

Node-Pore Coded Coincidence Correcting Microfluidic Channel Framework: Code Design and Sparse Deconvolution

*Michael Kellman
Michael Lustig, Ed.*



Electrical Engineering and Computer Sciences
University of California at Berkeley

Technical Report No. UCB/EECS-2017-213

<http://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-213.html>

December 14, 2017

Copyright © 2017, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

Acknowledgement

I would like to thank Professor Michael Lustig, Professor Lydia Sohn, Francois Rivest, and Alina Pechacek for their intellectual discussion and advance reading of my work and this thesis.

I would also like to thank my fellow graduate student colleagues for their support and intellectual discussions. Specifically, I would like to give special thanks to my lab mates in the Lustig & Waller Labs. What I have learned from them extends far beyond the pages of this thesis.

**Node-Pore Coded Coincidence Correcting Microfluidic Channel
Framework: Code Design and Sparse Deconvolution**

by Michael R. Kellman

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences,
University of California at Berkeley, in partial satisfaction of the requirements for the
degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

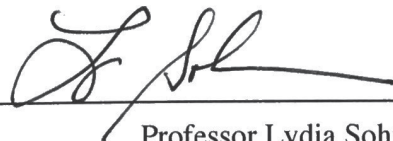
Committee:



Professor Michael Lustig
Research Advisor

12/13/17

(Date)



Professor Lydia Sohn
Second Reader

12/14/17

(Date)

**Node-Pore Coded Coincidence Correcting Microfluidic Channel Framework:
Code Design and Sparse Deconvolution**

Copyright 2017
by
Michael Kellman

Abstract

Node-Pore Coded Coincidence Correcting Microfluidic Channel Framework: Code Design
and Sparse Deconvolution

by

Michael Kellman

Master of Science in Electrical Engineering and Computer Science

University of California, Berkeley

Professor Michael Lustig, Chair

We present a novel method to perform individual particle (*e.g.* cells or viruses) coincidence correction through joint channel design and algorithmic methods. Inspired by multiple-user communication theory, we modulate the channel response, with Node-Pore Sensing, to give each particle a binary Barker code signature. When processed with our modified successive interference cancellation method, this signature enables both the separation of coincidence particles and a high sensitivity to small particles. We identify several sources of modeling error and mitigate most effects using a data-driven self-calibration step and robust regression. Additionally, we provide simulation analysis to highlight our robustness, as well as our limitations, to these sources of stochastic system model error. Finally, we conduct experimental validation of our techniques using several encoded devices to screen a heterogeneous sample of several size particles.

To my family and loved ones.

Contents

Contents	ii
List of Figures	iii
List of Tables	vi
1 Introduction	1
2 Methods	4
2.1 Coding Microfluidic Channels	4
2.2 Algorithmic Methods	6
2.3 Sources of Model Error	10
2.4 Experimental Methods	11
3 Results & Analysis	13
3.1 Simulations and Estimator Analysis	13
3.2 Experimental Microsphere Results	18
4 Discussion & Conclusion	21
4.1 Discussion	21
4.2 Conclusion	22
Bibliography	23
A Code	27
B Adaptive Detection Scheme	28
C Iterative Re-weighted Least-Squares	29
D Channel and Electrode Fabrication	30

List of Figures

1.1	Illustration of Coincidence Events: (a) a single particle transiting the channel, (b) horizontal coincidence of two particles, (c) partial coincidence of two particles, (d) vertical coincidence of two particles.	1
2.1	Device and Code Design: (a) A photograph of a pair of polydimethylsiloxane (PDMS) channels encoded with wider (nodes) and narrower (pores) regions bonded to a glass substrate with a pair of electrodes. The zoomed in region of photograph highlights a portion of the computer aided design schematic of the two channels' interleaved nodes and pores. (b) The top view of a channel as a single particle transits and the expected impedance signature from the Manchester-Barker length 13 encoding sequence. Vertical dashed lines highlight the correspondence between segments of the channel and the binary code. (c) Quasi-orthogonality (pulse compression) properties of the transitional Barker length 13 sequence and the Manchester-Barker length 13 sequence we utilize to encode our channels.	5
2.2	Coincidence Event and Forward Model Construction: (a) a traditional Coulter-counter channel and (b) a Coulter-counter MB length 13 encoded channel and their resultant impedance signals during coincidence event. In this coincidence event, a smaller faster particle #1 travels at a speed of v_1 and generates a pulse with height α_1 and duration τ_1 and a larger slower particle #2 travels at a speed of v_2 and generates a pulse with height α_2 and duration τ_2 . (c) An illustration of the decomposition of the signal resulting from the coincidence event in (a) into Equation 2.3: a forward model, A , a sparse vector, x , a baseline, b , and a noise term, n . The forward model is a dictionary of channel response signals parametrized by arrival and transit-time. x is a sparse vector representing individual particle's amplitudes at indices representing the arrival and transit time of these particles.	6

2.3	Successive Interference Cancellation Method: (a) Flow chart of our iterative SIC algorithm that solves the sparse deconvolution problem posed in Equation 2.4. (b) Illustration of coincidence event of a larger particle (#1) with two smaller particles (#2,#3) and a piece of experimental data (outlined in Section 2.4) from a MB length 13 encoded channel. (c) Step-by-step figures of three iterations of our algorithm applied to experimental data in part b, (top row) correlation with MB length 13 code’s matched filter bank and circled (red) detection of peak at arrival and transit time of the current iteration’s detected particle. Small black arrows indicate undetectable particles in first iteration. (middle row) overlaid fitted detected model components to experimental data, (bottom row) cancelled interference from iteration’s current model fit. The cancelled interference signal from previous iteration is used as input to the next iteration.	8
3.1	Pulse-Height Estimator Analysis: Effects of ideal (blue curves) versus simulated (red curves) system-model error on least-squares (LS) pulse-height estimation compared to the effects of simulated system-model error on the robust regression (yellow curves) pulse-height estimation. Estimator bias in percent error from the simulated is plotted as a function of the particle’s SNR (dB) for MB length 7, 11, 13 encoded channels. This is the observed range of SNR for the particles in our experiments outline in Section 2.4. Error bars represent one standard deviation of variability in the error of our estimate.	14
3.2	Histograms of Transit-Time Error: Histograms of transit-time error as a percent of true transit time for a range of transit times: 187.5 ms (left column), 150.0 ms (middle column), and 117.5 ms (right column) for each channel encoding: MB-length 7 (top row), 11 (middle row), 13 (bottom row). This is the observed range of transit times in our experiments outline in Section 2.4. The sample mean and variance of each distribution are as indicated.	16
3.3	Histograms of True Detections and False Alarms: Normalized histograms of false-alarm detections, orange, in the presence of manually selected experimental $10\mu m$, and $15\mu m$ true detections, blue, for MB length 7, 11, and 13 encoded channels. Sample mean and sample standard deviation of false-alarm distributions are as indicated. Black dashed vertical dashed lines represent denote two standard deviations greater than the mean of the false-alarm distributions.	17
3.4	Experimental Microsphere Results: (a-c) coded channel designs for MB length 7, 11, 13, respectively, (d-f) single particle detections for MB length 7, 11, 13, respectively, (g-i) coincidence particle detections for MB length 7, 11, 13, respectively, (j-l) pruned coincidence particle detections for MB length 7, 11, 13, respectively.	18
3.5	Experimental Matched Filtering: Example impedance responses from $5\mu m$, $10\mu m$, $15\mu m$ particles transiting a MB length 13 encoded channel (from experiments outlined in Section 2.4) and their matched-filter response. Red arrows emphasize the scale at which match filtering is able to boost SNR.	19

- 3.6 Experimental Temporal Resolution: Experimental channel responses (from experiments outlined in Section 2.4) for $15\mu m$ particles for MB length 7 (a), 11 (b), 13 (c) codes and the overlay of their matched filter responses (d). Curve colors in panel d correspond to the different code's colors in panels a through c. 20

List of Tables

3.1	Signal to Noise Ratio Analysis: Average experimental impedance-domain SNR for $5\mu m$, $10\mu m$, $15\mu m$ diameter particles for each encoding configuration, experimental matched-filter domain SNR for $5\mu m$, $10\mu m$, $15\mu m$ diameter particles for each encoding configuration. Particles from non-coincident and coincident events are included in these averages and counts are reported in Section 3.2.	19
-----	---	----

Acknowledgments

I would like to thank Professor Michael Lustig, Professor Lydia Sohn, Francois Rivest, and Alina Pechacek for their intellectual discussion and advance reading of my work and this thesis.

I would also like to thank my fellow graduate student colleagues for their support and intellectual discussions. Specifically, I would like to give special thanks to my lab mates in the Lustig & Waller Labs. What I have learned from them extends far beyond the pages of this thesis.

Chapter 1

Introduction

The Coulter principle [1] is a ubiquitous method for accurately measuring the number and size of particles in a solution [2]. The method is based on detecting and measuring a current pulse generated by a particle (*e.g.* cells or viruses), suspended in an electrolyte solution, as it flows through an aperture or channel that has an applied voltage potential across it. Since its inception in the 1950's, the Coulter principle has been used in a number of diverse applications, from biomedical (*e.g.* automated red blood cell counting [3]–[5], viral/pathogen detection [6]–[11], DNA detection [12]–[17]) to industrial applications (*e.g.* food [18], cosmetics [19]). Key to many of these applications is that the size of the aperture or channel width must be commensurate to that of the particles to be measured in order to ensure a high signal-to-noise ratio (SNR). In the case of measuring a heterogeneous mixture, when too small of an aperture or channel width is used, larger-sized particles can lead to clogging. Conversely, when too large an aperture or channel width is used, smaller-sized particles may go undetected due to low SNR. Thus, despite the overall simplicity and wide utility of the Coulter principle, challenges still exist.

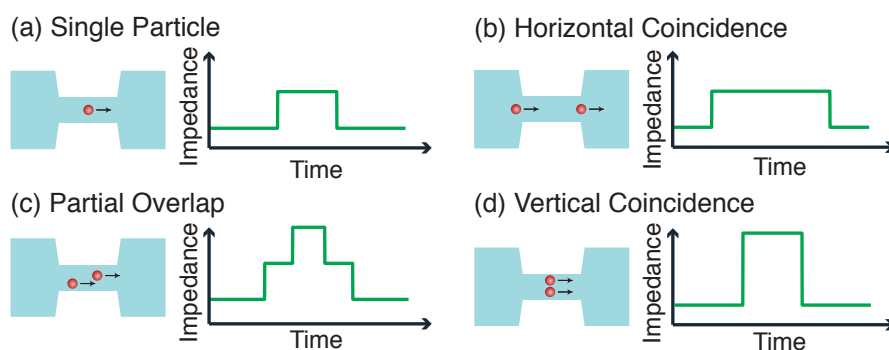


Figure 1.1: Illustration of Coincidence Events: (a) a single particle transiting the channel, (b) horizontal coincidence of two particles, (c) partial coincidence of two particles, (d) vertical coincidence of two particles.

Another well-recognized design challenge for any appropriately sized aperture or channel width is that arising from coincidence events [3], [20]–[22], *i.e.* when two or more particles enter the sensing region simultaneously. Such events lead to errors whose level of severity depends on the degree at which coincidence occur: almost no overlap, horizontal coincidence (Fig. 1.1b), to partial overlap (Fig. 1.1c), to extreme overlap, vertical coincidence (Fig. 1.1d). In general, coincidence events lead to ambiguities in detection, error in size (and, in turn, incorrect size distributions in a population), and missed rare events. Coincidence corrections range from simple removal of the particular detection to complex algorithms [20], [21], [23]–[25].

Correction of these errors can be compensated statistically by estimating the expected number of detection counts that were coincident events. This is achieved by modeling the expected number of particles in the sensing region as a Poisson process, parametric with the volume of the sensing region and number of detections [20], [21], [23]. Another correction involves screening the same sample at several known dilutions and then computing the actual count as well as the rate of coincidence events [24], [25]. Today, commercial Coulter counters utilize the first method to statistically perform coincidence correction [26]–[28], based on the number of particles detected and volume of the sensing region. While these methods might be accurate for correcting distributions, they do not resolve individual coincidence events. This is problematic when detecting specific and small sub-populations or rare cell events [6].

With the goal of achieving individual particle coincidence correction for applications of high-throughput screening, several impedance-based multi-channel designs have been proposed, each of which relies upon a unique signal-encoding mechanism (*e.g.* individual sensing electrodes [29], frequency-division multiplexing [30], orthogonal-code-division multiplexing with co-planar electrodes [31], [32]) that ultimately gives each channel of the device a signature enabling separation. While these designs allow the detection and separation of particles simultaneously transiting different channels, they fail to provide adequate coincidence correction when two or more particles transit the same channel simultaneously. Then to achieve the goal of a higher sensitivity to smaller particles, that would otherwise have SNR too low to detect, several single-channel designs have been proposed, each relying upon, again, a signal-encoding mechanism (*e.g.* channel sidewall modulation [33], electrodes on top and bottom of the channel [34], co-planar electrodes [35]). However, these methods do not enable individual particle coincidence correction.

Recently, a work by Liu et al. [31], [32] developed an approach that can perform the separation of coincidence particles transiting spatially-multiplexed channels based on the principles of code-division-multiple-access (CDMA) from communication theory. Their solution is part enabled by their electrode-encoding mechanism that differentially encodes each channel with a quasi-orthogonal signature and in part by their iterative interference cancellation algorithm. Their method is successful in separating coincidence corrections, but requires complex circuit-system modeling to properly calibrate the system [32] and their choice of binary codes have sub-optimal separation properties for coincidence correction within a single channel.

Here, we propose a joint channel design and processing framework that achieves high

sensitivity for small particles and individual particle coincidence correction, while maintaining only a single external measurement and a single channel configuration. By modeling the single microfluidic channel as a communication system that encodes passing particle signals like packets of information, we can apply the signal processing from multi-user communication theory [36] to detect and separate particles in coincidence settings. Specifically, we achieve this by amplitude modulating our channel response with Node-Pore Sensing [37], [38] (NPS) to generate a modified Barker [39] binary-code sequence. The Barker code's structure allows us to take advantage of its optimal quasi-orthogonal properties to resolve coincidences. Furthermore, with its special filtering properties, we can also achieve a gain in SNR, thereby providing a higher sensitivity to smaller particles.

We formulate the coincidence correction problem as a sparse deconvolution inverse problem and exploit the channel design and sparsity of particles flowing through the channel to provide an efficient solution via a successive interference cancellation (SIC) algorithm. With analysis, we highlight the robustness, as well as the limitations, of our method to several sources of stochastic system model error and experimental error. Additionally, we propose modifications to the SIC algorithm, which include a data-driven system model calibration step and a robust particle size estimate step, to mitigate the effects of system model error. Experimental validation is performed via the fabrication and measurement of several encoded devices and subsequent screening of heterogeneous samples. Finally, we discuss how to pick which length code to use, other coding schemes, and other sources of system modeling error.

Chapter 2

Methods

2.1 Coding Microfluidic Channels

Balakrishnan et al. [37], [38] previously demonstrated that impedance measurements of a microfluidic channel can be modulated by fabricating channels with locally wider (nodes) and narrower (pores) regions in specific sequences (Fig. 2.1a), hereafter referred to as NPS channels. As a particle transits from a node to a pore, the impedance response will increase proportionally to the ratio between the particle’s volume and the pore’s cross sectional area [40] as described in Equation 2.1 where R is the baseline impedance, ΔR is the increase in impedance, d is the diameter of the particle, L is the length of the channel, and D is the effective diameter of the channel.

$$\frac{\Delta R}{R} = \frac{d^3}{LD^2} \frac{1}{1 - 0.8(\frac{d}{D})^3} \quad (2.1)$$

Because a node’s cross sectional area is much larger than the particle, as a particle transits from a pore to a node, the impedance response will nearly return to baseline (Fig. 2.1b), thus enabling the impedance response’s binary amplitude modulation. Similar to the binary amplitude modulation used in communication systems [41], NPS coding can be used as the general encoding mechanism to provide structure in the system response of microfluidic channels[38].

We utilize the flexibility of NPS channel encoding to amplitude modulate our channel’s response with the Barker coding scheme: a binary sequence ubiquitous in communication systems (*e.g.* Direct Sequence Spread Spectrum in 802.11b Wi-Fi [41] and in radar for high-resolution detection and ranging [42]). To describe the encoding scheme’s unique properties, it is useful to draw on analogies between Barker-coded NPS channels and classic communication theory. Thus, in the terminology of communication theory, the measured signal response of a single node or pore, with its associated transit time (τ), is a symbol and the bandwidth (BW) of the channel’s encoding or symbol-rate is one over the symbol’s transit time (*i.e.* $BW = \frac{1}{\tau}$).

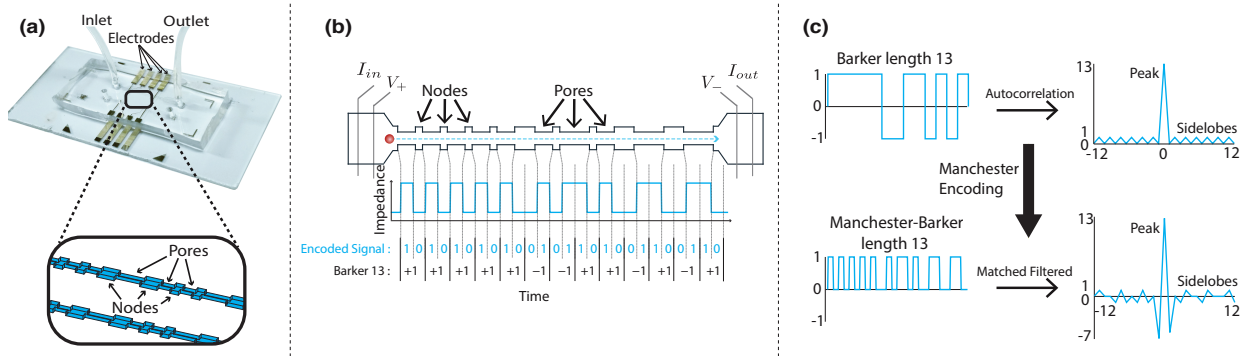


Figure 2.1: Device and Code Design: (a) A photograph of a pair of polydimethylsiloxane (PDMS) channels encoded with wider (nodes) and narrower (pores) regions bonded to a glass substrate with a pair of electrodes. The zoomed in region of photograph highlights a portion of the computer aided design schematic of the two channels' interleaved nodes and pores. (b) The top view of a channel as a single particle transits and the expected impedance signature from the Manchester-Barker length 13 encoding sequence. Vertical dashed lines highlight the correspondence between segments of the channel and the binary code. (c) Quasi-orthogonality (pulse compression) properties of the transitional Barker length 13 sequence and the Manchester-Barker length 13 sequence we utilize to encode our channels.

The central appealing property of the Barker-coding scheme is its quasi-orthogonality (referred to as pulse compression in communication theory), such that when correlated with its model, the response gives a focused high-energy peak at the point of perfect overlap and a minimal response at all other shifts (side-lobes) (Fig. 2.1c). The full width half max (FWHM) of the peak, referred to as the temporal resolution, is proportional to the transit time of a single symbol and represents the minimum distance by which two particles can be separated. Additionally, the height of the peak represents the gain in SNR we expect and is proportional to the square root of the number of symbols divided by the bandwidth (*i.e.* $\text{gain} = \sqrt{\frac{\# \text{ symbols}}{BW}}$). The bandwidth of our encoding scheme can be controlled via the speed of particles flowing through the channel and by the dimensions of the channel.

Due to the one-sided nature of our amplitude modulated impedance signal, we took a similar approach to Levanon et al. [43] and adapted the Barker code by encoding its bit-values in signal transitions, Manchester encoding [44] (*i.e.* +1 as high-to-low transition, -1 as low-to-high transition). Impedance transitions from high-to-low are achieved with a pore-node sequence and low-to-high with a node-pore sequence. The combined Manchester-Barker code, hereafter referred to as MB codes, is shown in Figure 2.1b. The result is a binary sequence that is double in length and has similar properties in terms of the quasi-orthogonality and SNR gain to that of the traditional Barker code when correlated with the model, hereafter referred to as matched filtering (highlighted in Figure 2.1c).

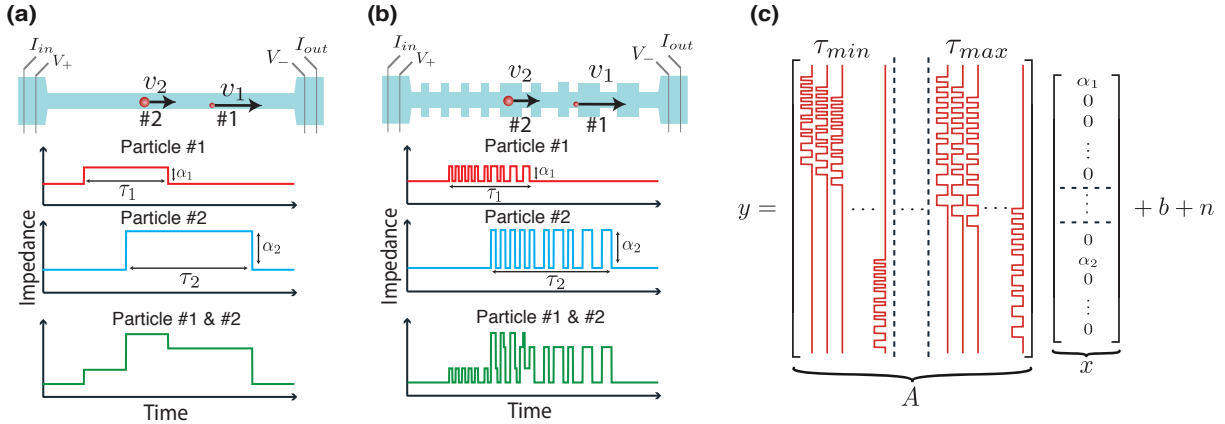


Figure 2.2: Coincidence Event and Forward Model Construction: (a) a traditional Coulter-counter channel and (b) a Coulter-counter MB length 13 encoded channel and their resultant impedance signals during coincidence event. In this coincidence event, a smaller faster particle #1 travels at a speed of v_1 and generates a pulse with height α_1 and duration τ_1 and a larger slower particle #2 travels at a speed of v_2 and generates a pulse with height α_2 and duration τ_2 . (c) An illustration of the decomposition of the signal resulting from the coincidence event in (a) into Equation 2.3: a forward model, A , a sparse vector, x , a baseline, b , and a noise term, n . The forward model is a dictionary of channel response signals parametrized by arrival and transit-time. x is a sparse vector representing individual particle's amplitudes at indices representing the arrival and transit time of these particles.

2.2 Algorithmic Methods

System Modeling

As a particle transits the channel at constant pressure, it generates an impedance signal response that depends on the node-pore encoding, the length of the channel, and the velocity of the particle. If a similar particle transits the channel at another time with the same velocity, it will produce a similar impedance signal response. If a particle has a different velocity, from following a different velocity streamline or due to interactions with the channel sides, the impedance response will have a similar shape, but will be dilated or compressed in time. Furthermore, if in a coincidence event, the impedance signal response will be the superposition of their respective impedances had each particle transited the channel individually. Thus, the impedance measurement of the channel response (y_t) for particles passing through the channel can be well approximated as a linear time-invariant (LTI) system that is parametrized by each particle's transit time. This allows particles transiting the channel to be modeled as a convolution of a time-dilated system response (h_τ) with a series of scaled impulse functions (x_t) (Fig. 2.2a-b). Each impulse represents the arrival time and the signal amplitude that is proportional to the associated particle's size. In addition, a

time-varying affine term is added to account for slow time-varying baseline drift (b_t). Finally, we include an additive Gaussian noise term (n_t), which we assume to have zero mean and variance corresponding to sensor noise.

$$y_t = h_\tau * x_t + b_t + n_t \quad (2.2)$$

A discretization of the problem can be formulated in matrix form as

$$y = Ax + b + n, \quad (2.3)$$

where the columns of A consist of unit-amplitude shifted and dilated dictionary of channel responses for a range of transit-times and x a sparse vector in which each non-zero element represents the signal amplitude of an individual particle and which indices represent an individual particle's arrival-time and transit-time. This system model is illustrated in Figure 2.2c.

Inverse Problem Formulation

The problem of estimating particles' signal amplitudes, arrival times, and transit times can be viewed as a deconvolution. While unconstrained deconvolution is in general a difficult problem, in this case, the number of particles passing through the channel is statistically bounded by the channel length and the particle concentration within the solution. Therefore, we can pose the deconvolution as the following cardinality constrained linear inverse problem,

$$\begin{aligned} \min_{x,b} \quad & \|Ax + b - y\|_2 + \lambda \|Db\|_2 \\ \text{s.t.} \quad & \text{cardinality}\{x \in \text{range}(t, \tau)\} < k \end{aligned} \quad (2.4)$$

in which the number of particles transiting the channel, $\text{cardinality}\{x \in \text{range}(t, \tau)\}$, in a fixed period of time is constrained by k . We solve Equation 2.4 for approximately sparse entries in x which correspond to arrival and transit time with amplitude proportional to the size of the particle. We also simultaneously solve for the baseline term, b , which is constrained to be smooth using the ℓ_2 regularization of its second-order difference. This difference operator is represented by D . The parameter λ controls the bandwidth of the estimated baseline and is empirically tuned to penalize the high-frequency spectrum of the baseline signal to not allow cross-talk between the values of x_i and b .

Implementation of Successive Interference Cancellation

To reduce the computational complexity of the process, we break the input signal into overlapping blocks. Each block is processed separately, and the results are consolidated at the end. For each block, we solve the sparse deconvolution using a greedy successive interference cancellation algorithm, similar to orthogonal matching pursuit [45]. Figure

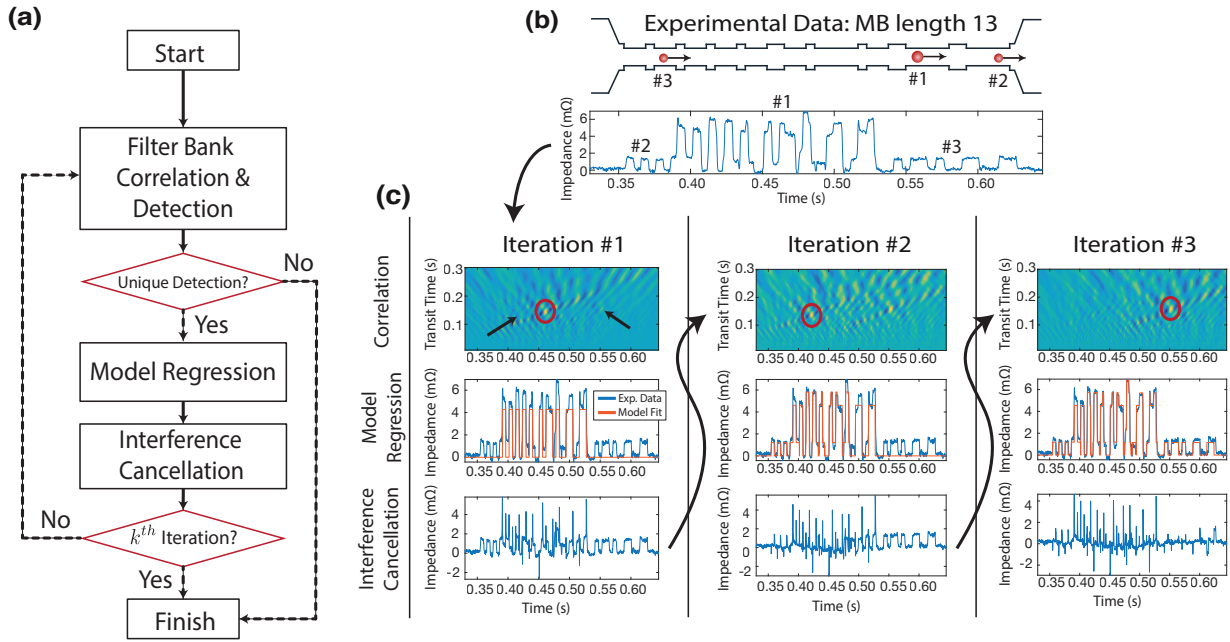


Figure 2.3: Successive Interference Cancellation Method: (a) Flow chart of our iterative SIC algorithm that solves the sparse deconvolution problem posed in Equation 2.4. (b) Illustration of coincidence event of a larger particle (#1) with two smaller particles (#2, #3) and a piece of experimental data (outlined in Section 2.4) from a MB length 13 encoded channel. (c) Step-by-step figures of three iterations of our algorithm applied to experimental data in part b, (top row) correlation with MB length 13 code’s matched filter bank and circled (red) detection of peak at arrival and transit time of the current iteration’s detected particle. Small black arrows indicate undetectable particles in first iteration. (middle row) overlaid fitted detected model components to experimental data, (bottom row) cancelled interference from iteration’s current model fit. The cancelled interference signal from previous iteration is used as input to the next iteration.

2.3a illustrates our method. It is outlined by iterating a sequence of correlations with a matched filter-bank (*i.e.*, the dictionary), detection, model fitting, and cancellation steps. Over the iterations we construct a list of the strongest detections, referred to as the list of detected signal components. We grow the list by adding the strongest unique detection every iteration. Each iteration we jointly fit all detected signal model components and a smooth baseline term to the data to estimate the particles’ impedance signal amplitudes. By cancelling/peeling the strongest signal in each iteration, we allow for the impedance response of smaller particles to be detected in successive rounds. Figure 2.3b is one example of experimental data (outlined in Section 2.4) of a coincidence of three particles (one of size $15\mu m$ and two of size $10\mu m$) transiting through a MB length 13 encoded channel and Figure 2.3c illustrates three iterations of the SIC algorithm.

Basic Methodology

In this section, we first describe a basic approach for implementing the SIC algorithm. Figure 2.3c demonstrates several iterations of the algorithm on a block of experimental data. First, we apply the transpose, *i.e.* A^T , matrix to the data, hereafter referred to as the matched filter-bank response, or correlation map $c(t, \tau)$ parametrized by arrival times t and transit times τ . This operation is equivalent to filtering with a matched filter-bank of time-dilated and time-contracted MB codes. The filter bank is comprised of normalized filters each with a unique transit time parameter such that the continuous parameter space is linearly discretized over a range of plausible particle transit times. The number of filters in the filter bank determines the resolution of transit-time parameter estimate and thus effects the computation complexity as well as accuracy of our method, this is discussed further in Section 3.1. The top row of Figure 2.3c shows the filter-bank response, in which the dominating peak indicates a particle's presence, and red circle indicates its detection. The small black arrow-heads indicate peaks that are initially obscured by the signal interference from the large particle, to be later revealed through the peeling process.

We adopt an adaptive threshold detection scheme [42] that is performed on the matched filter-bank response to localize a particle's arrival and transit time parameters. The large dynamic range of particle sizes could yield different levels of 'true-peaks' corresponding to actual particles, and 'false-peaks' corresponding to sidelobes of the matched filter-bank response. Therefore an adaptive scheme is necessary to manage false alarms and mis-detections in a signal with a wide dynamic range of particle sizes. Our adaptive detection scheme is motivated by the constant false alarm rate criteria [46], where each correlation value $c(t, \tau)$ is evaluated against a threshold derived from an estimate of the surrounding signal energy. Let $\mathcal{E}(t_0, \tau_0)$ be a function that computes a surrounding signal energy of the correlation function $c(t, \tau)$ in a window around $t = t_0$ and $\tau = \tau_0$. In this case, the criteria for detection at point (t_0, τ_0) is set to be

$$c(t_0, \tau_0) > \alpha \mathcal{E}(t_0, \tau_0), \quad (2.5)$$

where α is a tuning parameter. Robust estimates of the surrounding signal energy function as well as selections of α are further discussed in Appendix B.

The detection scheme chooses a list of possible candidate detections, from which a single unique detection, corresponding to the greatest correlation value, is selected for each iteration. The transit-time and arrival-time estimates for a detection are chosen to correspond to that of the best correlating filter in the matched filter bank and the time point when that matched filter response is maximum. Once a detection is added to the list, the amplitude of all the detected particles are re-evaluated using a least-square regression. The middle row of Figure 2.3c shows examples of resulting fits. More formally, we define \tilde{A}_i , to be a subset of columns of A corresponding to signal responses of detected particles by the i^{th} iteration, and $\tilde{x}_i \in \mathcal{R}^i$ to be the unknown signal amplitudes of the subset of detected peaks by the i^{th} iteration. In this case, the entire least-squares fit at the i^{th} iteration is formulated as,

$$\min_{\tilde{x}_i, b} \|\tilde{A}_i \tilde{x}_i + b - y\|_2 + \lambda \|Db\|_2 \quad (2.6)$$

In each iteration we only solve for a few variables and the baseline and hence the computation complexity of this portion of the algorithm is low.

Finally, the time signals of the detected particles are computed and subtracted from the acquired signal by computing $r = y - \tilde{A}_i \tilde{x}_i$. This is illustrated in the bottom row of Figure 2.3c. This residual is used as an input to the next iteration for the purpose of possibly detecting other particles of equal or lesser size.

This process is repeated until either of the stopping conditions, *i.e.*, no more significant detections in the block or the process reaches the k^{th} iteration of SIC are met. The second stopping criteria enforces the cardinality constraint from Equation 2.4. Finally, in post-processing, the particle sizes are computed as a function of channel dimensions and pulse height [40] via Equation 2.1.

2.3 Sources of Model Error

The bottom row of Figure 2.3c illustrates that the interference cancelled signals at each iteration have sparse outlier residuals. The presence of sparse residual outliers could possibly introduce biases into the least-square particle size estimates, errors in the arrival-time and transit-time estimates, and detections errors. These sparse outlier residuals cannot be fit by varying the amplitude or transit time of the signal components and thus are a result of modeling error (discrepancies between the ideal and measured MB signals). The discrepancies can be separated into two components; average and a stochastic. The average deviation affects all particles equally and could result from channel geometry variation or be due to miscalculations between the ideal code and the actual device. The stochastic deviations are small deviations from the model and could arise from random interactions between the particle and the channel. This could be caused by particles travelling along streamlines with different flow-rates; such streamlines could be the result of parabolic flow effects[47].

System Model Calibration

To model the average deviation effect, we use a data-driven calibration step in which we empirically recalculate the symbol lengths of our MB code to match the average impedance response of particles that go through the channel. The calibration is achieved by performing a first pass of detecting high SNR particles ($15\mu m$ particles in our experiments outlined in Section 2.4) using the ideal matched filter bank. For each of the detected signatures, we look at the transit time of each node and pore of the MB code, specifically the number of samples between signal transitions normalized by the total number of samples of the MB code signature. This scale invariant measure is then used to compare between the detected particle signatures and to generate a calibrated model through regression. In our list of detections, a

small number of the high SNR signals are also coincidence events, thus introducing possibly incorrect symbol timing measurements into the calibration. To enable our calibration to be robust to these outlier events, we perform a robust regression [48] rather than a least-squares regression, by utilizing an ℓ_1 -norm rather than an ℓ_2 -norm as seen in equation 2.7.

$$\arg \min_{z_{\text{cal}}} \sum_{i=0}^{N-1} \|z_i - z_{\text{cal}}\|_1. \quad (2.7)$$

Where z_i is the vector of normalized symbol timings for the i^{th} detection and we solve for the calibrated MB code timings, z_{cal} . Once the calibrated MB code model is estimated, a new matched filter bank is generated.

Robust Regression Formulation of SIC

Even with the steps described above, we will still have stochastic variation in the transit of particles that would result in model mismatches between our data and calibrated MB codes. These mismatches will manifest as outlier residuals, which can be seen clearly in Fig. 2.3c. These sparse outliers can introduce bias into the estimation of the pulse amplitudes. To reduce this bias, we turn again to a robust regression [48] approach with the following formulation,

$$\min_{\tilde{x}_i, b} \|\tilde{A}_i \tilde{x}_i + b - y\|_1 + \lambda \|Db\|_2. \quad (2.8)$$

In our implementation we solve the robust regression via an iterative re-weighted least-squares (IRLS) [49] approach (See App. C).

2.4 Experimental Methods

Device Fabrication

Similar to those outlined in Balakrishnan's et al.[38], we employed standard micro-fabrication techniques to fabricate planar electrodes onto glass substrates. These electrodes allow us to perform a four-probe measurement to sense the current through the entire MB-coded NPS device. Further, we employed standard soft-lithographic techniques to embed our MB-coded channel in polydimethylsiloxane (PDMS) slabs (See App. D). We implemented three different MB-length codes: 7, 11, and 13. We chose the channel height to be $20\mu\text{m}$ and overall length as 4 mm. We varied the width of the pores and nodes to maintain a 2.5 ratio of node-to-pore width to ensure the signal would return to baseline when particles transit from pore to node. Specifically, pores are designed to be $20\mu\text{m}$ wide and nodes are designed to be $50\mu\text{m}$ wide. We also kept a 2.5 pore-to-node length ratio to maintain an equal transit time of particles in nodes and pores (illustrated in Figure 3.4). This ensures that the recorded pulses have an equal number of samples in the impedance response for each symbol

of a particle’s signature (illustrated in Figure 2.1b). To seal the devices, we exposed both the PDMS slabs and glass substrates with the pre-fabricated electrodes (See App. D) to oxygen plasma (300mT, 80W, 30sec), aligned the two together, and then permanently bond the PDMS and glass substrate by heating at 80°C for 30 minutes.

Experimental Setup

We applied a 1V DC voltage across the device and passed the measured current signal through a current preamplifier (DL Instruments 1211, Brooktondale, NY, USA) to a data acquisition (DAQ) board (National Instrument PCI-6035E, Austin, TX, USA). The DAQ sampled the analog signal at a rate of 50 kHz.

We screened a 1 : 1 : 1 ratio mixture of polystyrene microspheres with mean and standard deviation diameter: $\mu = 4.9\mu m, \sigma = 0.44\mu m$ (Interfacial Dynamics 1-5000), $\mu = 9.98\mu m, \sigma = 1.12\mu m$ (Polysciences Inc. 64130), and $\mu = 14.73\mu m, \sigma = 1.36\mu m$ (Polysciences Inc. 64155) in phosphate-buffered saline at a concentration of 5×10^5 particles/mL. In addition, Bovine Serum Albumin is mixed into the particle solution at a final concentration of 1% to coat particles and prevent against clogging and multiparticle aggregates from forming. We flowed this mixture of microspheres through three device designs (Fig. 3.4 row one): MB length 7, MB length 11, and MB length 13. A flow controller (Elveflow OB1 MK3) was used to apply 70 mbar of pressure through the channel, resulting in a flow rate of $\sim 2\mu l/min$. Prior to applying our methods outlined in Section 2.2, we apply Ohm’s law to our current measurements to obtain the impedance response. We low pass filter and down-sample our signal by a factor of 15 to get samples at an equivalent rate of 3.33 kHz.

Chapter 3

Results & Analysis

3.1 Simulations and Estimator Analysis

In this section we analyze our processing system in terms of the bias and variance of the estimated pulse heights and transit-times of particles. In addition, we study how reliably our system can detect particles in a coincidence event. The practical issues governing these parameter estimators' performance are the various system model errors present in the measured signal as well as noise. These result in bias in the pulse-height estimates, variance in transit-time estimates, and decreased dynamic range of reliable detections as compared to an ideal system. To evaluate the performance of our system for the different non-idealities, we created a tool for simulating the system response of our NPS experimental setup.

Our simulation tool focuses on mimicking both the deterministic and stochastic components of our system. The tool is parametric with deterministic channel specific parameters (z_{geo} , a vector containing the node pore ordering and the length of each node or pore in the channel encoding) and experimental specific parameters (pulse height, α , transit time through the channel, τ). The resulting ideal system model simulation, $x(\alpha, \tau, z_{geo})$, is sampled at 50kHz. In addition, smooth transition response from nodes to pores (node-pore) and visa versa is accounted for by convolving $x(\alpha, \tau, z_{geo})$ with a normalized Hanning window, k , with a length that is a percentage of the total length of the system response.

Stochastic system model components for additive noise and random particle-channel interactions are included. Additive white Gaussian noise (AWGN) with zero mean and unit variance is scaled to experimentally observed levels, σ_{exp} , and added to the simulation. Particle-channel interactions manifest by manipulating the channel geometry, now represented z_{stoch} , which is a random vector according to a uniform distribution with mean z_{geo} and spread \pm a percentage of the total length of the system response. These factors are summarized below in Equation 3.1.

$$y_{t,sim} = k * x(t, \alpha, \tau, z_{stoch}) + \sigma_{exp} n_t \quad (3.1)$$

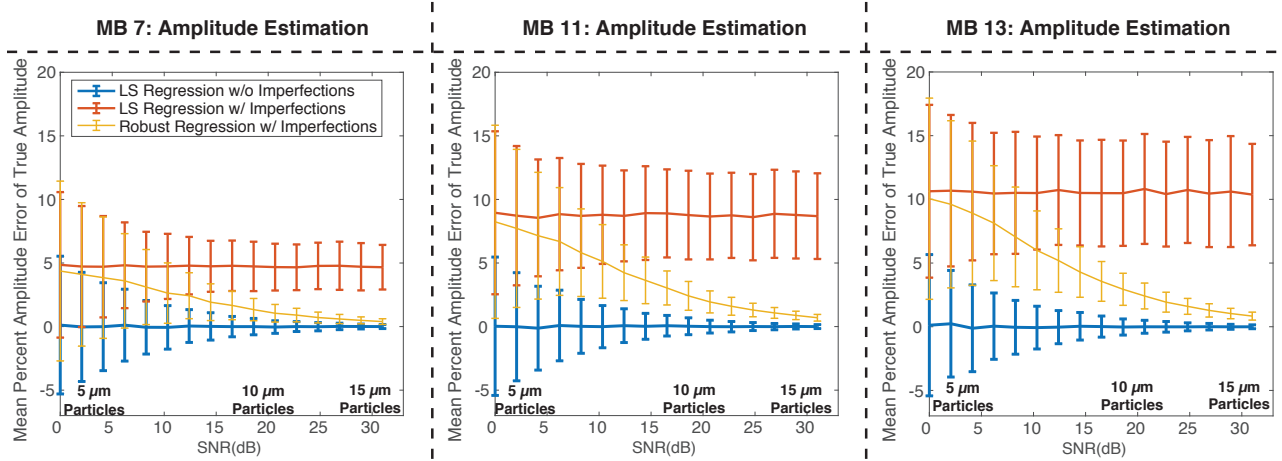


Figure 3.1: Pulse-Height Estimator Analysis: Effects of ideal (blue curves) versus simulated (red curves) system-model error on least-squares (LS) pulse-height estimation compared to the effects of simulated system-model error on the robust regression (yellow curves) pulse-height estimation. Estimator bias in percent error from the simulated is plotted as a function of the particle’s SNR (dB) for MB length 7, 11, 13 encoded channels. This is the observed range of SNR for the particles in our experiments outlined in Section 2.4. Error bars represent one standard deviation of variability in the error of our estimate.

Pulse-Height Analysis

To isolate and analyze the effect of MB code non-idealities, which we attempt to mitigate via robust regression, we performed a simulation experiment mimicking our experimental system. We simulated instances of non-overlapping particles going through MB length 7, 11, 13 encoded channels of the same length that yield a transit time of 150 ms (the average observed transit time for our experiments outlined in Section 2.4). The random particle-channel interactions are accounted by uniformly varying the channel geometry, z_{geo} by 1% the total response length and the node-pore transition kernel has length 0.5% the total response length. The simulated pulses were evaluated at 50KHz sampling rate and down-sampled to a rate of 3.333kHz. A range of relevant particle sizes were simulated by varying the pulse height of signatures simulated. In order to match our experimental data, the AWGN is scaled to have a standard deviation matching that of the experiments, $\sigma_{exp} = 1.24 \times 10^{-4}$ such that the range of pulse heights correspond to a range of SNRs from 0-30 dB. This is the observed range of SNR for the particle sizes in our experiments outlined in Section 2.4. Baseline drifts were not simulated here, as their effect is isolated in the regression performed in Equation 2.6 and Equation 2.8. For each particle size, a 1000 signatures were generated, from which the pulse heights were estimated with both the least-squares and the robust regression techniques. The sample bias and variance of the estimates were then computed.

Figure 3.1 highlights the imperfections’ effects on the bias and variance estimator metrics of the least-square pulse-height estimate side-by-side with our proposed robust regression

pulse-height estimate. The blue curves in Figure 3.1 show the bias and variance for ideal responses with varying noise and without stochastic and node-pore transition effects. In that case, a least-square regression would be suitable for pulse-height estimation. However, in the presence of the system model imperfections, the least-squares fit would induce significant additional bias and variance in pulse height estimates, as visualized by the red curves in Figure 3.1. In contrast, our method of robust regression (yellow curves in Figure 3.1) provides a reduction in estimation bias and variance over the least-square pulse height estimates on simulated signals with imperfections over a relevant range of SNR, which correspond to particle size. Another important observation that is seen in Figure 3.1 is that the robust regression plays a more important role in the longer codes, MB length 11 and 13. These more complex codes are more susceptible system model error as they exhibit many more node-pore transitions.

Transit-Time Analysis

To analyze the effect of MB code non-idealities on our ability to estimate the transit time through the device, we perform a simulation to empirically measure this estimator’s bias and variance. Again, we simulated instances of non-overlapping particles going through MB length 7, 11, 13 encoded channels, but now rather with the transit time of 150.0 ms, 25% faster (112.5 ms), and 25% slower (187.5 ms) (the observed range of transit time for our experiments outline in Section 2.4). The random particle-channel interactions and node-pore transitions are accounted for with the same parameters as in Section 3.1. The simulated pulses were evaluated at 50KHz sampling rate and down-sampled to a rate of 3.333kHz. The pulse height simulated was set be $4m\Omega$, corresponding to $15\mu m$ particles and held constant, as this parameter minimally effects the estimate of transit time. Again, baseline drifts were not simulated here because their effects are isolated in the regression performed in Equation 2.6 and Equation 2.8. The simulated signals’ transit times (a 1000 signatures per transit time) were then estimated by applying a matched filter bank with 500 filters evenly spaced from 30ms to 270ms and selecting the transit time corresponding to the matched filter with the maximum response.

Figure 3.2 highlights the variability of the transit-time estimator through histograms. As visualized, all encoding configurations have similar shaped distributions of error. Much as in the case for amplitude estimation, the longer codes demonstrated slightly more sensitivity to the stochastic variations, highlighted by the increased variances (upper right of each subplots). Additionally, we observe that slower particles have a greater absolute error in transit time than faster particles, but have the same relative error. This estimator analysis helps us to choose the correct number of filters to discretize our transit time parameter space. The variability due to non-idealities determines the upper bound on our estimator’s accuracy resolution, thus telling us the fewest number of filters we need to accurately estimate transit time. It also suggests that the transit time resolution of the filter bank can be coarser for slow transit time and finer for the faster transit times.

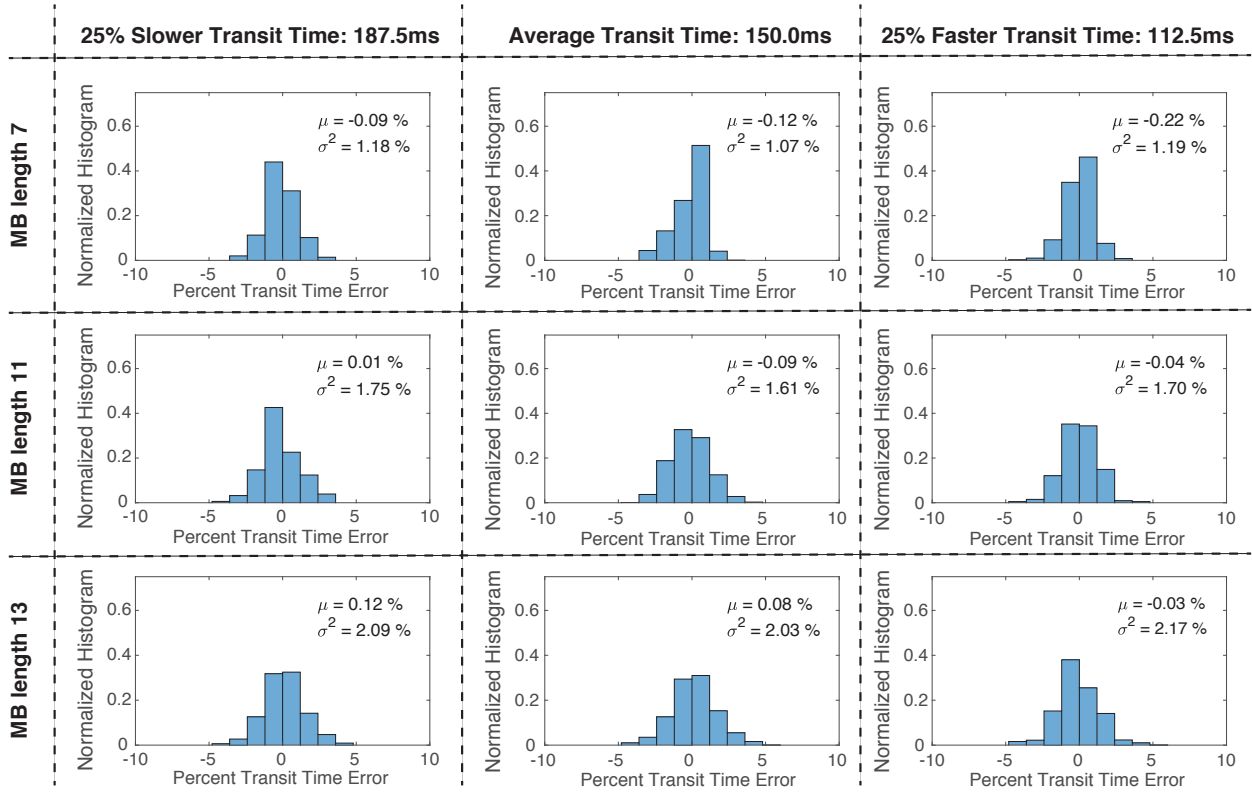


Figure 3.2: Histograms of Transit-Time Error: Histograms of transit-time error as a percent of true transit time for a range of transit times: 187.5 ms (left column), 150.0 ms (middle column), and 117.5 ms (right column) for each channel encoding: MB-length 7 (top row), 11 (middle row), 13 (bottom row). This is the observed range of transit times in our experiments outline in Section 2.4. The sample mean and variance of each distribution are as indicated.

Dynamic Range Analysis in Coincidence events

Any model mismatch in the SIC process would yield a residual error that could inhibit the detection of signals from smaller particles. The level of this residual, which could be modeled as stochastic interference in the correlations with the matched filter-bank would effectively determine the lower bound on the dynamic range of reliable detections in coincidence events. For example, the residuals from a mismatched estimate of a $15\mu m$ particle could potentially mask-out the peak correlation of particles smaller than a colloquial $5\mu m$ particle – thus limiting the dynamic range in a coincidence event with the $15\mu m$ particle to detecting particles larger than $5\mu m$. Alternatively, the same residual energy could correlate with the matched filter bank and cause a false alarm detection of a $5\mu m$.

To analyze the effect of MB code non-idealities on our ability to detect particles in coincidence settings, we manually selected $10\mu m$ and $15\mu m$ non-coincident detections from

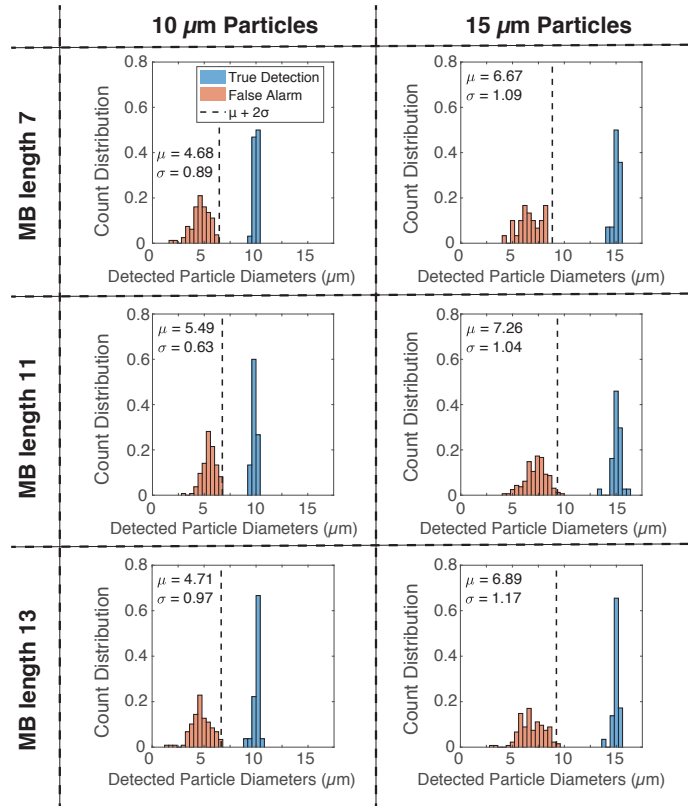


Figure 3.3: Histograms of True Detections and False Alarms: Normalized histograms of false-alarm detections, orange, in the presence of manually selected experimental $10\mu\text{m}$, and $15\mu\text{m}$ true detections, blue, for MB length 7, 11, and 13 encoded channels. Sample mean and sample standard deviation of false-alarm distributions are as indicated. Black dashed vertical dashed lines represent denote two standard deviations greater than the mean of the false-alarm distributions.

each experimental setup: MB length 7 (32 and 14 counts respectively), 11 (30 and 37 counts respectively), and 13 (27 and 29 counts respectively) in post processing. We perform the same processing as outlined in algorithmic methods on each selected detection with a stopping criteria of 6 SIC iterations. From the processing we expect exactly a single true detection to be present. All successive detections that are found are false alarms and can be used to estimate a distribution of false alarms.

Figure 3.3 shows the histogram of detections from an MB length 7 (top row), 11 (middle row), and 13 (bottom row) experiments for selected $10\mu\text{m}$ (left column) and $15\mu\text{m}$ (right column) particles. The mean and standard deviations for the false-alarm distributions are listed within each subplot and thresholds, marked by black dashed vertical lines, denote two standard deviations above the mean particle size of the false-alarm distribution (a false alarm probability of 2.275%). These thresholds represent possible lower bounds on the dynamic

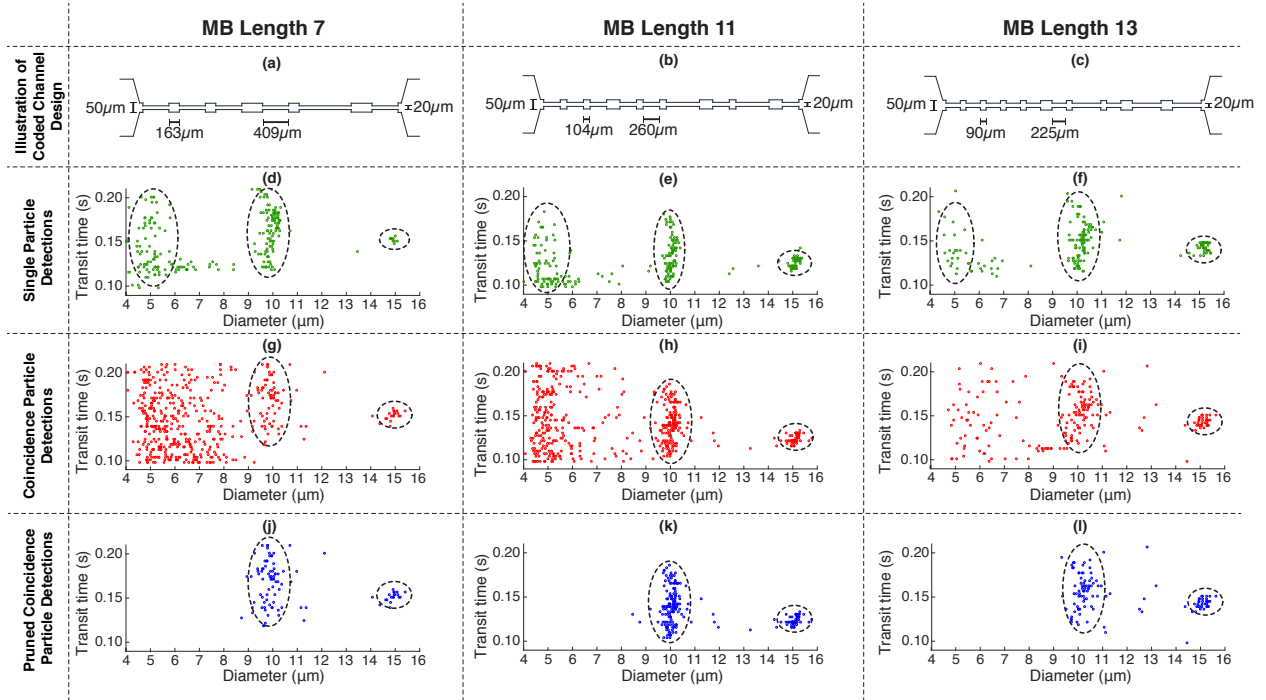


Figure 3.4: Experimental Microsphere Results: (a-c) coded channel designs for MB length 7, 11, 13, respectively, (d-f) single particle detections for MB length 7, 11, 13, respectively, (g-i) coincidence particle detections for MB length 7, 11, 13, respectively, (j-l) pruned coincidence particle detections for MB length 7, 11, 13, respectively.

range of particle sizes that we can detect in coincidence settings with an acceptable number of false alarms detections.

3.2 Experimental Microsphere Results

Experiments were conducted as outlined in Section 2.4, raw data was processed with methods outlined in Section 2.2, and results are presented in Figure 3.4d-l. We separated the detections to those with and without coincidence events. In the non-coincidence detections, visualized in Figure 3.4d-f, distinctive clusters of three differently sized particles are detected. In the coincident setting, visualized in Figure 3.4g-i, the lower range of particle sizes is flooded with spurious false alarms and only two of the larger distinctive clusters are present. Motivated by the unreliable detection of smaller particles, detection pruning based on pulse-height amplitude is performed on coincidence event detections, according to analysis in Section 3.1 and with results displayed in Figure 3.4j-l. Coincidence detections in the presence of $10\mu m$ and $15\mu m$ particles are pruned if their amplitude is less than thresholds derived from analysis in Section 3.1. From Figure 3.4, we report $5\mu m$, $10\mu m$, and $15\mu m$ di-

Particle Size (μm)	Impedance Domain SNR (dB)			Matched-Filter Domain SNR (dB)		
	5	10	15	5	10	15
MB length 7	1.77	18.89	30.52	22.93	39.98	51.55
MB length 11	1.61	19.33	30.94	24.10	39.69	51.08
MB length 13	2.18	18.47	30.33	23.08	39.57	51.23

Table 3.1: Signal to Noise Ratio Analysis: Average experimental impedance-domain SNR for $5\mu m$, $10\mu m$, $15\mu m$ diameter particles for each encoding configuration, experimental matched-filter domain SNR for $5\mu m$, $10\mu m$, $15\mu m$ diameter particles for each encoding configuration. Particles from non-coincident and coincident events are included in these averages and counts are reported in Section 3.2.

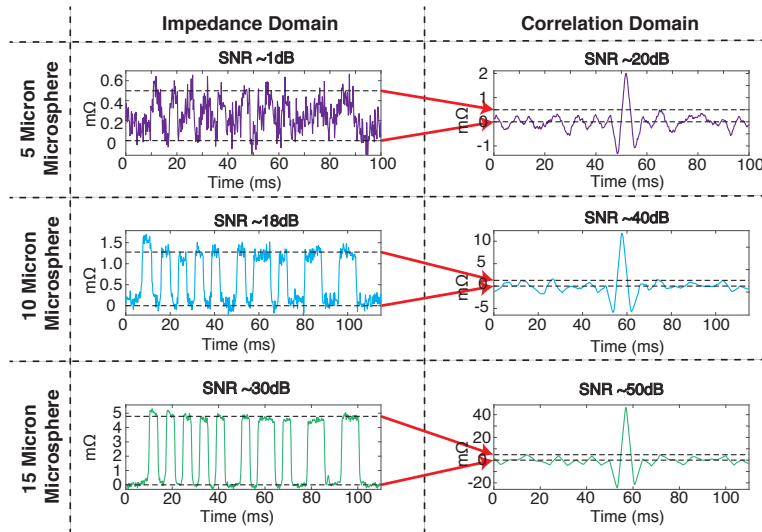


Figure 3.5: Experimental Matched Filtering: Example impedance responses from $5\mu m$, $10\mu m$, $15\mu m$ particles transiting a MB length 13 encoded channel (from experiments outlined in Section 2.4) and their matched-filter response. Red arrows emphasize the scale at which match filtering is able to boost SNR.

ameter non-coincidence and coincidence event detected particle counts for MB length 13 (27, 161, 86), MB length 11 (78, 228, 116), and MB length 7 (100, 158, 34). Further highlighted in Figure 3.4 is the wider spread of transit time for $5\mu m$ and $10\mu m$ particle distributions. This could be due to the parabolic flow profile [47] across the cross section of the channel. This would cause the flow-rate to vary from the central axis of the channel, so while the larger particles remain in the center, smaller particles could move off-axis and experience slower flow-rates.

As discussed earlier, the gain in SNR is due to the properties of matched filtering, and is specifically helpful when detecting $5\mu m$ diameter particles. These particles have low SNR

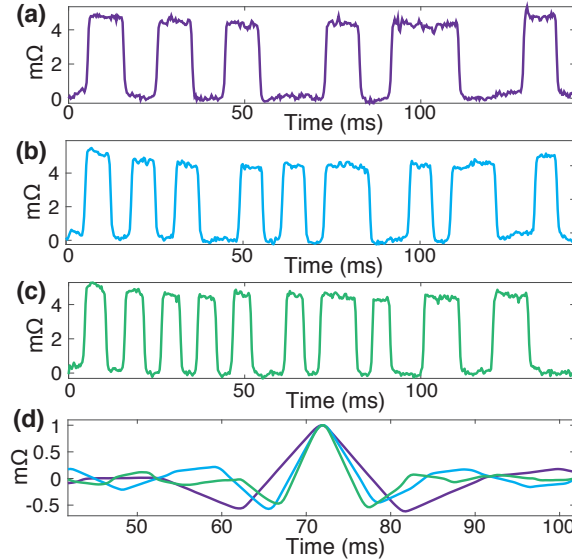


Figure 3.6: Experimental Temporal Resolution: Experimental channel responses (from experiments outlined in Section 2.4) for $15\mu m$ particles for MB length 7 (a), 11 (b), 13 (c) codes and the overlay of their matched filter responses (d). Curve colors in panel d correspond to the different code’s colors in panels a through c.

(~ 1 dB) in the measured signal and show improved SNR (~ 23 dB) after matched filtering. Average SNR for the detected $5\mu m$, $10\mu m$, and $15\mu m$ particles (Fig. 3.4) are reported in Table 3.1 in the measured impedance domain and in the matched-filter domain (counts for these averages were reported above). Further, Figure 3.5 visualizes this SNR gain with experimental impedance signatures in the left column and matched-filtered responses of the same experimental data in the right column. In addition, the left column of Figure 3.5 highlights the wide dynamic range of pulse heights we observe for particles of $5\mu m$ (top row), $10\mu m$ (middle row), $15\mu m$ (bottom row) diameters transiting our channels.

The gain in temporal resolution due to matched filtering is discussed theoretically in Section 2.1 and is observed experimentally and highlighted in Figure 3.6. For a constant length channel, MB length 7, 11, and 13 have increasingly finer temporal resolution inversely proportional to their code’s BW. Specifically, the FWHM of each correlation peak in Figure 3.6 decreases as their respective code’s BW increases.

Chapter 4

Discussion & Conclusion

4.1 Discussion

Our MB codes can be characterized by their temporal resolution and number of symbols. The temporal resolution of a code is determined by the symbol transit time and is controlled by the flow speed and channel's dimensions. These specifications are often application specific parameters, thus special consideration is required when deciding which code to use in a specific instance.

For a device of fixed length, the MB 13 encoding will have the finest temporal resolution. The minimum symbol length is limited by the diameter of the particles that are being sized. The symbol length must be several times longer than the particles' diameter, such that modulation of the signal when transiting from node to pore still occurs. Failure to do so will result in symbol pulses that are not sharp and pronounced, leading to induced system-modeling error. When the minimum symbol length cannot be met and shorter device are desired, then shorter MB codes (i.e. MB length 7 and MB length 11) should be used to encode the channel.

As channel length further increases and the proposed set of MB codes are dilated to fit the longer channels, desired temporal resolution will decrease and an increased number of coincidence events could be observed. Applications requiring significantly longer channels will require a new set of more advanced codes to achieve the same temporal resolution, as the longest known Barker sequence has length 13. These longer codes will have less favorable quasi-orthogonality properties than our proposed set of codes, but could still be effective in providing both gain a in SNR and temporal resolution. Possible candidates for longer sequences with similar but sub-optimal properties [50] include maximum length sequences [51], Gold sequences [52], and sequences that are comprised of non-binary symbols [50].

Experimental data has validated our methods and has driven important modifications and analysis to our algorithmic methods. We observed that decreased bias due to the robust regression (in agreement with our simulation) is the most important factor in mitigating the effect of system modeling error. Without this component, significant energy is left in

the residual signal that well correlates with the system model resulting in false alarms. We believe, based on our analysis in Section 3.1, that false alarms determine the dynamic range's lower limit of detectable particle size in coincidence events and are the limiting factor of channel-encoding methods for individual particle coincidence correction.

Our analysis, based on simulations with experimentally relevant parameters, suggests several trends that could influence code selection. While more complex codes (i.e. MB length 11 and 13) provide a finer temporal resolution and a less pronounced sidelobes than MB length 7 encoding, they are more susceptible to channel and flow non-idealities due to the increased number of node-pore transitions. These non-idealities manifest themselves as a combination of errors: increased pulse height estimator bias, increased transit time estimator variance, and an increased number of false alarm detections.

We have experimentally observed several other areas of deviation from the ideal system model: large particles' node responses do not completely return to baseline, pulse heights can vary from pore to pore within a single particle response, and the node-pore transitions are not instantaneous, but smooth. These issues are not completely handled by either of the calibration or the robust regression steps, thus motivating future work in system model design and calibration. Possible more complicated system models that could be incorporated would have more degrees of freedom, which may result in over-fitting the data, thus special consideration must be taken.

4.2 Conclusion

We have demonstrated that by encoding a microfluidic channel with a MB code arrangement of nodes and pores, we have the ability to increase the sensitivity of the device to smaller particles and to provide individual particle coincidence correction. Both are enabled by the joint design of the channel and the sparse deconvolution algorithm. We identify several sources of modeling error and mitigate most effects via a data-driven system model calibration step and robust model regression. We analyze the performance of our method to estimate particle size and transit time as well as the effective dynamic range of particle sizes we can reliably detect in coincidence settings. Finally, we experimentally validate several channel designs that fit in our framework by screening a heterogeneous sample of several size particles through them.

Bibliography

- [1] W. H. Coulter, “Means for Counting Particles Suspended in a Fluid,” pp. 1–9, Jul. 1953.
- [2] M. D. Graham, “The Coulter principle: foundation of an industry,” *Journal of the Association for Laboratory Automation*, vol. 8, no. 6, pp. 72–81, Dec. 2003.
- [3] W. H. Coulter, “High Speed Automatic Blood Cell Counter and Cell Size Analyzer,” pp. 1–11, Oct. 1956.
- [4] G. Brecher, M. Scheiderman, and G. Williams, “Evaluation of Electronic Red Blood Cell Counter,” pp. 1–11, Dec. 1956.
- [5] C. Mattern, F. Brackett, and B. Olson, “Determination of Number and Size of Particles by Electrical Gating: Blood Cells,” pp. 1–15, Jun. 1956.
- [6] D. Yang, S. Leong, A. Lei, and L. L. Sohn, “High-throughput microfluidic device for rare cell isolation,” in *SPIE Microtechnologies*, S. van den Driesche, Ed., SPIE, Jun. 2015.
- [7] M. R. Chapman and L. L. Sohn, *Label-Free Resistive-Pulse Cytometry*, 2nd ed., ser. Methods in Cell Biology. Elsevier Inc., May 2011, vol. 102.
- [8] L. Yang and R. Bashir, “Electrical/electrochemical impedance for rapid detection of foodborne pathogenic bacteria,” *Biotechnology Advances*, vol. 26, no. 2, pp. 135–150, Mar. 2008.
- [9] A. F. Sarioglu, N. Aceto, N. Kojic, M. C. Donaldson, M. Zeinali, B. Hamza, A. Engstrom, H. Zhu, T. K. Sundaresan, D. T. Miyamoto, X. Luo, A. Bardia, B. S. Wittner, S. Ramaswamy, T. Shioda, D. T. Ting, S. L. Stott, R. Kapur, S. Maheswaran, D. A. Haber, and M. Toner, “A microfluidic device for label-free, physical capture of circulating tumor cell clusters,” *Nature Methods*, vol. 12, no. 7, pp. 685–691, May 2015.
- [10] S. H. Au, B. D. Storey, J. C. Moore, Q. Tang, Y.-L. Chen, S. Javaid, A. F. Sarioglu, R. Sullivan, M. W. Madden, R. O Keefe, D. A. Haber, S. Maheswaran, D. M. Langenau, S. L. Stott, and M. Toner, “Clusters of circulating tumor cells traverse capillary-sized vessels,” *Proceedings of the National Academy of Sciences*, vol. 113, no. 18, pp. 4947–4952, May 2016.

- [11] D. Holmes, D. Pettigrew, C. H. Reccius, J. D. Gwyer, C. van Berkel, J. Holloway, D. E. Davies, and H. Morgan, "Leukocyte analysis and differentiation using high speed microfluidic single cell impedance cytometry," *Lab on a Chip*, vol. 9, no. 20, pp. 2881–10, 2009.
- [12] J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, "Characterization of individual polynucleotide molecules using a membrane channel," *Proc. Natl. Acad. Sci.*, vol. 93, pp. 13 770–13 773, Nov. 1996.
- [13] D. Kozak, W. Anderson, R. Vogel, and M. Trau, "Advances in resistive pulse sensors: Devices bridging the void between molecular and microscopic detection," *Nano Today*, vol. 6, no. 5, pp. 531–545, Oct. 2011.
- [14] R. Peng and D. Li, "Detection and sizing of nanoparticles and DNA on PDMS nanofluidic chips based on differential resistive pulse sensing," *Nanoscale*, vol. 9, no. 18, pp. 5964–5974, 2017.
- [15] F. Lan, B. Demaree, N. Ahmed, and A. R. Abate, "Single-Cell Genome Sequencing at Ultra-High-Throughput with Microfluidic Droplet Barcoding," *Nature Publishing Group*, pp. 1–10, May 2017.
- [16] B. M. Venkatesan and R. Bashir, "Nanopore sensors for nucleic acid analysis," *Nature Publishing Group*, vol. 6, no. 10, pp. 615–624, Sep. 2011.
- [17] C. H. Reccius, S. M. Stavis, J. T. Mannion, L. P. Walker, and H. G. Craighead, "Conformation, Length, and Speed Measurements of Electrostatically Stretched DNA in Nanochannels," *Biophysical Journal*, vol. 95, no. 1, pp. 273–286, Jul. 2008.
- [18] L. W. Phipps and F. H. S. Newbould, "Determination of leucocyte concentrations in cow's milk with a Coulter counter," *doi.org*, pp. 1–14, 1966.
- [19] J. W. Mullin and H. M. Ang, "Crystal Size Measurement: Comparison of the Techniques of Sieving and Coulter Counter," *Powder Technology*, pp. 153–156, Feb. 1974.
- [20] M. Wales and J. N. Wilson, "Theory of Coincidence in Coulter Particle Counters," *Review of Scientific Instruments*, vol. 32, no. 10, pp. 1132–1136, Oct. 1961.
- [21] —, "Coincidence in Coulter Counters," *Review of Scientific Instruments*, vol. 33, no. 5, pp. 575–576, May 1962.
- [22] J. C. Samyn and J. P. McGee, "Count Loss with the Coulter Counter," *Journal of Pharmaceutical Science*, vol. 54, no. 12, pp. 1794–1799, Nov. 1965.
- [23] L. H. Princen and W. F. Kwolek, "Coincidence Corrections for Particle Size Determinations with the Coulter Counter," *Review of Scientific Instruments*, vol. 36, no. 5, pp. 646–653, May 1965.
- [24] S. M. Lewis, J. M. England, and F. Kubota, "Coincidence Correction in Red Blood Cell Counting," *Physics in Medicine and Biology*, vol. 35, no. 8, pp. 1159–1162, Mar. 1989.

- [25] K. Kersting, “Specific Problems Using Electronic Particle Counters,” *Hydrobiological Bulletin*, pp. 5–12, 1985.
- [26] “Coulter Principle Short Course,” Tech. Rep., 2014.
- [27] “DxH 500 Hematology Series,” Tech. Rep., 2015.
- [28] *Beckman Coulter Z Series*, 1997.
- [29] J. Zhe, A. V. Jagtiani, P. Dutta, J. Hu, and J. Carletta, “A Micromachined High Throughput Coulter Counter for Bioparticle Detection and Counting,” *Journal of Micromechanics and Microengineering*, vol. 16, no. 8, pp. 1530–1539, Jun. 2006.
- [30] A. V. Jagtiani, J. Carletta, and J. Zhe, “A Microfluidic Multichannel Resistive Pulse Sensor Using Frequency Division Multiplexing for High Throughput Counting of Micro Particles,” *Journal of Micromechanics and Microengineering*, vol. 21, no. 4, pp. 045 036–11, Mar. 2011.
- [31] R. Liu, N. Wang, F. Kamili, and A. F. Sarioglu, “Microfluidic CODES: a scalable multiplexed electronic sensor for orthogonal detection of particles in microfluidic channels,” *Lab on a Chip*, vol. 16, no. 8, pp. 1350–1357, Apr. 2016.
- [32] R. Liu, W. Waheed, N. Wang, O. Civelekoglu, M. Boya, C.-H. Chu, and A. F. Sarioglu, “Design and modeling of electrode networks for code-division multiplexed resistive pulse sensing in microfluidic devices,” *Lab on a Chip*, vol. 17, pp. 2650–2666, Jul. 2017.
- [33] M. Javanmard and R. W. Davis, “Coded Corrugated Microfluidic Sidewalls for Code Division Multiplexing,” *IEEE Sensors Journal*, vol. 13, no. 5, pp. 1399–1400, Mar. 2013.
- [34] D. Polling, S. C. Deane, M. R. Burcher, C. Glasse, and C. H. Reccius, “Coded Electrodes for Low Signal-Noise Ratio Single Cell Detection in Flow-Through Impedance Spectroscopy,” pp. 1–3, Jul. 2010.
- [35] P. Xie, X. Cao, Z. Lin, N. Talukder, S. Emaminejad, and M. Javanmard, “Processing gain and noise in multi-electrode impedance cytometers: Comprehensive electrical design methodology and characterization,” *Sensors & Actuators: B. Chemical*, vol. 241, pp. 672–680, Mar. 2017.
- [36] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [37] K. R. Balakrishnan, J. C. Whang, R. Hwang, J. H. Hack, L. A. Godley, and L. L. Sohn, “Node-Pore Sensing Enables Label-Free Surface-Marker Profiling of Single Cells,” *Analytical Chemistry*, vol. 87, no. 5, pp. 2988–2995, Mar. 2015.
- [38] K. R. Balakrishnan, G. Anwar, M. R. Chapman, T. Nguyen, A. Kesavaraju, and L. L. Sohn, “Node-pore sensing: a robust, high-dynamic range method for detecting biological species,” *Lab on a Chip*, vol. 13, no. 7, pp. 1302–6, 2013.
- [39] R. H. Barker, “Group synchronizing of binary digital systems,” *Communication Theory*, pp. 273–287,

- [40] R. W. Deblois, C. P. Bean, and R. K. A. Wesley, "Electrokinetic measurements with submicron particles and pores by the resistive pulse technique," *Journal of Colloid and Interface Science*, vol. 61, no. 2, pp. 323–335, Sep. 1977.
- [41] F. D. Ohrtman and K. Roeder, *Wi-Fi Handbook: Building 802.11b Wireless Networks*, 1st ed. New York: McGraw-Hill Inc, 2003.
- [42] M. A. Richards, *Fundamentals of radar signal processing*. Tata McGraw-Hill Education, 2005.
- [43] N. Levanon, "Noncoherent Pulse Compression," *IEEE Transactions on Aerospace and Electronic Systems*, pp. 756–765, Apr. 2006.
- [44] W. Stallings, *Data and Computer Communications*, 8th ed. Pearson Education, 2007.
- [45] J. A. Tropp and A. C. Gilbert, "Signal Recovery From Random Measurements Via Orthogonal Matching Pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, Dec. 2007.
- [46] S. Blake, "OS-CFAR Theory for Multiple Targets and Nonuniform Clutter," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 24, no. 6, pp. 785–790, Nov. 1987.
- [47] O. A. Saleh and L. L. Sohn, "Correcting off-axis effects in an on-chip resistive-pulse analyzer," *Review of Scientific Instruments*, vol. 73, no. 12, pp. 4396–4398, Dec. 2002.
- [48] R. Andersen, *Modern Methods for Robust Regression*. Sage, 2008.
- [49] I. Daubechies, R. Devore, M. Fornasier, and C. S. Gunturk, "Iteratively Reweighted Least Squares Minimization for Sparse Recovery," *Pure and Applied Mathematics*, pp. 1–38, Oct. 2009.
- [50] J. Jedwab, "What can be used instead of a Barker sequence?" *Contemporary Mathematics*, pp. 1–26, Feb. 2008.
- [51] S. W. Golomb and G. Gong, *Signal Design for Good Correlation*, ser. For Wireless Communication, Cryptography, and Radar. Cambridge University Press, 2005.
- [52] R. Gold, "Optimal Binary Sequences of Spread Spectrum Multiplexing," *IEEE Transactions on Information Theory*, vol. 13, pp. 619–621, 1967.

Appendix A

Code

Software for demonstration and reproducibility purposes are associated with this digital object identifier:

`10.5281/zenodo.861990`

Some simulation and experimental results have a run-time that is on the order of hours. Thus, preprocessed results have been included in the compressed folder labeled Results.

Datasets for demonstration and reproducibility purposes are associated with this digital object identifier:

`10.5281/zenodo.860191`

Appendix B

Adaptive Detection Scheme

Adaptive detection has been used for a wide range of applications in radar and is neatly explained in the Detection Processing chapter of the book Fundamentals of Radar Signal Processing [42]. Classic detection methods require the noise and interference statistics of the signal to be known to select a detection threshold that guarantees a specific false alarm rate and mis-detection rate. If the noise and interference statistics are unknown, then they can be estimated from the signal to appropriately select the detection threshold. In addition if the noise and interference statistics are time-varying then the estimate of the statistics must be re-estimated often.

A common method to dynamically estimate the noise and interference statistics is cell-averaging adaptive detection, where a range of samples around the detection being evaluated are averaged to estimate the noise and interference energy. The energy is then scaled, α , to determine the detection threshold for a specific false alarm rate. This method makes the assumption that detections are isolated, however in coincidence settings this is not a valid assumption, motivating an adaptive detection scheme that is robust to multiple detections. Among several, Ordered Statistic (OS) adaptive detection is a method for robustly estimating the surrounding signal energy when many detections may be present. The OS adaptive detection [46] scheme takes the samples from the previously specified range around the detection being evaluated, sorts them in ascending order, and takes the k^{th} greatest value as the noise and interference energy level around the detection. Similar to before, the energy estimate is then scaled, α , to determine the detection threshold associated with a specific false alarm rate.

Our detection scheme implements the OS adaptive detection where we use the median, $k = \frac{N}{2}$ of the surrounding samples to be a robust estimate of the surrounding energy. We then tuned the α parameter to yield low false alarm rates over the dynamic range of particle sizes we wish to detect. In practice the false alarm detections in the lower SNR regime are considered for appropriate α selection.

Appendix C

Iterative Re-weighted Least-Squares

In our implementation, the robust regression performed to fit the data to the model is solved via iterative re-weighted least-squares (IRLS) [49]. Specifically, by iterating between solving a weighted least-squares problem (equation C.1) and update step where the weighting term, W , is recomputed (equation C.2)

$$\min_{\tilde{x}, b} \|W(\tilde{A}\tilde{x} + b - y)\|_2 + \lambda \|Db\|_2 \quad (\text{C.1})$$

$$\text{where, } W_{(j,j)} = \frac{1}{|\tilde{A}_{(j)}\tilde{x} + b_{(j)} - y_{(j)}|} \quad (\text{C.2})$$

Appendix D

Channel and Electrode Fabrication

We use standard soft lithography to create a negative-relief master of our MB encoded microfluidic device using SU-8 resist (MicroChem, Santa Clara, CA, USA) on a polished silicon wafer. We then cast a polydimethylsiloxane (PDMS) mold with the channel embedded.

To make the negative-relief master on the silicon wafer, we first treat the substrate with hexamethyldisilazane (HMDS, Spi-Chem, West Chester, PA, USA) and then spin coat SU-8 photoresist. For a final height of $20\mu\text{m}$, we use SU8 3025 at 4500 rpm for 30 seconds, with a 5 seconds ramp at 500 rpm. We then place the wafer on a hot plate at 95°C for a 10 minute soft bake and expose it with a mylar mask, that has our device design, to UV light (22 seconds at $9 \frac{\text{mW}}{\text{cm}^2}$). After UV exposure, we place back the wafer on the hotplate for a post exposure bake (1 minute at 65°C and 10 minutes at 95°C). We use SU-8 developer to develop and obtain a negative-relief master of our desired design and measure the height of the features using a profilometer.

The electrodes, to perform a four-point measurement, are patterned onto a glass substrate (VWR, Radnor, PA, USA). We use photolithography to create a negative pattern the electrodes on the glass substrate. First, we spin Shipley S1813 resist (Fisher Scientific, Pittsburgh, PA, USA) at 3000 rpm for 30 seconds, bake it on a hot plate at 100°C for 1 minute and then expose the slides with a specific mask under UV light in the same mask aligner (25 seconds at $9 \frac{\text{mW}}{\text{cm}^2}$). We develop the photoresist in MF 321 and thoroughly rinse it with deionized (DI) water. We perform a three layer deposition, titanium (100) platinum (250) gold (250) with an electron-gun evaporator, and then an acetone lift-off to remove excess metal and photoresist. The last step is etching the gold layer on the primary part of the electrodes. We use a gold etchant (Transene Company, Danvers, MA, USA) for 30s to expose the platinum layer. Electrodes are then cleaned in a 1:1:10 RCA solution, rinsed with DI water, and stored in methanol until used in order to ensure they stay clean.