EXHIBIT A

Interview with Examiners
Ex parte Reexamination of U.S. Patent 6,201,839

August 22, 2014

Key Points

- **Zeng's Sections 4.4 and 5.2 Do Not Have Any, Let Alone a "Set," of Signal-Dependent Branch Metric Functions.**
  - Zeng assumed away the signal-dependent noise problem solved by the CMU inventors
  - Zeng's functions in Sections 4.4 and 5.2 are not signal-dependent
  - Dr. Lee, the Requester's expert, admitted this in 1992 and again in 1995
  - Zeng's own thesis advisor said the CMU inventors were first

- **Litigation Confirms that Zeng's Thesis Does Not Invalidate.**
  - Marvell originally asserted that Zeng's Thesis anticipated claim 4, but Marvell dropped (and now has waived) any and all invalidity claims based on the thesis
  - Marvell's expert reviewed Zeng's Thesis but did not say it invalidated claim 4
  - Zeng is an employee of Marvell, but Marvell never called him as a witness
  - Marvell lost $1.5 billion judgment after four-week trial

Agenda

- **Roadmap to CMU's Response**

- **The Claimed Invention**
  - Invention requires "a set of signal-dependent branch metric functions" that are applied to "a plurality of signal samples" from different time instances to compute branch metric values for branches of a trellis.
  - Solved long-felt need in HDD industry -- Media noise increases with data density and prior art detectors could not adequately account for it.

- **Zeng's Thesis**
  - Does not disclose a "set of signal-dependent branch metric functions" that are applied to "a plurality of signal samples."
  - Not enabling

- **The *CMU v. Marvell* litigation**
  - Evidence that Marvell requested this reexamination
  - Marvell's treatment (or lack thereof) of Zeng's Thesis in the litigation
  - Evidence of secondary considerations

Roadmap to CMU's Response

- **Inventor Declarations**
  - Dr. Jose Moura – Professor at Carnegie Mellon University
    - Describes problem inventors sought to solve and how they solved it
  - Dr. Aleksandar Kavcic – Professor at University of Hawaii
    - Shows why Zeng's channel model, simulation results, and detector equations are fatally flawed
- **Declaration of Dr. Christopher Bajorek**
  - Managed IBM's HDD business unit and Komag's HDD disk business
    - Instrumental in developing technology that increased data density, leading to the need for the Kavcic-Moura invention
  - Describes the nature and structure of noise in HDDs, including signal-dependent noise
  - Shows why Zeng's "random jitter" is not signal-dependent noise
- **Declaration of Dr. Steven McLaughlin**
  - Chair of School of Electrical and Computer Engineering at Georgia Tech
  - Describes Viterbi detection, including branch metric functions
  - Shows why Zeng's BMFs in Sections 4.4 and 5.2 are not signal-dependent, let alone a "set" of such functions
- **Underlying Record in *CMU v. Marvell***

Background

Signal-Dependent Branch Metric Functions

- **Key issue is whether Zeng's functions in Section 4.4 and 5.2 are signal-dependent**
  - Signal-dependent branch metric functions account for the signal-dependent structure of the media noise
  - Signal-dependent noise (SDN) has ***structure*** that is attributable to a ***specific*** sequence of symbols written to the disk

- **Requester says Zeng's functions in Sections 4.4 and 5.2 account for SDN because of Zeng's "random jitter" (Δ_k or its variance σ_Δ^2). That is wrong.**
  - Zeng's "random jitter" is not signal-dependent noise
  - Lee publicly confirmed this in 1992 (and again in 1995)

- **To demonstrate this we need to discuss:**
  - How data are written to and read from a hard disk
  - Signal-dependent noise
  - Viterbi detectors and branch metric functions
  - The invention of claim 4
  - Why Zeng's "random jitter" is not signal-dependent

Magnetic Hard Disk Drive

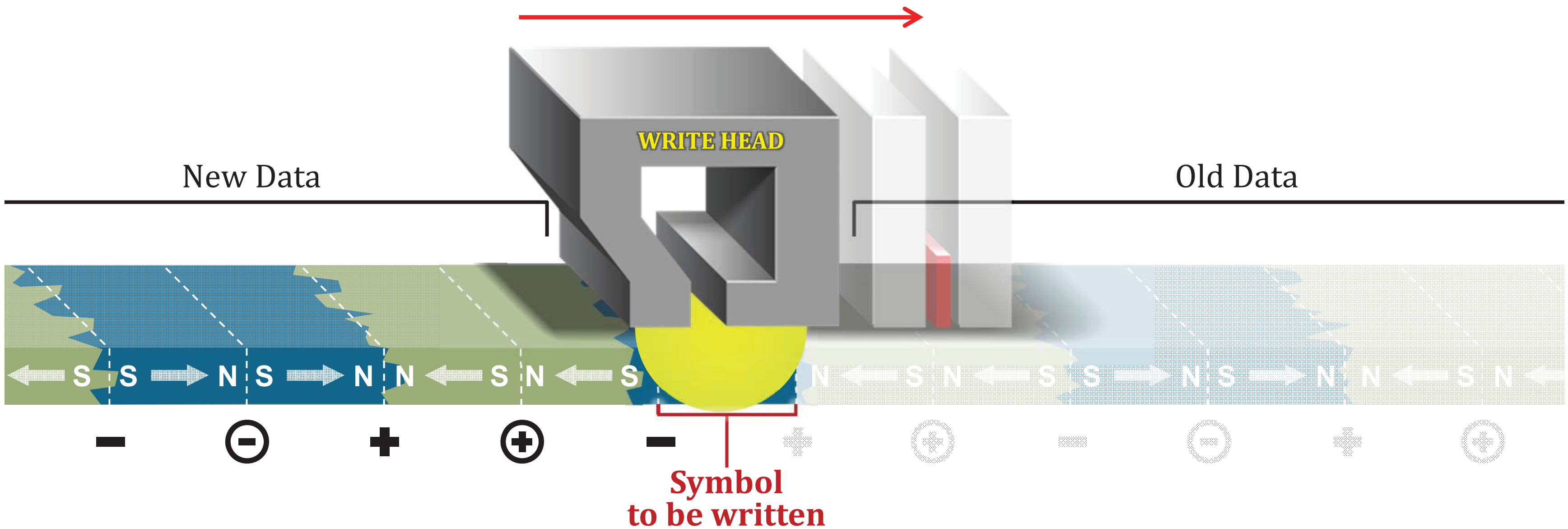
Platter/Disk
*Stores **data** in millions of concentric **tracks***

One Track
*Contains tens of millions of symbol **regions***

Read/Write Head
- *Writes digital data as **magnetic symbol regions***
- *Reads **magnetic symbol regions** from the **tracks***

Writing Data to the Hard Disk

The WRITE HEAD generates an alternating magnetic field
to set the polarity of each magnetic symbol region

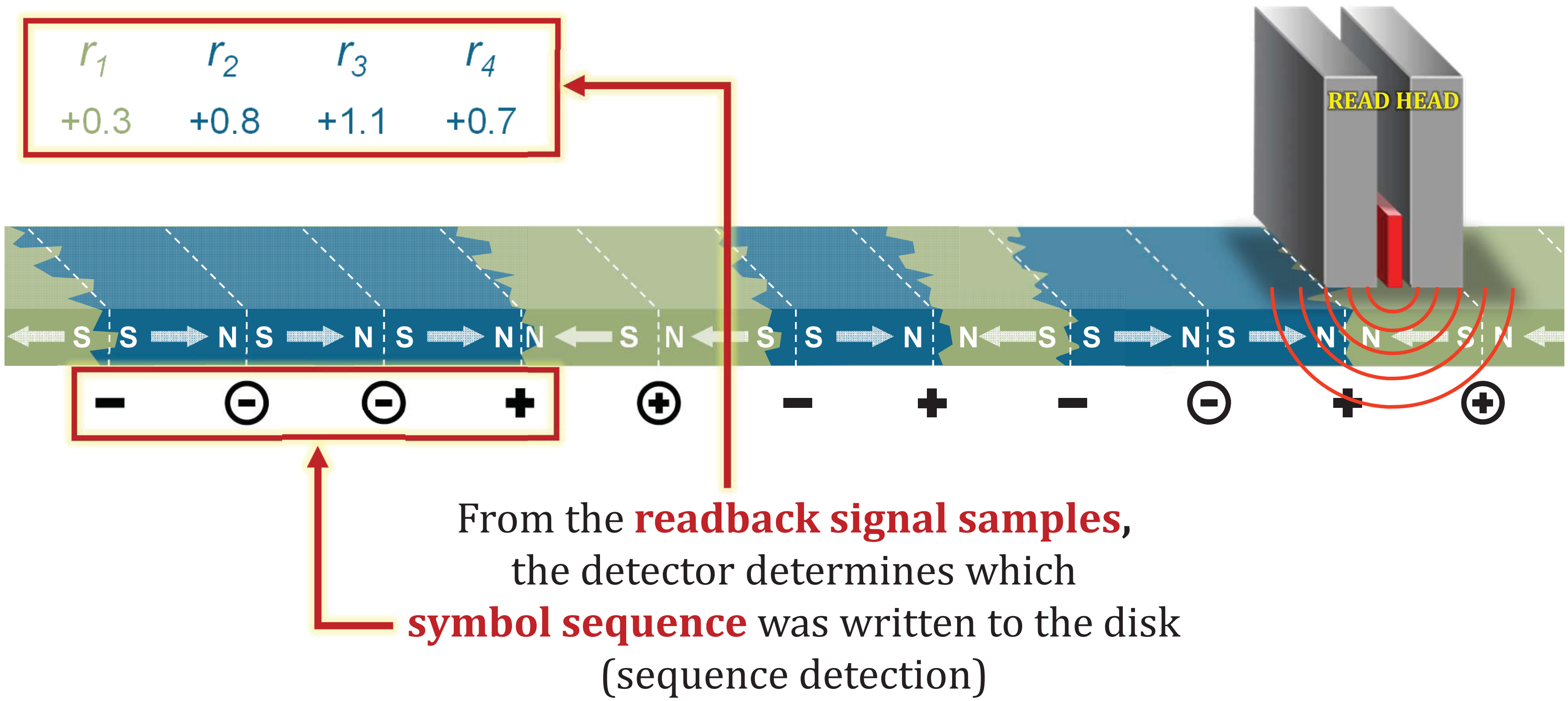


+	Indicates a positive transition	S → N N ← S
−	Indicates a negative transition	N ← S S → N
⊕	Indicates a non-transition whose nearest preceding transition was a positive transition	N ← S N ← S
⊖	Indicates a non-transition whose nearest preceding transition was a negative transition	S → N S → N

Source: '839 col. 3:49-64; Fig. 4

Reading Data from the Hard Disk

Read head generates a readback signal based on magnetic fields from the symbols written to the disk.

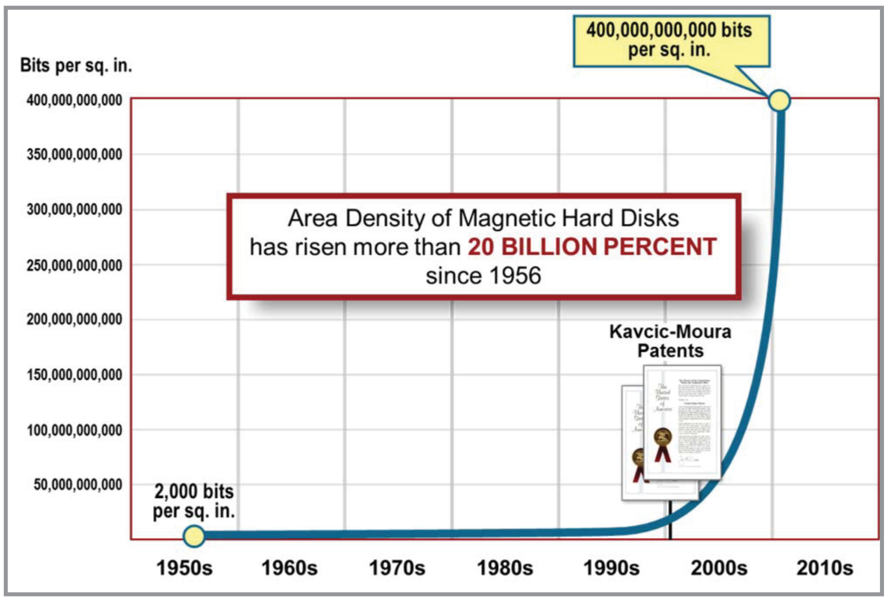


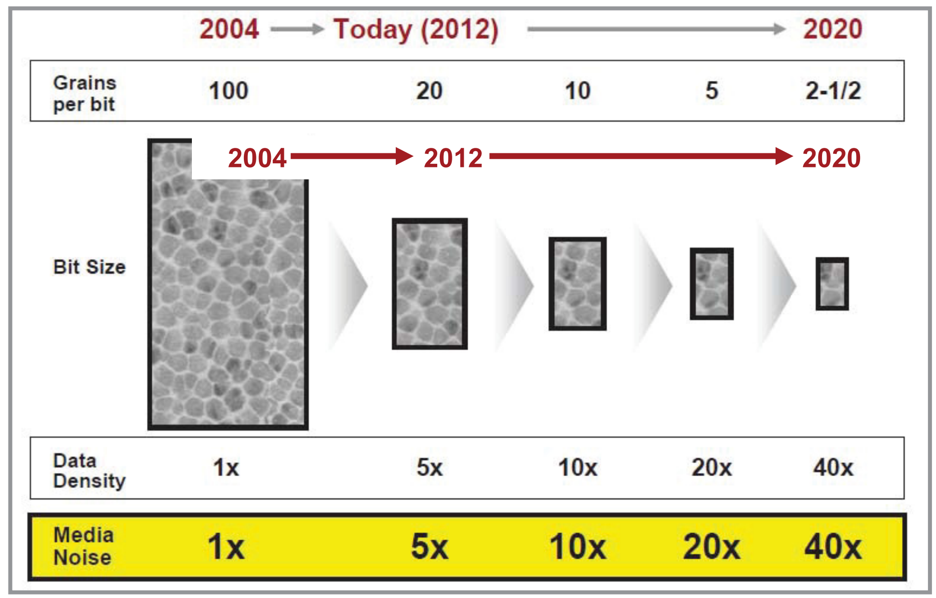
From the **readback signal samples**, the detector determines which **symbol sequence** was written to the disk (sequence detection)

Complications from Media Noise

The Problem

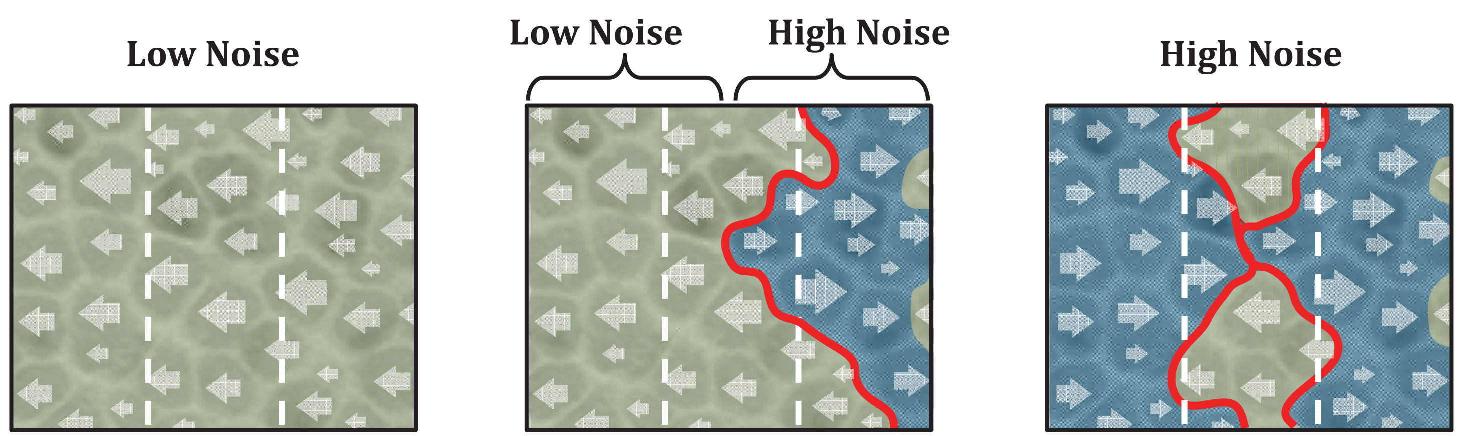
- Data density has risen more than 20 billion percent since 1956
- Media noise, which is correlated and signal-dependent, dominates as data density increases
- Prior art detectors could not account for this correlated, signal-dependent media noise, which limited how densely data could be written





Complications from Media Noise

- **Media noise is "signal-dependent"**
- **Has noise structure that is attributable to a specific sequence of symbols written to the disk**
  - Each specific sequence of symbols has its own noise structure

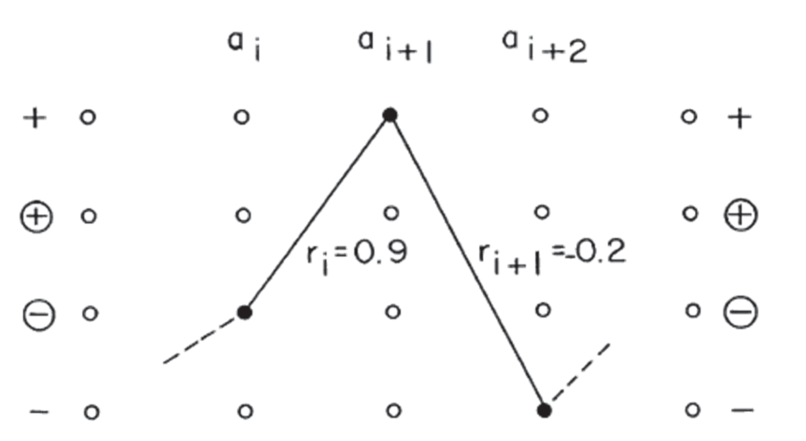


- **More details later when demonstrating Zeng's failure to address signal-dependent noise.**

Viterbi Detectors and Branch Metric Functions

- **A Viterbi detector determines the most likely sequence of symbols written to the disk based on the signal samples by:**
  - Computing branch metric values for branches of a trellis
  - Summing the branch metric values for branches along a path to get path metric values
    - The paths correspond to the possible symbol sequences
  - The path with lowest path metric value corresponds to best estimate of symbols written to disk

FIG. 5



Example Prior Art Branch Metric Function

$$M_i = N_i^2 = (r_i - m_i)^2$$

- Euclidean branch metric function
- Eq. 8 of '839 Patent
- ***Assumes noise is white***

The Invention

The Invention

’839 Patent, Claim 4

4. A method of determining branch metric values for branches of a trellis for a Viterbi-like detector, comprising:

selecting a branch metric function for each of the branches at a certain time index from a set* of signal-dependent branch metric functions; and

applying each of said selected functions to a plurality of signal samples to determine the metric value corresponding to the branch for which the applied branch metric function was selected, wherein each sample corresponds to a different sampling time instant.

*** Both the Requester and Lee acknowledge that the claimed “set” of signal-dependent branch metric functions requires at least two such functions.**
See Lee Dec. at ¶ 13; Request at p. 20.

NOVELTY!

- At least two different signal-dependent branch metric functions
- Each function operates on a plurality of signal samples
- For each branch (for which a branch metric value is computed), the corresponding function is selected

Important Considerations that Inventors Addressed

- **Signal-dependent functions must be neighborhood-sensitive**

$$= \arg\left[\min_{all\ a_i} \sum_{i=1}^{N} M_i(r_i, r_{i+1}, \ldots, r_{i+L}, a_{i-K_l}, \ldots, a_{i+L+K_t})\right]$$

M_i represents the branch metric of the trellis/tree in the Viterbi-like algorithm. The metric is a function of the observed samples $r_i, r_{i+1}, \ldots, r_{i+L}$. It is also dependent on the postulated sequence of written symbols $a_i - K_1, \ldots, a_i + L + K_t$, which ensures the signal-dependence of the detector. As a consequence, the branch metrics for every branch in the tree/trellis is based on its corresponding signal/noise statistics.

'839 patent at col. 5:49-52

- **Signal-dependent functions must be polarity-sensitive**

The symbols '+' and '–' denote a positive and a negative transition, respectively. The symbol '⊕' denotes a written zero (no transition) whose nearest preceding non-zero symbol is a '+' while '⊖' denotes a written zero whose nearest preceding transition is a negative one, i.e., '–'. This notation is used because a simple treatment of transitions as '1's and no transitions as '0's is blind to signal asymmetries (MR head asymmetries and base line drifts), which is inappropriate for the present problem.

'839 patent at col. 3:56-64

The Solution – Claim 4

- **Equation 13 – a "set" of signal-dependent branch metric functions, because the covariance matrix C_i is different for different branches**

$$M_i = \log\frac{\det\ C_i}{\det\ c_i} + \underline{N}_i^T C_i^{-1} \underline{N}_i - \underline{n}_i^T c_i^{-1} \underline{n}_i$$

- **FIR Filter Embodiment – a "set" because there is a separate filter (i.e., a separate function), with its own tap weights, for each branch**

FIG. 3A

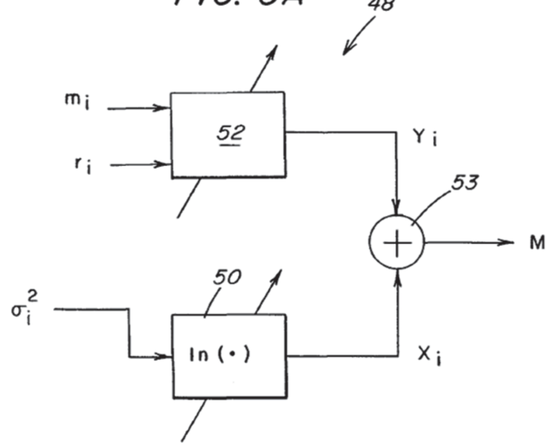
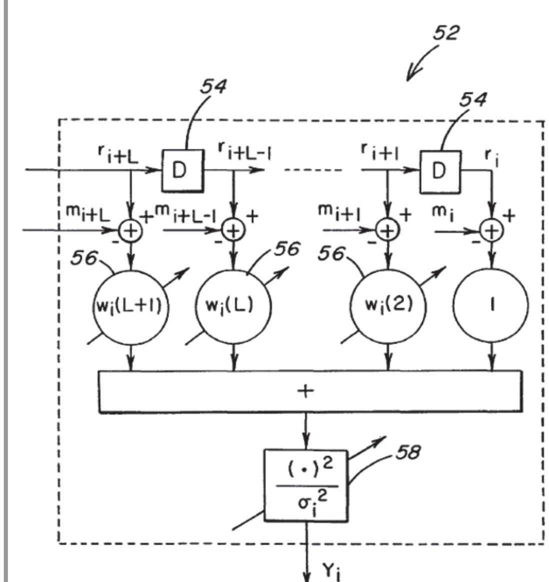


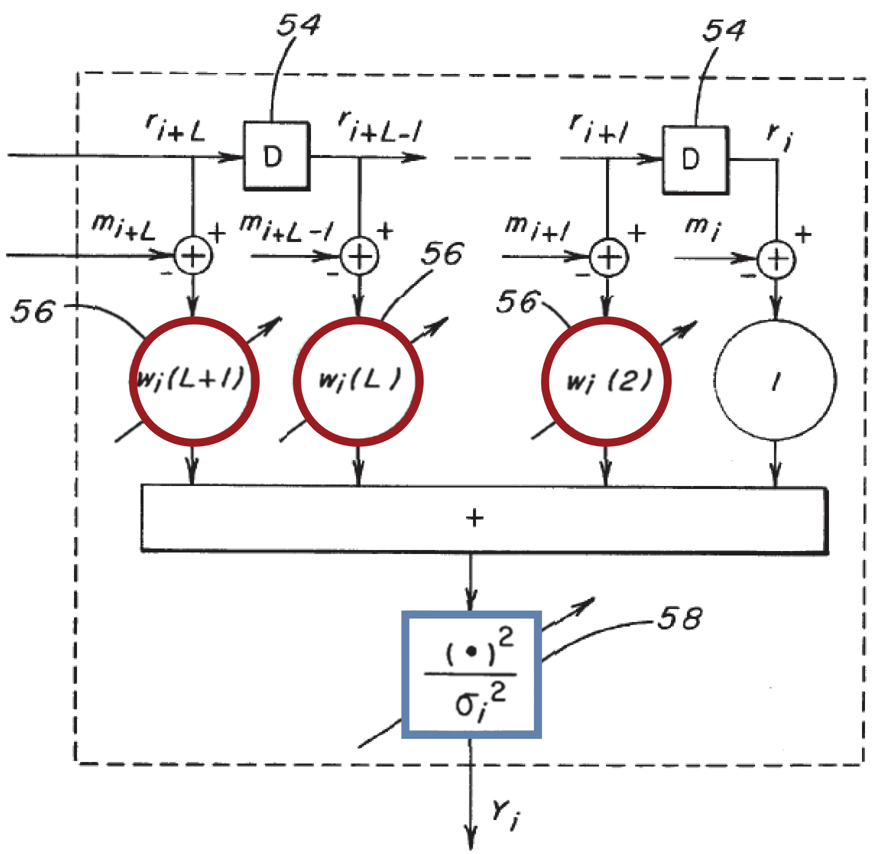
FIG. 3B

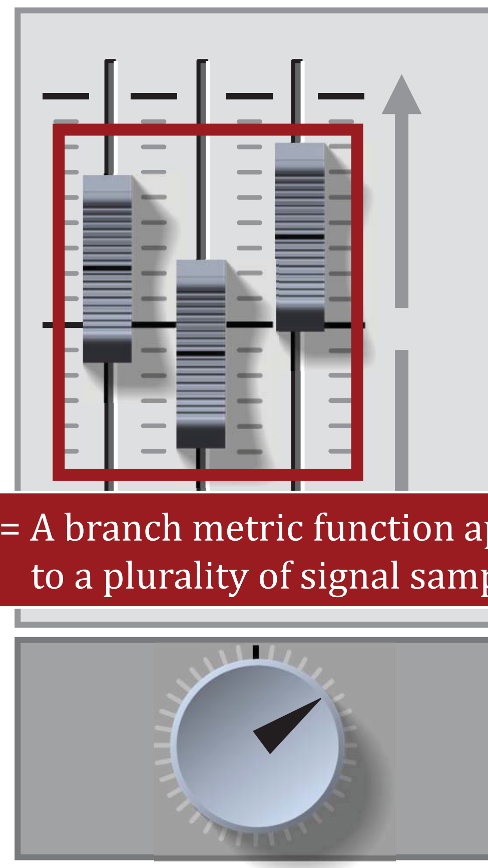


The Invention

- **Apply a weight (w_i) to signal samples of different time instances**
- **Each branch has its own FIR filter (and its own tap weights)**

'839 Patent, Fig. 3B

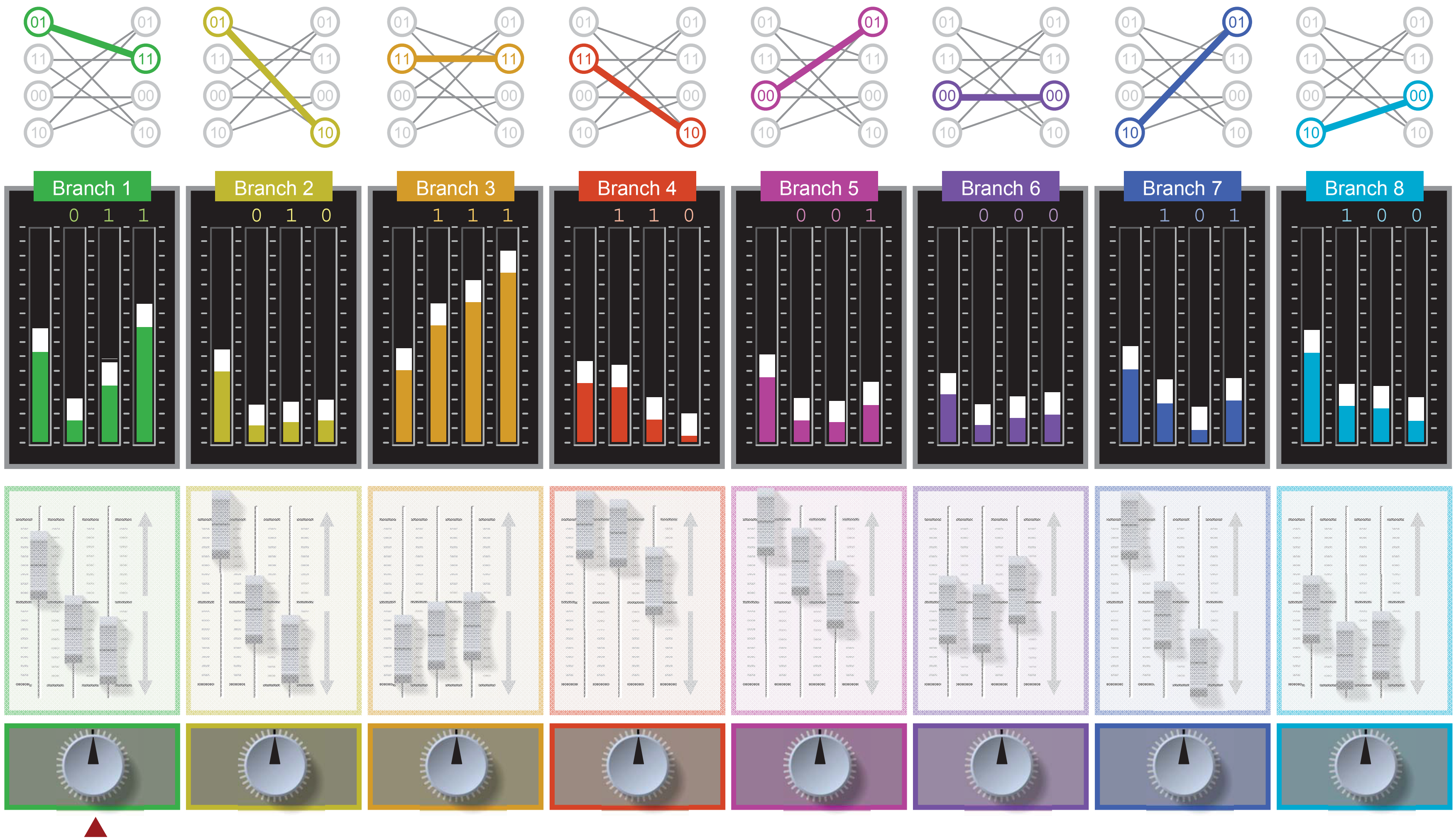




Sliders = A branch metric function applied to a plurality of signal samples

The Invention

- A branch metric function applied to a plurality of signal samples



Zeng's Thesis

Zeng's Random Jitter is White, Not Signal-Dependent

- **Zeng's Channel Model – equation 4.1**

aka, random peak shift

Jitter noise is modeled as a random shift in the transition position. The overall channel response with jitter can be written as

$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t), \qquad (4.1)$$

Zeng's Thesis, p. 52

"random jitter" term

- **Zeng's "random jitter" is "independent and identically distributed random variables with zero mean" - in other words, "white"**

Sampling $Z(t)$ at $t = kT$ and keeping only Δ_k and Δ_{k-1} terms (the most important terms) in Equation (4.3), we obtain the simplified discrete-time channel model with both jitter noise and additive white noise

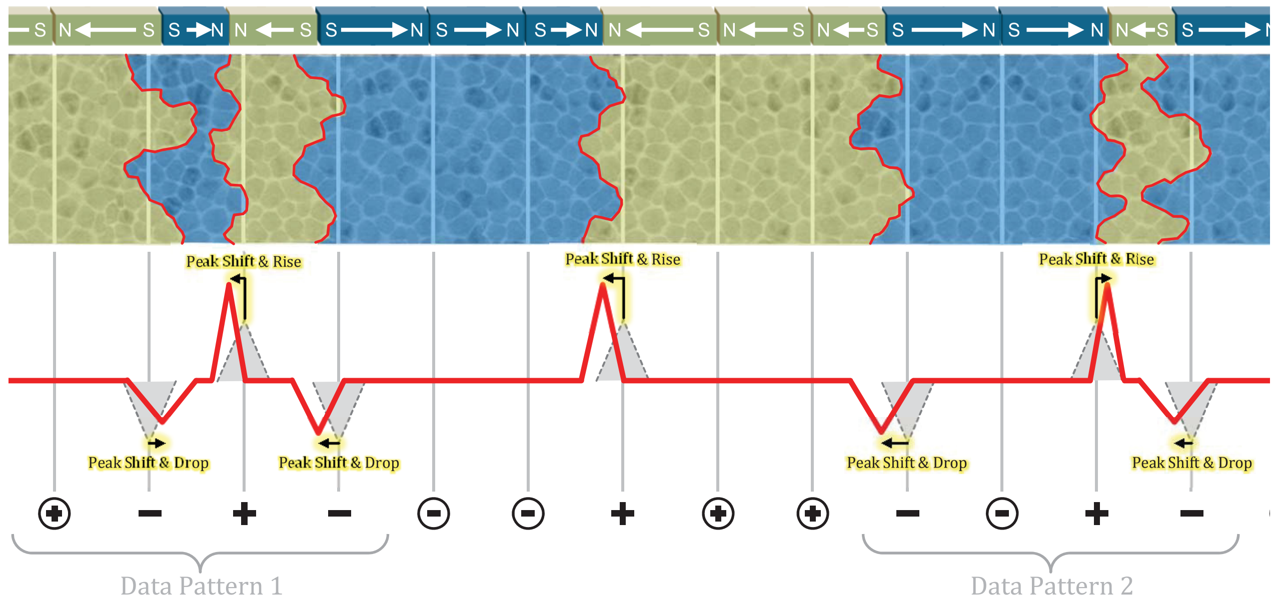
$$\begin{aligned} Z_k &= y_k + w_k - a_k\Delta_k + a_{k-1}\Delta_{k-1} && (4.5) \\ &= y_k + n_k, \qquad k = 1, 2, \ldots, N && (4.6) \end{aligned}$$

where $n_k = w_k - a_k\Delta_k + a_{k-1}\Delta_{k-1}$ and $y_k = b_k - b_{k-2}$, and Δ_k is expressed as a fraction of T. We assume that Δ_k's are independently and identically distributed random variables with zero mean and that jitter noise is independent of additive noise. However, the overall noise term n_k may still be correlated with n_{k-1} or n_{k+1}. n_k is a Gaussian random variable if we assume both jitter and white noise are Gaussian.

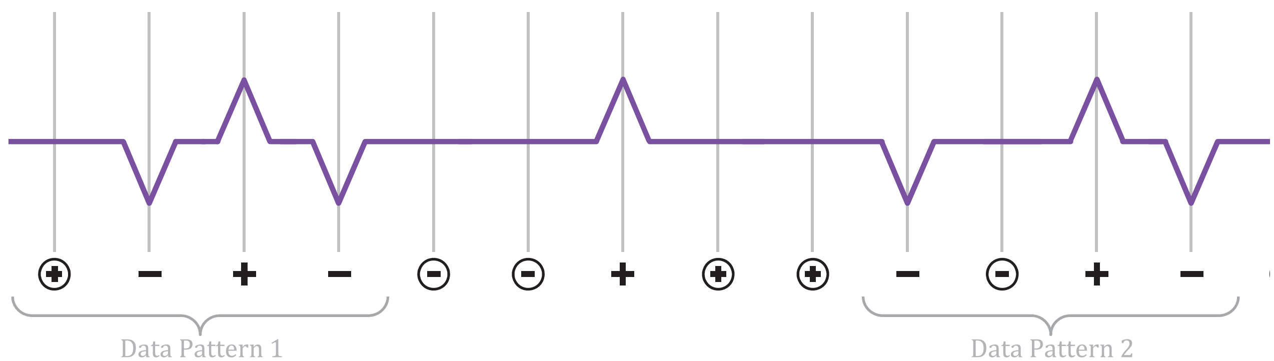
Zeng's Thesis, p. 52

Signal dependent noise which includes noises from polarity and neighborhood effects.

Polarity Dependent



ng's hypothetical
ter noise model.
ot signal dependent)

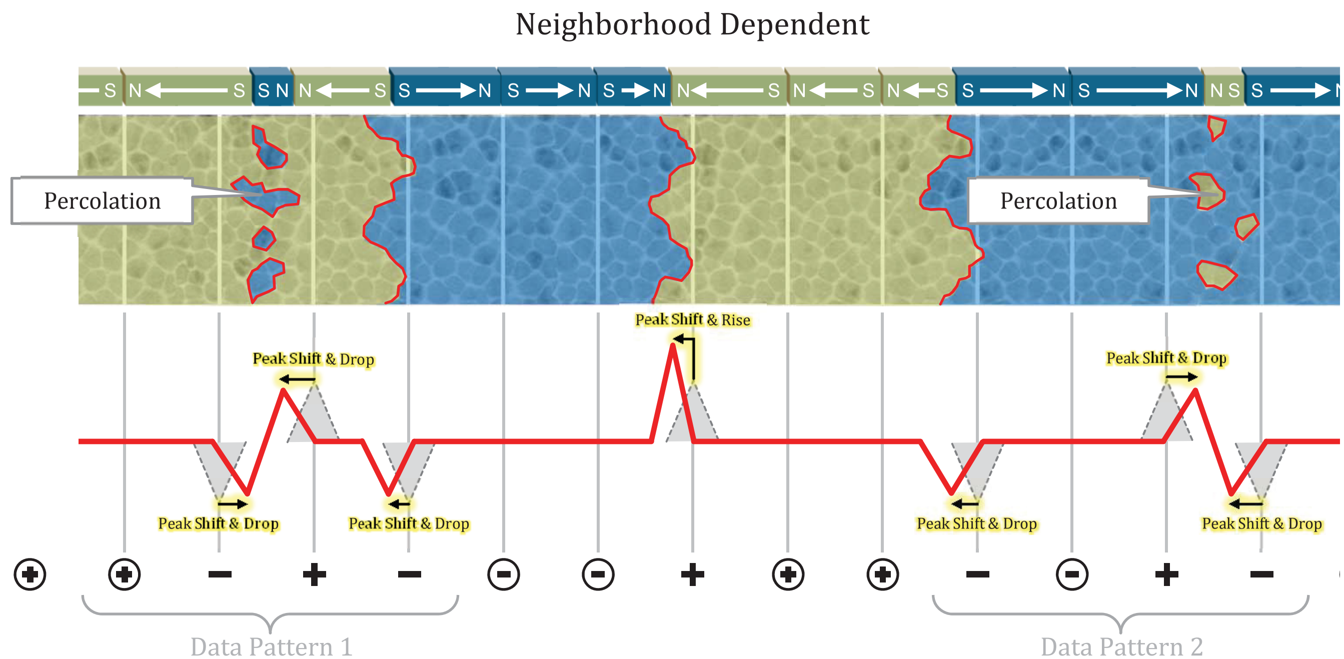


$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

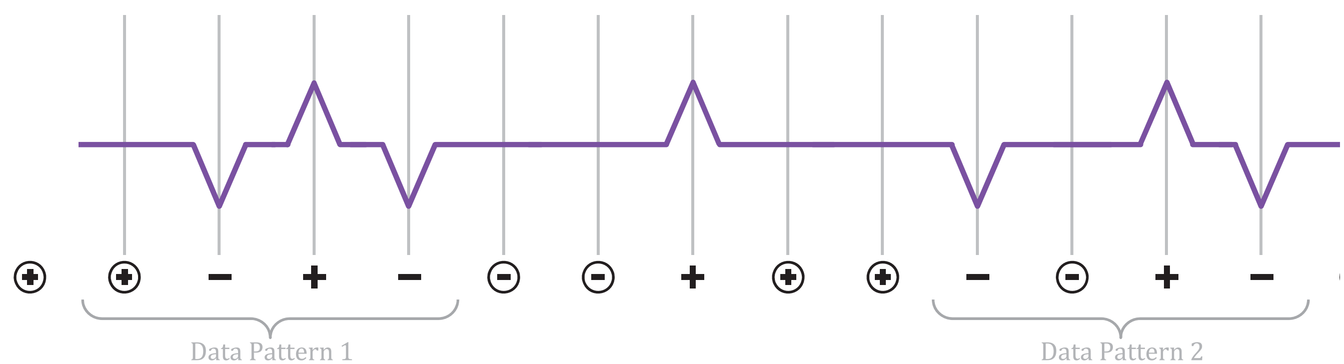
To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.



Zeng's hypothetical jitter noise model. (Not signal dependent)



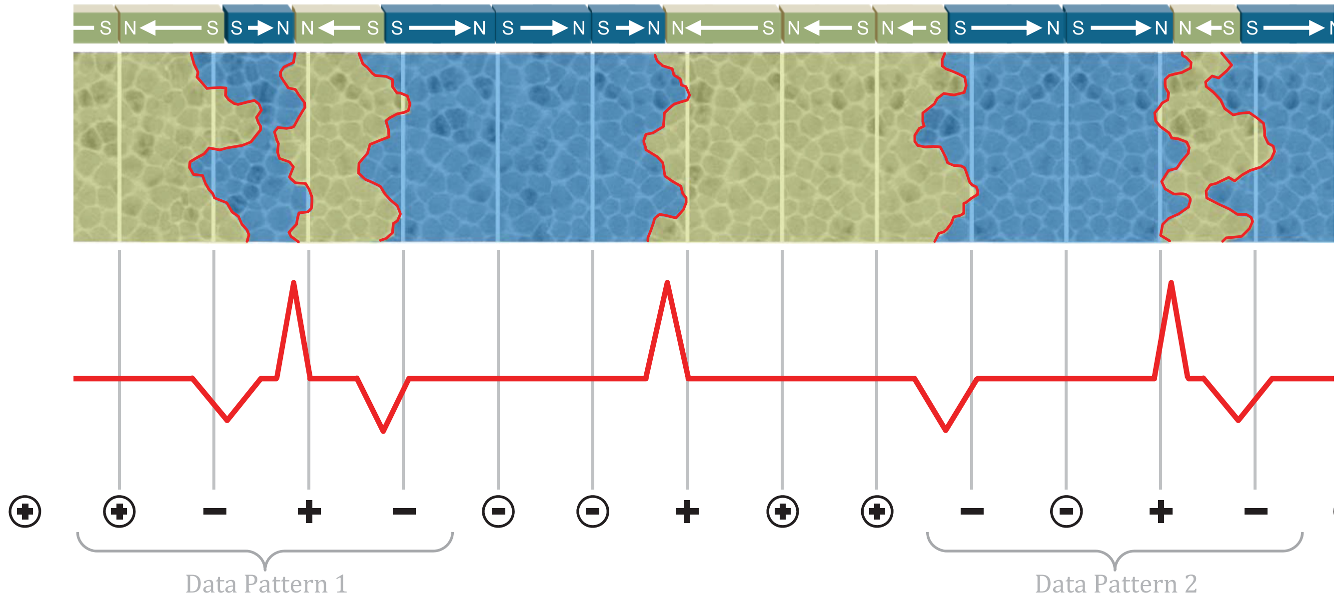
$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

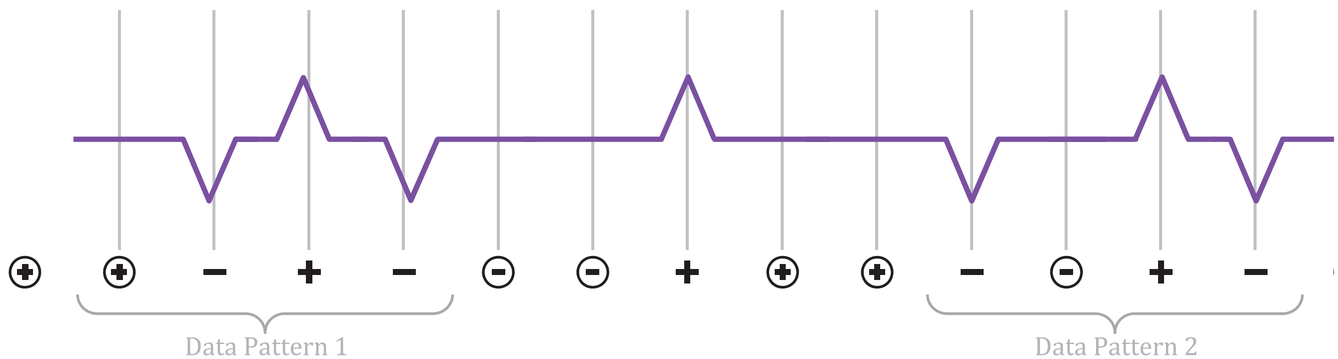
Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.

Neighborhood Dependent



Zeng's hypothetical jitter noise model.
(Not signal dependent)

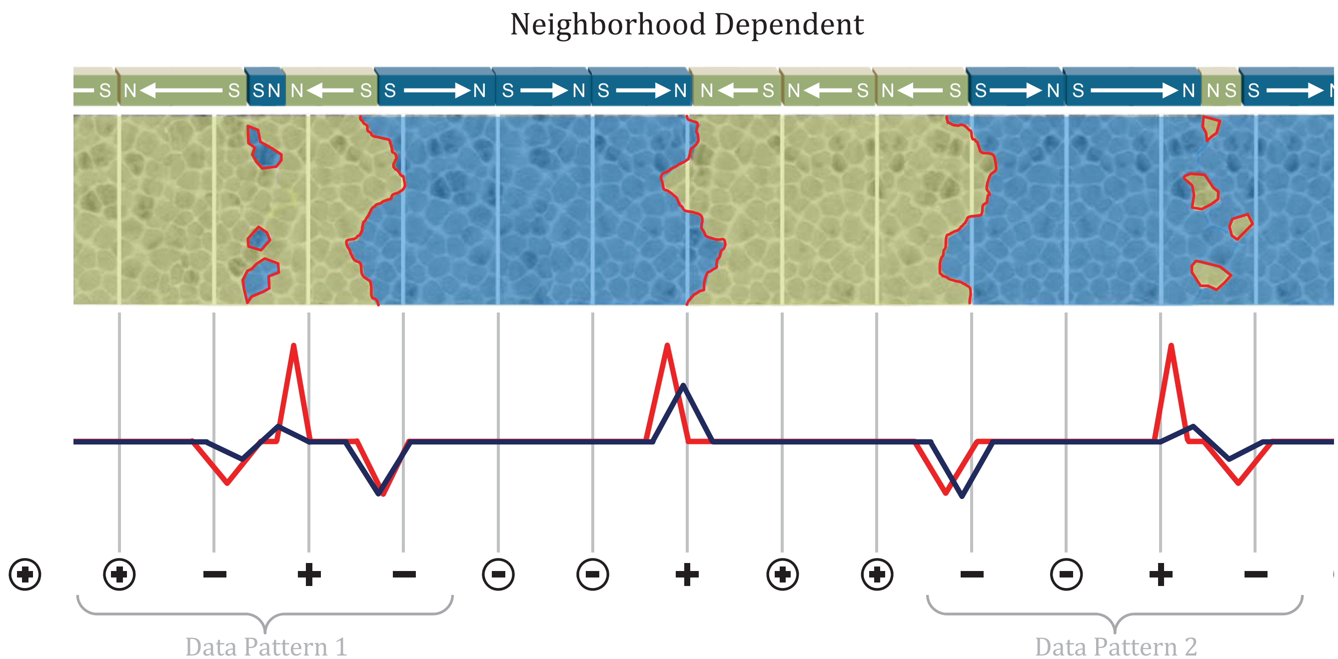


$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

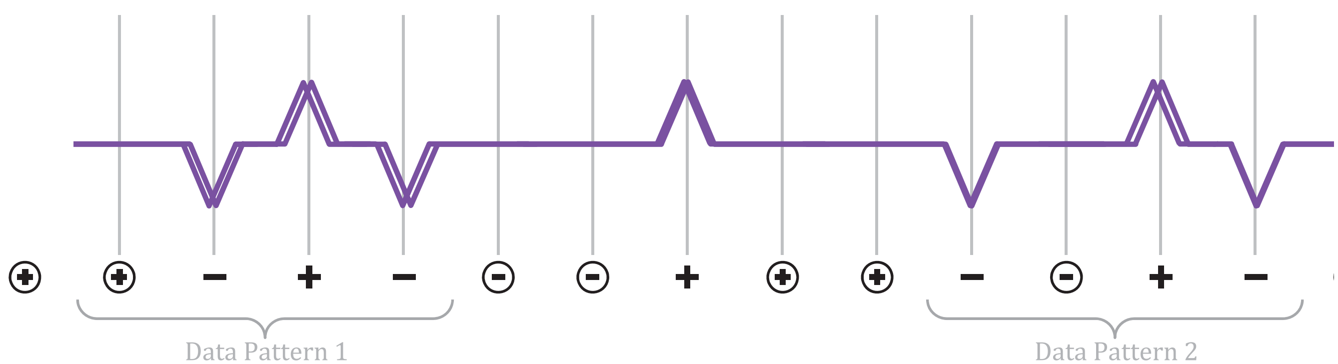
To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.



Zeng's hypothetical jitter noise model.
(Not signal dependent)

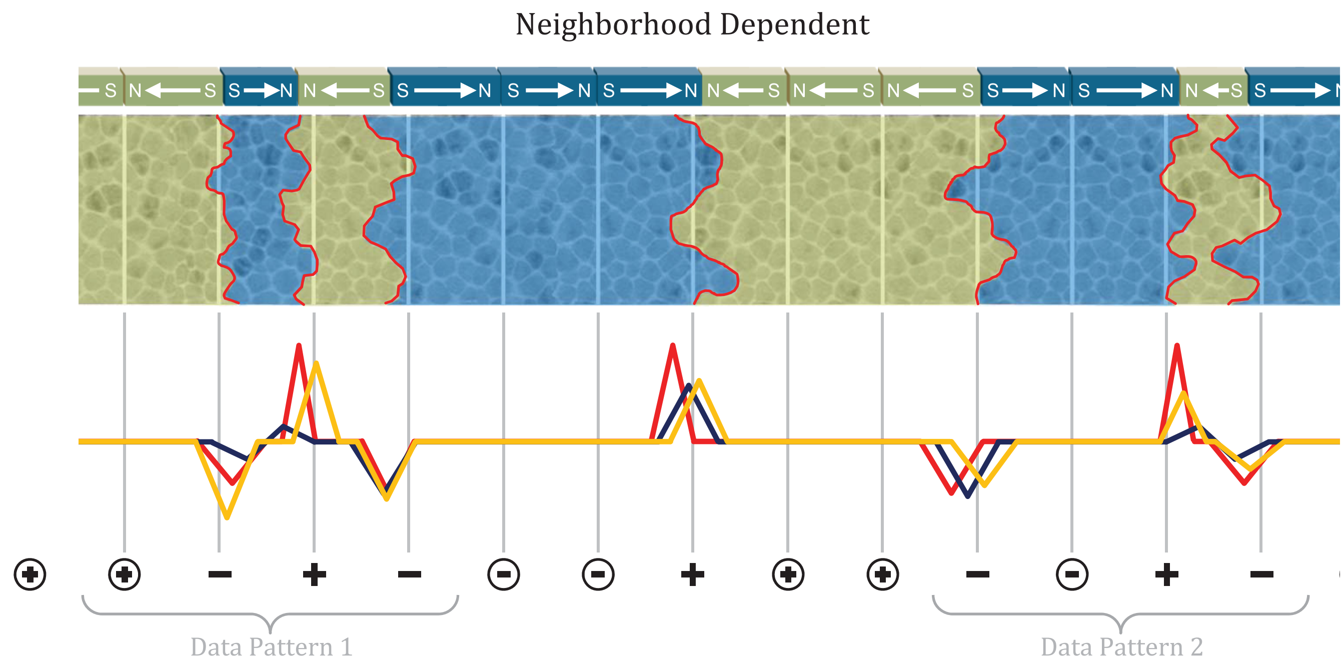


$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

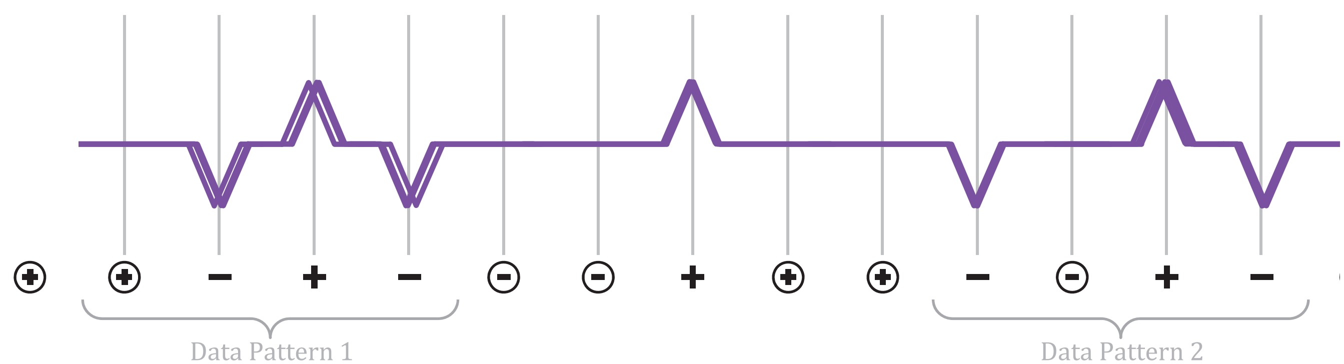
To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.



Zeng's hypothetical jitter noise model.
(Not signal dependent)



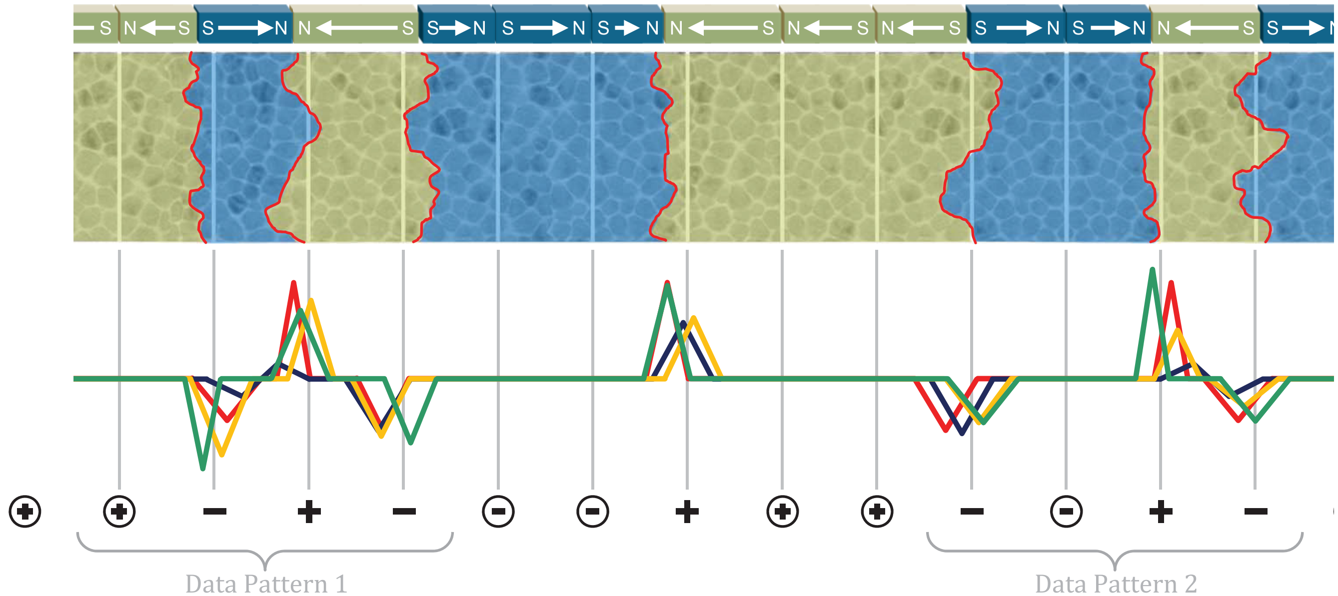
$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

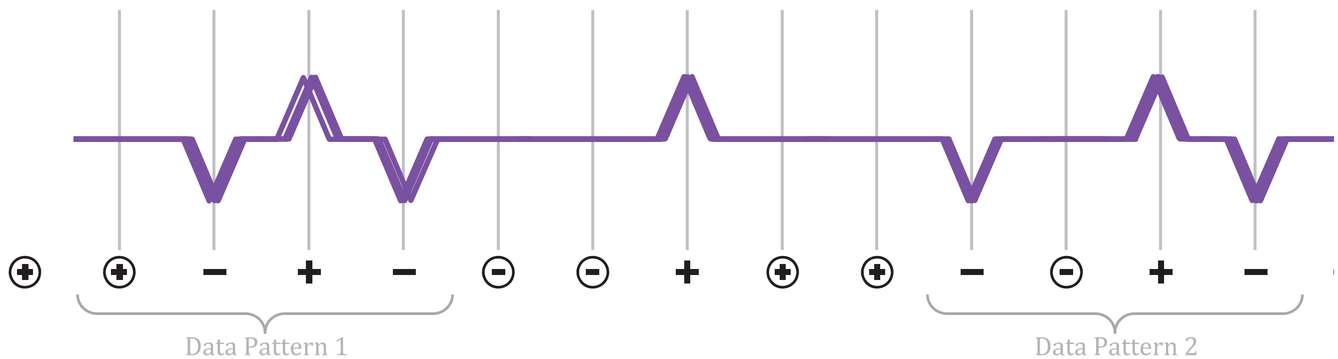
Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.

Neighborhood Dependent



Zeng's hypothetical jitter noise model.
(Not signal dependent)



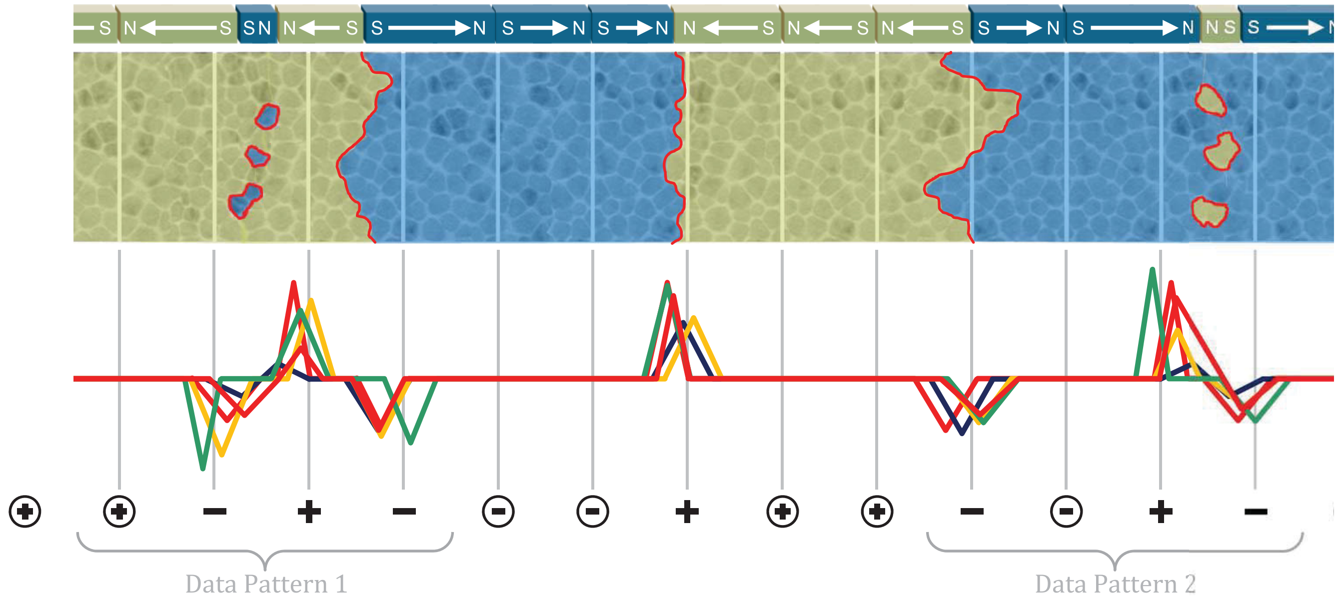
$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

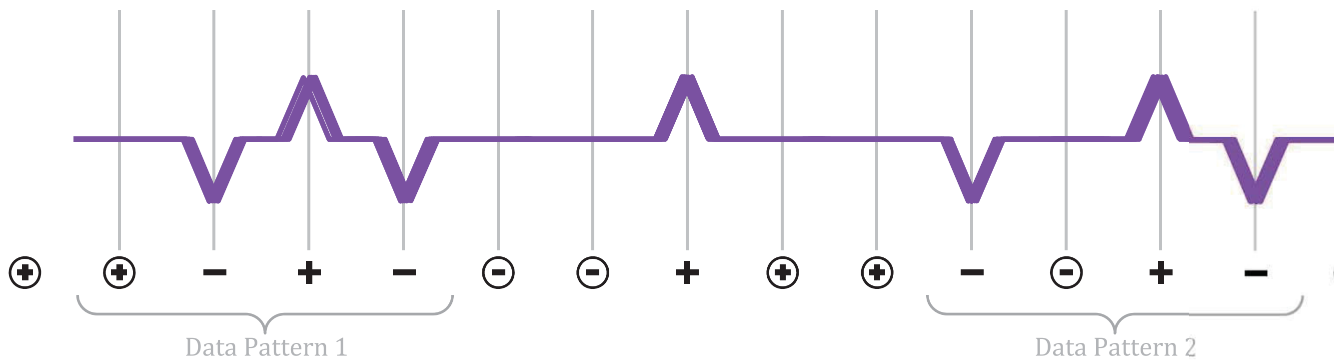
Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.

Neighborhood Dependent



ng's hypothetical
ter noise model.
ot signal dependent)

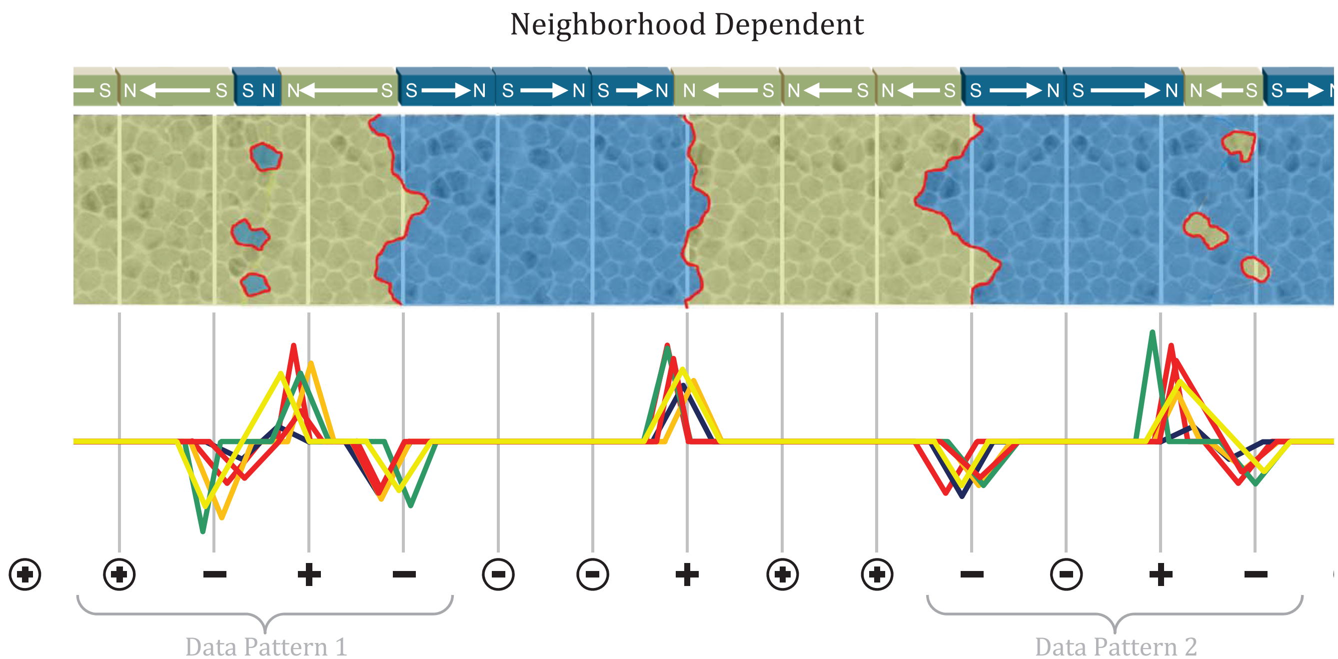


$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

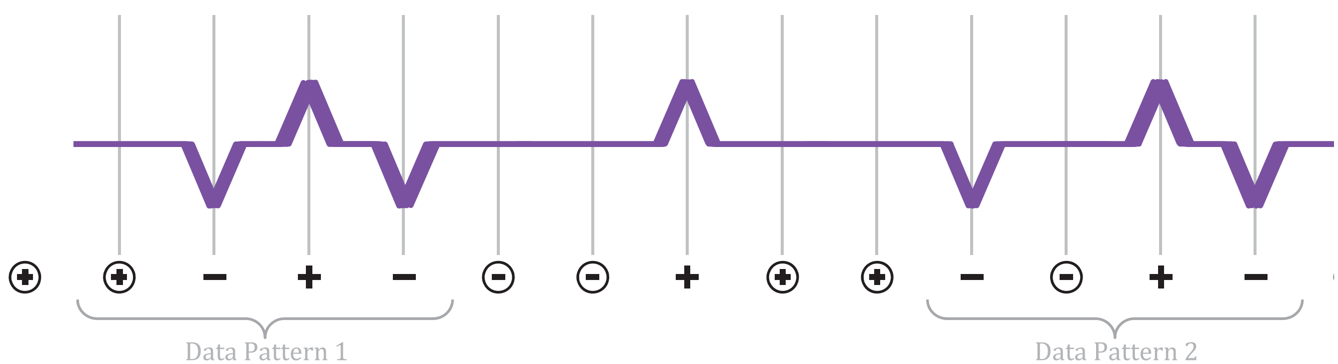
To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.



Zeng's hypothetical jitter noise model.
(Not signal dependent)

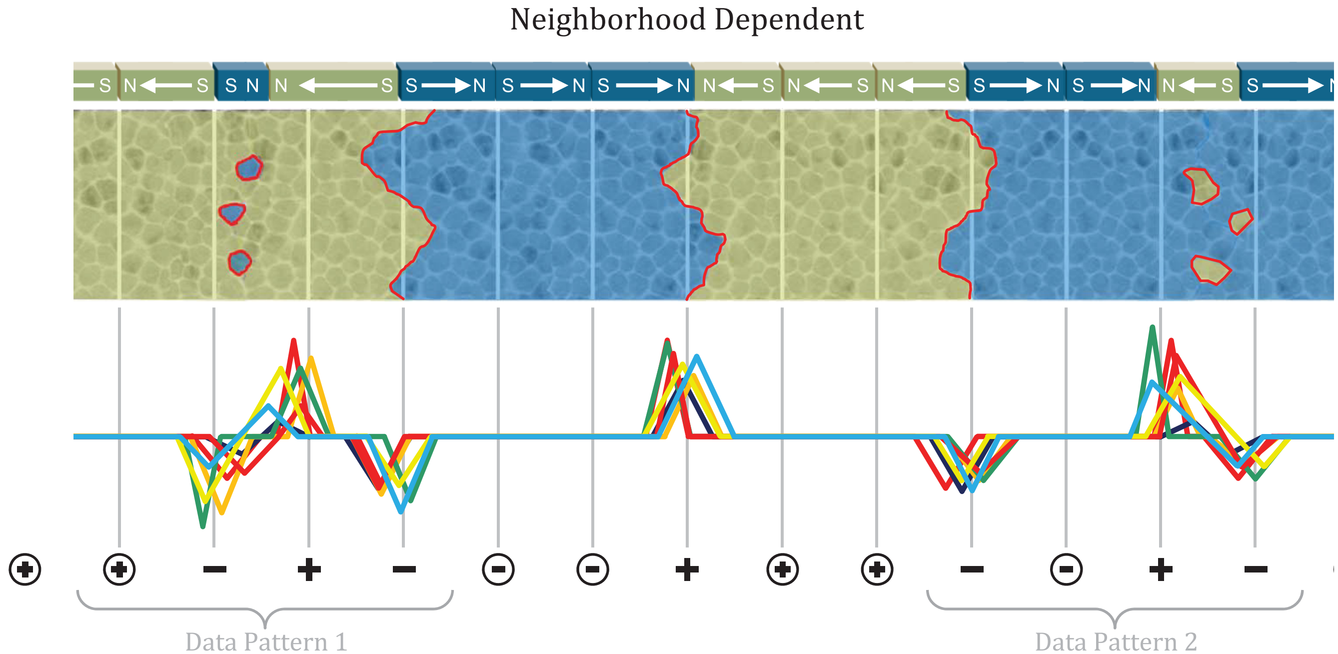


$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

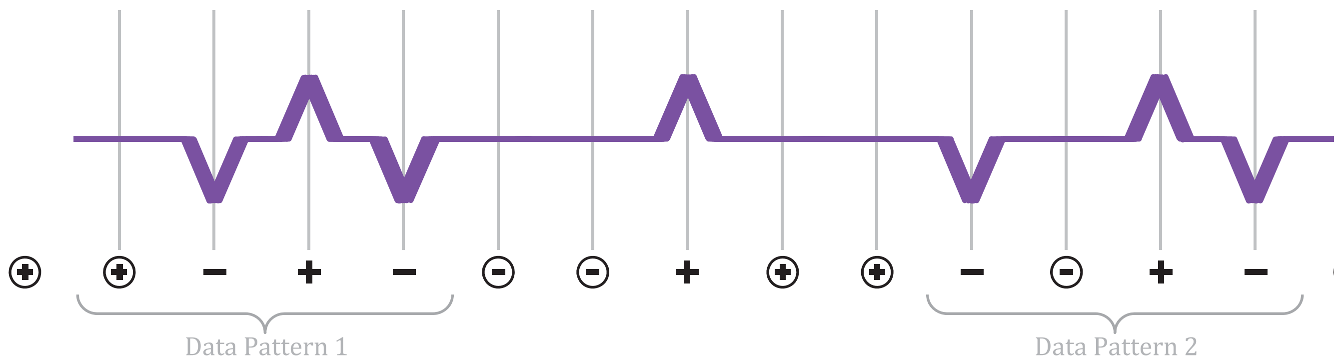
To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.



Zeng's hypothetical jitter noise model.
(Not signal dependent)



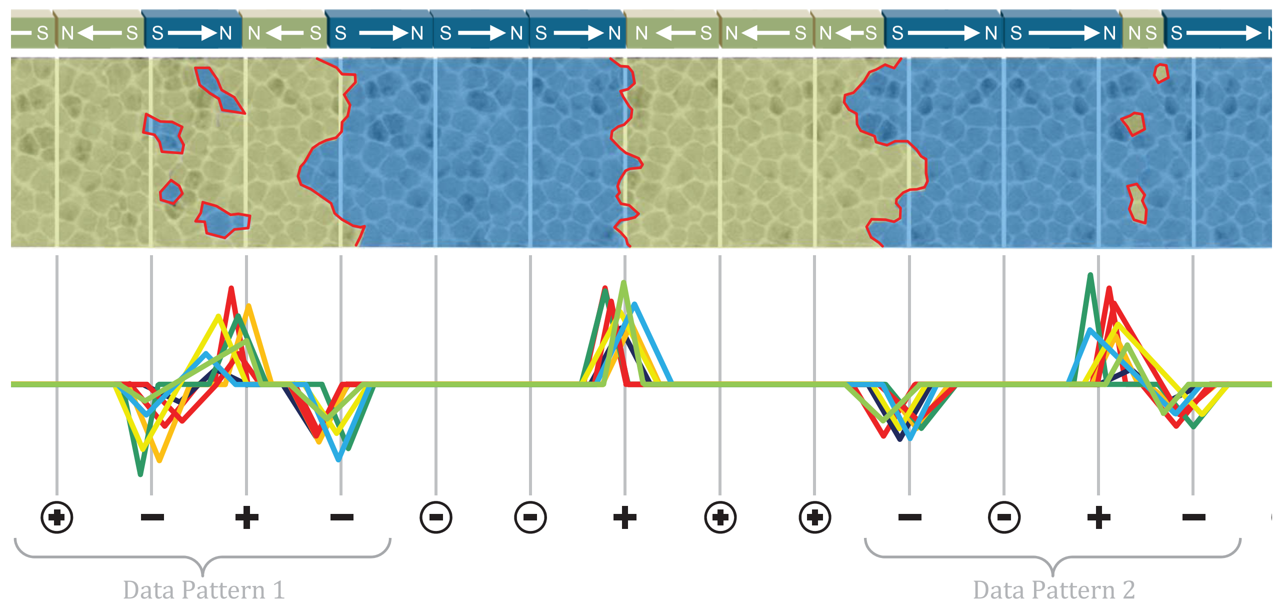
$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

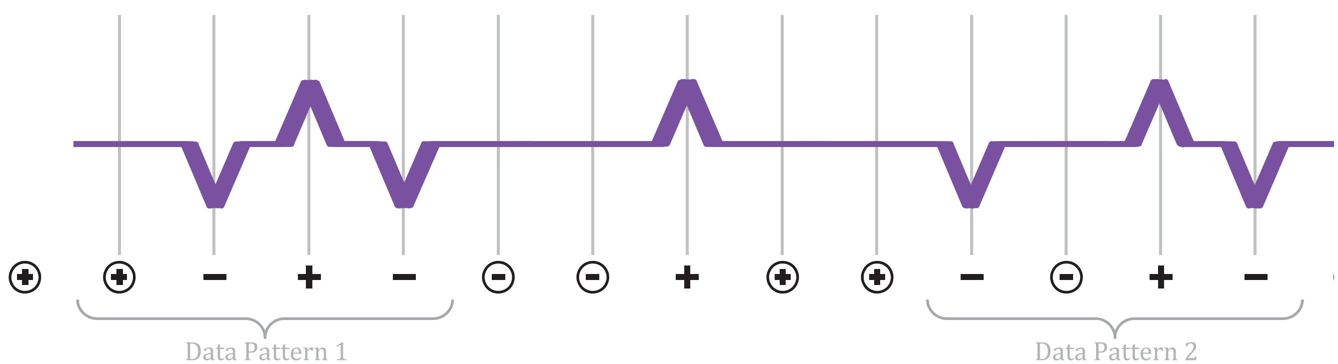
Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.

Neighborhood Dependent



Zeng's hypothetical jitter noise model.
(Not signal dependent)

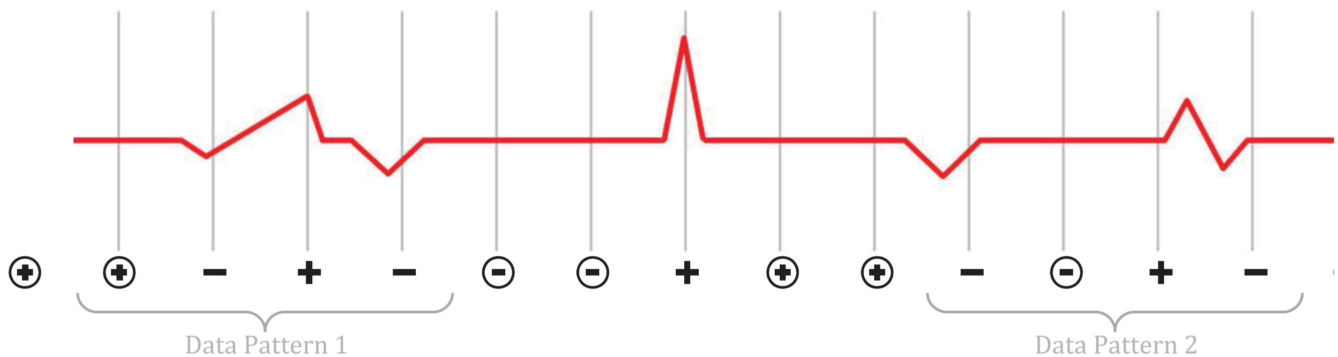


$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

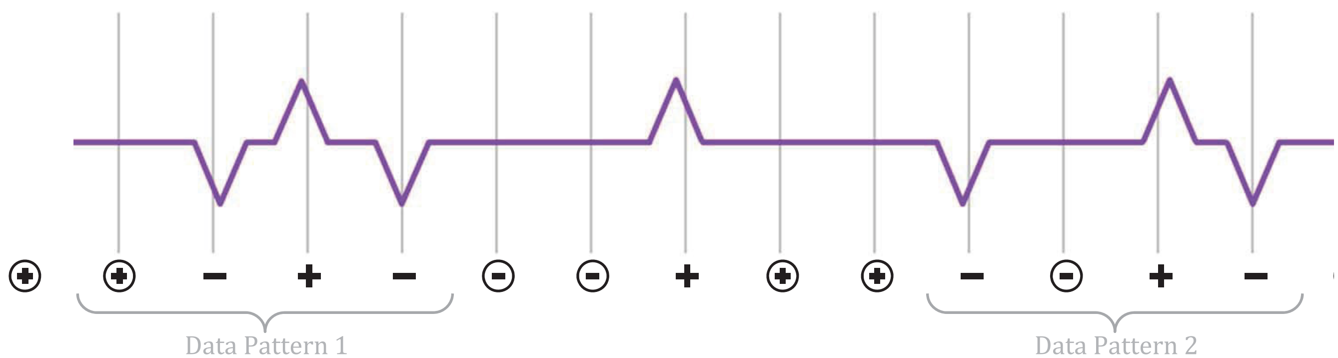
To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.



Zeng's hypothetical jitter noise model.
(Not signal dependent)



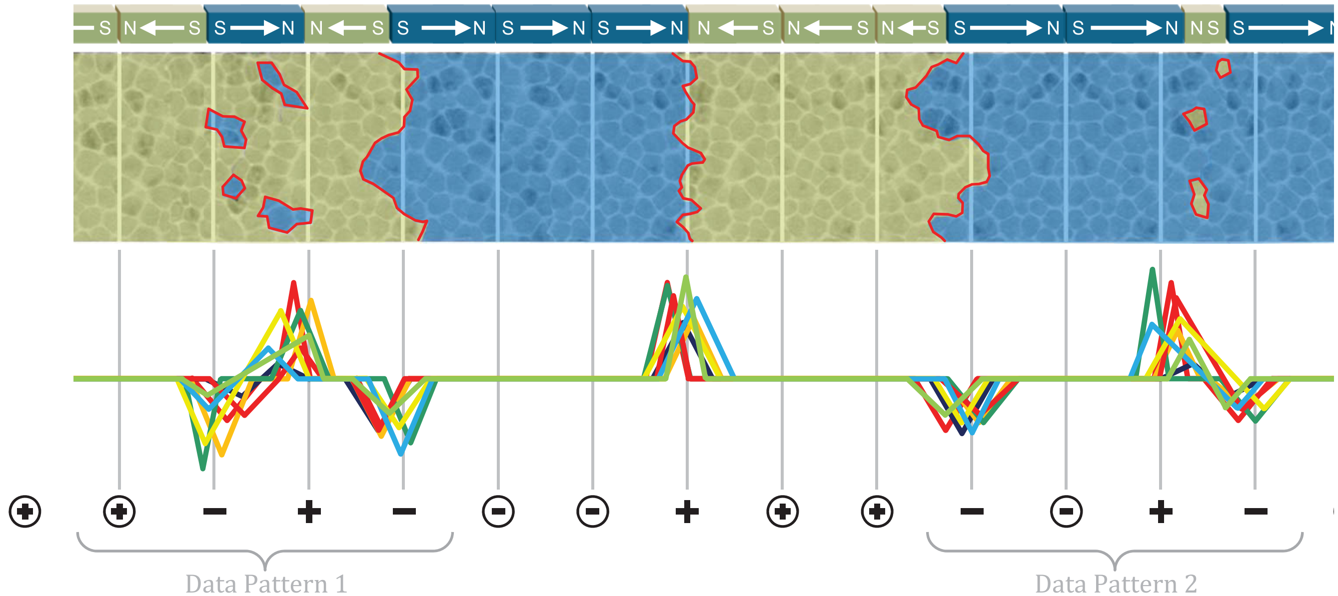
$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

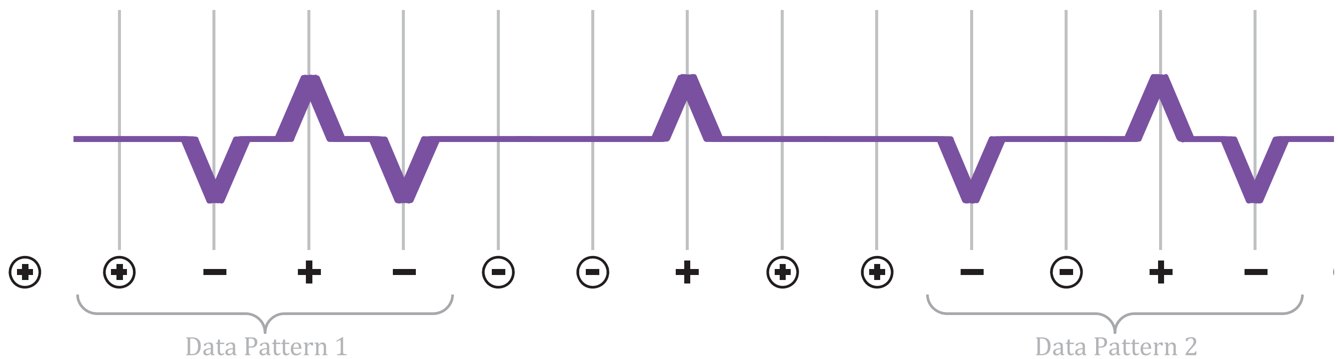
Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.

Neighborhood Dependent



Zeng's hypothetical jitter noise model.
(Not signal dependent)



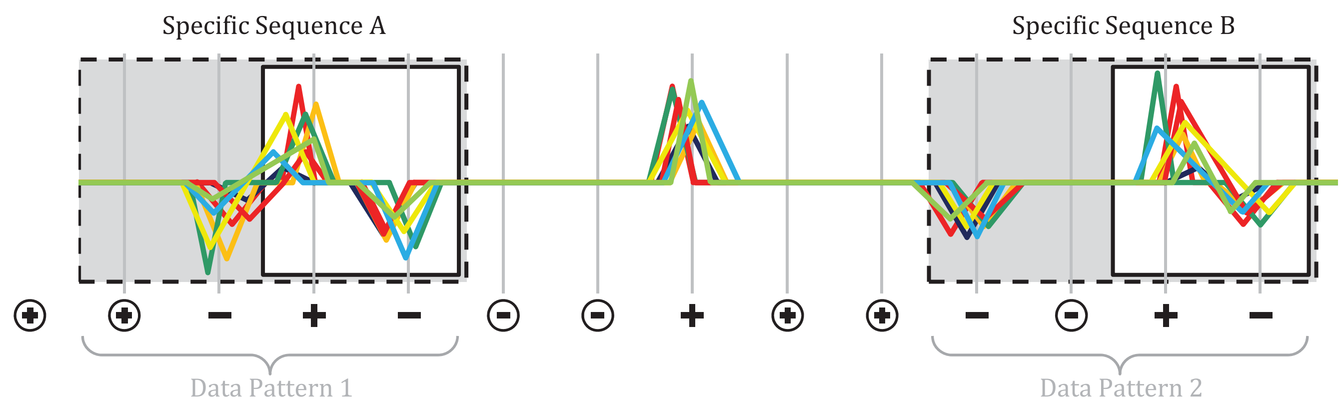
$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

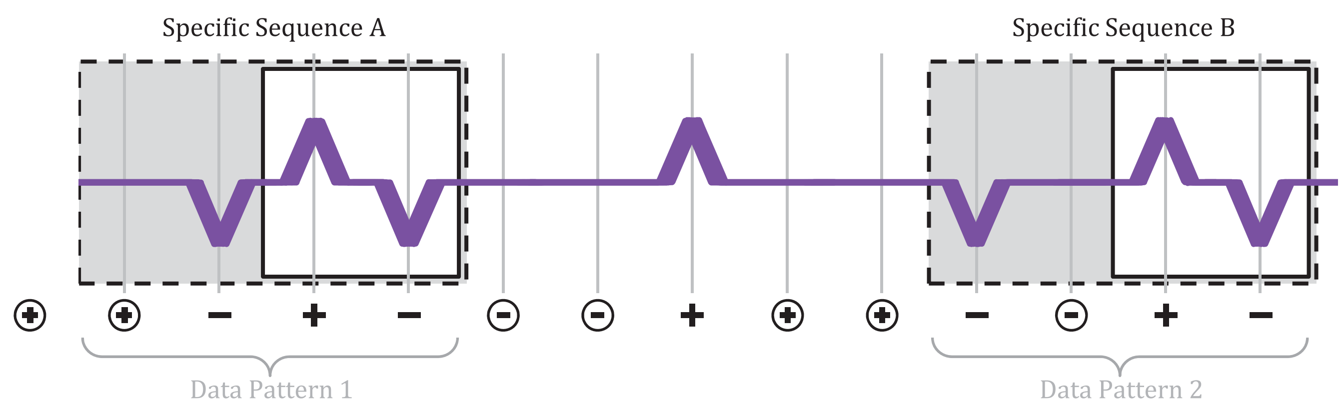
Not Addressing Further Flaw of Zeng's D=1 Constraint

Neighborhood Dependent

Signal dependent noise which includes noises from polarity and neighborhood effects.



Zeng's hypothetical jitter noise model.
(Not signal dependent)



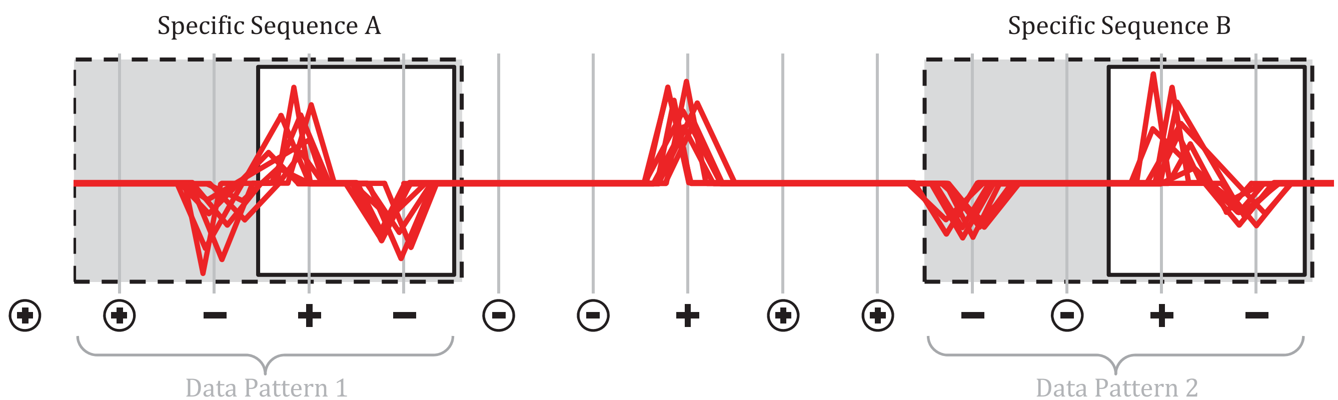
$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

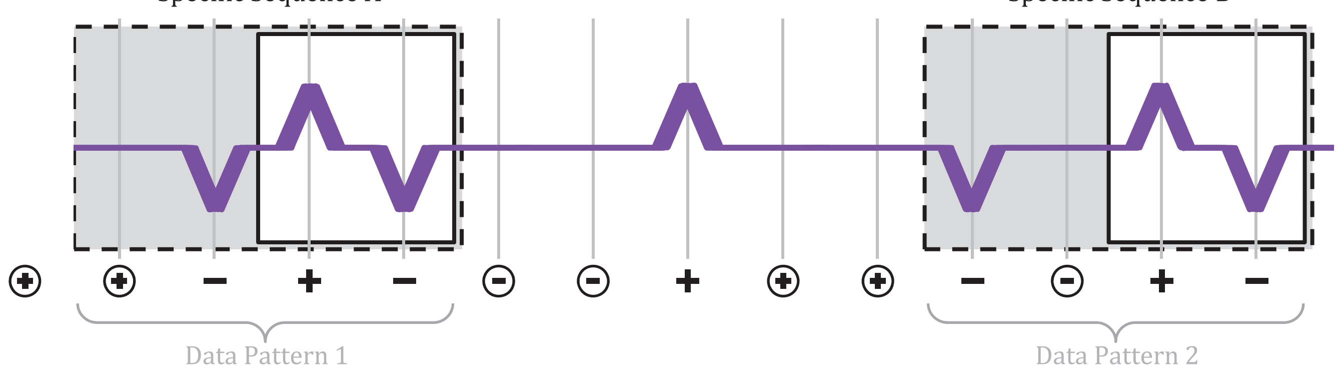
Signal dependent noise which includes noises from polarity and neighborhood effects.

Neighborhood Dependent



Zeng's hypothetical jitter noise model.
(Not signal dependent)

Not Neighborhood Dependent



$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

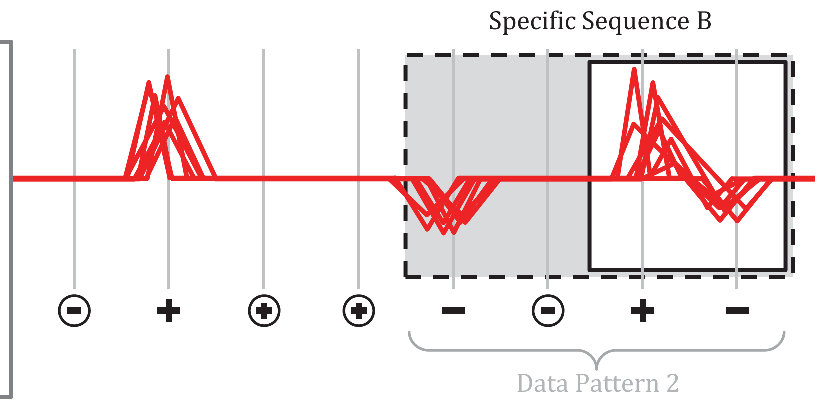
Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal dependent noise which includes noises from polarity and neighborhood effects.

Neighborhood Dependent

'839 Patent, Column 5:49-55

The metric is a function of the observed samples $r_i, r_{i+1}, \ldots, r_{i+L}$. It is also dependent on the postulated sequence of written symbols $a_i-K_1, \ldots, a_i+L+K_t$, which ensures the signal-dependence of the detector. As a consequence, the branch metrics for every branch in the tree/trellis is based on its corresponding signal/noise statistics.

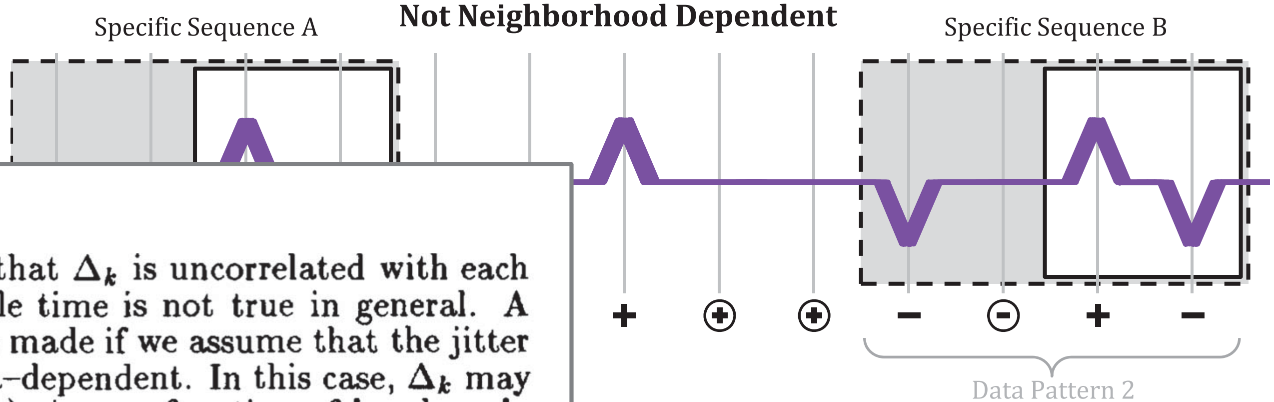


Not Neighborhood Dependent

Zeng's hypothetical jitter noise model.
(Not signal dependent)

Lee's 1992 Paper

The assumption that Δ_k is uncorrelated with each other at each sample time is not true in general. A better model can be made if we assume that the jitter term Δ_k is also data-dependent. In this case, Δ_k may be replaced by $\Delta(\mathbf{i}_k)$, *i.e.* a function of $\mathbf{i}_k$ where $\mathbf{i}_k$ indicates past data history.



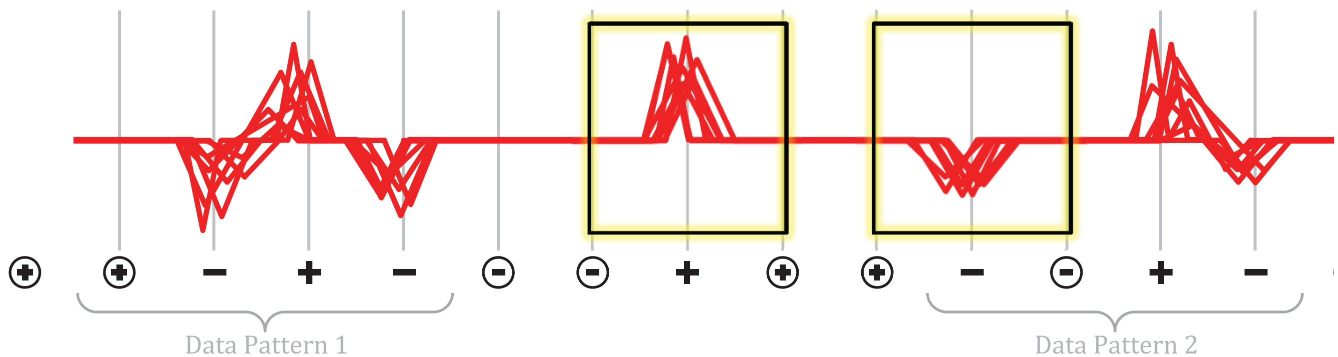
$Z(t) = \ldots$ $k=1$

Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

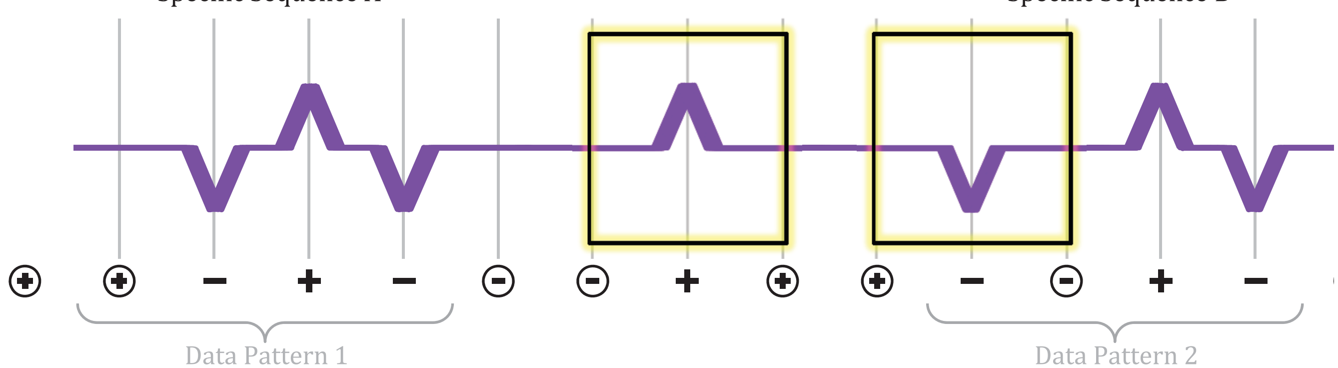
Signal dependent noise which includes noises from polarity and neighborhood effects.

Polarity Dependent



Zeng's hypothetical jitter noise model.
(Not signal dependent)

Not Polarity Dependent



$$Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t),$$

To be described in Dr. Bajorek's Declaration

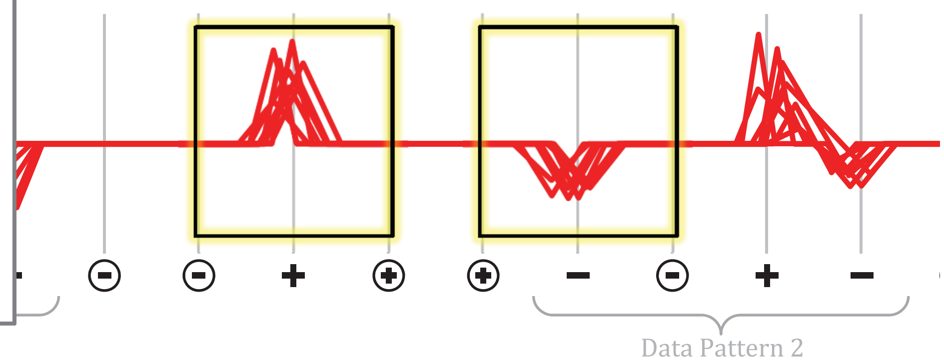
Not Addressing Further Flaw of Zeng's D=1 Constraint

Signal d… which i… rom po… ıeighbo…

'839 Patent, Column 3:51-64

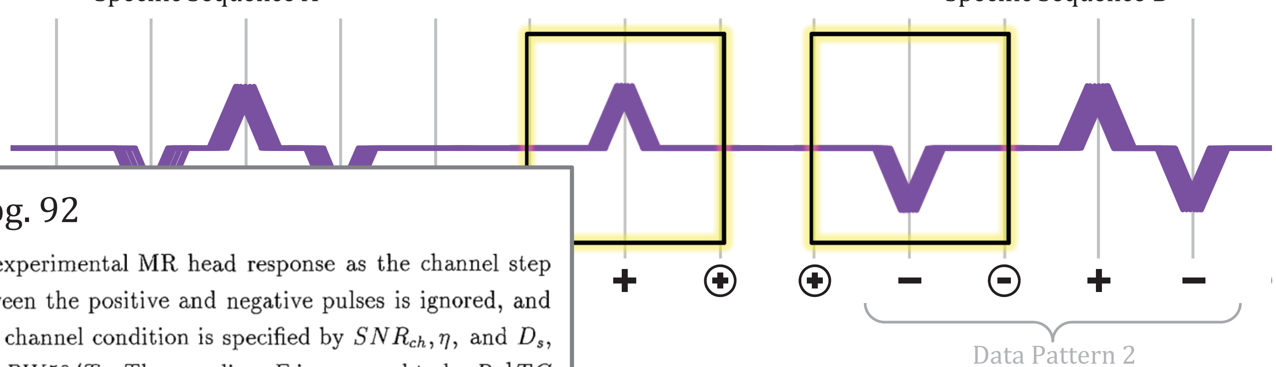
The correlation-sensitive maximum likelihood sequence detector (CS-MLSD) **28** is described hereinbelow. Assume that N>1 channel bits (symbols), $a_1, a_2, \ldots, a_N$, are written on a magnetic medium. The symbols a_i, i=1, . . . , N, are drawn from an alphabet of four symbols, $a_i \epsilon \{+, \oplus, -, \ominus\}$. The symbols '+' and '–' denote a positive and a negative transition, respectively. The symbol '⊕' denotes a written zero (no transition) whose nearest preceding non-zero symbol is a '+' while '⊖' denotes a written zero whose nearest preceding transition is a negative one, i.e., '–'. This notation is used because a simple treatment of transitions as '1's and no transitions as '0's is blind to signal asymmetries (MR head asymmetries and base line drifts), which is inappropriate for the present problem. In FIG. **3** a sample waveform

Polarity Dependent



Not Polarity Dependent

Zeng's hypothetical jitter noise model.
(Not signal dependent)



Zeng's 1994 Thesis, pg. 92

In our analysis, we use an experimental MR head response as the channel step response. The asymmetry between the positive and negative pulses is ignored, and a linear model is assumed. The channel condition is specified by SNR_{ch}, η, and D_s, the symbol density expressed as $PW50/T_s$. The equalizer F is assumed to be $R^{-1}TG$ once the target G is chosen and, thus, the search variables are the L elements of the vector G. Examining the expression for SNR_{eff}, we see that SNR_{eff} is a function of $SNR_{ch}, \eta, D_s, G, L, C_k$, and the error pattern.

$z(t) =$

To be described in Dr. Bajorek's Declaration

Not Addressing Further Flaw of Zeng's D=1 Constraint

Lee Confirmed that Zeng's Random Jitter is White, Not Signal-Dependent

What Dr. Lee said in 1990s:

4 Channel model

In this section, the same channel model as in [5, 6] has been defined for performance analysis. The output of class–IV partial response (PR4) channel corrupted by jitter and additive noise is expressed as

$$z(t) = \sum_{k=1}^{M} a_k h(t - kT - \Delta_k T) + n(t) \quad (10)$$

The jitter term Δ_k is assumed to be a white Gaussian random variable with variance σ_Δ^2 and to be small. The input sequence $\{a_k\}$ is converted via the formular $a_k = (b_k - b_{k-1})/2$ where b_k is i.i.d. with value ± 1. Assume $h(t)$ to be the response of the channel to a single transtion. Then,

$$h(t) = 2\left[\frac{\sin(\pi t/T)}{\pi t/T} + \frac{\sin(\pi(t-T)/T)}{\pi(t-T)/T}\right]$$

Zeng Paper

Zeng's Channel Model Eq. 4.1

Random jitter is "white"

The assumption that Δ_k is uncorrelated with each other at each sample time is not true in general. A better model can be made if we assume that the jitter term Δ_k is also data–dependent. In this case, Δ_k may be replaced by $\Delta(\mathbf{i}_k)$, *i.e.* a function of $\mathbf{i}_k$ where $\mathbf{i}_k$ indicates past data history. But in this paper a simple

Dr. Lee criticizing Zeng's model for not being "data-dependent"

Lee's 1992 peer-reviewed IEEE paper, p. 963
See also Lee's Thesis at pp. 99-100

What Dr. Lee says in 2014:

- Zeng's "random jitter" "reflects the signal-dependent noise"

52. In addition, the branch metric functions for many of the branches in the 16-state detector expressly incorporate variables that represent signal dependent noise. For example, the branch metric functions for branches 5, 6, 11, and 12 utilize the variable σ_Δ^2, which is defined as the variance of Δ_k. *Id.* at 65 ("Here σ_Δ^2 is the variance of Δ_k."). And, the variable Δ_k represents the random jitter in the position of the transition response, and thus reflects the signal-dependent noise. *See id.* at 51 ("Δ_k represents the random jitter in the position of the transition response."); *see also id.* at 10 ("The media noise is believed to be largely caused by transition noise, which is the random fluctuations of the transition region. . . . This fluctuation, usually called jitter, is also reflected in the playback signal. It can cause position shift or width variation in the transition

Lee Declaration, para. 52, p. 18

"Data-dependent" and "signal-dependent" are terms used in the field to indicate that the noise is dependent on the specific sequence of symbols written to the disk.

Zeng Assumed Away the Problem Solved by the Kavcic-Moura Invention

- **Assumes jitter is IID (independently and identically distributed) random variables**
- **Ignores the polarities of the transitions**
  - "The asymmetry between the positive and negative pulses is ***ignored*** and a linear model is assumed." Zeng Thesis at p. 92.
- **d = 1 RLL constraint**
  - Zeng does not permit consecutive transitions in Sections 4.4 and 5.2 because he says it is "impossible" to make the detector otherwise. Zeng Thesis at p. 65.
  - Consequence is that Zeng's channel model ignores that transitions interact – the neighborhood effect.
- **Zeng assumes "precompensation" when *writing* two transitions in a sequence.** Zeng Thesis at 9.
  - Allows Zeng to ignore neighborhood-dependence in BMFs when ***reading*** data.
- **Zeng assumes the disk is AC-erased ("degaussed") prior to each write to eliminate "overwrite effect."** *See* Zeng Thesis at p. 9.
- **Recognizes but *ignores* that non-transition sequences have signal-dependent noise.** Zeng Thesis at 10.
- **Ignores all other identified signal-dependent noise.** Zeng Thesis at p. 9.

Explained in detail in Dr. Bajorek's declaration.

Zeng's Functions in Section 4.4 are Not Signal-Dependent

No.	Function	Target	Branch	Signal-Dep?	Involves Jitter Variance Term σ_Δ^2
1	$\ln \sigma_w^2 + \dfrac{[z_k - y_k]^2}{\sigma_w^2}$	$y_k = 0$	1	No	No
		$y_k = 0$	6		
2	$\ln \sigma_w^2 + \dfrac{\left[\frac{1}{\sqrt{2}}(z_k - y_k) + \frac{1}{\sqrt{2}}(z_{k-1} - y_{k-1})\right]^2}{\sigma_w^2}$	$y_k = 2$, $y_{k-1} = 2$	2	No	No
		$y_k = -2$, $y_{k-1} = -2$	5		
3	$\ln[\sigma_w^2 + 8\sigma_\Delta^2] + \dfrac{\left[\frac{-1}{\sqrt{2}}(z_k - y_k) + \frac{1}{\sqrt{2}}(z_{k+1} - y_{k+1})\right]^2}{\sigma_w^2 + 8\sigma_\Delta^2}$	$y_k = 2$, $y_{k+1} = 2$	3	No	Yes
		$y_k = -2$, $y_{k-1} = -2$	4		

- **None of these functions account for the structure of the media noise associated with a <u>specific</u> sequence of written symbols.**
- **Only one of them (No. 3) has a term related to Zeng's "random jitter"**
- **Even if Zeng's "random jitter" accounted for signal-dependent noise, at most only one such function, so *<u>not</u>* a "set"**

Explained in detail in Prof. McLaughlin's declaration

Zeng's Functions in Section 5.2 are Not Signal-Dependent & Not a "Set" of Such Functions

No.	Function	Target	Branch	Signal-Dep?	Involves Jitter Variance Term σ_Δ^2
1	$\ln\sigma_w^2 + \frac{[ZD_k - t_a]^2}{\sigma_w^2}$	$t_a = 0$	1	No	No
		$t_a = 2g_2$	2		
		$t_a = -2g_{-1}$	3		
		$t_a = -2g_{-1}+2g_2$	4		
		$t_a = 2g_{-1}-2g_2$	13		
		$t_a = 2g_{-1}$	14		
		$t_a = -2g_2$	15		
		$t_a = 0$	16		
2	$\ln\sigma_w^2 + \frac{\left[\frac{j_0}{j_o^2+j_1^2}[ZD_k - t_a] - \frac{j_1}{j_o^2+j_1^2}[ZD_{k-1} - t_b]\right]^2}{\sigma_w^2}$	$t_a = 2g_{-1}-2g_1$ $t_b = -2g_0$	7	No	No
		$t_a = -2g_1$ $t_b = 2g_0$	8		
		$t_a = 2g_1$ $t_b = 2g_0$	9		
		$t_a = -2g_{-1}+2g_1$ $t_b = 2g_0$	10		
3	$\ln[\sigma_w^2 + 4(j_o^2+j_1^2)\sigma_\Delta^2] + \frac{\left[\frac{j_1}{j_o^2+j_1^2}[ZD_k - t_a] + \frac{j_0}{j_o^2+j_1^2}[ZD_{k+1} - t_b]\right]^2}{\sigma_w^2 + 4(j_o^2+j_1^2)\sigma_\Delta^2}$	$t_a = -2g_0$ $t_b = -2g_1$	5	No	Yes
		$t_a = -2g_0+2g_2$ $t_b = -2g_1$	6		
		$t_a = 2g_0-2g_2$ $t_b = 2g_1$	11		
		$t_a = 2g_0$ $t_b = 2g_1$	12		

- **None of these functions account for the structure of the media noise associated with a specific <u>sequence</u> of written symbols.**
- **Only one of them (No. 3) has a term related to Zeng's "random jitter"**
  - Cannot constitute the claimed "set"
- **Contrary to Dr. Lee's assertions, Section 5.2 only has 3 BMFs.**
  - Dr. Lee wrongly believes that changing the *input* to a function (like the target value) changes the function.
  - Marvell made this same mistake in the litigation.

Explained in detail in Prof. McLaughlin's declaration

Zeng's Own Thesis Advisor Said CMU Inventors Were First

- **Zeng's Thesis advisor was Prof. Jaekyun Moon.** *See* Zeng Thesis at p. i.
- **In 2001 peer-reviewed paper, Prof. Moon said the signal-dependent detector was "first derived" by the CMU inventors**

IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, VOL. 19, NO. 4, APRIL 2001

Pattern-Dependent Noise Prediction in Signal-Dependent Noise

Jaekyun Moon, *Senior Member, IEEE*, and Jongseung Park

I. INTRODUCTION

INCORPORATING noise prediction into the branch metric computation of the Viterbi algorithm has been shown to improve performance of partial response maximum likelihood (PRML) detectors. The performance improvement comes from the effective whitening of the noise samples that became correlated at the detector input due to the equalization constraint [1]–[3]. This approach, called the noise predictive maximum likelihood (NPML) method, also has been extended to signal-dependent medium noise channels, where the noise characteristics depend highly on the local bit patterns [4]. This paper is intended to provide a more formal and general treatment of the NPML as applied to channels subject to signal-dependent noise. The maximum likelihood sequence detector (MLSD) for signal-dependent Gaussian noise has been first derived in [5] under the assumption that the noise can be modeled as an autoregressive (AR) process (or Markov process). The resulting structure is a Viterbi algorithm that also incorporates signal-dependent noise prediction into branch metric computation [5], [6]

[5] A. Kavcic and J. M. F. Moura, "Correlation-sensitive adaptive sequence detection," *IEEE Trans. Magn.*, vol. 34, pp. 763–771, May 1998.

IEEE TRANSACTIONS ON MAGNETICS, VOL 34, NO 3, MAY 1998 763

Correlation-Sensitive Adaptive Sequence Detection

Aleksandar Kavčić, *Student Member, IEEE*, and José M. F. Moura, *Fellow, IEEE*

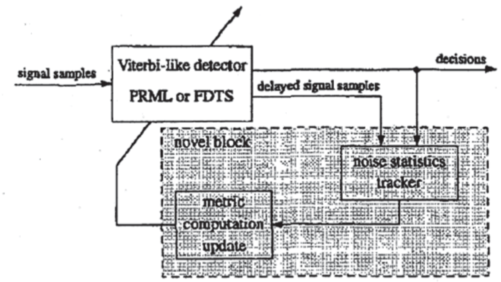


Fig. 1. Block diagram of the Viterbi-like CS-SD.

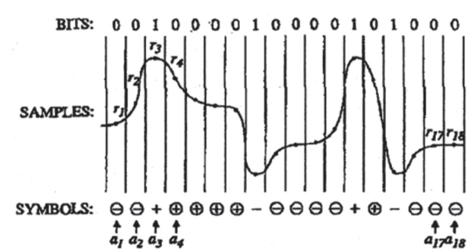


Fig. 2. Sample signal waveform, its samples and written symbols.

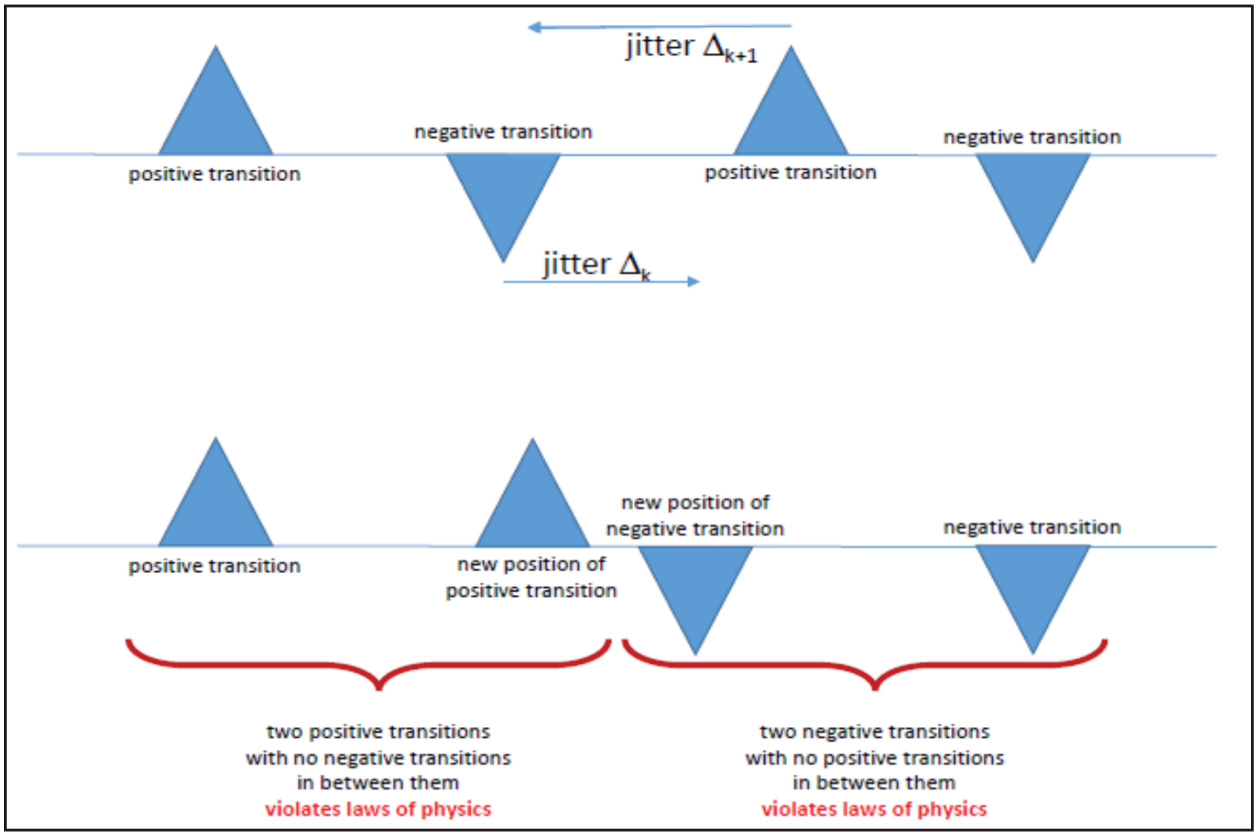
This 1998 IEEE paper by Profs. Kavcic and Moura is virtually identical to the '839 patent.

Zeng's Thesis is Not Enabling

Zeng's Thesis

Not Enabling

1. **Physically impossible channel model**
   - Permits consecutive positive and consecutive negative transitions, which is physically impossible

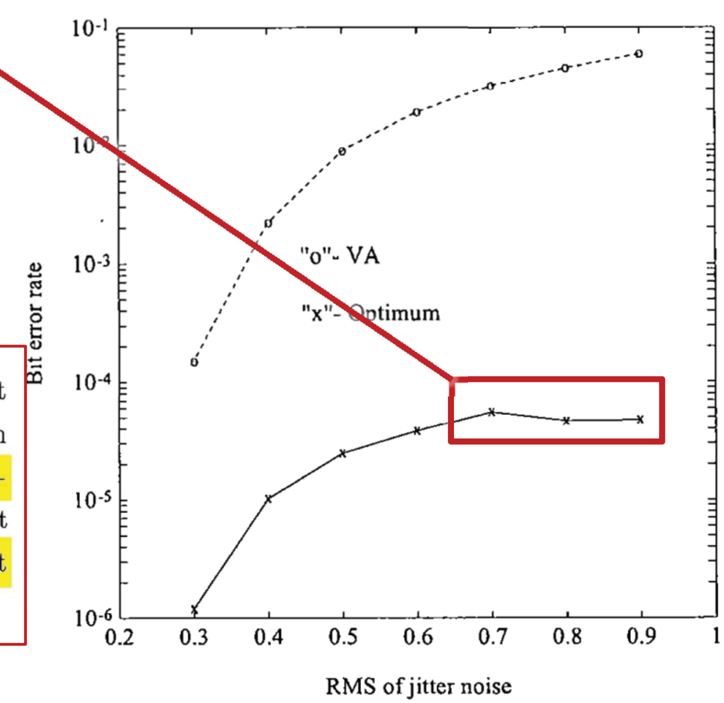


Explained in detail in Prof. Kavcic's declaration

Zeng's Thesis

Not Enabling

1. Physically impossible channel model
2. **Physically impossible simulation results**
   - Performance does ***not*** deteriorate with increasing noise
   - Confounding Analysis: Zeng's says this performance is not better than prior art, so Zeng developed model in Section 5.2, but Section 5.2 is full of math errors.



Zeng Thesis at p. 69

to that in Fig. 4.6 based on the channel model Equation (4.3) including all significant jitter terms. Unfortunately, the optimum detector developed in this section based on the simplified channel model Equation (4.22) does not improve the performance significantly upon the conventional Viterbi algorithm (less that two times improvement in terms of bit error rate). This leads us to study the more realistic model in next chapter.

Zeng Thesis at p. 71

Explained in detail in Prof. Kavcic's declaration

Zeng's Thesis

Not Enabling

1. Physically impossible channel model
2. Physically impossible simulation results
3. **Disabling mathematical errors in Section 5.2**
   - Consider Eq. 5.13 of Zeng's Thesis (p. 78)
   - Used for branch 5 in Zeng's Section 5.2 trellis

Mixed-up *j* terms

Wrong target

Entire sum in numerator not squared

$$\ln(\sigma_w^2 + 4(j_0^2 + j_1^2)\sigma_\Delta^2) + \frac{j_1(ZD_k + 2g_0) + j_0(ZD_{K+1} + 2g_1)^2}{(\sigma_w^2 + 4(j_0^2 + j_1^2)\sigma_\Delta^2)(j_0^2 + j_1^2)}$$

- These ***errors are repeated*** in Section 5.2 for each of branches 6 through 12

Explained in detail in Prof. Kavcic's declaration

The *CMU v. Marvell* Litigation

CMU v. Marvell

- **Ropes & Gray represents Marvell before the USPTO**
- **Requester requested reexamination of only litigated claims**
- **Requester did not seek *Inter Partes* Review**
  - Marvell is only entity in the world barred from seeking IPR
- **Marvell's litigation counsel would neither admit nor deny that Marvell is behind the Request**
- **Dr. Lee (Requester's expert) and Zining Wu (Marvell's CTO) both had the same thesis advisor at Stanford at the same time -- Prof. John Cioffi, a former Marvell board member**
- **Even if Marvell did not request reexamination, its conduct in the litigation is highly probative given its expertise, access to Zeng and incentive to turn over every stone.**
  - The views of persons of ordinary skill are probative regarding the content of a reference. *See Dayco Prods., Inc. v. Total Containment, Inc.*, 329 F.3d 1358, 1368-69 (Fed. Cir. 2003)

CMU v. Marvell

- **Ropes & Gray represents Marvell before USPTO, *see, e.g.*, Ser. No. 14/166,428**

14/166,428 — ITERATIVE DECODER SYSTEMS AND METHODS

Select New Case | Application Data | Transaction History | Image File Wrapper | Continuity Data | Fees | Published Documents | Address & Attorney/Agent | Display References

Bibliographic Data

Application Number:	14/166,428
Filing or 371 (c) Date:	01-28-2014
Application Type:	Utility
Examiner Name:	BUTLER, SARAI E
Group Art Unit:	2114
Confirmation Number:	4151
Attorney Docket Number:	MP2333C1C1/04048-0167-103
Class / Subclass:	714/795
First Named Inventor:	Panu Chaichanavong , Mountain View, CA (US) all inventors
First Named Applicant:	MARVELL WORLD TRADE LTD. St. Michael, (BB) all applicants
Entity Status:	Undiscounted
Title of Invention:	ITERATIVE DECODER SYSTEMS AND METHODS

14/166,428 — ITERATIVE DECODER SYSTEMS AND METHODS

Select New Case | Application Data | Transaction History | Image File Wrapper | Continuity Data | Fees | Published Documents | Address & Attorney/Agent | Display References

Correspondence Address

Name:	ROPES & GRAY LLP
Address:	PATENT DOCKETING 39/361 1211 AVENUE OF THE AMERICAS NEW YORK NY 10036-8704
Customer Number:	-

Attorney/Agent Information

Reg #	Name	Phone
71406	Amodeo, Gabriele	212-596-9496
30818	Badke, Bradford	212-596-9031
47414	Baughman, Jon	202-508-4600
51452	Bellomy, Mark	617-951-7000
61129	Bertenthal, Matthew	650-617-4007

14/166,428 — ITERATIVE DECODER SYSTEMS AND METHODS

Select New Case | Application Data | Transaction History | Image File Wrapper | Continuity Data | Fees | Published Documents | Address & Attorney/Agent | Display References

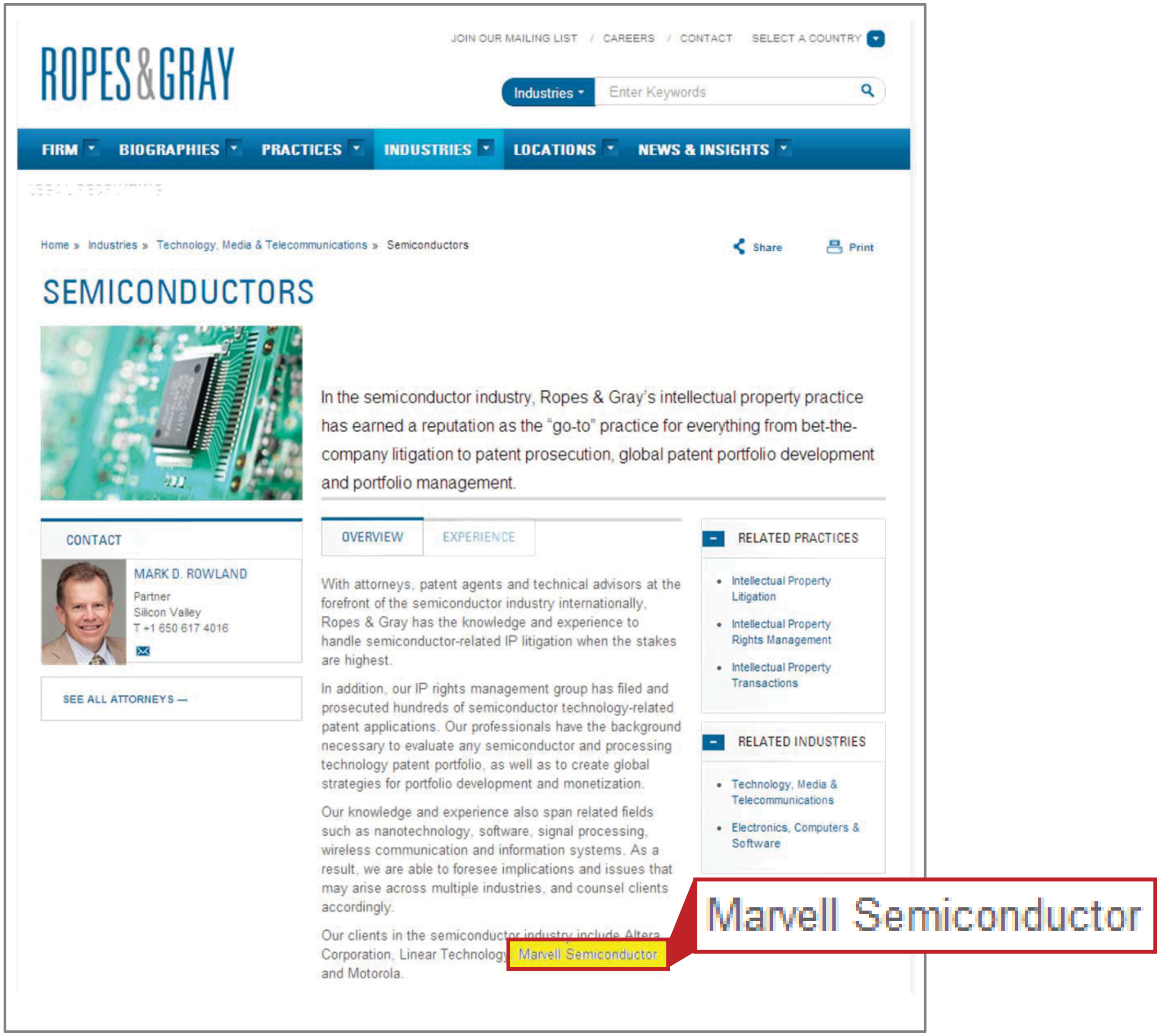
Inventor Information

Name	Address
Panu Chaichanavong	Mountain View, CA (US)
Nedeljko Varnica	San Jose, CA (US)
Nitin Nangare	Santa Clara, CA (US)
Gregory Burd	San Jose, CA (US)
Zining Wu	Los Altos, CA (US)

Witnesses for Marvell in CMU case

CMU v. Marvell

- **Ropes & Gray lists Marvell as a client on its website**



CMU v. Marvell

- **CMU sued Marvell on March 6, 2009**
- **Marvell mounted a vigorous defense**
  - "armies of attorneys" according to Judge Fischer
  - Marvell produced Zeng's Thesis to CMU early in case
  - **Marvell cited Sections 4.4 and 5.2 of Zeng Thesis in its invalidity contentions**
  - Lead lawyer had Ph.D in Electrical Engineering
  - **Marvell's expert, Prof. John Proakis, reviewed Zeng's Thesis in preparing his expert report but did *not* opine that it invalidates claim 4.**

Prof. John Proakis

- PhD from Harvard
- Former professor at Northeastern Univ.
- 20 yrs consulting in HDD read channel industry
- Author of numerous textbooks

Invalidity of U.S. Patent No. 6,201,839

by Weining Zeng, "Effective Detection Schemes for Magnetic Recording Channels with Severe Nonlinearities and Media Noise," Ph.D. Dissertation, University of Minnesota (Oct. 1994) ("Zeng")

Based upon the claim interpretations Plaintiff appears to be asserting and the applications of those interpretations to Defendants' products in Plaintiff's Preliminary Infringement Contentions, Zeng anticipates and/or renders obvious, alone or in combination with other prior art identified in Defendants' Preliminary Invalidity Contentions, the asserted claims as described in part below. These invalidity contentions are not an admission by Defendants that the accused products, including any current or past versions of these products, are covered by, or infringe these claims, particularly when they are properly construed. Nothing in these contentions should be interpreted as an acquiescence to or assertion of a particular claim construction by Defendants.[1]

'839 Claim Language	Disclosure
1. A method of determining branch metric values for branches of a trellis for a Virterbi-like detector, comprising:	*See, e.g.*, pp. 50-54.
selecting a branch metric function for each of the branches at a certain time index; and	*See, e.g.*, pp. 52, 68; Eqs. (4.7), (4.25)-(4.28), (5.3)-(5.21) and accompanying text.
applying each of said selected functions to a plurality of signal samples to determine the metric value corresponding to the branch for which the applied branch metric function was selected, wherein each sample corresponds to a different sampling time instant.	*See, e.g.*, pp. 52, 68; Eqs. (4.7), (4.25)-(4.28), (5.3)-(5.21) and accompanying text.
2. The method of claim 1 further comprising the step of receiving said signal samples, said signal samples having signal-dependent noise, correlated noise, or both signal-dependent and correlated noise associated therewith.	*See, e.g.*, p. 50.
3. The method of claim 1 wherein said branch metric functions for each of the branches are selected from a set of signal-dependent branch metric functions.	*See, e.g.*, pp. 52, 68; Eqs. (4.7), (4.25)-(4.28), (5.3)-(5.21) and accompanying text.

[1] To the extent Zeng is argued by Plaintiff or found by the Court not to explicitly teach certain limitations in the asserted claims, such limitations would have been inherent and/or obvious as described in Defendants' Preliminary Invalidity Contentions.

Marvell's Invalidity Contentions, Nov. 16, 2009

Marvell Had Every Incentive to Put on Best Possible Defense

- **Despite asserting Zeng's Thesis in 2009, Marvell Ignored Zeng's Thesis at Trial**
  - Zeng, a Marvell employee, did not testify that he was first
  - Professor Proakis ignored Zeng's Thesis in his trial testimony, even he reviewed it in preparing his report
  - Despite filing two summary judgment motions on validity and several motions for reconsideration, Marvell never relied on Zeng's Thesis
  - Not raised by Marvell's appeal

THE WALL STREET JOURNAL.

A Record Patent Verdict

By Don Clark

A federal jury in Pittsburgh found Wednesday that chipmaker **Marvell Technology Group** Ltd. should pay nearly $1.17 billion for infringing patents held by Carnegie Mellon University.

The award ranks as the third largest ever in a patent case, according to Lex Machina Inc., a firm that provides data and analytic services used in intellectual-property litigation.

The verdict comes amid a surge in patent litigation around the world, much of it conducted by rivals in the smartphone industry or by companies that buy up and enforce patents.

Marvell, by contrast, was accused by a university of infringing patents related to disk drives used in personal computers, server systems and other products. The damages set by the jury would top the award of about $1 billion issued in August against **Samsung Electronics** Co. in a closely watched case against **Apple** Inc.

In response, Marvell's shares plunged 10% to $7.40 at 4 p.m. on the Nasdaq Stock Market. The company ended Wednesday with a market value of nearly $4 billion. Marvell, which is known for chips used in data-storage applications, had $2 billion in cash and investments as of Oct. 27.

A spokesman for Marvell said the Santa Clara, Calif., company was disappointed by the verdict but hopeful the judge would reverse the finding based on post-trial motions. If unsuccessful in that effort, Marvell will appeal, he said.

K&L Gates LLP, a law firm representing Carnegie Mellon, said the jury agreed with its allegations that Marvell sold billions of chips incorporating the patented technology without a license to do so.

The jury's finding of what amounts to willful infringement provides a legal foundation for the judge to increase the damages by as much as three times the amount awarded by the jury, the firm said.

Large jury verdicts in patent cases are often sharply reduced following appeals. On the other hand, interest typically accrues during that process.

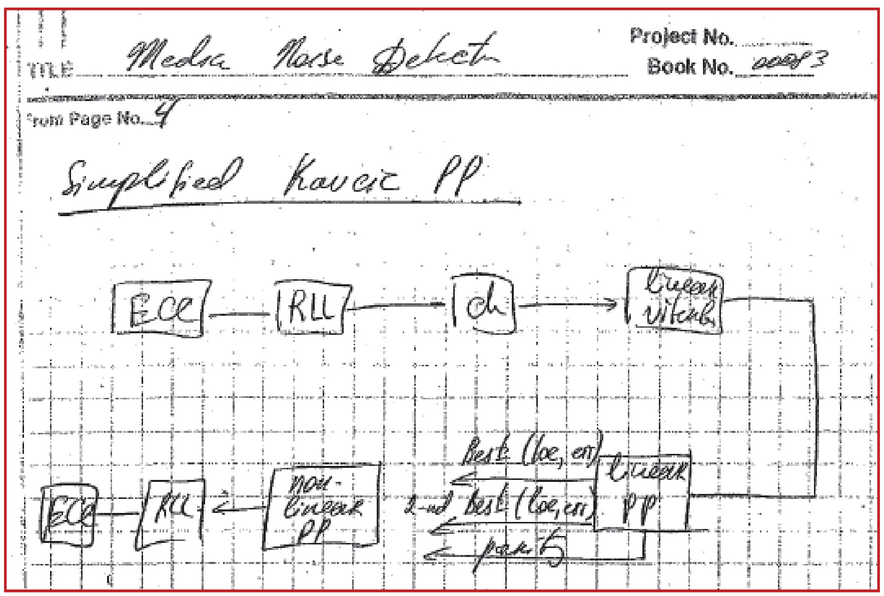
"By any measure this is a substantial damages award, which could become even larger," said

Please turn to the next page

Wall Street Journal, Dec. 27, 2012, p. B1

Secondary Indicia of Nonobviousness

- **Unprecedented secondary indicia of nonobviousness revealed at trial**
  - Copying
  - Industry praise
  - Failure by others
  - Commercial Success
  - Satisfaction of long-felt need
- **Copying**
  - Marvell's CTO, Zining Wu, admitted that he reviewed Zeng's work when Marvell started to develop its first signal-dependent detector, so-called "MNP"
  - Yet Marvell initially named the MNP after Prof. Kavcic – "kavcicPP"



Excerpt from lab notebook of Marvell engineer, Gregory Burd (Trial Ex. P-196)

Quotes from Judge Fischer

- The "Court finds that the equities clearly favor CMU… rather than Marvell, which ***copied CMU's patents consciously and deliberately for an entire decade.***" Dkt. 920 at 70.
- The "Court holds that the credible evidence presented at trial sufficiently establishes that ***Marvell deliberately copied CMU's Patents.***" Dkt. 933 at 19-20.

CMU v. Marvell

Secondary Indicia of Nonobviousness

- **Copying**
  - On its 2nd generation detector, Marvell went "all in" on the Kavcic-Moura design

From: Zi-Ning Wu
Sent: Friday, January 10, 2003 3:54 PM
To: Toai Doan
Cc: Runsheng He; Ravi Narasimhan; Hui-Ling Lou
Subject: Weekly status: 1/6/03 -- 1/10/03

Marvell Confidential

1. MNP enhancement: Greg and I discussed the approach of using a different noise whitening filter for each branch. It turns out to be the original structure that Kavcic proposed in his paper.
We also found a way to move the noise whitening filter out of the Viterbi. Therefore, the speed bottleneck would be the (y-y_hat)^2/sigma^2 operation in the branch metric calculation.
This method has a potential gain of 0.2 dB over our current MNP.

IEEE TRANSACTIONS ON MAGNETICS, VOL. 34, NO 3, MAY 1998 763

Correlation-Sensitive Adaptive Sequence Detection

Aleksandar Kavčić, *Student Member, IEEE*, and José M. F. Moura, *Fellow, IEEE*

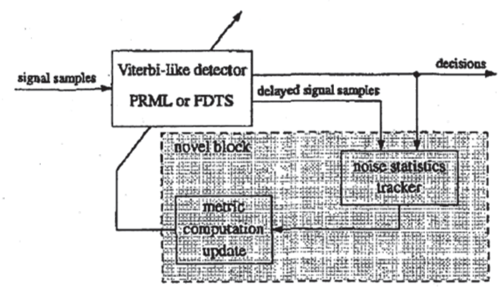


Fig. 1. Block diagram of the Viterbi-like CS-SD.

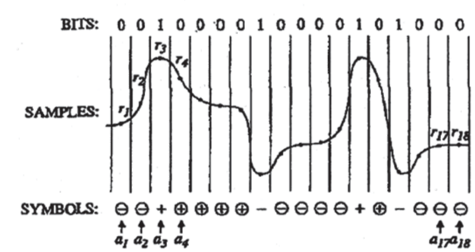
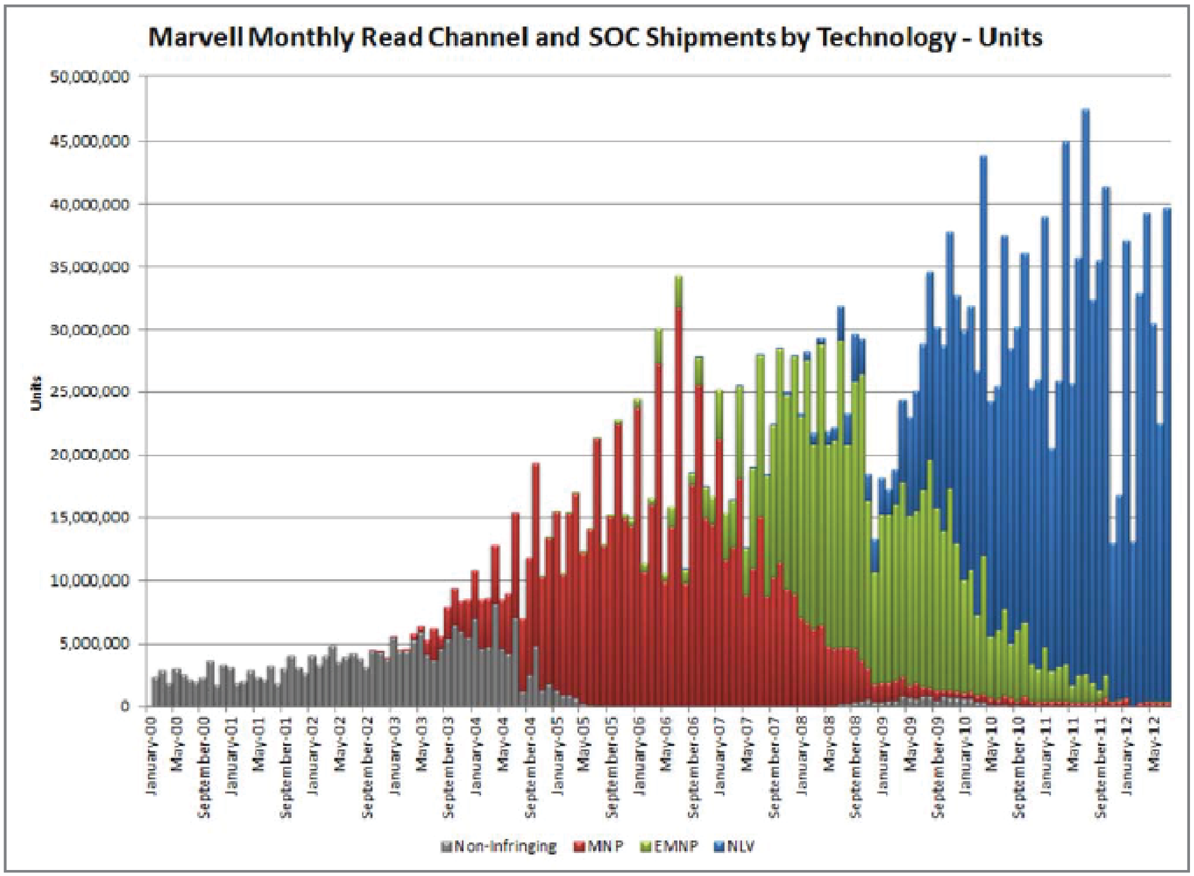


Fig. 2. Sample signal waveform, its samples and written symbols.

This 1998 IEEE paper by Profs. Kavcic and Moura is virtually identical to the '839 patent.

Secondary Indicia of Nonobviousness

- **Commercial Success**
  - Sales of prior non-infringing chips effectively ceased after infringing chips introduced
  - From March 2003 to July 2012, Marvell sold over 2.3 billion chips, with an average revenue $4.42/chip ($10.3 billion total revenue), and with an average profit of $2.16/chip, for a total profit of $4.97 billion



Secondary Indicia of Nonobviousness

- **A Nexus Between Marvell's Copying and its Commercial Success**
  - 7 years after copying claim 4 in designing its MNP chips, Marvell cited the MNP as one of two technologies that "helped firmly establish Marvell as the market leader"

PLAINTIFF'S TRIAL EXHIBIT P-703

From: Peggy Fang <pfang@marvell.com>
Sent: Friday, August 8, 2008 4:50 PM

To all,

I would like to announce the promotion of Zi-Ning Wu to VP of Data Storage Technology. Zi-Ning will be reporting directly to Sehat Sutardja starting today. In his new role, he and his team will be in charge of the development of the Read Channel IP and other IPs for use in the Data Storage Business Unit.

Zi-Ning has been with Marvell for the past 9 years working in the Data Storage Signal Processing team. In the past few years, Zi-Ning has helped me in the definition of our Read Channel roadmap along with his main responsibility of developing our Read Channel architectures and algorithms. In addition, Zi-Ning has been involved in many technical engagements with our Data Storage customers to strengthen Marvell's position with existing customers and to establish new relationships with potential customers. Working with our Read Channel VLSI team and our Data Storage SOC design teams, Zi-Ning and his DSP team have been instrumental in the development of the Media Noise Processor (MNP) and Advance ECC (AECC) for our Data Storage products. The introduction of these technologies has helped firmly establish Marvell as the market leader in the HDD IC business.

Secondary Indicia of Nonobviousness

- **Praise and Acclaim in the Industry**
  - Marvell uses Kavcic's method as its "Benchmark" because Prof. Kavcic is "VIP" and on the "Cutting Edge" of the Field

Burd Testimony

Q. Okay. And why use the Kavcic approach, as the yardstick?

MR. RADULESCU: Objection as to form.

THE WITNESS: Well, that is -- you know, you probably don't want to ask myself, necessarily, just because he's a name, kind of became a yardstick. I don't know why. I mean, people use it. It's like when you say -- you know, there are certain people which get associated with -- with some event. So you say why did they became -- become associated with those events? I don't know.

Ronald Reagan is credited with breaking down the wall. Well, I didn't see him break any bricks. Right? But yet, he is the one. So same thing.

BY MR. GREENSWAG:

Q. Did you ever use -- I mean -- I mean, did you ever have any other yardsticks that you used for either the media noise processor or the non-linear detector?

MR. RADULESCU: Objection as to form.

THE WITNESS: Well, no, because like I said, he is kind of VIP which everybody tries to cite and everybody is citing, even in the papers. Right?

We do have other yardsticks for other developments. Like, for example, we have been

working on soft Reed Solomon ECC decoder, and there is, like, very conventional Berlekamp-Massey algorithms, and then everybody is comparing themselves to that.

So it's a natural thing to compare yourself to, you know, people whose work considered to be, you know, on a leading edge, or on the cutting edge of a field. Right?

Burd Tr. at 934-936

Secondary Indicia of Nonobviousness

- **Satisfaction of Long-felt Need**
  - Many years after introducing the infringing technology, it is still “a must” at Marvell
  - Even after being found a willful infringer, Marvell has still not take the infringing technology out its chips

From: Gregory Burd
Sent: Tuesday, February 6, 2007 9:28 AM
To: Alex Nazari <anazari@marvell.com>
Subject: RE: ECC setting

Alex,

If I remember correctly, you have sent Samsung a report on AECC performance in the past. This report has plots of MNP+PECC vs. MNP+AECC. We did not do comparison with linear Viterbi since now days the drives are dominated by media noise, and MNP or NLV is a must. I will try to dig up some slides on NLV performance. But basically, everything is the same as with MNP. The AECC algorithm has not changed, we just switch the source of soft info (from MNP to NLV),

Bets, greg

Trial Ex. P-607

Secondary Indicia of Nonobviousness

- **Failure by Others**
  - Prior to copying the CMU patents in late 2001, Marvell considered and rejected a few noninfringing alternatives, but they could not get any of them to work satisfactorily.
  - Marvell spent most its time and resources on its so-called the "iterative detector," but Marvell's executives called it:

- a "coffee warmer"

Alan Armstrong, VP of Marketing

Q. What was that chip, the iterative channel you produced during that time?
A. You know, I'm going to have to -- I could find that out. I -- if I said a number, it might -- it might be wrong, but I could find that out and get back to you. It's kind of an infamous chip within Marvell. We call it the "coffee warmer chip."

(379:23–380:4)

- "the Corvair, … unsafe at any speed"

Bill Brennan, VP of Sales

I call it the Corvair, you know, unsafe at any speed. (95:18–19)

So it killed us on power, killed us on data rate, killed us on -- on size and cost, and -- it was a great idea, super idea. It simply just didn't work.

(95:24–96:1)

Key Points

- **Zeng's Sections 4.4 and 5.2 Do Not Have Any, Let Alone a "Set," of Signal-Dependent Branch Metric Functions.**
  - Zeng assumed away the signal-dependent noise problem solved by the CMU inventors
  - Zeng's functions in Sections 4.4 and 5.2 are not signal-dependent
  - Dr. Lee, the Requester's expert, admitted this in 1992 and again in 1995
  - Zeng's own thesis advisor said the CMU inventors were first

- **Litigation Confirms that Zeng's Thesis Does Not Invalidate.**
  - Marvell originally asserted that Zeng's Thesis anticipated claim 4, but Marvell dropped (and now has waived) any and all invalidity claims based on the thesis
  - Marvell's expert reviewed Zeng's Thesis but did not say it invalidated claim 4
  - Zeng is an employee of Marvell, but Marvell never called him as a witness
  - Marvell lost $1.5 billion judgment after four-week trial