}^{j}-1)$, subject to the constraints of (9.11) and (9.12). Since the flow model imposes no other constraints on these parameters, they are referred to as the *independent discretionary split fractions (IDSF's)*. The terminology is useful when it is desired to exclude $\alpha_{\eta n_{\eta}^{j}}$.

In Figure 10 is repeated the structure of the tree shown in Figure 7, which is that of the hypothetical circuit, but with different labeling. In Figure 10, the discretionary branches of the tree are labeled with the set of α variables appropriate for the quantification of the splits at the discretionary nodes.

There are various extensions of the back-end flow model which, contrary to the conservation relation of (9.10), allow for loss of circuits in the back-end treatments. One simple extension is conveniently characterized in terms of DSF's. It is to require the split fractions to sum to some value (less than one) which simulates a certain average loss rate in the treatments due to random effects which do not depend on design parameter values.

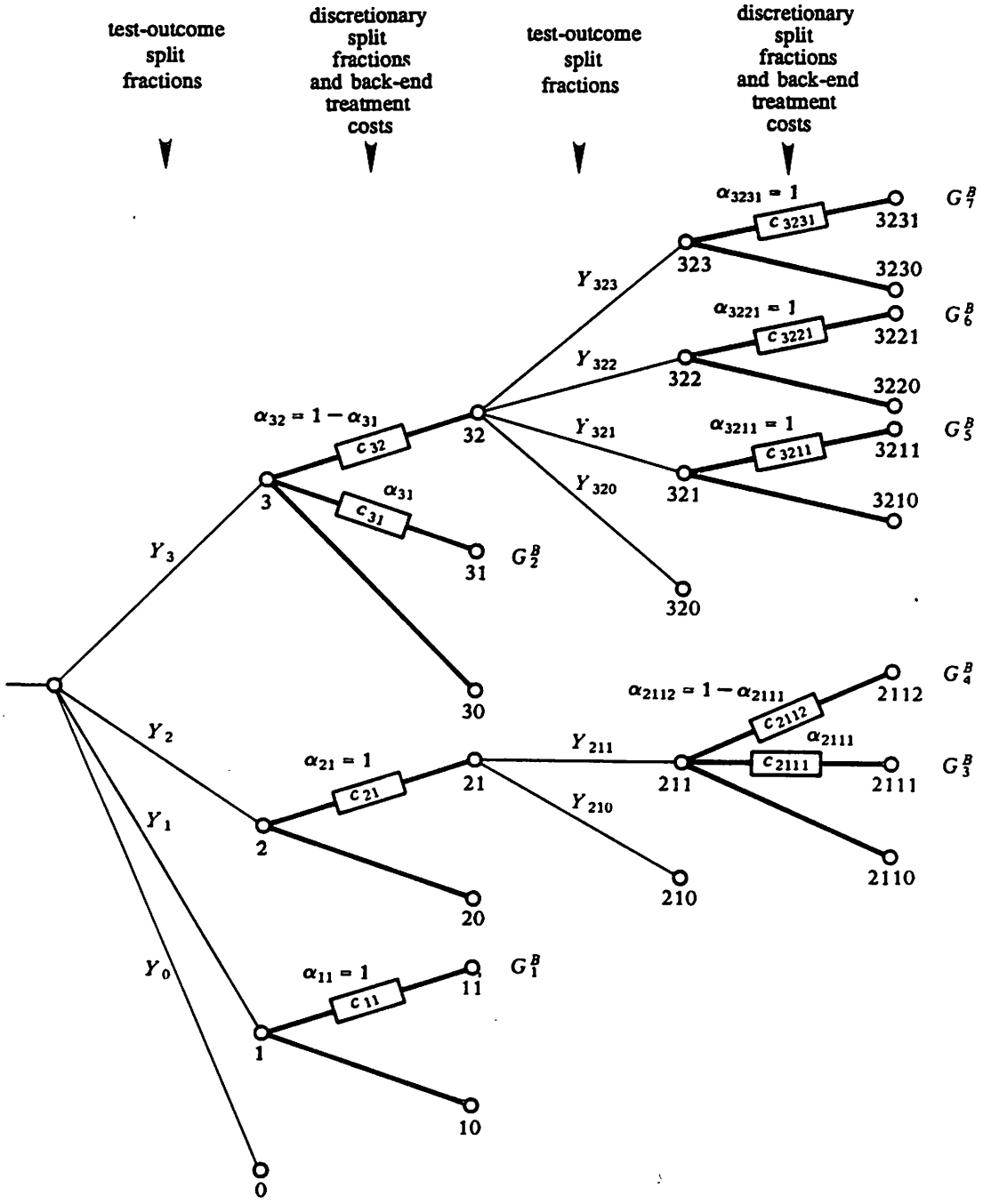


Fig. 10. Back-end processing tree for the hypothetical circuit, showing variables quantifying the splits.

(Such loss rates are generally low - summing to .97 might be typical.)

Since the values of the IDSF's at all the non-leaf discretionary nodes of the tree in general affect how well the proportions of treatments in the sellable back-end outcomes matches with the proportions demanded in the market, they in general have a first-order effect on the product's contribution profit. They also have the other properties listed at the beginning of Section 1.2 enabling them to be treated as design parameters in the methodology, and they are so treated. The *discretionary split fraction vector*, denoted α, is the vector with elements the IDSF's of all the non-leaf discretionary nodes in the tree. Note that there exists a matrix, say $\mathbf{A}^{\alpha}$, with more rows than columns and with each row consisting either of a single element -1 and the remaining elements zero, or of some elements 0 and some +1, and a vector, say $\mathbf{k}^{\alpha}$, with elements either 0 or -1, such that the inequalities of (9.11) and (9.12)) can be written as the linear constraint,

$$\mathbf{A}^{\alpha}\,\alpha + \mathbf{k}^{\alpha} \leq 0. \tag{9.13}$$

Having modeled the discretionary splits, the discussion returns to the more fundamental objective of determining the back-end outcome circuit count $\tilde{N}^{c}_{ow}$. To accomplish this, it is necessary to have a more precise notation to index the treatments at the discretionary node. Let j_b generically index the required treatments at nodes in stage b. Similarly, let g_b generically index the test-outcome groups associated with test-outcome nodes in stage b.

The number of circuits with characteristic combinations in a particular back-end outcome group depends on the values of all the DSF's encountered by the circuits in traversing the path through the tree associated with that outcome group. Let

n_o^{α} the number of discretionary splits in the path from root to leaf node for back-end outcome o.

With arbitrary DSF's the expression for the desired back-end circuit count of (9.3) must be modified to,

$$\tilde{N}^c_{ow} = [\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o^\alpha} j_{n_o^\alpha};o}\ ] \times [1^{BE}_w \sum_{c=1}^{N^{cl}(\Lambda,\mathbf{x})} 1_{Z_{wc} \in Z(G^B_o)}], \tag{9.14}$$

where the "$;o$" appendage on the subscripts of the α's is used to signify that all the preceding subscripts depend on the outcome index o (since the outcome determines the path through the tree). Note that the expression is in the form of a product of a discretionary factor, and a test-outcome factor containing indicator functions. This expression is called the *indicator-based* form of the back-end outcome counts.

If the use of ratios of circuit counts at nodes is applied to the test-outcome nodes as it has been to the discretionary nodes, it enables a more intuitive, *yield-based* expression for the above test-outcome factor, and, in the next chapter, for the back-end cost. Let

$\tilde{N}^c_{\eta;w}$ the number of circuits from wafer w impingent on test-outcome node η, assuming no circuits have been discarded due to catastrophic defects.

n^g_η highest numbered index g of the test-outcome groups at test-outcome node η.

$\tilde{N}^c_{\eta g;w}$ the number of circuits from wafer w, which move to son g of node η, assuming no circuits have been discarded due to catastrophic defects.

Then the *test-outcome split fraction*, or *yield of test-outcome group g and node η, for wafer w*, denoted $Y_{\eta g;w}$, is defined as,

$$Y_{\eta g;w} = \tilde{N}^c_{\eta;w} / \tilde{N}^c_{\eta g;w}, g = 0, 1, \ldots, n^g_\eta\ .$$

Both numerator and denominator in the definition depend on the processing disturbances $\{\mathbf{D}_{wcd}\}_{d=1\ c=1}^{n^{dc}\ N^{cl}(\Lambda,\mathbf{x})}$, the device dimension vector $\mathbf{x}$, the specification level vector $\mathbf{e}$, and the

discretionary split fractions of the discretionary splits encountered by the circuits between the root and the test-outcome node in question. The latter quantities, however, appear as common factors in both numerator and denominator, so that the yields are independent of them. Hence,

$$Y_{\eta g;w} = Y_{\eta g;w}(\{\mathbf{D}_{wcd}\}_{d=1\ c=1}^{n^{dc}\ N^{cl}(\Lambda,\mathbf{x})}, \mathbf{x}, \mathbf{e}), \ g = 0, 1, \ldots, n_{\eta}^{g} .$$

In Figure 10, the test-outcome branches of the tree for the hypothetical circuit are labeled with the set of Y variables, with the subscripting by $;w$ omitted, appropriate for the quantification of the splits at the test-outcome nodes.

Now, assume again that wafer w has survived the front-end processing. The number of circuits from the wafer impingent on the root is N^{cl}. The number of circuits impingent on son g_1 of the root is then $N^{cl}\, Y_{g_1;w}$. The number of circuits impingent on test-outcome son j_1 of son g_1 of the root is then $N^{cl}\, Y_{g_1;w}\, \alpha_{g_1 j_1}$. Continuing this process, it is not difficult to see that, with the survival assumption removed, the desired outcome counts can be written in one of the two "chronologically natural" forms, depending on whether the circuits exit the tree in a test-outcome or discretionary substage,

$$\tilde{N}_{ow}^{c} = 1_{w}^{BE} N^{cl}\, Y_{g_1;ow}\, \alpha_{g_1 j_1;o} Y_{g_1 j_1 g_2;ow} \cdots \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o^{\alpha}} j_{n_o^{\alpha}};o} ,$$

or

$$\tilde{N}_{ow}^{c} = 1_{w}^{BE} N^{cl}\, Y_{g_1;ow}\, \alpha_{g_1 j_1;o} Y_{g_1 j_1 g_2;ow} \cdots Y_{g_1 j_1 g_2 j_2 \cdots g_{n_o^{Y}};ow} .$$

where,

n_o^{Y} number of test-outcome splits in the path from root to leaf node for back-end outcome o.

These two cases can be combined to give the yield-based form of the expression for the outcome counts for outcome o and wafer w,

$$\begin{aligned}\tilde{N}^c_{ow} = &[\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o \alpha} j_{n_o \alpha};o}\ ] \\ &\times [N^{cl}\ 1^{BE}_w Y_{g_1;ow}\ Y_{g_1 j_1 g_2;ow}\ \cdots Y_{g_1 j_1 g_2 j_2 \cdots g_{n_o Y};ow}],\end{aligned} \tag{9.15}$$

The assumption made near the outset of this section, that there is no loss of circuits due to catastrophic failures resulting from local defects, can now be addressed. In Chapter 7, the wafer w defect yield random variable Y^{df}_w was introduced. Its definition suggests that it should have a simple multiplicative effect on the circuit count for wafer w. That is,

$$N^c_{ow} = Y^{df}_w \tilde{N}^c_{ow}.$$

With, this, the indicator-based form of the back-end outcome count of (9.14) becomes,

$$\begin{aligned}N^c_{ow} = &[\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o \alpha} j_{n_o \alpha};o}\ ] \\ &\times [1^{BE}_w Y^{df}_w \sum_{c=1}^{N^{cl}} 1_{Z_{wc} \in Z(G^B_o)}],\end{aligned}$$

and the yield-based form of (9.15) is modified to,

$$\begin{aligned}N^c_{ow} = &[\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o \alpha} j_{n_o \alpha};o}\ ] \\ &\times [1^{BE}_w Y^{df}_w N^{cl}\ Y_{g_1;ow}\ Y_{g_1 j_1 g_2;ow}\ Y_{g_1 j_1 g_2 j_2 \cdots g_{n_o Y};ow}].\end{aligned}$$

As an illustrative example of this equation, referring to Figure 10, it is not difficult to write for back-end outcome $o = 6$ of the hypothetical circuit,

$$N^c_{6w} = \alpha_{32}\ \alpha_{3221}\ 1^{BE}_w Y^{df}_w N^{cl}\ Y_{3;w}\ Y_{322;w}$$

Now, finally, carrying out the summation of (9.2), and recalling that o was arbitrary, gives the indicator-based expression,

$$N_o^c = [\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o^\alpha} j_{n_o^\alpha};o}]$$
$$\times [\sum_{w=1}^{n^w} 1_w^{BE} Y_w^{df} \sum_{c=1}^{N^{cl}} 1_{Z_{wc} \in Z(G_o^B)}], \ o = 1,2,\ldots,n^o.$$

Similarly, the yield-based expression is,

$$N_o^c = [\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o^\alpha} j_{n_o^\alpha};o}]$$
$$\times [N^{cl} \sum_{w=1}^{n^w} 1_w^{BE} Y_w^{df} Y_{g_1;ow}\ Y_{g_1 j_1 g_2;ow} \cdots Y_{g_1 j_1 g_2 j_2 \cdots g_{n_o^Y};ow}], \ o = 1,2,\ldots,n^o.$$

The N_o^c are random variables which depend on design parameters and previously introduced random variables. These dependencies can be explicitly displayed for the indicator-based form as,

$$\begin{aligned} N_o^c = {} & [\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o^\alpha} j_{n_o^\alpha};o}] \\ & \times [\sum_{w=1}^{n^w} 1_w^{BE}(\{\mathbf{D}_{wtd}\}_{d=1\ t=1}^{n^{dt}\ n^t}, 1)\ Y^{df}(\mathbf{Y}_w, \Lambda, \mathbf{x}) \\ & \times \sum_{c=1}^{N^{cl}(\Lambda,\mathbf{x})} 1_{Z_{wc} \in Z(G_o^B)}(\{\mathbf{D}_{wcd}\}_{d=1}^{n^{dc}}, \mathbf{x}, \mathbf{e})], \ o = 1,2,\ldots,n^o. \end{aligned} \tag{9.16}$$

and for the yield-based form as,

$$\begin{aligned} N_o^c = {} & [\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o^\alpha} j_{n_o^\alpha};o}] \\ & \times [N^{cl}(\Lambda, \mathbf{x}) \sum_{w=1}^{n^w} 1_w^{BE}(\{\mathbf{D}_{wtd}\}_{d=1\ t=1}^{n^{dt}\ n^t}, 1)\ Y^{df}(\mathbf{Y}_w, \Lambda, \mathbf{x}) \\ & \times Y_{g_1;ow}\ Y_{g_1 j_1 g_2;ow} \cdots Y_{g_1 j_1 g_2 j_2 \cdots g_{n_o^Y};ow}], \ o = 1,2,\ldots,n^o, \end{aligned}$$

where all the parametric yields depend on $\{\mathbf{D}_{wcd}\}_{d=1}^{n^{dc}}$, $\mathbf{x}$, and $\mathbf{e}$. Even more simply, the dependence of the N_o^c on design parameters and random variables can be symbolized as,

$$N_o^c = N_o^c(\{\mathbf{D}_{wtd}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}, 1, \{\mathbf{D}_{wcd}\}_{d=1\ c=1\ w=1}^{n^{dc}\ N^{cl}(\Lambda,\mathbf{x})\ n^w}, \mathbf{x}, \Lambda, \{\upsilon_w\}_{w=1}^{n^w}, \mathbf{e}, \alpha).$$

As will be seen in Chapter 12, it is necessary to determine the *expected back-end circuit counts* n_o^c, which are the expected values of the N_o^c. Before an expression for these can be written, some comments and definitions pertaining to the distributions of the pertinent random variables are needed.

First, let $f_\Lambda(\lambda)$ and $f_Y(\upsilon)$ denote the density funtcions of the random variables Λ and Y, respectively. The distributions of the disturbance random vectors in the sets $\{\mathbf{D}_{wid}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}$ and $\{\mathbf{D}_{wcd}\}_{d=1\ \ c=1\ \ w=1}^{n^{dc}\ N^{cl}(\Lambda,\mathbf{x})\ n^w}$ are all independent multivariate normal distributions, but there is a rather complex hierarchical system by which they are parametrized (see [MAL82]). Expectation integrals of random variables depending on these disturbances are very unwieldy, due to the many density functions and differential elements which must be present. Therefore, for simplicity, the entire multivariate density for the disturbances associated with wafer w will be symbolized by $f_{\mathbf{D}^w}(\mathbf{d}_w)$. Since the symbol $\mathbf{d}$ is allocated to the representation of disturbances, the differential element will be symbolized unconventionally here with ∂. Also, a single integration symbol will suffice to indicate multiple integration. With all of this, then, the expected circuit counts can be written as,

$$n_o^c = \int N_o^c(\{\mathbf{d}_{wid}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}, \mathbf{l}, \{\mathbf{d}_{wcd}\}_{d=1\ \ c=1\ \ w=1}^{n^{dc}\ N^{cl}(\lambda,\mathbf{x})\ n^w}, \mathbf{x}, \lambda, \{\upsilon_w\}_{w=1}^{n^w}, \mathbf{e}, \alpha)$$
$$\prod_{w=1}^{n^w} [f_{\mathbf{D}^w}(\mathbf{d}_w)\ \partial\mathbf{d}_w f_Y(\upsilon_w)\ \partial\upsilon_w] f_\Lambda(\lambda)\ \partial\lambda \tag{9.17}$$

Of course in applying any of these forms, the design parameter constraints of (9.1),

$$\mathbf{A}_{bs}^e\ \mathbf{e}_{bs} + \mathbf{k}_{bs}^e \le 0\ ,\ s = 1,2,\ldots,n_b^s\ ,\ b = 1,2,\ldots,n^b\ . \tag{9.1}$$

and of (9.13),

$$\mathbf{A}^\alpha\ \alpha + \mathbf{k}^\alpha \le 0\ , \tag{9.13}$$

must be included.

Chapter 10

MODELING OF BACK-END COST

In the context of the back-end model structure described in the previous chapter, all the costs of the back-end processing are incurred in performing the various back-end treatments associated with the design product. Specifically, as was described in Sections 9.5 and 9.7, each processing stage b has n_b^τ distinct treatments to which circuits passing through it may be subjected, indexed by $\tau_b \in \{1,2,\ldots,n_b^\tau\}$. Let

$c_{\tau_b}^b$ (back-end) cost per circuit of performing treatment τ_b.

These costs are usually adequately modeled as constants. However for the sake of generality, they are modeled as having a dependence on accumulated production experience with the design product, w^h. That is,

$$c_{\tau_b}^b = c_{\tau_b}^b(w^h), \tau_b = 1,2,\ldots,n^b.$$

Explicit display of this dependence, however, will be temporarily omitted for notational

simplicity. At each discretionary node η, for each $j \in \{1, 2, \ldots, n_\eta^j\}$ there is in addition to a non-leaf son and a DSF $\alpha_{\eta j}$, a treatment $\tau_{\eta j}^r$ and its associated cost $c_{\tau_{\eta j}^r}^b$. The costs of the back-end treatments for the hypothetical circuit are labeled in the rectangles representing the treatments, in Figure 10.

As in the modeling of circuit counts, there are two alternative forms of the expression for back-end cost - indicator-based, and yield-based. The former is developed first.

There are two key ideas underlying the indicator-based expression. The first was articulated in Section 9.7: all circuits represented by characteristic combination vectors belonging to the same back-end outcome group share the same back-end cost. Accordingly, let

c_o^b back-end cost associated with outcome o.

Second, trivially, this cost is just the sum of costs incurred in the discretionary substages. Recalling that n_o^α is the number of discretionary splits in the path from root to leaf node for back-end outcome o, this is expressed as,

$$c_o^b = c_{g_1 j_1 ; o} + c_{g_1 j_1 g_2 j_2 ; o} \cdots + c_{g_1 j_1 g_2 j_2 \cdots g_{n_o^\alpha} j_{n_o^\alpha} ; o}.$$

The indicator-based expression for total back-end cost is obtained simply by summing over the outcome groups. The result is

$$C^{BE} = \sum_{o=1}^{n^o} c_o^b N_o^c.$$

which with dependencies on design parameters, random variables, and w^h displayed, is, from (9.16),

$$C^{BE} = \sum_{o=1}^{n^o} c_o^b(w^h)[\alpha_{g_1 j_1;o}\ \alpha_{g_1 j_1 g_2 j_2;o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o} j_{n_o};o}]$$
$$\times [\sum_{w=1}^{n^w} 1_w^{BE}(\{\mathbf{D}_{wtd}\}_{d=1\ t=1}^{n^{dt}\ n^t}, 1)\ Y^{df}(\mathbf{Y}_w, \Lambda, \mathbf{x}) \sum_{c=1}^{N^{cl}(\Lambda,\mathbf{x})} 1_{Z_{wc} \in Z(G_o^b)}(\{\mathbf{D}_{wcd}\}_{d=1}^{n^{dt}}, \mathbf{x}, \mathbf{e})]. \tag{10.1}$$

The yield-based expression for cost is determined as the sum of the costs incurred by the circuits originating with the individual wafers. There are (again) two key ideas underlying the yield-based expression for the back-end cost of a particular wafer. First, it is possible to associate with every node a *cost remaining*, which is simply the total cost which will be incurred by the circuits impingent on the node in question, in completing their path through the tree. Second, the cost remaining at a particular node is just the sum of the costs remaining at its sons, weighted by the split fraction associated with the son. Recalling that test-outcome nodes (or, equivalently, discretionary branches) but not discretionary nodes have inherent treatment costs associated with them, it is possible to write the following nested expression for back-end cost.

$$C^{BE} = N^{cl} \sum_{w=1}^{n^w} 1_w^{BE} Y_w^{df} \sum_{g_1=1}^{n^{g_1}} Y_{g_1;w} \sum_{j_1=1}^{n^{j_1}} \alpha_{g_1 j_1} [c_{g_1 j_1}$$
$$+ \sum_{g_2=1}^{n^{g_2}_{g_1 j_1}} Y_{g_1 j_1 g_2;w} \sum_{j_2=1}^{n^{j_2}_{g_1 j_1 g_2}} \alpha_{g_1 j_1 g_2 j_2} [c_{g_1 j_1 g_2 j_2} + \cdots$$
$$+ \sum_{g_{n^b}=1}^{n^{g_{n^b}}_{g_1 j_1 \cdots g_{(n^b-1)}}} Y_{g_1 j_1 \cdots g_{n^b};w} \sum_{j_{n^b}=1}^{n^{g_{n^b}}_{g_1 j_1 \cdots g_{n^b}}} \alpha_{g_1 j_1 \cdots j_{n^b}} [c_{g_1 j_1 \cdots j_{n^b}}] \cdots]]].$$

Note that the expression, unlike any preceding it, reflects and exploits the tree structure of the back-end flow. It is straightforward to display the dependence on design parameters and random variables in this expression, and this step is omitted.

Note that the dependence of C^{BE} on design parameters, random variables, and w^h can be represented in short form as,

$$C^{BE} = C^{BE}(w^h, \{\mathbf{D}_{wtd}\}_{d=1}^{n^{dt}}{}_{t=1}^{n^t}{}_{w=1}^{n^w}, \mathbf{l}, \{\mathbf{D}_{wcd}\}_{d=1}^{n^{dc}}{}_{c=1}^{N^{cl}(\Lambda,\mathbf{x})}{}_{w=1}^{n^w}, \mathbf{x}, \Lambda, \{\mathbf{Y}_w\}_{w=1}^{n^w}, \mathbf{e}, \alpha) \,. \quad (10.2)$$

Chapter 11

ELEMENTS OF REVENUE MODELING

Modeling the costs incurred in the manufacture an IC product requires modeling the economic behavior of the producer. In contrast, modeling the revenue derived from producing an IC product requires modeling the economic behavior not only of the producer, but also of decision-makers other than the producer, i.e., of potential customers who autonomously decide whether to buy some units of the design product, and if so, how many. This difficult problem is at the core of modeling revenue.

In the revenue model the customer decision-making model is embedded in a larger model structure. This chapter will focus primarily on the latter.

The discussion in Chapter 6 centering on (6.8) suggests that it is of interest to model $r^{n^w} = r^{dn^w} + r^{bn^w}$, corresponding to the starting in production of a batch consisting of n^w wafers. However, as has been stated repeatedly, the ultimate goal of the modeling

presented in this dissertation is to model an appropriate profit quantity associated with the entire n^{wl} wafers started in production during the design lifetime. Toward this end the design and brother product revenues associated with the design lifetime are defined as,

r^d revenue from the sale of all sellable circuits resulting from starting in production n^{wl} wafers of the design product during its design lifetime.

r^b revenue from the sale of all sellable circuits that are realizations of brother products, during the design lifetime of the design product.

Note that these were implicitly defined in Chapter 2, but not in a terminology appropriate for the discussion of this chapter. For example, the definition of design product revenue was based on "the design lifetime of the product line". Note also that in this definition of brother product revenue, as well as for the remainder of the dissertation, the terminology "brother products" is implicitly intended to refer to those of the design product.

The goal of the revenue modeling is to model the revenue called by abuse of terminology the *total revenue* denoted by r and given in terms of the above revenue components by,

$$r = r^d + r^b. \tag{11.1}$$

As was described in words early in Chapter 6, this quantity can be related to the revenue associated with n^w wafers by,

$$r = \frac{n^{wl}}{n^w} r^{n^w}. \tag{11.2}$$

reflecting the fact that in this static model of revenue, revenues can simply be scaled by the number of wafers in terms of which they are defined. The modeling of r could be carried out by defining variables corresponding to the starting of n^w wafers, writing an expression for r^{n^w}, and applying (11.1). However, it is notationally more economical to directly

model r and this is the approach taken here.

11.1. High-level Model Decomposition

The revenue deriving from the sale of a collection of finished units of a product is trivially the sum of the revenue resulting from the sale of each individual unit, which is just its selling price. A statement of additivity more useful in decomposing the modeling problem, however, is the following elementary principle. The total revenue derived from the sale of a set of circuits is the sum of the revenues of the elements of any proper partition of the set. Before the partition of choice in this modeling can be described, some aspects of the purchase process must be discussed.

The first observation pertains to the modeling of the economic behavior of the potential, mentioned at the outset. Specifically, the reaction of a potential customer to the availability of a product may be viewed as a set of decisions, one for each variant of the product, on whether to buy one or more units of the variant. Buying some units of a particular variant represents a desire to have circuits with characteristics no worse than those defining the variant.

Decomposing the revenue modeling problem rests on defining two types of variants. In the beginning of Section 9.3 the creation of product variants was discussed, briefly. To define the two types of variants of interest here, it is necessary to do so in greater detail.

In Section 9.3, what can now be referred to as the *variant structure* of the product, i.e., the number of variants and the characteristics of each, except possibly for some electrical specification levels, was essentially described as resulting from a largely non-parametric design task involving the collaboration of employees with specialized areas of expertise

within the IC business. In applications of the methodology to the development of new products, the objective of this task can be described as attempting to maximize the profit expected from the product by defining a set of variants which achieve a good match of the capabilities of the firm to manufacture the product with the total market for it, as best it can be anticipated. There are no variants with characteristics tailored to the needs of just one customer. Variants which are created by the process which has just been summarized are called *standard variants*. The subset of the resultant characteristics which are to be the published characteristics, are used to make a product specification document which is widely distributed to potential customers, and included in the next printing of the producer's databook.

After customers become aware of the standard variants of a product, some minority of them might decide they would like to be able to buy circuits with product characteristics similar but not identical to those of one of the standard product variants. They might then enter into negotiations with the producer leading to an agreement that the producer will supply a certain number of circuits at a certain price, with minimum characteristics defining a special variant peculiar to the customer. Variants created by this process are called *non-standard variants*.

For the case of IC design in IC houses, which is the primary interest in this dissertation, all products have standard variants. A design product in general has non-standard variants unless it is a new design, in which case it does not. A given brother product usually has non-standard variants. The likelihood of this being the case increases with the length of time by which the market introduction of the brother product precedes the beginning of the design lifetime of the design product.

Customers express their interests in purchasing circuits of either type of variant by placing an order with the producer. It may happen that a customer wishes to buy circuits of more than one product variant, which would result an order for more than one variant. Assume instead for convenience in the discussion that the term "order" means a request for only one variant. There is no loss of generality in this assumption because an order for multiple variants can be considered to be a set of orders each for a single variant.

At any particular time the producer has a set of unfilled orders. A particular finished circuit may meet or exceed the characteristics of more than one of the variants on order. If the circuit is to be sold, the producer must select which order he wishes to assign the circuit to. (If the order requests only one unit, the circuit fills the order.) The *purchase variant assignment* of a finished circuit is the variant of the order to which the producer selects to assign the circuit. In other words, the purchase variant assignment of a circuit is the variant "as which" it is sold.

Now finally can be described the partition of the set of circuits sold which is the basis for the decomposition of the modeling at the highest level. Referring to (11.1), the circuits have already been partitioned according to whether they are realizations of the design product, or its brothers. Both of these sets are in turn partitioned according to whether they are sold as standard or non-standard variants. Specifically, every design product circuit which is sold, is sold either as a standard variant of the design product, or a non-standard variant of the design product, and every brother product circuit which is sold, is sold either as a standard variant of the brother product, or a non-standard variant of the brother product. For the quantitative modeling, let the *standard revenue of the design product*, denoted r^{ds}, be defined by

r^{ds} revenue from the sale of all circuits resulting from the production of n^{wl} wafers of the design product, and sold with purchase variant assignments which are standard variants.

the *non-standard revenue of the design product*, denoted r^{dn}, by

r^{dn} revenue from the sale of all circuits resulting from the production of n^{wl} wafers, and sold with purchase variant assignments that are non-standard variants

the *standard revenue of the brother products*, denoted r^{bs}, by

r^{bs} revenue from the sale of all circuits that are realizations of brother products, sold during the lifetime of the design product with purchase variant assignments that are standard variants

and the *non-standard revenue of the brother products*, denoted r^{bn}, by

r^{bn} revenue from the sale of all circuits that are realizations of brother products, sold during the lifetime of the design product with purchase variant assignments that are non-standard variants.

Then, since the purchase variant assignments of each circuit sold is unique, the product revenue can be written as,

$$r = r^{ds} + r^{dn} + r^{bs} + r^{bn} \tag{11.3}$$

This is the structure of the revenue model at the highest level.

It is appropriate to introduce here some notation which will be used throughout the remainder of the chapter, in particular to index all the variants of products contributing to revenue. Let,

n^{vds} number of standard variants of the design product
n^{vd} total number of variants of the design product
n^{vbs} number of standard variants of brother products
n^{vb} total number of variants of brother products

Variants are indexed by v, in four ranges associated with the four variant types, as follows.

1 to n^{vds}:	standard variants of the design product
$n^{vds}+1$ to n^{vd}:	non-standard variants of the design product
$n^{vd}+1$ to $n^{vd}+n^{vbs}$:	standard variants of the brother products
$n^{vd}+n^{vbs}+1$ to $n^{vd}+n^{vb}$:	non-standard variants of the brother products

Note the suggestion here, as turns out to be the case throughout the modeling of revenue, that there is no need to distinguish between variants that are associated with different brother products. All the variants of all the brother products of any design product can instead be thought of as being associated with a single brother product, if so desired.

As suggested by the descriptions of standard and non-standard variants in this section, the modeling of these two types of revenue is different. Non-standard revenue is discussed first.

11.2. Non-standard Revenue

At the present time no technique has been developed to apply to the difficult problem of predicting the details of non-standard sales agreements which come into existence after the application of the new methodology. Hence the modeling of non-standard revenue herein includes only that deriving from sales of design or brother products for which sales agreements are in existence at the beginning of the lifetime of the design product.

The agreements associated with these variants generally call for a specified number of circuits with the minimum characteristics of the variant, to be supplied by the producer over a specified time interval, at a specified unit price. The time interval could conceivably range from a few months to two years, and thus may or may not exceed the design lifetime of the design product. In any case, a quantity of the variant to be sold during the design lifetime can be estimated. If the life of the agreement is less than the design lifetime, the quantity to be sold is set to the quantity demanded in the agreement. If not, the delivery schedule of the agreement can be used to determine the quantity expected in effect to be sold during the design lifetime. The quantitative modeling proceeds as follows.

Let

q_v number of circuits sold with purchase variant assignment v, $v \in \{n^{vds}+1, n^{vds}+2, \ldots, n^{vd}, n^{vd}+n^{vbs}+1, \ldots, n^{vd}+n^{vb}\}$.

Each known non-standard variant has a unit price associated with it. Let,

p_v unit selling price at which circuits with purchase variant assignment v, $v \in \{n^{vds}+1, n^{vds}+2, \ldots, n^{vd}, n^{vd}+n^{vbs}+1, \ldots, n^{vd}+n^{vb}\}$ are sold.

Then the revenue from the sale of the circuits resulting from the starting in production of n^{wl} wafers, and sold with purchase variant assignment v, $v \in \{n^{vds}+1, n^{vds}+2, \ldots, n^{vd}, n^{vd}+n^{vbs}+1, \ldots, n^{vd}+n^{vb}\}$ is just $p_v\, q_v$ and therefore,

$$r^{dn} = \sum_{v=n^{vds}+1}^{n^{vd}} p_v\, q_v. \tag{11.4}$$

and,

$$r^{bn} = \sum_{v=n^{vd}+n^{vbs}+1}^{n^{vd}+n^{vb}} p_v\, q_v. \tag{11.5}$$

Note that the the set of circuits which contributes to the non-standard revenue has in effect

been partitioned into disjoint sets each consisting of the circuits which are sold as one of the non-standard variants. Note also that all the p_v and q_v quantities introduced above are constants in the model. Therefore, so are r^{dn} and r^{bn}. Defining the non-standard revenue, denoted r^n, as,

$$r^n = r^{dn} + r^{bn}, \tag{11.6}$$

it follows that r^n is a constant in the model.

11.3. The Fully Parametrized Model for Standard Revenue

11.3.1. The Revenue Summation, and the Identification of Design Parameters

The modeling of r^n has shown that revenue can be expressed as a sum of price-quantity products. This is the case for the standard revenue as well, as will be seen shortly. However, little else remains the same in the modeling of standard revenue.

The contrast of most immediate import is that partitioning the circuits which are sold as standard variants on the basis of their purchase variant assignment alone is not enough to model the revenue precisely. The number of circuits requested in an order is called the *order quantity*. According to universal practice of IC producers, the price at which a circuit is sold depends not only on its purchase variant assignment, but also on the order quantity under which the circuit is purchased. Each producer has a set of *order quantity ranges* which apply not only to the design product, but in general, incidentally, to his entire product line. The order quantity ranges constitute a proper partition of the natural numbers, with each range consisting of a string of consecutive natural numbers. A typical example is the set of quantity ranges [1-9], [10-99], [100-999], and [above 999]. Associ-

ated with each (variant, order quantity range) pair is a unique selling price, so that the assignment of a circuit to an order determines the price at which it is sold. In Figure 8 is represented the selling price of units of the hypothetical circuit, as a function of variant and order quantity range.

Note that if there are orders with order quantities in the highest range (extending theoretically to infinity), and if the revenue from such customers is expected to have a significant influence on the optimal parametric design, additional special prices, associated with individual customers, may need to be introduced. However, for simplicity, this circumstance is not discussed here.

The quantitative modeling begins as usual with some definitions of notation. Let

n^r number of order quantity ranges in which the producer's IC's are sold.

Let r index the order quantity ranges. (Distinguishing the use of the letter r as an index and as representing a revenue quantity is easily done based on the manner of its use.) Then the prices and quantities useful in modeling the standard revenue deriving from the design product and its brothers can be defined as follows. Let

p_{vr} price of circuits sold with purchase variant assignment v and order quantity range r, $r = 1,2,\ldots,n^r$, $v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs}$,

and

q_{vr} quantity of circuits sold with purchase variant assignment v and order quantity range r, $r = 1,2,\ldots,n^r$, $v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs}$.

Clearly the restrictions

$$0 \le p_{vr}\ , r = 1,2,\ldots,n^r, v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs} \qquad (11.7)$$

and

$$0 \leq q_{vr}\ , r = 1,2,\ldots,n^r, v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots, n^{vd}+n^{vbs} \qquad (11.8)$$

are needed here to avoid meaningless interpretations of the variables. Then the standard revenue of the design product is easily seen to be the sum of price-quantity products given by,

$$r^{ds} = \sum_{v=1}^{n^{vds}} \sum_{r=1}^{n^r} p_{vr} q_{vr} \qquad (11.9)$$

and that of its brothers by,

$$r^{bs} = \sum_{v=n^{vd}+1}^{n^{vd}+n^{vbs}} \sum_{r=1}^{n^r} p_{vr} q_{vr} \qquad (11.10)$$

It remains to describe how the $(n^{ds}+n^{bs})n^r$ prices and quantities on the right-hand sides in the above two expressions are determined in the methodology. In the framework of the criteria for design parameter appropriateness of Section 1.2, it is not difficult to see that they are directly implementable, have in general a first-order effect on the goodness criterion, and form an independent set (thus not destroying the independence of the total set of design parameters). The situation regarding the properties "producer-controlled", and "specific to the design product optimization" is less obvious. The price quantities are discussed first.

It is clear that the prices of all the standard variants are producer-controlled, unlike those of the non-standard variants. It is also clear that the prices of the standard variants of the design product are specific to the design product optimization. Hence, in accordance with the criteria of Section 1.2 just mentioned, they are among the design parameters of the methodology.

As to whether the standard-variant prices of the brother products are specific to the design product optimization, the standard variant prices of each brother product are initially determined at the time the product is designed. If these prices are to be set optimally at the beginning of the lifetime of the design product, it is necessary to have in addition to the revenue modeling elements presented herein, a moderately accurate model for the effects of changes in brother product prices - changes that could occur an arbitrarily short time after introduction of the brother product. Such price changes are usually not in the interest of the producer, for a variety of reasons, and it is not appropriate to attempt to account for the effects of brother product standard variant price changes in the development of a goodness criterion for the design product. Therefore, this is not done. Hence the prices of standard variants of brother products are parameters that, in the terminology of Section 1.2, have side effects, or equivalently, are not specific to the design product optimization. Accordingly, they are not among the design parameters of the methodology. Instead, quite naturally, their existing values are taken as given constants.

Turning to the determination of the purchase quantities on the right-hand side of (11.9) and (11.10), since the completion of a purchase requires the acquiescence of the producer (taking the form of shipping the circuit(s) ordered), he has direct control over at least one (and usually all) of the purchase quantities, at least in some neighborhood on the positive side of zero. To say that this were not the case would be to claim there is an aspect of the economic system being model requiring that each purchase quantity be exactly zero. Hence the purchase quantities are producer-controlled. Furthermore the design product standard variant purchase quantities are clearly specific to the design product optimization. Hence, they are among the design parameters of the methodology.

Although changes to the brother product standard variant purchase quantities affects the desirability of the design of those products, the effect is accounted for by the inclusion of brother product revenue in the profit model. Therefore, these quantities are specific to the design product optimization, as defined in Section 1.2. Hence they also are among the design parameters of the methodology.

It is a fairly elementary observation that since prices tend not to be zero, the profit random variable is a monotonically increasing function of the purchase quantity design parameters, taken one at a time. It is therefore essential that all upper bounds on purchase quantities present in the economic system be known, i.e. modeled with appropriate accuracy. Such upper bounds arise from two other elementary observations:

(1) The total purchase quantity of a particular variant cannot exceed the expected number of circuits resulting from the starting in production of n^{wl} wafers and having characteristics meeting or exceeding those which define the variant.

(2) The purchase quantity of a particular variant in a particular order quantity range cannot exceed the number of circuits of that description which customers are willing to buy during the design lifetime.

The first of these observations leads to a set of constraints on the purchase quantities which are called *supply constraints* and the second to a set of constraints that belong to a set called *demand constraints*. Supply constraints form a bridge between the revenue and cost models, and are discussed in the next chapter. Demand constraints are discussed next.

11.3.2. Market Bounds on Purchase Quantities

As articulated in item (2) above, the market imposes upper bounds on purchase quantities, which must be adequately modeled. Let

q_{vr}^{D} maximum quantity of circuits of variant v that potential customers of the design product, are willing to purchase in order quantity range r, $r = 1,2,\ldots,n^{r}$, $v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots, n^{vd}+n^{vbs}$, during the design lifetime.

In conventional economic parlance, this would be called the *expected demand for variant v in quantity range r*. With these quantities formally defined, the upper bounds in question can of course be written as,

$$q_{vr} \leq q_{vr}^{D}, r = 1,2,\ldots,n^{r}, v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots, n^{vd}+n^{vbs}. \qquad (11.11)$$

The role of the quantities which have been introduced to this point can be described as follows. All the prices and characteristics of the standard variants of the design product are generated by the producer with the goal of maximizing expected profit. The prices and characteristics of the standard variants of the brother products are given constants. All these prices and characteristics are inputs to the decisions of the potential customers. Their collective response determines a set of q_{vr}^{D} values. This output of the customers is an input to the producer, who then selects the q_{vr} values, subject to (11.11), again with the goal of maximizing expected profit. The resultant standard revenues are of course given by (11.9) and (11.10).

11.3.3. Limitations of the Fully Parametrized Model

The treatment of standard revenue which has been described above is referred to as the *fully-parametrized* model of standard revenue. All of the elements of this model are theoretically valid. However it would not be advantageous to implement it. This is essentially because in the interest of addressing the effect of order quantity ranges, n^r distinct purchase quantity and price design parameters are used for each variant. This parametrization of purchase quantities makes the estimation of the expected demands q_{vr}^D, for $r = 1,2,\ldots,n^r$, $v = 1,2,\ldots,n^{vds}$, $n^{vd}+1,\ldots,n^{vd}+n^{vbs}$ difficult, and generally offers little gain in expected profit.

The difficulty in the estimation of expected demands can be appreciated by considering some aspects of their modeling. They depend on the autonomous behavior of individual potential customers. Specifically, the preferences of customers among all variants and order quantity ranges of all products in the market which fit their application, must be modeled. The preferences which determine q_{vr}^D depend on the characteristics and price not only of variant v and order quantity range r, but in general on those of all other variants and order quantity ranges of the product, as well as those of other products of the same class that are on the market, including those of other producers.

It is quite sensible to assume that most potential customers when faced with a decision of this complexity make the decision hierarchically, in particular first choosing a variant (and thereby a product) without regard to information pertaining to choice of order quantity range, then choosing an order quantity range from among those available for the selected variant. In making the latter choice, the potential customer must consider budgetary pressures, purchasing procedures of his firm, and the stage of development of

the electronic system in which the IC's are to be used (e.g., prototype, production). The difficulty of modeling such a choice is evident. In fact it would require modeling, for example, the decision of an individual customer who has decided to buy 15 circuits of a particular variant, and is faced with the quantity range example presented before ([1-9], [10-99], [100-999], [above 999]), whether to buy all 15 at once, or 5 first, and 10 more later.

The reason that the gain in profit from the use of multiple price parameters for each variant can be expected to be small, can be explained through the introduction of an equivalent parametrization for price. Let the indexing of the order quantity ranges by r be in natural ascending order. By this is meant that $r = 1$ represents the order quantity range in which single units are sold, $r = 2$ the order quantity range adjacent the $r = 1$ range, and so forth. The prices p_{v1}, $v = 1,2,\ldots,n^{vs}$, which are referred to as the *unit quantity prices*, are like any of the other prices in that they are considered to be real-valued variables which in the application of the new profit model would never be exactly zero. Hence The n^r price parameters $p_{v1}, p_{v2}, \ldots, p_{vn^r}$ associated with arbitrarily chosen variant v, can be re-parametrized as follows. Let

$$\rho_{vr1} = p_{vr}/p_{v1},\ r = 2,3,\ldots,n^r. \tag{11.12}$$

Then clearly the parameter set $\{p_{v1}, \rho_{v21}, \rho_{v31}, \ldots, \rho_{vn^r1}\}$ could be used as design parameters associated with variant v, instead of the original prices. Consider the decomposition of the optimization of the prices of the variant into the separate optimizations of p_{v1} and of the ratios ρ_{vr1}, $r = 2,3,\ldots,n^r$. These two optimizations are to an extent decoupled, in the sense that the p_{v1} value which maximizes revenue tends to depend on prices and characteristics of variants competing with variant v, while the ρ_{vr1} values which maximize

revenue tend to depend on the probabilistic distributions of the quantity needs of the pool of potential customers. Regarding the selection of the latter values, in current practice each producer has evolved a set of constants ρ_{vr1}, $r = 2,3,\ldots,n^r$ that are independent of p_{v1}, and of variant, product, and time as well. There are several reasons to believe these constants are nearly optimal, i.e., that the increase in profit that would result from replacing them with optimized ratios would be small:

(1) the setting of the ratios is readily recognized by marketing and management personnel as an optimization problem which is largely decoupled from the setting of other product parameters, whether they have the familiarity with optimization theory to characterize it as such or not.

(2) there is great economic incentive to have near-optimal ratios.

(3) if the producer considered the ratios to be far from optimal, and wanted to study their effect on revenue, he would not suffer whatsoever for lack of data.

Consequently, the capability of the fully parametrized revenue model to optimize the ratios between the prices of each variant is not considered a significant advantage in its use.

In view of these arguments against the use of the fully parametrized revenue model, a simplified treatment has been developed, which uses the decomposed optimization of price ratios just mentioned, and uses only one price and one quantity design parameter for each variant.

11.4. The Simplified Model for Standard Revenue

The basis for the simplified model for standard revenue is a mathematically exact transformation of the expression for revenue in the fully-parametrized model. Equation (11.9) for design product revenue and Equation (11.10) for brother product revenue, the right-hand sides of which differ only in their limits of summation, are each transformed in a like manner. The case of design product revenue is considered first.

Consideration of this case begins with the observation that it can be assumed that,

$$\sum_{r=1}^{n^r} q_{vr} \neq 0 \, , \, v = 1,2,\ldots,n^{vds} \, .$$

Otherwise, there would exist a particular variant index v^*, say, such that

$$\sum_{r=1}^{n^r} q_{v^*r} = 0,$$

which, since the purchase quantities are non-negative, would imply that,

$$q_{v^*r} = 0, r = 1,2,\ldots,n^r.$$

But if this were the case, variant v^* could be excluded from the summation of (11.9). This type of condition is degenerate, and for notational simplicity will be neglected. Then it is permissible to multiply and divide inside the first summation of (11.9) by $\sum_{r=1}^{n^r} q_{vr}$, which gives,

$$r^{ds} = \sum_{v=1}^{n^{vds}} \frac{\sum_{r=1}^{n^r} q_{vr} p_{vr}}{\sum_{r=1}^{n^r} q_{vr}} \sum_{r=1}^{n^r} q_{vr} \qquad (11.13)$$

Now define the *purchase quantity of standard variant* v, denoted q_v by

$$q_v = \sum_{r=1}^{n^r} q_{vr}\,,\ v = 1,2,\ldots,n^{vds},$$

and the *average selling price of standard variant* v, denoted p_v by

$$p_v = \frac{\sum_{r=1}^{n^r} q_{vr} p_{vr}}{\sum_{r=1}^{n^r} q_{vr}},\ v = 1,2,\ldots,n^{vds}, \tag{11.14}$$

Then from (11.13) the design product standard revenue can be written as,

$$r^{ds} = \sum_{v=1}^{n^{vds}} p_v\, q_v. \tag{11.15}$$

Turning to the case of brother product standard revenue, it not difficult to see that the argument just presented for design product standard revenue applies directly for brother product standard revenue if the above definitions of purchase quantities and average selling prices are applied to brother product variants, and the index of the variant summations are changed to run from $v = n^{vd} + 1$ up to $n^{vd} + n^{vbs}$. In particular, corresponding to (11.15) can be written the expression for brother product standard revenue,

$$r^{bs} = \sum_{v=n^{vd}+1}^{n^{vd}+n^{vbs}} p_v\, q_v. \tag{11.16}$$

In the simplified model, design and brother product standard revenue are given by (11.15) and (11.16), respectively. All standard variant purchase quantities are design parameters, the average selling prices of all standard variants of the design product are design parameters, and the average selling prices of all standard variants of the brother products are given constants. The argument supporting the status of the purchase quantities as design parameters is straightforward, but the discussion of the treatment of prices of standard variants is not. Purchase quantities are discussed first.

11.4.1. The Treatment of Purchase Quantity Variables

The argument that variant purchase quantities are correctly treated as design parameters can parallel that made for the purchase quantity variables of the fully parametrized model, which were dependent on order quantity range. To avoid repetition, the argument is omitted here.

There is further parallelism between the purchase quantity variables used in the fully parametrized model and those used in the simplified model. First, note also that the non-negativity constraints of (11.8) imply the analogous constraints that,

$$0 \leq q_v \ , v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots, n^{vd}+n^{vbs}.$$

Let the *design product standard variant purchase quantity vector*, denoted $\mathbf{q}^{ds}$ be defined by,

$$\mathbf{q}^{ds} = [q_1\ q_2\ \cdots\ q_{n^{vds}}]',$$

and the *brother product standard variant purchase quantity vector*, denoted $\mathbf{q}^{bs}$ be defined by,

$$\mathbf{q}^{bs} = [q_{n^{vd}+1}\ q_{n^{vd}+2}\ \cdots\ q_{n^{vd}+n^{vbs}}]'.$$

Then these constraints can be written as,

$$0 \leq \mathbf{q}^{ds}, \tag{11.17}$$

and,

$$0 \leq \mathbf{q}^{bs}. \tag{11.18}$$

For more compact notation, let the *standard variant purchase quantity vector*, denoted $\mathbf{q}^{s}$ be defined by,

$$\mathbf{q}^{s} = [\mathbf{q}^{ds\prime}\ \mathbf{q}^{bs\prime}]'.$$

Then the inequalities of (11.17) and (11.18) can be written together as,

$$0 \leq \mathbf{q}^s. \tag{11.19}$$

Also parallel is the imperative that all upper bounds on purchase quantities q_v, for $v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs}$ be modeled. In particular, for each standard variant, there is a maximum quantity of circuits that customers are willing to buy. Let the *demand for standard variant* v, denoted q_v^D, be defined by,

$$q_v^D = \sum_{r=1}^{n^r} q_{vr}^D. \tag{11.20}$$

Then the inequalities of (11.11), summed over r, imply the following set of constraints on the standard variant purchase quantities,

$$q_v \leq q_v^D, v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs}.$$

Defining the *design product standard variant demand vector* $\mathbf{q}^{Dd}$ by

$$\mathbf{q}^{Dd} = [q_1^D \; q_2^D \; \cdots \; q_{n^{vds}}^D]',$$

the *brother product standard variant demand vector* $\mathbf{q}^{Db}$ by

$$\mathbf{q}^{Db} = [q_{n^{vd}+1}^D \; q_{n^{vd}+2}^D \; \cdots \; q_{n^{vd}+n^{vbs}}^D]',$$

and the *standard variant demand vector* $\mathbf{q}^D$ by

$$\mathbf{q}^D = [\mathbf{q}^{Dd\prime} \; \mathbf{q}^{Db\prime}]'.$$

allows these inequalities to be written as,

$$\mathbf{q}^s \leq \mathbf{q}^D. \tag{11.21}$$

11.4.2. The Treatment of Price Variables

The modification of the treatment of price variables in the fully parametrized model is based on relinquishing the objective of achieving optimal ratios between the order quantity range prices of each variant. The discussion is in terms of the parametrization introduced

in (11.12).

For the brother product variants, the parameters $\rho_{vr1}, r = 1,2,\ldots,n^r$ are fixed constants, since the prices on the right-hand side of (11.12) are fixed constants in the fully parametrized model. The first step in deriving the simplified treatment of prices is the assumption that these price ratios are also fixed constants for design product variants. In the expectation of achieving a reduction of the dimensionality of the "price design parameter subspace" to n^{vds}, the optimal setting of these price ratios is omitted from the scope of the new methodology, with the assumption that the producer's best knowledge about setting them is to be used. Most producers have a single set of price ratios that apply to all products in the product class of interest, and, for simplicity, this is assumed here. Note that given $p_{v1}, v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs}$, all other prices of the fully-parametrized model are given by,

$$p_{vr} = \rho_{r1}p_{v1}, r = 2,3,\ldots,n^r, v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs}. \quad (11.22)$$

(Note the subscript v has been dropped from the price ratios to signify their independence of product and variant.)

The next step in deriving the simplified treatment of prices pertains to the determination of the average selling prices of the variants. In the case of brother products, it is assumed that the fractions of the purchase quantities in each order quantity range, i.e. $(q_{vk}/\sum_{r=1}^{n^r} q_{vr})$ for $k = 1,2,\ldots,n^r$, are independent of the values of the design parameters of the design product. It is quite reasonable, in view of the hierarchical approach of the customers decision-making, discussed earlier, to expect that changes in design parameter values would affect choice of variant, and not order quantity range. From the definition of the average selling prices in (11.14) and the fact that the brother product prices are fixed in

the fully parametrized model, it follows that p_v for $v = n^{vd}+1,\ldots,n^{vd}+n^{vb}$ can be considered fixed. These average prices can be calculated based on actual market data. (In fact, it is currently the practice in the industry to calculate them, albeit for other purposes, such as monitoring of product sales performance.)

Note that the non-negativity constraints of (11.7) and (11.8) imply the analogous constraints that,

$$0 \le p_v \ , v \ = 1,2,\ldots,n^{vds}.$$

Letting the *brother product standard variant average price vector*, denoted $\mathbf{p}^{bs}$ be defined by,

$$\mathbf{p}^{bs} = [p_{n^{vd}+1}\ p_{n^{vd}+2}\ \cdots\ p_{n^{vd}+n^{vbs}}]',$$

the above constraints can be written more compactly as,

$$0 \le \mathbf{p}^{bs}.$$

Regarding the determination of average selling prices of variants of the design product, it is assumed the average selling price of variant v can be modeled with sufficient accuracy as a function of its unit quantity price. That is, that there exist functions $p_v(\ )$ for $v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots, n^{vd}+n^{vbs}$ not necessarily all the same, such that,

$$p_v = p_v(p_{v1}), \tag{11.23}$$

This assumption is supported by the following observations:

(1) For any v, p_v is theoretically bounded by the smallest and largest prices, i.e.

$$\rho_{n^r 1} p_{v1} \le p_v \le p_{v1}.$$

(2) Producer's values of $\rho_{n^r 1}$ are generally no less than .6, and frequently in the vicinity of .7.

(3) Marketing personnel can within some tolerable error predict the average selling price of a variant v based on its order quantity range prices p_{vr}, $r = 1,2,\ldots,n^r$, and familiarity with the types of organizations which tend to buy the variant, e.g. low-purchase-volume military development organizations.

The functions of (11.23) might be of no greater complexity than that of a quadratic, or even a simple proportionality, and nevertheless achieve adequate accuracy over some maximum expected range of p_{v1} variation.

Note that as in the case for brother products, from the definition of (11.14),

$$0 \leq p_v \ , v = n^{vd}+1,\ldots,n^{vd}+n^{vb}.$$

Letting the *design product standard variant average price vector*, denoted $\mathbf{p}^{ds}$ be defined by,

$$\mathbf{p}^{ds} = [p_1 p_2 \cdots p_{n^{vds}}]',$$

the above constraints can be written more compactly as,

$$0 \leq \mathbf{p}^{ds}. \tag{11.24}$$

Taking stock of the development of the simplified model, note in terms of the calculation of revenues in (11.15) and (11.16) that the determination of all purchase quantities on the right-hand sides has been described (they are design parameters). So also has the determination of the average selling prices, except for the role of the variables p_{v1} for $v = 1,2,\ldots,n^{vds}$, which is in need of attention.

Note that these prices satisfy the criteria of Section 1.2 for treatment as design parameters, and if they are so treated, the result is a valid simplified revenue model, with a price design parameter space dimensionality of only n^{vds}. Average selling prices are computed from (11.23), and order quantity range prices, on which variant demands in general

depend, from (11.22). Instead, however, one last modification to the treatment is made, primarily for aesthetic purposes. Note that the functions $p_v(p_{v1}), v = 1,2,\ldots,n^{vs}$ would most certainly be monotonic (monotonically increasing, incidentally). Let $p_{v1}(p_v), v = 1,2,\ldots,n^{vs}$ denote their inverses. Although the model of average selling price which had been described represents what is in the real world a "response" to actual purchase prices $p_{v1}, v = 1,2,\ldots,n^{vs}$, the existence of the inverses allows the roles of the p_v and p_{v1} to be reversed. Instead of considering the unit quantity prices to be design parameters from which the average selling prices are calculated, equivalently the average selling prices are considered to be design parameters from which the unit quantity prices are calculated.

Having introduced and described the roles of the principle variables of the simplified model, which is the model actually used in the methodology, it is desirable to introduce an additional feature to the model that enhances its realism. There are some perverse psychological effects which may enter into the way customers view the price of an IC variant and which are difficult to model in the demand modeling methodology. These effects are generally triggered in a customer when a variant price seems to be "too low" or "too high", raising suspicions about its quality. A price which is "too low" can raise such suspicions in the minds of the minority of customers who are, in effect, judging the quality of an IC by its price. A price for a relatively low-performance variant which is "too low" can make a customer who is considering buying a higher priced variant suspicious that in so doing he would not be getting a good value. A price for a variant that is higher than it was in the past can make customers think that the price increase implies the IC house has begun to have some sort of difficulties producing the product which will eventually affect them if

they buy the product. Note that this latter effect is generally the most important effect of this kind.

Generally these undesirable effects can be prevented by placing bounds on prices, i.e., an upper or lower bound, or both, on each variant price. The selection of such bounds is a form of parametric decision making. However, without the capability of explicitly modeling the effects, the bounds cannot be automatically and optimally determined. On the other hand, frequently these effects are well enough understood by marketing personnel that the long-term profitability of the product is better served by allowing them to subjectively set bounds than by setting no bounds. This is especially true when it is in fact considered to be unwise to raise the price of one or more variants of a product being redesigned. Therefore the capability to impose an upper or lower bound, or both, on each variant price is appended to the revenue model. If $\mathbf{A}^p$ is a matrix with n^v columns and with each row having n^v-1 zeros and its remaining element +1 or −1, and $\mathbf{p}^k$ a column vector of constants, the bounds of the type just described can be written as,

$$\mathbf{A}^p \mathbf{p}^{ds} + \mathbf{p}^k \leq 0. \tag{11.25}$$

11.4.3. Highlights of the Demand Model

The determination of the key variables in the simplified treatment of standard revenue as it has been described to this point, is as follows. Design and brother product revenue are given by (11.15) and (11.16), respectively. As suggested following those equations, q_v for $v = 1,2,\ldots,n^{vds}, n^{vd}+1,\ldots,n^{vd}+n^{vbs}$ and p_v for $v = 1,2,\ldots,n^{vds}$ are design parameters, and p_v for $v = 1,2,\ldots,n^{vds}$ are given constants. p_{v1} for $v = 1,2,\ldots,n^{vds}$ are calculated from the inverse functions, $p_{v1}(p_v), v = 1,2,\ldots,n^{vds}$. From these are determined the remaining $n^{vds}(n^r-1)$ design product prices according to

(11.22). The brother product prices are given constants. Based on all these prices, and the characteristics of all the variants of the design and brother products, and the same data for all variants in the relevant market, the variant demand vector $\mathbf{q}^D$ is determined. This serves as the upper bound for the variant purchase quantity design parameter vector $\mathbf{q}^s$ according to (10.15).

The aspect of the above summary that is the least adequately described is the problem that was described at the outset of the chapter as being at the core of revenue modeling - the modeling of the standard variant demands. These were formally defined in (11.20) in terms of the demands that depend on order quantity range, which were defined in turn immediately preceding (11.11). In words, combining these two definitions, for $v = 1, 2, \ldots, n^{vds}, n^{vd}+1, \ldots, n^{vd}+n^{vbs}$, q_v^D is the maximum quantity of circuits of variant v that potential customers of the design product, are willing to purchase during the design lifetime. Although some brief comments have been made concerning the dependence of these demands on various types of information, precise identification of types of information on which they depend has not been done.

The demands depend on the choices of individual potential customers, among all the variants available. In order to save time and effort, individual customers may of course choose to ignore some available information relevant to their choice, but to be perfectly general, it must be assumed that customers consider three types of information for every available variant: qualitative characteristics associated with particular variants (including, among others, package type, temperature range upon which performance specs are based, and reliability measures performed); performance specs; prices. The qualitative characteristics associated with variants are specified once a particular variant is identified. For a

given variant, the subset of the large amount of information referred to in the above list that depends on variables in the methodology are the prices of the standard variants of the design product, and the specs of the design product, e. The prices depend on p_{v1} for $v = 1,2,\ldots,n^{vds}$. These in turn depend on the design product standard variant average prices through the functions $p_{v1}(p_v)$ for $v = 1,2,\ldots,n^{vds}$. In mathematical terms, then, using the vector of design product standard variant average prices $\mathbf{p}^{ds}$,

$$\mathbf{q}^D = \mathbf{q}^D(\mathbf{p}^{ds}, \mathbf{e}),$$

and the constraint of (11.21) becomes

$$\mathbf{q}^s \leq \mathbf{q}^D(\mathbf{p}^{ds}, \mathbf{e}). \tag{11.26}$$

Note that any valid model of the functions in $\mathbf{q}^D$ would necessarily utilize in some way all of the information of the three types listed above. This includes qualitative characteristics, specs, and prices of brother products and of all products of other producers that are in the product class.

The demand functions in $\mathbf{q}^D$ as a group constitute an element of the profit model that is referred to as the *demand model.* This is standard economics terminology, although many demand models in economics are dynamic models, rather than static, and most which include the effects of product attributes on demand involve less well-defined attributes. Development of a methodology to determine this static demand function for IC products taken one at a time has been a major goal of this research, entailing a level of effort somewhat less than, but comparable with that spent on all other aspects of the profit model. The work has been carried out in a joint effort with Paul Messinger, formerly with the Group for the Application of Mathematics and Statistics in Economics at Berkeley, and who has undertaken major initiatives for and made essential contributions to the work. At

the time of this writing, essentially all of the details of this methodology have been proposed, and data pertaining to a particular circuit selected for study collected, in a large-scale survey of practicing electronic systems design engineers. However, the detailed results of this component of the research are expected to be presented in later writings.

It is of interest, however, to state here two characteristics of the proposed demand models. First, the demand is subject to significant random effects arising from inherent uncertainties in human decision-making and omission of some factors that might influence purchase decisions. Therefore in sophisticated demand models the variant demands would be modeled as random variables, depending not only on $\mathbf{p}^{ds}$ and $\mathbf{e}$, but in general on other more elementary random variables. In this case, in view of the weighted sum form of (11.15) and (11.16), and of the role played by the variant demands in (10.19), the summary of the variant v demand random variable which is appropriately substituted for q_v^D, in (10.19) for example, is the expectation value of the variant v demand random variable, with respect to the appropriate underlying probability distributions.

The second characteristic of the proposed demand modeling methodology which is relevant here is that the expected values of the demand random variables are differentiable functions of their arguments $\mathbf{p}^{ds}$ and $\mathbf{e}$.

11.5. Summary of the Revenue Model

It now remains only to collect and integrate the model elements that have been derived in this chapter. Note that equations appearing below with reference numbers lower than (11.27) are exact repetitions of equations presented earlier.

The total revenue, modeling of which has been the goal in this chapter, is given by,

$$r = r^{ds} + r^{dn} + r^{bs} + r^{bn} \tag{11.3}$$

In view of the fact that the design and brother product non-standard revenues, given, respectively, by

$$r^{dn} = \sum_{v=n^{vds}+1}^{n^{vd}} p_v\, q_v. \tag{11.4}$$

and,

$$r^{bn} = \sum_{v=n^{vd}+n^{vbs}+1}^{n^{vd}+n^{vb}} p_v\, q_v, \tag{11.5}$$

are constants, it is slightly preferable to de-emphasize them by using the definition of (11.6) to rewrite (11.3) as,

$$r = r^{ds} + r^{bs} + r^{n} \tag{11.20}$$

Inserting the summations of (11.15) and (11.16) in this equation yields,

$$r = \sum_{v=1}^{n^{vds}} p_v\, q_v + \sum_{v=n^{vd}+1}^{n^{vd}+n^{vbs}} p_v\, q_v + r^{n} \tag{11.27}$$

where the design parameters in this expression are constrained by,

$$0 \le \mathbf{p}^{ds}, \tag{11.24}$$

$$\mathbf{A}^{p}\,\mathbf{p}^{ds} + \mathbf{p}^{k} \le 0, \tag{11.25}$$

$$0 \le \mathbf{q}^{s}. \tag{11.19}$$

and,

$$\mathbf{q}^{s} \le \mathbf{q}^{D}(\mathbf{p}^{ds}, \mathbf{e}). \tag{11.26}$$

Chapter 12

MODELING THE SUPPLY CONSTRAINTS

As was stated in Section 11.3.1, in introducing the supply constraints, the total purchase quantity of a particular variant cannot exceed the expected number of circuits resulting from the starting in production of n^{wl} wafers and having characteristics meeting or exceeding those which define the variant. Also as mentioned there, upper bounds on purchase quantities arising from the finite quantities of circuits in all categories resulting from starting limited numbers of wafers in production are called supply constraints. Naturally, purchase quantities of both design and brother product have supply constraints. It is necessary to express these in mathematical form for incorporation into the profit model.

First note that producers when faced with a short supply of a variant of one product cannot in good faith substitute units of a different product, even if it is a brother of the first. Hence each product has a separate set of supply constraints that are do not involve

numbers of circuits of other products. Consequently, modeling of supply constraints is discussed in terms of a single product - the design product. Supply constraints for brother products are easily obtained from those for the design product by holding appropriate design parameters of the brother products fixed.

Modeling of supply constraints for the design product begins with some preliminary definitions. Note that for notational simplicity, the superscript d denoting design is temporarily omitted. Let the set of indices of the variants be called V^v. That is,

$$V^v = \{1,2,\ldots,n^v\}.$$

Similarly, let the set of indices of the sellable back-end outcomes be called O^s. That is,

$$O^s = \{1,2,\ldots,n^{os}\}.$$

Finally, for any $v \in V^v$, let χ_v be the characteristic combination vector of variant v, and define the set of outcome indices of groups with elements covering variant v, denoted $O(v)$, by,

$$O(v) = \{o \in O^s \mid \chi_v \lesssim \chi\,,\ \chi \in G_o^B\}.$$

Note that this definition uses the fact, established in Section 9.7, that if one χ vector belonging a particular back-end outcome group covers a particular variant χ vector, so do the others in the group.

The purchase quantity of variant v is by definition q_v. It remains to determine for variant v the expected number of circuits resulting from the starting in production of n^{wl} wafers and having characteristics meeting or exceeding those which define the variant. Recall from the definition of N_o^c at the beginning of Section 9.8 and that of n_o^c as its expected value, that n_o^c is the expected number of circuits originating from n^w wafers, represented by χ vectors in back-end outcome group o. Then the expected number of

such circuits that would originate from n^{wl} wafers is just $\frac{n^{wl}}{n^{w}} n_o^c$. Since a circuit can be sold as a particular variant if and only if its χ vector covers that of the variant, the expected number of circuits resulting from the starting in production of n^{wl} wafers and having characteristics meeting or exceeding those which define the variant is just $\frac{n^{wl}}{n^{w}} \sum_{o \in O(v)} n_o^c$. Therefore, the mathematical form of the constraint condition described at the outset of this section is,

$$q_v \leq \frac{n^{wl}}{n^{w}} \sum_{o \in O(v)} n_o^c,$$

or, defining ρ^w by,

$$\rho^w = \frac{n^{wl}}{n^{w}}, \tag{12.1}$$

it is,

$$q_v \leq \rho^w \sum_{o \in O(v)} n_o^c. \tag{12.2}$$

This of course must be imposed for all $v \in V^v$.

For example, for the hypothetical circuit, by referring to Figure 7, it can be seen that $O(1) = \{1,2\}$, so that one supply constraint that must hold for the circuit is,

$$q_1 \leq \rho^w [n_1^c + n_2^c].$$

The constraints referred to at the outset of this section and expressed in (12.2) are, however, not all the upper bounds on purchase quantities facing the producer who decided to start n^{wl} wafers in production during the design lifetime. Considering the hypothetical circuit again, it could be that for a given mix of circuits in the various outcome groups, revenue is maximized with q_2 non-zero. But since circuits in sellable outcome group G_2^B can be sold either as variant 1 or variant 2, suggesting a competition between these

variants, the constraint, stronger than that of (12.2),

$$q_1 + q_2 \leq \rho^w [n_1^c + n_2^c].$$

must be imposed. In fact the determination of the complete set of supply constraints is a matter of some complexity, as may be evident from a display of the supply constraints for the hypothetical circuit:

$$q_2 \leq \rho^w [n_2^c]$$

$$q_1 + q_2 \leq \rho^w [n_1^c + n_2^c]$$

$$q_4 \leq \rho^w [n_4^c + n_7^c]$$

$$q_6 \leq \rho^w [n_6^c + n_7^c]$$

$$q_4 + q_6 \leq \rho^w [n_4^c + n_6^c + n_7^c]$$

$$q_3 + q_4 \leq \rho^w [n_3^c + n_4^c + n_7^c]$$

$$q_5 + q_6 \leq \rho^w [n_5^c + n_6^c + n_7^c]$$

$$q_3 + q_4 + q_6 \leq \rho^w [n_3^c + n_4^c + n_6^c + n_7^c]$$

$$q_4 + q_5 + q_6 \leq \rho^w [n_4^c + n_5^c + n_6^c + n_7^c]$$

$$q_3 + q_4 + q_5 + q_6 \leq \rho^w [n_3^c + n_4^c + n_5^c + n_6^c + n_7^c]$$

All the required supply constraints consist of upper bounds on a sum (coefficients +1) of a set variant purchase quantities. Hence the maximum number of constraints that can possibly be required for a circuit with n^v variants is the number of possible combinations of variants that might be represented on the left-hand side of the equation (2^{n^v}), minus 1 because the combination of variants which is no variants at all cannot give rise to to a constraint. With this, the problem of determining the supply constraints can be decomposed into two problems: determining which of the $2^{n^v}-1$ possible constraints should be present; determining which back-end outcome groups should be represented on the right-hand side

of each of the constraints.

The solution to the latter of these two subproblems is intuitive. For a given set of variants, the sum of the purchase quantities of the variants must not exceed the expected number of circuits resulting from the starting in production of n^{wl} wafers and having characteristics meeting or exceeding those of at least one of the variants. Again it is necessary to express this in mathematical form. First, the definition of $O(\cdot)$ is extended to set arguments as follows. Let V denote any subset of V^v. Then $O(V)$ is defined by,

$$O(V) = \bigcup_{v \in V} O(v).$$

Then if V^* is a subset of V^v which is to be represented by the sum of its purchase quantities on the left-hand side of a supply constraint, the constraint should be,

$$\sum_{v \in V^*} q_v \leq \rho^w \sum_{o \in O(V^*)} n_o^c.$$

Therefore, if there are to be $n^i \leq 2^{n^v}-1$ inequalities comprising the supply constraints, and the subsets of V^v to be represented on the left-hand sides are denoted $V_i^*, i = 1,2,\ldots,n^i$, then the entire set of supply constraints can be written as,

$$\sum_{v \in V_i^*} q_v \leq \rho^w \sum_{o \in O(V^*)} n_o^c, \quad i = 1,2,\ldots,n^i.$$

In terms of this notation, the former of the two subproblems described above is that of determining the sets $V_i^*, i = 1,2,\ldots,n^i$. A procedure for solving this subproblem for a given product has been developed. Considerations of space preclude its inclusion here. See [RIL86] for further details. However, in simple cases a convincing solution to the problem can be obtained intuitively. Some comments on this approach are presented here to enable a start in the development of intuition sufficient to carry through the approach.

An essential first step in the problem is to summarize the set relationships between the various back-end outcome groups and variants for the particular circuit. There are two useful such summaries. One is the $n^v \times n^{os}$ matrix **H** the (v,o) element of which is defined by,

$$h_{vo} = \begin{cases} 1 & \text{if } o \in O(v) \\ 0 & \text{otherwise} \end{cases}$$

For the hypothetical circuit, analysis of the sellable back-end outcomes and variant vectors depicted in Figure 7 yields the following H matrix.

$$\mathbf{H} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

From the H matrix, it is possible to generate the other type of summary of the set relationships of interest, which is graphic and provides considerable intuition. This is the Venn diagram with universe Ω in which for each $v \in V^v$, the set $\bigcup_{o \in O(v)} G_o^B$ is depicted using a simple geometric figure (usually a circle will do). In Figure 11 is shown this Venn diagram representation for the hypothetical circuit. In the figure the simple geometric forms representing the sets $\bigcup_{o \in O(v)} G_o^B$ are circles for $v = 1,2$, and triangles for $v = 3,4,5,6$. Circles do not suffice for all the variants because of the especially "interesting" set relationships present in the hypothetical circuit example. Back-end outcome groups are represented by intersections and set differences of the simple geometric forms. For example, the group $G_5^B = \{[1\ 1\ 2\ 0\ 1\ 1]', [1\ 1\ 2\ 0\ 1\ 1]'\}$ is represented by the trapezoid at the bottom of the collection of triangles.

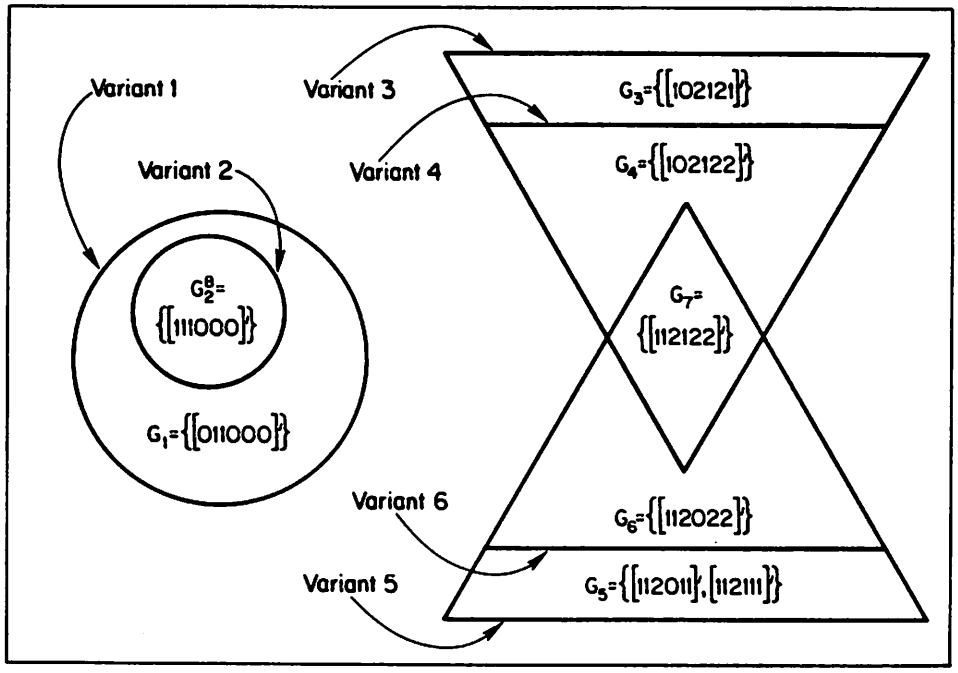


Fig. 11. Venn diagram representation of the set relationship between the back-end groups and the characteristic combinations of the variants, for the hypothetical circuit example.

There is another interpretation of the same diagram in which the universe is the set of circuits with possible characteristic combinations. The simple geometric figures represent sets of circuits which can be sold as the various variants, and their intersections represent circuits with characteristic combinations in the various sellable back-end outcome groups. It is this interpretation that is most helpful in determining which supply constraints are required. The sets V_i^*, $i = 1,2,\ldots,n^i$ can be determined by listing all the $2^{n^w}-1$ possibilities and by inspection of the Venn diagram, eliminating the unneeded ones. Among the principles useful in this elimination are the following, described by example. For the hypothetical circuit, the constraint on q_1 of (12.2) is not needed, because

$$q_1 + q_2 \leq \rho^w [n_1^c + n_2^c]$$

and the non-negativity constraint on q_2,

$$-q_2 \leq 0$$

imply (12.2), which is therefore redundant. No constraint on $q_1 + q_3$ is needed because there are no circuits which can be sold as variant 1 and also as variant 3 (these variants have different package treatments).

For the purposes of the profit model, it suffices to observe that there exist two matrices $\mathbf{A}^q$, with n^v columns, and $\mathbf{A}^n$, with n^{os} columns, and with elements 0 or 1 ($\mathbf{A}^n$ has rows which are formed by the Boolean non-exclusive "or" operation on sets of rows of the **H** matrix interpreted as Boolean vectors), such that if

$$\mathbf{n}^c = [n_1^c \; n_2^c \; \cdots \; n_{n^{os}}^c]', \tag{12.3}$$

and,

$$\mathbf{q} = [q_1 \; q_2 \; \cdots \; q_{n^v}]',$$

the required supply constraints can be written as,

$$\mathbf{A}^q \mathbf{q} \leq \frac{n^{wl}}{n^w} \mathbf{A}^n \mathbf{n}^c,$$

or,

$$\mathbf{A}^q \mathbf{q} - \frac{n^{wl}}{n^w} \mathbf{A}^n \mathbf{n}^c \leq 0. \tag{12.4}$$

This completes the modeling of the supply constraints for the design product (and the suppression of the superscript d). Before extending these results to the brother products, it is desirable to rewrite (12.4) in notation that will distinguish it from the brother product constraints to be written momentarily. The self-explanatory renaming of variables chosen is,

$$\mathbf{A}^{qd} \mathbf{q}^d - \frac{n^{wl}}{n^w} \mathbf{A}^{nd} \mathbf{n}^{cd} \leq 0. \tag{12.5}$$

For a particular arbitrarily chosen brother product, say product p of the set of brother products Ψ^b of Section 2.3, it is clear that there exist selection matrices $\mathbf{A}_p^{qb}$ and $\mathbf{A}_p^{nb}$ with appropriate numbers of rows and columns, and a ratio of wafer numbers ρ_p^w such that if $\mathbf{q}_p^b$ and $\mathbf{n}_p^{cb}$ denote its purchase quantity and back-end circuit count vectors, respectively, the supply constraints for the product can be written as,

$$\mathbf{A}_p^{qb} \mathbf{q}_p^b - \rho_p^w \mathbf{A}_p^{nb} \mathbf{n}_p^{cb} \leq 0, \; p \in \Psi^b. \tag{12.6}$$

Note that $\mathbf{n}^{cd}$ has been modeled in Chapter 9, and depends on design parameters the brother product analogs of which are held constant in the formulation. Therefore the $\mathbf{n}_p^{cb}$ are vectors of constants.

The above two equations are the supply constraints in the methodology. However since the decoupled nature of the brother product supply constraints is not a particularly valuable feature of the model, it is desirable to combine them into a single inequality, which can then be combined with (12.5) to give a single supply constraint for the

formulation. From (12.6) it is not difficult to see that there exist selection matrices $\mathbf{A}^{qb}$ and $\mathbf{A}^{nb}$ such that if, in the notation of Chapter 11, q^b is defined as,

$$\mathbf{q}^b = [q_{n^{wd}+1}\ q_{n^{wd}+2}\ \cdots\ q_{n^{wd}+n^{vb}}]'.$$

then the supply constraints for all brother products can be written as,

$$\mathbf{A}^{qb}\mathbf{q}^b - \frac{n^{wl}}{n^w}\mathbf{A}^{nb}\mathbf{n}^{cb} \leq 0. \tag{12.7}$$

where $\mathbf{n}^{cb}$ is a vector of brother product back-end circuit counts ordered to correspond appropriately to the order of elements in $\mathbf{q}^b$, and scaled to absorb the ratio between the corresponding ρ_p^w and the design product wafer ratio $\frac{n^{wl}}{n^w}$.

Now if the *design and brother product quantity vector*, denoted $\mathbf{q}$, is defined as,

$$\mathbf{q} = [\mathbf{q}^{d\prime}\ \mathbf{q}^{b\prime}]',$$

which, incidentally, is just the vector of all purchase quantities,

$$\mathbf{q} = [q_1\ q_2\ \cdots\ q_{n^{wd}+n^{vb}}]',$$

$\mathbf{n}^c$ is defined by,

$$\mathbf{n}^c = [\mathbf{n}^{cd\prime}\ \mathbf{n}^{cb\prime}]',$$

then there exist matrices (most conveniently block diagonal) $\mathbf{A}^q$ and $\mathbf{A}^n$ such that, equivalent to (12.5) and (12.10),

$$\mathbf{A}^q\mathbf{q} \leq \frac{n^{wl}}{n^w}\mathbf{A}^n\mathbf{n}^c.$$

Chapter 13

THE PROFIT MODEL EQUATIONS

In Chapter 6 was derived formula (6.8),

$$\pi^{n^w} = -n^w c^{O/w} - C^{FE} - C^{BE} + r^{n^w}, \tag{6.8}$$

which showed that the contribution profit random variable could be modeled by separately modeling opportunity cost, front-end cost, back-end cost, and revenue. Let

π contribution profit associated with the starting in production of n^{wl} wafers during the design lifetime.

Now the final expressions for $E\pi$ can be derived.

The opportunity cost associated with the entire lifetime is trivially $n^{wl} c^{O/w}$ and the revenue associated with the entire lifetime is by definition the quantity r, modeled in Chapters 11 and 12. However, the second and third terms on the right of this expression are random variables. Since the goal of the methodology is the maximization of the

expectation value of contribution profit, it is necessary to write expressions for the expected value of these costs. First, define the *actual cost* random variable associated with the starting in production of n^w wafers as,

$$C^A = C^{FE} + C^{BE}. \tag{13.1}$$

Now let

$$c^A = EC^A.$$

The quantities on which C^{FE} and C^{BE} depend are displayed in (8.3) and (10.2), respectively. From these, if the dependence on design parameters is suppressed, the dependencies of C^A can be represented as,

$$C^A = C^A(w^h, \{\mathbf{D}_{wtd}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}, \{\mathbf{D}_{wcd}\}_{d=1\ \ c=1\ \ w=1}^{n^{dc}\ N^{cl}(\Lambda,\mathbf{x})\ n^w}, \{\mathbf{Y}_w\}_{w=1}^{n^w}, \Lambda). \tag{13.2}$$

Clearly c^A depends on both w^h and n^w, i.e.,

$$c^A = c^A(w^h, n^w),$$

which is defined for $\frac{n^w}{2} \le w^h \le n^{wl} - \frac{n^w}{2}$. The quantity remaining to be determined is c^{Al}, defined as,

c^{Al} expected actual costs associated with the starting in production of n^{wl} wafers during the design lifetime.

With this, we can write,

$$E\pi = -n^{wl}c^{O/w} - c^{Al} + r. \tag{13.3}$$

An exact expression for c^{Al} can be determined, as was described in Chapter 6, by considering the special case when the batch consists of a single wafer. That is,

$$c^{Al} = \sum_{w^h\,=\,1}^{n^{wl}} c^A(w^h, 1).$$

The evaluation of the expectations with respect to the random variables on the right of

(13.2) has already been demonstrated, in Section 9.8. There, in (9.17), the expectation value of the circuit count N_o^c was written. The expectation of C^A can obtained by a simple replacement of the integrand in (9.17) by C^A. Therefore,

$$c^{Al} = \sum_{w^h=1}^{n^{wl}} \int C^A(w^h, \{\mathbf{d}_{w^h td}\}_{d=1\,t=1}^{n^{dt}\;n^t}, \{\mathbf{d}_{w^h cd}\}_{d=1\quad c=1}^{n^{dc}\;N^{cl}(\lambda,\mathbf{x})}, \upsilon_{w^h}, \lambda)$$

$$[f_{\mathbf{D}^w}(\mathbf{d}_{w^h})\,\partial\mathbf{d}_{w_h} f_Y(\upsilon_{w^h})\,\partial\upsilon_{w_h}]\, f_\Lambda(\lambda)\,\partial\lambda$$

This expression, although exact, would require the estimation of the electrical performance of all the circuits manufactured during the lifetime of the product, which is prohibitive. Also, to apply the supply constraints to the entire collection of circuits produced during the design lifetime is an inferior treatment of them, as described in Chapter 6.

The expression for c^{Al} which is approximate but avoids these disadvantages is obtained as follows. The expression uses the model for C^A for $n^w \neq 1$, in which case the model itself is approximate. Recall that the model assumes that for each manufacturing operation the cost to perform the operation on every wafer in the batch is the value of the learning curve cost for that operation, evaluated at $\hat{w}^h = w^h$. Consequently, the corresponding expectation c^A is approximate. Now recall that in (12.1), ρ^w was defined as,

$$\rho^w = \frac{n^{wl}}{n^w}. \tag{12.1}$$

The second approximation made in the derivation is that ρ^w is an integer. Then it is not difficult to see that

$$c^{Al} = \sum_{g=1}^{\rho^w} c^A((g - \tfrac{1}{2})n^w, n^w).$$

Now consider the approximation,

$$c^{Al} = \frac{1}{n^w} \int_0^{n^{wl}} c^A(w^h, n^w)\, \partial w^h. \tag{13.4}$$

This approximation would be valid except for the fact that for

$$1 \le w^h \le \frac{n^w}{2}$$

or

$$n^{wl} - \frac{n^w}{2} \le w^h \le n^{wl},$$

the model formally requires values of disturbance and defect yield random variables with indices outside the domain of the first n^{wl} natural numbers. Hence C^A and c^A are not formally defined in this case. However, this difficulty is only notational. All the model needs is a set of those random variables for n^w wafers, and how they might be indexed is immaterial. With this, it is not difficult to extend the domain of definition of c^A to the first n^{wl} natural numbers. Let $\tilde{c}^A$ be defined as,

$$\tilde{c}^A = \begin{cases} \dfrac{w^h + \dfrac{n^w}{2}}{n^w}\, c^A & \text{if } 1 \le w^h \le \dfrac{n^w}{2} \\ c^A & \text{if } \dfrac{n^w}{2} \le w^h \le n^{wl} - \dfrac{n^w}{2} \\ \dfrac{n^{wl} - w^h + \dfrac{n^w}{2}}{n^w}\, c^A & \text{if } n^{wl} - \dfrac{n^w}{2} \le w^h \le n^w \end{cases}$$

Then

$$c^{Al} = \frac{1}{n^w} \int_0^{n^{wl}} \tilde{c}^A(w^h, n^w)\, \partial w^h \tag{13.5}$$

is a valid approximation. On the other hand, if n^w is sufficiently small relative to n^{wl}, the overestimate of (13.4) may be tolerated. In the interest of not adding to the complication of some of the equations to follow which are based the expression for c^{Al}, (13.4) will be used instead of (13.5).

If the expectation integral is inserted in (13.4) for c^A, the result is,

$$c^{Al} = \frac{1}{n^w} \int_0^{n^{wl}} \int C^A(w^h, \{\mathbf{d}_{wtd}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}, \{\mathbf{d}_{wcd}\}_{d=1\ \ c=1\ \ w=1}^{n^{dc}\ N^{cl}(\lambda,\mathbf{x})\ n^w}, \{\upsilon_w\}_{w=1}^{n^w}, \lambda)$$
$$\prod_{w=1}^{n^w} [f_{\mathbf{D}^w}(\mathbf{d}_w)\, \partial \mathbf{d}_w f_Y(\upsilon_w)\, \partial \upsilon_w] f_\Lambda(\lambda)\, \partial\lambda\, \partial w^h . \tag{13.6}$$

This is the desired approximate expression for c^{Al}.

Now all that remains to arrive at the final form of the relationships at the center of the methodology is to collect equations. If the definition of C^A in (13.1) is used in (13.6), and the roles of the design parameters are displayed explicitly, c^{Al} can be written as,

$$c^{Al} = \frac{1}{n^w} \int_0^{n^{wl}} \int \{C^{FE}(w^h, \{\mathbf{d}_{wtd}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}, \mathbf{l}, \lambda, \mathbf{x})$$
$$+\ C^{BE}(w^h, \{\mathbf{d}_{wtd}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}, \mathbf{l},$$
$$\{\mathbf{d}_{wcd}\}_{d=1\ \ c=1\ \ w=1}^{n^{dc}\ N^{cl}(\lambda,\mathbf{x})\ n^w}, \mathbf{x}, \lambda, \{\upsilon_w\}_{w=1}^{n^w}, \mathbf{e}, \alpha)\}$$
$$\prod_{w=1}^{n^w} [f_{\mathbf{D}^w}(\mathbf{d}_w)\, \partial \mathbf{d}_w f_Y(\upsilon_w)\, \partial \upsilon_w] f_\Lambda(\lambda)\, \partial\lambda\, \partial w^h . \tag{13.7}$$

If the actual expressions for C^{FE} and C^{BE} of (8.3) and (10.1) are substituted into the above integrand, and the resultant c^{Al} substituted into (13.3), along with the revenue expression of (11.27) with its last term dropped through another application of the principle of irrelevant constants, it is possible to write the problem formulation of the new methodology as follows. (Note that equations numbered below (13.8) are exact repetitions of equations appearing earlier, and are labeled with their original equation numbers.)

$$\max \{ E\pi(n^{wl}, \mathbf{l}, \mathbf{x}, \mathbf{e}, \alpha, \mathbf{p}, \mathbf{q}) = -n^{wl} c^{O/w} \tag{13.8}$$

$$- \frac{1}{n^w} \int_0^{n^{wl}} \int \{ \sum_{w=1}^{n^w} \{ c_0(w^h) + c_1(w^h) + 1_{1w}(\{\mathbf{d}_{wtd}\}_{d=1\ t=1}^{n^{dt}\ n^t}, \mathbf{l}) [c_2 + 1_{2w} [c_3(w^h) + \cdots$$

$$1_{(n^\sigma-2)w}(\{\mathbf{d}_{wtd}\}_{d=1\ t=1}^{n^{dt}\ n^t}, \mathbf{l}) [c_{(n^\sigma-1)w}(w^h) + 1_{(n^\sigma-1)w}(\{\mathbf{d}_{wtd}\}_{d=1\ t=1}^{n^{dt}\ n^t}, \mathbf{l}) [C^{\#}(\Lambda, w^h, \mathbf{x})]] \cdots]] \}.$$

$$- \sum_{o=1}^{n^o} c_o^b(w^h) [\alpha_{g_1 j_1; o}\ \alpha_{g_1 j_1 g_2 j_2; o}\ \cdots\ \alpha_{g_1 j_1 g_2 j_2 \cdots g_{n_o^\alpha} j_{n_o^\alpha}; o}]$$

$$\times [\sum_{w=1}^{n^w} 1_w^{BE}(\{\mathbf{d}_{wtd}\}_{d=1\ t=1}^{n^{dt}\ n^t}, \mathbf{l}) Y^{df}(\mathbf{Y}_w, \Lambda, \mathbf{x}) \sum_{c=1}^{N^{cl}(\lambda,\mathbf{x})} 1_{\mathbf{Z}_{wc} \in Z(G_o^B)}(\{\mathbf{d}_{wcd}\}_{d=1}^{n^{dc}}, \mathbf{x}, \mathbf{e}) \}$$

$$\prod_{w=1}^{n^w} [f_{\mathbf{D}^w}(\mathbf{d}_w)\, \partial \mathbf{d}_w f_{\mathbf{Y}}(\upsilon_w)\, \partial \upsilon_w] f_\Lambda(\lambda)\, \partial\lambda\, \partial w^h \quad + \sum_{v=1}^{n^{vds}} p_v q_v \quad + \sum_{v=n^{vd}+1}^{n^{vd}+n^{vbs}} p_v q_v \quad \},$$

subject to the design product supply constraints,

$$\mathbf{A}^{qd} \mathbf{q}^d - \frac{n^{wl}}{n^w} \mathbf{A}^{nd} \mathbf{n}^{cd} \leq 0. \tag{12.5}$$

and the brother product supply constraints,

$$\mathbf{A}_p^{qb} \mathbf{q}_p^b - \rho_p^w \mathbf{A}_p^{nb} \mathbf{n}_p^{cb} \leq 0,\ p \in \Psi^b. \tag{12.6}$$

where $\mathbf{n}_p^{cb}$ are vectors of constants, and where

$$\mathbf{n}^c = [n_1^c\ n_2^c\ \cdots\ n_{n^{\alpha s}}^c]', \tag{12.3}$$

$$n_o^c = \int N_o^c(\{\mathbf{d}_{wtd}\}_{d=1\ t=1\ w=1}^{n^{dt}\ n^t\ n^w}, \mathbf{l}, \{\mathbf{d}_{wcd}\}_{d=1\ c=1\ w=1}^{n^{dc}\ N^{cl}(\lambda,\mathbf{x})\ n^w}, \mathbf{x}, \lambda, \{\upsilon_w\}_{w=1}^{n^w}, \mathbf{e}, \alpha)$$
$$\prod_{w=1}^{n^w} [f_{\mathbf{D}^w}(\mathbf{d}_w)\, \partial \mathbf{d}_w f_{\mathbf{Y}}(\upsilon_w)\, \partial \upsilon_w] f_\Lambda(\lambda)\, \partial\lambda \tag{9.17}$$

the demand constraint,

$$\mathbf{q}^s - \mathbf{q}^D(\mathbf{p}^{ds}, \mathbf{e}) \leq 0, \tag{11.26}$$

the price constraint,

$$\mathbf{A}^p \mathbf{p}^{ds} + \mathbf{p}^k \leq 0, \tag{11.25}$$

and to the design parameter constraints:

$$\mathbf{A}^e_{bs}\ \mathbf{e}_{bs} + \mathbf{k}^e_{bs} \leq 0\ ,\ s = 1,2,\ldots,n^s_b\ ,\ b = 1,2,\ldots,n^b \tag{9.1}$$

$$\mathbf{A}^\alpha\ \boldsymbol{\alpha} + \mathbf{k}^\alpha \leq 0 \tag{9.13}$$

$$0 \leq \mathbf{p}^{ds} \tag{11.2}$$

$$0 \leq \mathbf{q}^s\ . \tag{11.19}$$

Chapter 14

CONCLUSIONS AND FUTURE WORK

14.1. Conclusions

In this dissertation has been defined a new methodology for the optimal parametric design of integrated circuits, that also provides a means for optimally deciding among limited numbers of proposed IC designs with differing qualitative characteristics. The methodology, as it has been described here, is appropriate for producers who sell their IC's in the open market. It is based on the maximization of the expectation value of an appropriate profit variable associated with the circuit being designed, accumulated over the lifetime of the design. A model for the profit random variable as a function of all product parameters which significantly affect it, and of the fundamental random variables used in modeling the important statistical phenomena of the fabrication process, has been derived, in Chapters 2 through 13.

The model for the profit random variable is represented in block diagram form in Figures 12 and 13. Figure 13 contains the model details for the block in Figure 12 labeled "Back-end Model". Both figures attempt to depict the functional dependencies among all the major model variables. In reference to both figures, note that the coefficients of all quantities impingent on summing nodes are +1 unless otherwise specified. In reference to Figure 12, note that $1_1, 1_2, \ldots, 1_{n^o-1}$, $\mathbf{Y}$, and $\mathbf{Y}^{df}$ are all vectors with elements the corresponding quantities for the n^w wafers analyzed. Also $\mathbf{Z}$ is a vector containing all the vectors $\mathbf{Z}_{wc}$ occurring in the model. Each dotted-line region present in the figure indicates the chapter of the dissertation which derives the relationships depicted inside the region. Figure 13 uses the variables $\mathbf{Z}_1$, $\mathbf{Z}_2$, and $\mathbf{Z}_{n^w}$. These are vectors containing all the $\mathbf{Z}_{wc}$ vectors pertaining to the particular wafer. The figure also uses some non-standard symbols. First, the N^{cl} "drives" three identical symbols resembling dimension markings used in 2-D drawings. This is to indicate that the number of circuit locations on the wafers is a variable. Second, the circuits of the wafers are represented below the horizontal curly brackets as short vertical lines. On the basis of its electrical performance, each circuit may "fall into" any of the n^o back-end outcome bins shown sitting in a horizontal row below the vertical circuit lines, increasing the circuit count of the outcome bin into which it falls.

The final modeling step is integration of the profit random variable equations represented by the block diagram with respect to the density functions of the underlying random variables. This has been carried out in Chapter 13, along with a transformation of the resultant integral to yield a convenient means of accounting for learning curve effects on costs. The complete set of equations on which the methodology is based are at the end

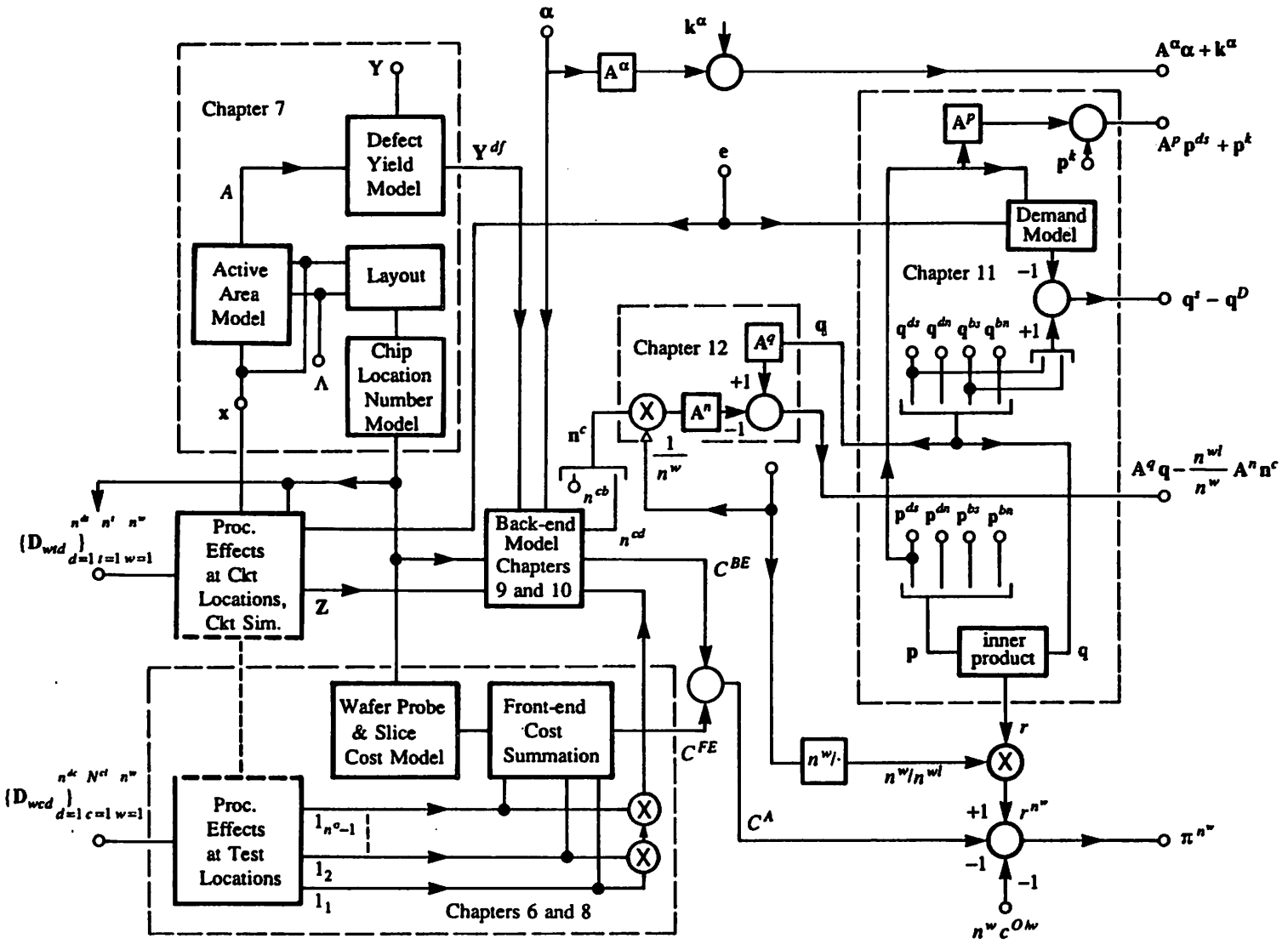


Fig. 12. Block diagram representation of the functional dependencies in the model of the profit random variable.

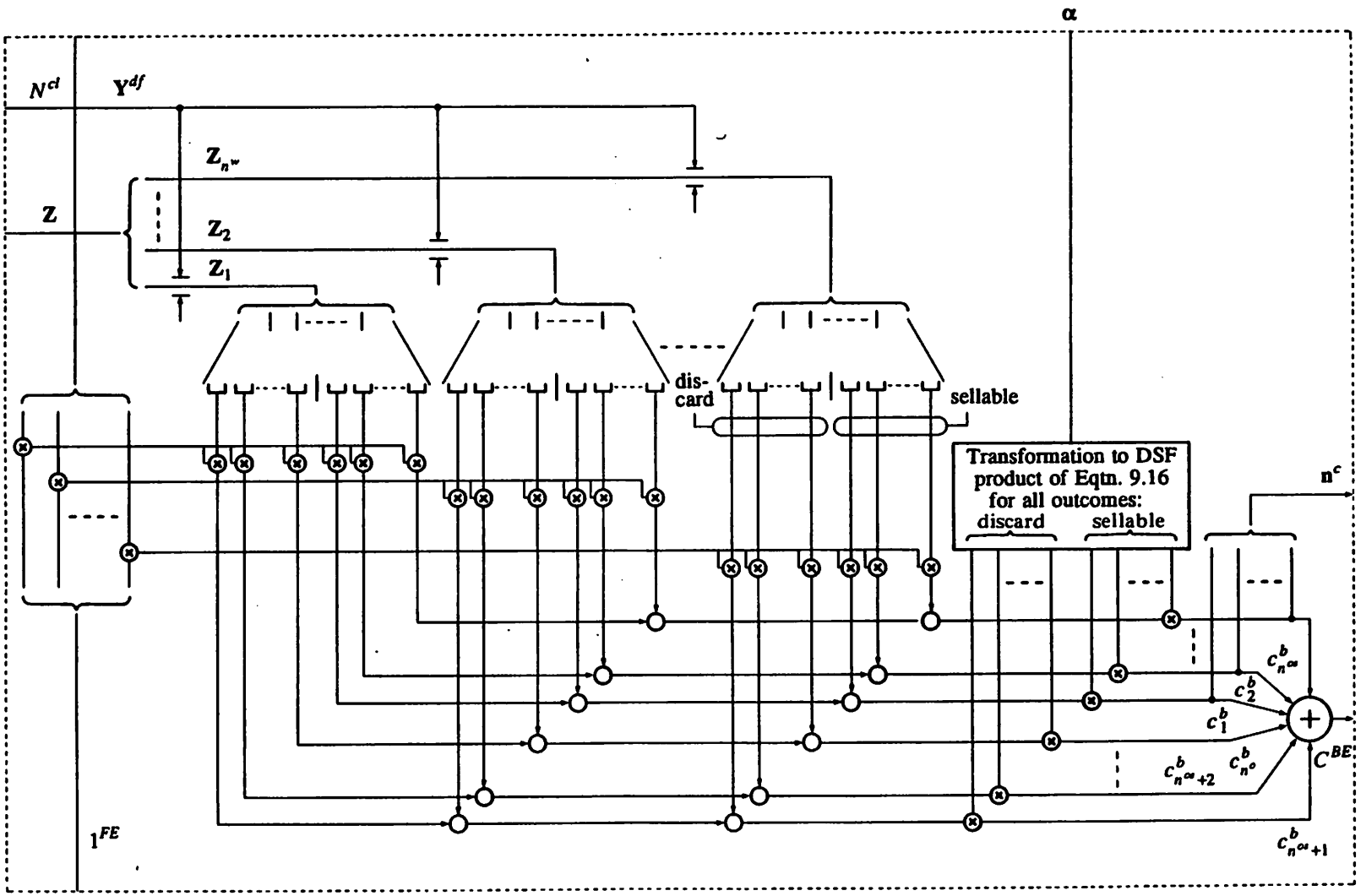


Fig. 13. Block diagram representation of the functional dependencies in the back-end cost model.

of that chapter.

Having presented the entire methodology in both equation and block diagram form, it is now feasible and appropriate to discuss some noteworthy observations about the methodology as a whole.

First, one might expect the methodology to have a major weakness in its apparent need to estimate demand for the design product throughout its design lifetime, since demand for IC's is notoriously unpredictable. However, the profit model has a property which can be expected to considerably reduce the impact of error in demand estimation on the product design. If the estimated demand for all the variants of a product is in error by some multiplicative factor, then the optimal value of n^{wl} and $\mathbf{q}$ will be in error by the same multiplicative error (which can be quickly compensated for by a change in production rate of the product), but the errors in the optimal values of all the other design parameters is zero. Their values are sensitive only to the ratios of demand among the variants of the product. This decoupling is a major feature of the model.

It is appropriate now to return the focus to the application of the proposed parametric design methodology, introduced in Section 1.1, and clarified in Section 1.3, to selection among limited numbers of discrete design alternatives that differ qualitatively rather than parametrically. The methodology represents a powerful tool for such selections, where essentially no objective tools have existed in the past. The appropriateness of profit as a goodness criterion for such selections does not require elaboration, but it is instructive to cite some significant examples of decisions to which the methodology may be applied.

One type of decision to which the discrete decision-making methodology could be applied is the choice of technology in which to fabricate a given circuit class. That is, the

producer could decide on the basis of expected profit whether to fabricate a circuit in, for example, a bipolar or MOS technology. Of course, there are many circuit classes for which more primitive analyses coupled with intuition based on experience provide a clear answer as to which technology should be more profitable. But there are also circuit classes, such as in the area of digital-analog interfacing, where an objective tool would be very valuable.

Given a technology in which a product is to be fabricated, there may remain a choice of fabrication process or fabrication line in which the wafers are to be produced. This is an another type of decision to which the discrete decision-making capability of the methodology may apply. Intuition and experience are less helpful here than in choice of technology.

An activity that is essentially universal in the development of an IC product is the consideration of alternative circuit topologies. This frequently entails tradeoffs between area and performance that are difficult to address, because among other things it is marketing personnel rather than circuit designers that are generally in the best position to judge the revenue implications of various levels of performance. The new methodology is well suited to this type of application, and in particular of all the various types of data used in the profit model, only the circuit description need be changed to provide the desired profit criterion values.

It is worth digressing to observe that it is in the area of topology that a type of secondary benefit of the proposed methodology may be realized. Consideration of statistical parameter variations has a major influence on choice of topology, especially in the case of analog products. This is because frequently subcircuits are added to an initial design to

make it more tolerant of statistical variations. The new methodology, in providing a more effective means to address statistical variations parametrically, should eventually lead to topologies with fewer devices, and to the additional profit improvement that this implies.

Finally, a type of discrete decision that is less obvious than those already discussed, but crucial to the success of an IC product, is the design of its variant structure. The concept of the variant structure entails not only the more obvious attributes such as number and types of packages and reliability procedures, but also set relationships between the circuits of the different variants defined for a product. The proposed methodology offers for the first time a systematic objective tool for making this type of decision.

The greatest potential benefits of the models presented in this dissertation are in the determination of optimal values for design parameters and the optimal selection among discrete design options, just discussed. However there are also significant potential benefits in applying insights that can be obtained from the model without carrying out any numerical computations. These benefits would result from partial shifts in design and product development practice, toward the methodology proposed here.

First and foremost, the model asserts the desirability of integrating parameter-setting functions typically carried out by individuals in different departments of an IC-producing firm, into a unified optimization framework. Second, the model illustrates that the design of IC products to be sold on the open market can be driven directly by the needs of that market. Most probably these two properties have never been achieved in the parametric design of any product, electronic or otherwise.

More detailed insights are numerous. One example pertains to variant structures, discussed above. It has been found that a major influence on the design of variant structures

is that they be sufficiently simple to enable employees in a range of capacities within the firm to work with them on a primarily intuitive basis. The new methodology makes a number of contributions toward allowing IC producers to exploit more complex variant structures. One contribution is that the concept of back-end outcomes and the Venn diagram representation of their set relationship with product variants forms a qualitative theoretical structure for systematic analysis of problems in the manufacture and sale of IC products with multiple variants. Another is that a procedure for generating the appropriate back-end flow tree for a variant structure of arbitrary complexity has been developed.

Another detailed insight pertains to the standard practice among marketing personnel of using fabrication cost data to set prices. (In Figure 12, this would be represented by a path feeding the expectation value of C^A to some operator which outputs prices $\mathbf{p}^{ds}$.) In the ideal, the only connections between the cost and revenue sides of the design process should be the performances **e** and the supply constraints.

14.2. Future Work

There are four broad classes of work remaining relating to the proposed methodology and its implementation: development of the demand model component of the revenue model; high-level modeling of design criteria for economic environments other than the open market; development of an algorithm that is effective for the numerical maximization of the profit criterion that has been modeled herein; software implementation of the methodology with appropriate user interfacing.

In Section 11.4.3 was discussed the status of the demand modeling. It bears restating that plans exist to carry it through to completion and report the results.

Regarding high-level modeling for economic environments other than the open market, the environment for which an appropriate comprehensive methodology would be most valuable, aside from the open market, is that in which IC-supplying divisions operate. (See the discussion surrounding Assumption 8N in Section 1.4.2.) Even though to develop an economic gain model for this case would require substantial new work only on the economic benefit side of the model, and not the cost side, in this researcher's opinion, further modeling is not the most important class of work to be pursued at this time. Because the potential benefits of carrying out actual designs based on numerical computations are undoubtedly greater than those from modeling insights, the most important classes of work are algorithm development, and software implementation, discussed next.

First note that although these classes of work were separately listed above, they are expected to overlap considerably, including in time, and hence are discussed together.

It is well known that in design methodologies less comprehensive than that of this dissertation, such as parametric yield maximization, there is considerable difficulty in obtaining satisfactory numerical results with reasonable computational cost, and that as a consequence, the development of efficient algorithms is the central problem. Computational requirements in maximizing the profit criterion can be expected to be at least as great. This is not due to the apparent functional complexity of the profit criterion, but to the need to generate statistically significant results for the selection of larger numbers of design parameters. Hence successful development of an effective algorithm is crucial to the success of the methodology.

Some work in reviewing algorithms proposed for conventional statistical design, and generating ideas for use in an algorithm for the profit-based methodology, has been done.

However more work of the latter type is needed, possibly starting with identification and study of simplified forms of the profit criterion that have similar computational properties.

As the literature in statistical design makes plain, it is essential to evaluate alternative algorithms experimentally, and the amount of experimentation required may call for a flexible and convenient software system for such experimentation. The simplified profit criteria just mentioned may be useful not only for the theoretical work, but in this experimentation as well.

Once an algorithm demonstrating acceptable performance has been determined, prototype software for actual design work could be developed. Although the new methodology, through its quantitative connecting of the needs of IC users with the producer's design process, greatly reduces the need for user interaction for the purpose of judging the merit of alternative designs, it is likely that as the methodology is further developed, estimates will increase as to the importance of interaction for checking for gross errors in model data, facilitating the optimization process, and enabling users to be confident of the results.

References

[ANT81] K. J. Antreich and R. K. Koblitz, "An interactive procedure to design centering," *Proc. 1982 IEEE Int. Symp. Circuits and Systems*, (Houston, TX), pp. 139-142, 1981.

[BRA81] R. K. Brayton, G. D. Hachtel, and A. L. Sangiovanni-Vincentelli, "A survey of optimization techniques for integrated-circuit design," *Proc. of IEEE*, vol.69, no. 10, pp.1334-1362, Oct. 1981.

[COO82] A. S. Cook and T. Downs, "Estimating manufacturing yield by means of Jacobian transformation," *IEE Proc.*, vol. 129, Pt. G, No. 4, pp. 169-180, Aug. 1982.

[DIV78A] D. A. Divekar, R. W. Dutton, and W. J. McCalla, "Experimental study of Gummel-Poon model parameter correlations for bipolar junction transistors," *IEEE J. Solid-State Circuits*, vol. SC-12, No. 5, pp. 552-559, Oct. 1977.

[DIV78B] D. A. Divekar and R. W. Dutton, "Model Parameter Correlations in Statistical Circuit Simulation," *Proc. IEEE 21st Midwest Symposium on Circuits and Systems*, Ames, Iowa, Sept. 1978.

[DIV84] D. A. Divekar, DC Statistical Circuit Analysis for Bipolar IC's Using Parameter Correlations - An Experimental Example," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-3, No. 1, pp. 101-103, Jan. 1984.

[FER85] A. Ferris-Prabhu, "Modeling of critical area in yield forecasts,: *IEEE J. Solid-State Circuits*, vol. SC-20, No. 4, pp. 874-878, Aug. 1985.

[GUP72] A. Gupta and J. Lathrop, "Yield analysis of large integrated-circuit chips," *IEEE J. Solid-State Circuits*, vol. SC-7, No. 5, pp. 389-394, Oct, 1972.

[GUP74] A. Gupta, W. Porter, and J. Lathrop, "Defect analysis and yield degradation of integrated circuits," *IEEE J. Solid-State Circuits*, vol. SC-9, No. 3, pp. 96-103, June, 1974.

[HEM81] R. Hemmert, "Poisson process and integrated circuit yield prediction," *Solid-State Electronics*, vol. 24, pp. 511-515, June 1981.

[HOC83] D. E. Hocevar, M. R. Lightner, and T. N. Trick, "Monte Carlo based yield maximization with a quadratic model," *Proc. 1983 IEEE Int. Symp. Circuits and Systems*, (Newport Beach, CA), pp. 558-561, April 1983.

[HOC84] D. E. Hocevar, M. R. Lightner, and T. N. Trick, "Monte Carlo based yield approximation technique for use in yield maximization," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-3, No. 4, pp. 279-287, Oct. 1984.

[HU79] S. Hu, "Some considerations in the formulation of IC yield statistics," *Solid-State Electronics*, vol. 22, pp. 205-211, Feb. 1979.

[MAL81] W. Maly, A. J. Strojwas, and S. W. Director, "Fabrication-based statistical design of monolithic IC's," in *Proc. ISCAS*, pp. 135-138, Apr. 1981.

[MAL82] W. Maly, and A. J. Strojwas, "Statistical simulation of the IC manufacuring process," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-1, No. 3, July 1982.

[MAL85] W. Maly, and Z. Pizlo, "Tolerance assignment for IC selection tests," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-4, No. 3, pp. 156-162, April 1985.

[NAS83] S. R. Nassif, A. J. Strojwas, and S. W. Director, *FABRICS II, A Statistical Simulator of the IC Fabrication Process: Users Manual*, SRC-CMU Center for Computer-Aided Design, 1983.

[NAS84] S. R. Nassif, A. J. Strojwas, and S. W. Director, "FABRICS II, A Statistically Based IC Fabrication Process Simulator", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-3, No. 1, January, 1984.

[NYE88] W. T. Nye, A. Tits, D. C. Riley, and A. Sangiovanni-Vincentelli, "DELIGHT.SPICE: An optimization-based system for the design of integrated circuits", *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-7, No. 4, pp. 501-519, April, 1988.

[OPA85] L. J. Opalski and M. A. Styblinski, "A new formulation of yield optimization problem: maximum profit approach," *Proc. 1985 IEEE Int. Symp. Circuits and Systems*, (Kyoto, Japan), pp. 1273-1276, June 1985.

[OPA86] L. J. Opalski and M. A. Styblinski, "A new formulation of yield optimization problem: maximum income approach," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-5, No. 2, pp. 346-360, April 1986.

[PRI70] J. Price, "A new look at yield of integrated circuits," *Proc. of IEEE*, vol. 58, pp.1290-1291, Aug. 1970.

[RIL84] D. C. Riley, "Methods for the statistical design of integrated circuits," in *EECS/ERL 1984 Research Summary*, J. Cook, ed., p. 118, U.C. Printing, Berkeley, CA, Jan. 1984.

[RIL85] D. C. Riley and A. Sangiovanni-Vincentelli, "A new, profit maximization methodology for statistical design of integrated circuits _ part 1: problem formulation," Electronics Research Laboratory Memo, May 27, 1985.

[RIL86A] D. C. Riley and A. Sangiovanni-Vincentelli, "Models for a new, profit-based methodology for statistical design of integrated circuits," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-5, No. 1, Jan. 1986.

[RIL86B] D. C. Riley and A. Sangiovanni-Vincentelli, "Models for a new, profit-based methodology for statistical design of integrated circuits," *Digest of Technical Papers, IEEE Int. Conf. Computer-Aided Design,*, (Santa Clara, CA), pp. 232-235, Nov. 11-13, 1986.

[SHY88] J.-M. Shyu and A. Sangiovanni-Vincentelli, "ECSTASY: a new environment for IC Design Optimization,", submitted for publication in *Digest of Technical Papers, IEEE Int. Conf. Computer-Aided Design,*, (Santa Clara, CA), Nov. 7-10, 1988.

[SIN81] K. Singhal and J. F. Pinel, "Statistical design centering and tolerancing using parametric sampling," *IEEE Transactions on Circuits and Systems*, vol. CAS-28, pp. 692-702, July 1981.

[STA81] C. Stapper, "Comments on "Some considerations in the formulation of IC yield statistics"," *Solid-State Electronics*, vol. 24, pp. 127-132, Feb. 1981.

[STA83] C. Stapper, F. Armstrong, and K. SAJI, "Integrated circuit yield statistics," *Proc. of IEEE*, vol.71, no. 4, pp.453-469, April, 1983.

[STE86] M. L. Stein, "An efficient method of sampling for statistical circuit design," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-5, No. 1, pp. 23-29, Jan. 1986.

[STY81] M. A. Styblinski, J. Ogrodski, L. Opalski, and W. Strasz, "New methods of yield estimation and optimization and their application to practical problems," *Proc. 1982 Int. Symp. Circuits and Systems*, (Houston, TX), pp. 131-134, April, 1981.

[STY83] M. A. Styblinski, "Stochastic approximation - a new tool for production yield optimization," *Proc. 1983 IEEE Int. Symp. Circuits and Systems*, (Newport Beach, CA), pp. 546-549, April 1983.

[STY84] M. A. Styblinski, and L. Opalski, "Software Tools for IC Yield Optimization with Technological Process Parameters," *Digest of Technical Papers, IEEE Int. Conf. Computer-Aided Design,*, (Santa Clara, CA), pp. 158-160, Nov. 12-15, 1984.

[STY86] M. A. Styblinski and L. J. Opalski, "Algorithms and software tools for IC yield optimization based on fundamental fabrication parameters," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-5, No. 1, pp. 79-89, Jan. 1986.

[VID82] L. M. Vidigal and S. W. Director, "A design centering algorithm for nonconvex regions of acceptability," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-1, No. 1, pp. 13-24, Jan. 1982.

[WAR74] R. Warner, "Applying a composite model to the IC yield problem," *IEEE J. Solid-State Circuits*, vol. SC-9, No. 3, pp. 86-95, June, 1974.

[WAR81] R. Warner, "A note on IC-yield statistics," *Solid-State Electronics*, vol. 24, No. 11, pp. 1045-1047, Dec. 1981.

[YU87] T. Yu, S. M. Kang, I. N. Hajj, and T. N. Trick, "Statistical performance modeling and parametric yield estimation of MOS VLSI," in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. CAD-6, No. 6, pp. 1013-22, Nov. 1987.

[ZEI84] D. A. Zein, C. W. Ho and A. J. Groudis, "A new interactive circuit design program," in *Proc. IEEE Int. Symp. Circuits and Systems*, (Houston, TX), pp. 913-917, 1980. R. Warner, "A note on IC-yield statistics," *Solid-State Electronics*, vol. 24, No. 11, pp. 1045-1047, Dec. 1981.