IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Reexamination of: U.S. Patent No. 6,201,839	)	**Examiner:** Nguyen, Linh M.
Inventors: Aleksandar Kavcic et al.	)	**Art Unit:** 3992
Reexamination Control No.: 90/013,125	)	**Atty. Docket No**. 97168REX
Reexamination Filing Date: January 21, 2014	)	

Title: METHOD AND APPARATUS FOR CORRELATION-SENSITIVE ADAPTIVE SEQUENCE DETECTION

RESPONSE TO OFFICE ACTION

K&L Gates LLP
Pittsburgh, PA 15222
September 3, 2014

VIA EFS-Web

Mail Stop *Ex Parte* Reexam
ATTN: Central Reexamination Unit
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Sir:

In response to the Office Action mailed June 4, 2014 in connection with the above-referenced *ex parte* reexamination of U.S. Patent 6,201,839 ("the '839 patent"), the patent owner, Carnegie Mellon University ("CMU"), responds as follows, wherein:

Remarks begin on page 1 following the Table of Contents.

TABLE OF CONTENTS

I. Introduction 1

II. Concurrently Filed Supporting Declarations 6

III. Notice of Concurrent Proceedings 7

IV. Background of the Relevant Technology 8

A. Summary of the Invention Described in the '839 Patent 8

1. Hard Disk Drives 8

a. Writing Data to the Disk 8

b. Reading Data from the Disk 12

2. Viterbi Sequence Detection 13

3. Background on Branch Metric Functions 15

4. Complications from Noise 18

a. Correlated Noise 19

b. Signal-Dependent Noise 19

5. Important Factors in Developing the Invention of Claim 4 25

6. The Invention of the CMU Patents 26

a. Equation 13 27

b. FIR Filter Embodiment 28

B. Summary of the Detectors in Sections 4.4 and 5.2 of Zeng's Thesis 31

1. Zeng's Channel Models and Branch Metric Functions 31

a. Zeng's General Channel Model 31

b. Section 4.4 35

c. Section 5.2 38

2. Zeng's Acknowledgement and Treatment of Signal-Dependent Noise 42

C. Peer-Reviewed and Other Commentary on Work by Zeng and Inventors of CMU Patent 45

V. Summary of the Request and the Office Action 51

VI. Applicable Legal Principles 52

A. Claim Construction 52

B. Anticipation 53

C. Obviousness 53

VII. Zeng Does Not Anticipate Claim 4 …………………………………………………. 55

A. Zeng Does Not Disclose Any, Much Less a "Set," of Signal-Dependent Branch Metric Functions ……………………………………………………………... 55

1. Zeng's "Random Jitter" is White and Not Signal-Dependent …………….. 55

2. Zeng's Mathematical Notation Confirms that his Channel Model and BMFs are Not Signal-Dependent …………………………………………….. 56

3. Zeng's BMFs in Sections 4.4 and 5.2 are Blind to the Polarities of the Transitions ………………………………………………………………. 57

4. Zeng Uses the d=1 RLL Constraint in Sections 4.4 and 5.2, Which Eliminates the Signal-Dependent Neighborhood Effects from Zeng's BMFs …………………………………………………………………. 59

B. Even if Zeng's "Random Jitter" were Signal-Dependent, Zeng Still Does Not Disclose a Set of Signal-Dependent Branch Metric Functions …………………. 59

C. Drs. Lee and Moon Confirm that Zeng's Thesis Does Not Anticipate ………… 62

D. Marvell's Actions Both Before and During the CMU Case Confirm that Zeng's Thesis Does Not Anticipate ……………………………………………………….. 63

1. Marvell is the Requester ……………………………………………………. 63

2. Marvell's Actions in the CMU Case Confirm CMU's Assertion that Zeng Does Not Invalidate Claim 4 ………………………………………………… 64

a. Summary of the Litigation …………………………………………… 65

b. Despite Facing Substantial Damages and Raising the Zeng Thesis as an Anticipatory Reference During the Litigation, Marvell Decided Not To Use Zeng or His Thesis at Trial ……………………………… 67

3. Prior to Being Sued, Marvell's Engineers Copied Claim 4, Not Zeng's Thesis ………………………………………………………………… 70

4. The Evidence in the CMU Case Showed that Claim 4 is Very Valuable …. 74

a. CMU's Invention was "Must Have" for Marvell …………………….. 74

b. Before Introducing the MNP, Marvell's Sales Were Declining, But Afterwards Sales Spiked and Sales of Noninfringing Chips Dropped Almost Immediately to Zero …………………………………….. 75

c. CMU's Technology Became an Industry Standard …………………. 78

d. The Nexus Between Marvell's Copying and its Commercial Success ……………………………………………………………….. 78

e. The Method of Claim 4 Continues to be "Must Have" for Marvell …. 79

E. Zeng is Not Enabling and Therefore Does Not Qualify as Prior Art …………... 79

		1. Zeng's Detectors are Derived from a Channel Model that Violates Laws of Physics ...	80
		2. Section 5.2 of Zeng's Thesis Has Pervasive Mathematical Errors that Make the Detector of Section 5.2 Inoperative and Unusable ...	85
	F.	There is Insufficient Evidence that Zeng's Thesis is a Printed Prior Art Publication ...	87
VIII.	Claim 4 Would Not Have Been Obvious in View of the Cited References ...		88
	A.	None of the References Teach or Suggest a Set of Signal-Dependent Branch Metric Functions that Are Applied to a Plurality of Time Variant Signal Samples ...	88
	B.	A Person Having Ordinary Skill in the Art Would Not Have Been Motivated to Modify Zeng's Detectors in Sections 4.4 and 5.2 Based on Either Lee's Thesis or the Coker Patent ...	89
		1. Zeng's Thesis Should be Accorded Little Weight Because of its Technical Flaws ...	89
		2. Modifying a Detector Like Zeng's That Uses a d=1 RLL Constraint Defies Logic in Light of the Demand for Increased Data Density ...	90
		3. A PHOSITA Would Not Be Motivated to Modify Zeng's Asynchronous Trellis Detector Based on the Synchronous Trellis Detectors in Lee's Thesis and the Coker Patent ...	91
		4. A PHOSITA Would Realize that Zeng's Reported Simulation Results are Physically Impossible and Would Not Be Motivated to Modify Teachings of a Physically Impossible Device ...	93
		5. A PHOSITA Would Not Modify Zeng's Parameters Based on Either Lee's Thesis or the Coker Patent Because Zeng's Model is Mathematically Incompatible with Both Lee's Thesis and the Coker Patent ...	94
		6. It is Commercially Unacceptable and Impractical to Degauss the Disk Prior to Each Write as Zeng Does ...	95
	C.	Overwhelming Secondary Considerations Show that Claim 4 Would Not Have Been Obvious ...	96
		1. Copying ...	96
		2. Commercial Success ...	98
		3. Praise and Acclaim in the Industry ...	101
		4. Satisfaction of a Long-felt, Unresolved Need ...	104
		5. Failure by Others ...	104
IX.	Dr. Lee's Declaration Should Be Given Little Weight ...		107
X.	Written Statement of Interview ...		110

XI. Service on Requester 111
XII. Conclusion 111

REMARKS

I. INTRODUCTION

Contrary to the assertions of the Requester and its expert declarant, Dr. Inkyu Lee, Sections 4.4 and 5.2 of Zeng's Thesis do not anticipate claim 4 of U.S. Patent 6,201,839 ("the '839 patent"), nor render it obvious in view of the other references cited in the Request and Office Action. The Office should confirm its patentability for at least the following reasons:

First, claim 4 is for a method of "computing branch metric values for branches of a trellis of a Viterbi-like detector" that accounts for correlated signal-dependent noise through the use of a "set of signal-dependent branch metric functions" applied to a plurality of signal samples from different sampling time instances. In contrast, Zeng assumed away all the signal-dependent noise in his channel models in Sections 4.4 and 5.2 of his thesis, and Zeng's branch metric functions (BMFs) that are derived from those channel models do not account for signal-dependent noise. Zeng's BMFs ignore the polarity of transitions and the neighborhood of symbols in a sequence of written symbols (data written to the disk). In so doing, Zeng's BMFs operate in a manner exactly contrary to the CMU patents, which explain that signal-dependent BMFs must depend on the sequence (or neighborhood) of symbols written to the disk (*see* '839 patent at col. 5:49-64), including being sensitive to the polarity of the transitions (*see* '839 patent at col. 3:54-64).

Second, squarely contradicting his opinions in this reexamination, Dr. Lee criticized Zeng's channel models in the 1990s, first in a peer-reviewed paper and later in his Ph.D. thesis, stating in those earlier papers that Zeng's "random jitter" is "***white***," which is mutually exclusive of signal-dependent noise. *See* Ex. 1 ("Lee '92 paper") at 963[1] and Exhibit E to the Request (Lee's Ph.D. thesis) at 99-100. A brief comparison of Dr. Lee's statements in his 1992 paper and his statements in his reexamination declaration is shown below.

[1] "Ex. __" herein refers to exhibits attached the Declaration of Mark G. Knedeisen in Support of CMU's Response to Office Action, filed herewith.

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 $\pm$ 1. Assume $h(t)$ to be the response of the channel to a single transtion. Then,

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

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

Zeng Paper

Zeng's Channel Model Eq. 4.1

Random jitter is "white"

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

Ex. 1 (Lee's 1992 paper) at p. 963; *See also* Lee 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

Also, Zeng's own thesis advisor, Prof. Jaekyun Moon, later confirmed Dr. Lee's criticism of Zeng's work, writing in two separate papers that the CMU invention -- and not Zeng -- was the first detector that accounted for signal-dependent noise, i.e., "noise characteristics [that] depend highly on the local bit patterns." Ex. 2 at 730; *see also* Ex. 3 at 101-12. These statements by Drs. Lee and Moon confirm CMU's position that Zeng's Thesis is not invalidating art.

Third, claim 4 requires a "set" of signal-dependent branch metric functions, i.e., more than one, applied to a plurality of signal samples from different sampling time instances. Even giving all benefits of the doubt to the Requester and Dr. Lee on the foregoing arguments about the effect of Zeng's "random jitter" terms, ***only one of Zeng's functions*** in each of Sections 4.4 and 5.2 contains a term related to Zeng's "random jitter" term, which the Requester and Dr. Lee assert makes the functions signal-dependent. This single function cannot constitute the "set" required by claim 4.[2]

Fourth, the '839 patent is involved in ongoing litigation (referred to herein as the "CMU case") and the actions of the defendants (including their team of scientists and lawyers with greater than ordinary skill) confirm that Zeng's Thesis does not invalidate claim 4. CMU asserted claim 4 of the '839 patent and a claim of a related patent, specifically claim 2 of U.S. Patent 6,438,180 ("the '180 patent")[3], against Marvell Technology Group, Ltd. and Marvell Semiconductor, Inc. (collectively "Marvell"). Judge Nora Barry Fischer presided over a trial that was held in this matter in November and December 2012. Following the 4-week trial, the jury unanimously concluded that Marvell willfully infringed the Asserted Claims, and that Marvell failed to prove that the Asserted Claims were invalid. The jury awarded CMU $1.169 billion in damages, one of the largest patent infringement verdicts ever. In post-trial orders, Judge Fischer increased the damages to approximately $1.536 billion "to penalize Marvell for its egregious behavior…." Ex. 4 (Dkt. 933) at 45.

In the context of this reexamination, there is strong evidence that Marvell is the true proponent of the reexamination request, yet Marvell's behavior during the CMU case confirms that Zeng's Thesis does not anticipate. Marvell produced a copy of Zeng's Thesis from its files at the outset of the litigation, and identified Dr. Zeng (who was and is a Marvell employee) as a person with relevant knowledge of the prior art. Marvell's technical invalidity expert expressly reviewed Zeng's Thesis in preparing his expert report but did not opine in his expert report that Zeng's Thesis invalidated claim 4 in any way. Moreover, Marvell never called Zeng to testify at trial or otherwise that his work anticipated or otherwise invalidated claim 4. Further demonstrating that Zeng's Thesis does not anticipate claim 4, when Marvell set out to design its

---

[2] Both the Requester and Dr. Lee agree that a "set" of signal-dependent BMFs must contain "more than one" signal-dependent BMF. *See* Request at 20; Lee Dec. at ¶13.

[3] The '180 patent is a continuation-in-part of the '839 patent. The inventors for the '839 and '180 patents are the same, Prof. Aleksandar Kavcic and Prof. José Moura. Collectively the two patents are referred to herein as the "CMU patents" and the two asserted claims are referred to as the "Asserted Claims."

first signal-dependent detector in the early 2000s, its engineers studied Zeng's work, but instead copied the CMU patents. Marvell even named its first signal-dependent detector "KavcicPP" after Prof. Kavcic, the first-named inventor for the CMU patents. Marvell's conscious decision to ignore Zeng and his thesis when contesting validity in a billion-dollar trial, and its decision to knowingly copy the CMU patents despite studying Zeng's work, is powerful evidence that neither Marvell nor its experts believe that Zeng invalidates the Asserted Claims.

Fifth, Sections 4.4 and 5.2 of Zeng's Thesis are not enabling. They are based on a channel model that defies the laws of physics in that Zeng's model permits consecutive positive or consecutive negative transitions and the parameters of the branch metric functions (BMFs) of Zeng's detectors in Section 4.4 and 5.2 cannot be set. This defect in Zeng's model is manifested in his reported simulation results, which also defy the laws of physics since the detector's performance does not deteriorate with increasing noise. On top of that, there are so many mathematical errors in Zeng's equations in Section 5.2 that the equations in that section are unusable in a detector. The Requester and Dr. Lee did not discuss the physics-defying channel model, the physically impossible simulation results, or the pervasive mathematical errors; to the contrary, the Requester and Dr. Lee quoted the erroneous equations verbatim in their papers. But the physics-defying channel models and pervasive errors make Zeng's detectors in Sections 4.4 and 5.2 inoperative even to a person having more than ordinary skill in the art. Therefore, these sections of Zeng's Thesis are nonenabling and neither constitutes an anticipatory reference. Nor would a person having ordinary skill in the art be motivated to modify Zeng's detectors in these sections because of the fatal errors.

Sixth, claim 4 is not obvious because none of the cited references (i.e., Zeng's Thesis, Lee's Thesis, or the Coker patent)[4] teach or suggest "a set of signal-dependent branch metric functions" that are applied to a plurality of signal samples from different sampling time instances. Moreover, the Requester and Dr. Lee failed to undertake the proper obviousness analysis. The Requester and Dr. Lee did not address the numerous reasons that a person skilled

[4] Zeng's Thesis is Exhibit D of the Request; Lee's Thesis is Exhibit E; and the Coker patent is Exhibit F.

in the art would not modify Zeng's non-functioning detectors in Sections 4.4 and 5.2 of the Zeng Thesis based on either Dr. Lee's Thesis or with the Coker patent. Furthermore, the trial record in the CMU case (among other things) contains compelling evidence of secondary considerations of nonobviousness that the Requester and Dr. Lee failed to disclose despite having access to the trial record. These secondary considerations include:

- Marvell copied the CMU patents -- not Zeng's Thesis -- even though Marvell's engineers studied Zeng's work at the time and even though Zeng is an employee of Marvell;
- Marvell named its infringing products after Dr. Kavcic because, according to one of Marvell's own engineers and its current Chief Technology Officer, "it's common practice" to name something after the author who discovered the solution, just like "Gaussian noise" is named after Gauss and the "Viterbi detector" is named after Dr. Viterbi;
- Marvell realized approximately ***$5 billion in profit*** (and $10 billion in total revenue) over nine years (2003-2012) from sales of HDD read channel chips specifically designed to copy of the CMU patents; it cited the "KavcicPP" circuit as one of the few technologies that made it the "market leader"; and it continues to keep the infringing circuit in its products even after the court in the CMU case found that Marvell willfully infringed the CMU patents and ordered Marvell to pay 50 cents per infringing chip on an on-going basis through the term of the CMU patents;
- After Marvell commercially introduced its infringing "KavcicPP" detector to the market, the sales of its prior-generation, noninfringing detector ceased almost immediately because the industry demanded the solution provided in the CMU patents and after more than ten years it is still "a must";
- Similar to Marvell's internal acclaim for the CMU invention, others in the field, including Zeng's own thesis advisor, Prof. Jaekyun Moon, recognized the CMU inventors -- not Zeng -- as "first deriving" the solution; and

- Others, including Marvell, were working on the problem of improving detector performance in view of increasingly dominant media noise in magnetic recording channels at the same time, but they all failed.

Finally, Dr. Lee's declaration is so unreliable that the Office should not and cannot reasonably rely upon it or the Request.[5] For example, Dr. Lee's declaration contradicts his prior technical writings with no explanation for the reversal in positions; ignores well-known, universally-accepted mathematical principles; did not discuss the pervasive math errors in Zeng's Thesis; relies on demonstrably faulty logic; is internally inconsistent; and relies on out-of-context and irrelevant statements in Zeng's Thesis to support his erroneous assertions.

II. CONCURRENTLY FILED SUPPORTING DECLARATIONS

In addition to the declaration referred to above (*see* footnote 1), this response is being filed concurrently with four other supporting declarations:

- A declaration from Prof. Steven W. McLaughlin (referred to herein as "McLaughlin Dec."). Prof. McLaughlin is the Chair of the School of Electrical and Computer Engineering at The Georgia Institute of Technology ("Georgia Tech.") and one of the world's leading experts in the areas of communications and information theory with decades of experience working with Viterbi detectors.

- A declaration from Dr. Christopher H. Bajorek ("Bajorek Dec."). Dr. Bajorek managed IBM's HDD business unit and Komag's disk business for many years, and is one of the world's leading experts on magnetic recording and the HDD industry.[6]

[5] As shown below at page 64, Dr. Lee has significant contacts with Marvell. His thesis advisor was Dr. John Cioffi, a former Marvell board member who guided Marvell's IPO, and he was studying with Dr. Cioffi at the same time as Marvell's current CTO, Dr. Zining Wu.

[6] After this Response, CMU recommends reviewing Dr. Bajorek's declaration because it serves as a foundation for the other declarations.

- A declaration from Prof. Aleksandar Kavcic ("Kavcic Dec."). Prof. Kavcic is the first named inventor of the '839 patent and currently is a professor at the University of Hawaii in the Department of Electrical Engineering.

- A declaration from Prof. José Moura ("Moura Dec."). Prof. Moura is the other named inventor of the '839 patent and is currently a professor at Carnegie Mellon University in Pittsburgh, Pennsylvania in the Department of Electrical and Computer Engineering.

III. NOTICE OF CONCURRENT PROCEEDINGS

As mentioned above, the '839 patent is involved in pending litigation (the CMU case), captioned *Carnegie Mellon University v. Marvell Technology Group, Ltd. et al.*, No. 2:09-cv-00290-NBF (W.D. Pa.). CMU asserted that Marvell Technology Group, Ltd. and Marvell Semiconductor, Inc. (collectively "Marvell") infringed claim 4 of the '839 patent and claim 2 of the '180 patent. *See* Ex. 5 (an updated copy of the docket from the CMU case).[7] The litigation is summarized below in Section VII.D of this response.

The same Requester that requested reexamination of claim 4 of the '839 patent, J. Steven Baughman of Ropes & Gray, also requested ex parte reexamination of claim 2 of the '180 patent on the same day (January 21, 2014) that he requested reexamination of claim 4 of the '839 patent. The ex parte reexamination of the '180 patent has been assigned Control No. 90/013,124. The Office granted the Request for the '180 patent and issued a first, non-final Office Action on July 31, 2014.

[7] References to "Dkt." herein refer to docket entries in the CMU case. References to "P-" refer to a CMU trial exhibit in the CMU case and references to "P-Demo" refer to a trial demonstrative used by a CMU witness in the CMU case. References to "Tr." filed by a date refer to trial transcripts in the CMU case on the indicated date.

IV. BACKGROUND OF THE RELEVANT TECHNOLOGY

A. Summary of the Invention Described in the '839 Patent

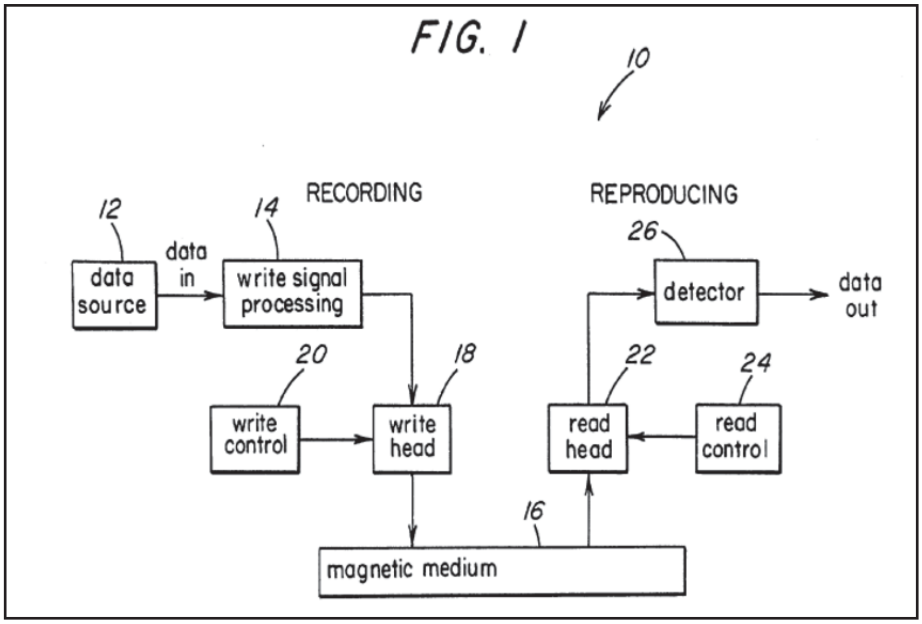
Claim 4 is directed to a "method of determining branch metric values for branches of a trellis for a Viterbi-like detector." To facilitate a discussion of the claimed invention, background on hard disk drives, Viterbi sequences detectors, and branch metric functions is provided below.

1. Hard Disk Drives

Data in a hard disk drive (HDD) are stored on disks that are organized in concentric tracks and coated with magnetic material. Arranged linearly along each track are tiny regions or cells that are used for storing the data through magnetic polarization of the regions. A "read/write" head is used to read data from and write data to the disk. *See* Bajorek Dec. at ¶¶ 31-35.

a. Writing Data to the Disk

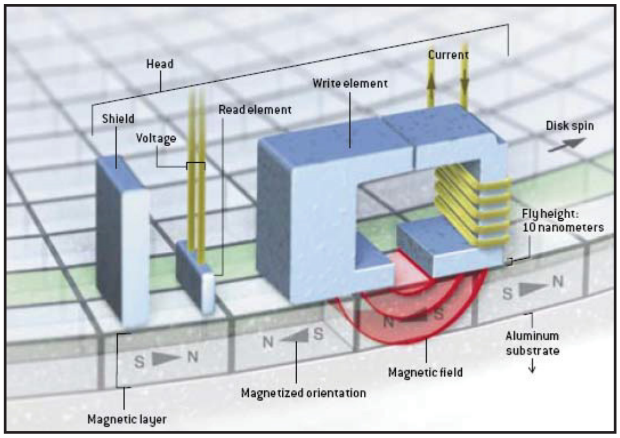
When user data are to be written by, for example, a computer to the disk, the data are typically modified prior to writing the data to the disk. First, the user data are typically encoded with an error correction code to protect the data against commonly encountered error modes, such as disk defects. Second, a separate processor converts the error-correction-encoded user data into an analog signal that is applied to the write head, in accordance with a waveform encoding technique or a modulation scheme. These operations can be performed by the "write signal processing" block 14 in Figure 1 of the '839 patent, reproduced below. The write head 18 records the resulting signal on the magnetic medium by magnetically polarizing the regions on the disk in accordance with the signal. *See* Bajorek Dec. at ¶ 36.



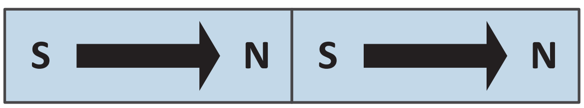
Each polarized region on the magnetic recording layer of the disk has a magnetic polarization that, once written by the write head, is oriented in a particular direction. The magnetic polarity of these regions can be changed from one direction to its opposite by the write head in order to write the data to the disk. The write head produces a variable magnetic field used to magnetize the individual regions of the recording layer. By varying the polarity of the magnetic field emitted by the write head, the magnetic polarity of the individual regions can be set as needed. Once the write head sets the magnetic polarity of a particular polarized region, that region's magnetic polarity is retained until it is overwritten later by the write head with new data. *See* Bajorek Dec. at ¶ 37.

In so-called "longitudinal magnetic recording" (LMR), illustrated in the diagram below, the regions on the disk are polarized to the left or to the right (as shown by the "N" and "S" references reflecting the positive and negative polarity of each individual region), in the plane of the magnetic layer. *See* Bajorek Dec. at ¶ 39.[8]

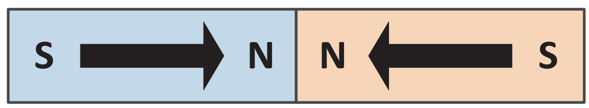
[8] The illustration is from Mark Fischetti, "Going Vertical," Scientific American, 2006. In the early-to-mid 2000s, the HDD industry migrated to so-called "perpendicular magnetic recording" (PMR), where the bit regions are polarized perpendicular to the plane of the magnetic layer. *See*



As shown in the above diagram, the polarized regions can be conceptualized as bar magnets having north (N) and south (S) ends. Adjacent polarized regions can have the same polarization or opposing polarizations, such as shown below:



SAME POLARIZATION

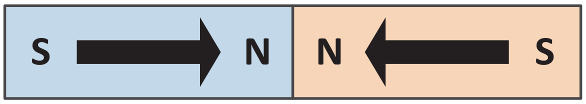


OPPOSING POLARIZATION

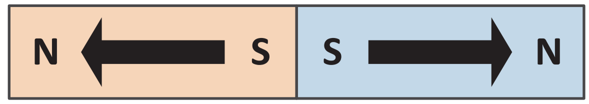
Bajorek Dec. at ¶ 40.

When adjacent polarized regions are magnetized in opposing directions, there is a "transition" in the polarity of the regions that can be detected by the read head. Necessarily, there are two types of transitions. In LMR, in one type of transition the N ends of the bar magnets are abutting, and in the other type of transition the S ends of the magnets are abutting, as shown below. The first type can be considered a "positive transition" and the second type can be considered a "negative transition."

Bajorek Dec. at p. 12, footnote 11. This response uses LMR to illustrate the relevant concepts without loss of generality.



Positive Transition

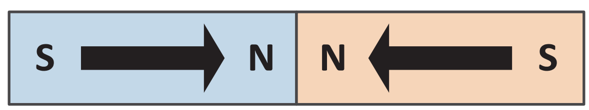
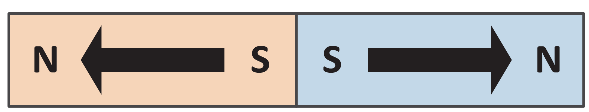
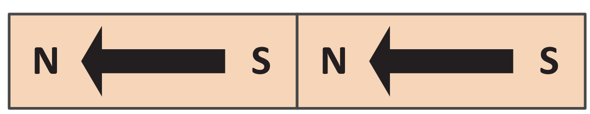
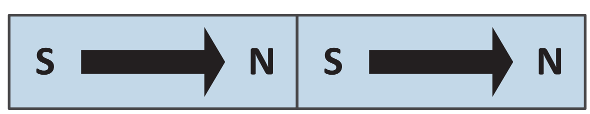


Negative Transition

See Bajorek Dec. at ¶ 41.

A second positive transition cannot immediately follow a first positive transition, and a second negative transition cannot immediately follow a first negative transition. Instead, a negative transition must follow a positive transition (with any number of non-transitions therebetween) and a positive transition must follow a negative transition (with any number of non-transitions therebetween). *See* Bajorek Dec. at ¶ 42.[9]

To account for and explain the differences between the polarity of the transitions, the CMU patents disclose and adopt a specific polarity-sensitive nomenclature using the following notation to illustrate the four possible sequences of two consecutive polarized regions or data ***symbols*** (reading the polarized regions from left to right):

+ indicates a positive transition	
− indicates a negative transition	
⊕ indicates a non-transition whose nearest proceeding transition was a positive transition	
⊖ indicates a non-transition whose nearest proceeding transition was a negative transition	

[9] Dr. Lee concedes this point. Lee Dec. at ¶ 25.

See ‘839 patent at col. 3:51-60. As specifically set out in the CMU patents, ignoring the polarity of transitions, i.e., simply treating all “transitions as ‘1’s and [all] no transitions as ‘0’s [as Zeng does] is blind to signal asymmetries…, ***which is inappropriate for the present problem***.” ‘839 patent at col. 3:60-63 (emphasis added); *see also* Bajorek Dec. at ¶ 43.

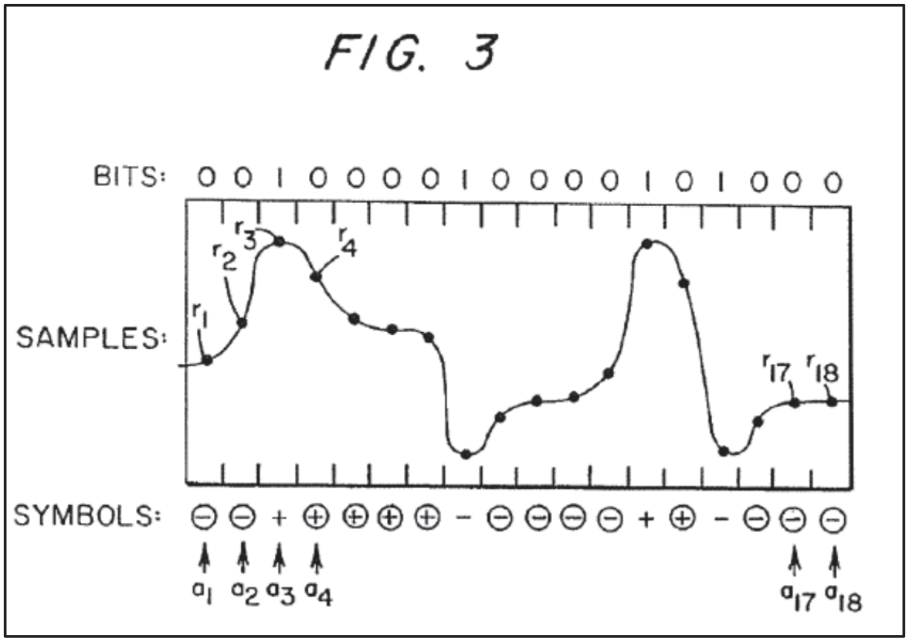
b. Reading Data from the Disk

In the reading operation, the read head (component 22 in Fig. 1 of the CMU patents), flying above the rotating disk, “retrieves the variations in the magnetic flux that are stored on the medium,” *see* ‘839 patent at col. 3:14-15, and converts the sensed magnetic fields into a continuous, analog electrical signal called the “readback signal.” Modern read heads use magnetoresistors to convert the variations in the magnetic flux from the disk to the readback signal. *See* Bajorek Dec. at ¶ 48. The sensed magnetic fields change the resistance of the magnetoresistor, which changes the voltage of the readback signal, which can vary between positive and negative voltages. For example, a positive transition on the disk can induce a positive pulse in the readback signal and a negative transition can induce a negative pulse. *See* Bajorek Dec. at ¶ 49.

The function of the detector (component 26 in in Fig. 1 of the CMU patents) is to determine the symbols written to the disk based on the readback signal. *See* Bajorek Dec. at ¶ 48. Early HDDs used “peak” or “threshold” detectors; if the signal was above or below a threshold voltage, a transition was detected. *See* McLaughlin Dec. at ¶ 19; Moura Dec. at ¶ 26. Peak detection works adequately only at low data densities when the intersymbol interference (ISI) effects are small. *See* McLaughlin Dec. at ¶ 20; Moura Dec. at ¶ 27. Beginning in the early 1990s, more sophisticated detectors were required to accommodate increasingly higher data storage densities. The predominant type of sophisticated detector that replaced peak detectors is the so-called “partial response maximum likelihood” or “PRML” detector. PRML detectors are digital sequence detectors that use digitized samples of the readback signal and some form of the Viterbi algorithm to detect the data read from the disk. *See* McLaughlin Dec. at ¶ 20; Moura Dec. at ¶ 27; Bajorek Dec. at ¶¶ 50-51.

2. Viterbi Sequence Detection

In the process of reading data from the disk, there are two sets of data streams: known and unknown. These two sets can be explained in connection with Figure 3 of the CMU patents, reproduced below.



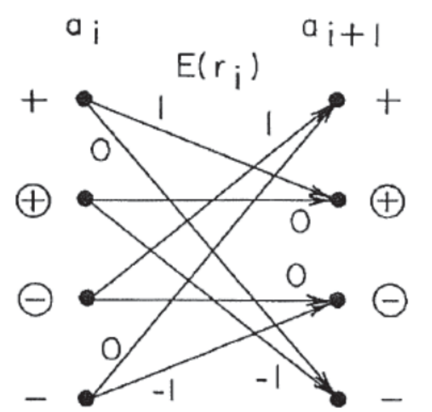
The "known" stream of information consists of the digitized samples of the readback signal, i.e., the $r_1,\ldots, r_{18}$ values in Figure 3. The "unknown" stream of information consists of the specific sequence of symbols written to the disk, represented by the *a* terms, i.e., $a_1, \ldots, a_{18}$. *See* '839 patent at col. 2:64 - col. 3:2. The function of the Viterbi detector (a detector that uses a form of the Viterbi algorithm) is to determine the most likely sequence of the *a* terms[10] (the symbols written to the disk) based on the signal sample values (i.e., the *r* terms) picked up by the read head that is flying over the surface of the disk. *See* McLaughlin Dec. at ¶ 21.

The Viterbi detector uses a trellis to detect the recorded symbols, where each specific symbol sequence (i.e., a "neighborhood" of symbols a_k) corresponds to a different branch (or

[10] From the *a*-stream, the data bits (the "1's" and "0's") of the bit stream can be determined. *See* McLaughlin Dec. at note 5.

sequence of branches) of the trellis. *See e.g.*, '839 patent at col. 10:26-34.[11] Figure 4 of the CMU patents illustrates an example trellis, in this case one cell of a PR4 trellis,[12] having 4 states (or nodes) at each time instance and two branches leaving and entering each state/node, for a total of 8 branches at a time instance of the trellis. *See* '839 patent at col. 2:49 and col. 10:26-34; *See* McLaughlin Dec. at ¶ 22.

FIG. 4



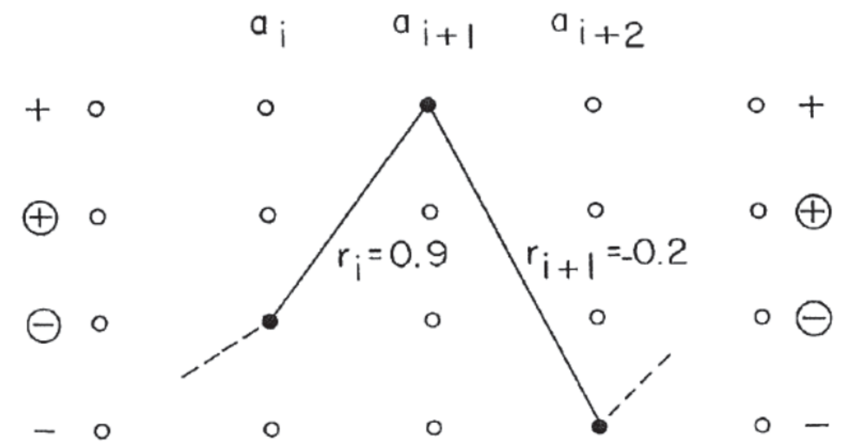
To detect the sequence of symbols written to the disk, the Viterbi detector computes so-called "branch metric values" for branches of the trellis, and sums the branch metric values for branches along a path through the trellis to compute a so-called "path metric value." Figure 5 of the CMU patents, reproduced below, shows an example of a path through a trellis, where the path is made up of a series of branches end-to-end. *See* McLaughlin Dec. at ¶ 24.

---

[11] Dr. Lee concedes this point. Lee Dec. at ¶ 29 ("one of eight different transitions, each of which is represented by a different branch in the trellis").

[12] PR4, or Partial Response Class IV, is a channel that is characterized by a D-transfer function of $(1\text{-}D^2)$. *See* McLaughlin Dec. at p. 9, footnote 6.

FIG. 5



The branch metric value for a branch is the measure of the distance (or error) between the signal sample value(s) (i.e., one or more *r* readings) and the expected (or "target") value(s) for the branch. These "target" or "expected" values are the ideal signal values that would be received by the detector for a particular branch of the trellis (a sequence of two symbols) if there were no noise in the read-back signal. *See* '839 patent at col. 8:48-50 ("[f]or example, in a PR4 system, the signal samples, if there is no noise in the system, fall on one of three target values 1, 0 or -1"); *see also* col. 8:59 - col. 9:17. A branch metric value for a branch is computed using a "branch metric function," described in the subsequent section. *See* McLaughlin Dec. at ¶ 25.

A "path metric value" is the sum of the branch metric values for the branches along a path through the trellis, so a path metric value can be thought of as the cumulative distance (error) for the sequence of branches along a particular path. The symbol sequence corresponding to the path with the best (lowest) path metric value is the detector's determination of the most likely sequence of symbols (a_1, a_2, …) written to the disk. *See* McLaughlin Dec. at ¶ 26.

3. Background on Branch Metric Functions

In mathematics, a distance function (also known sometimes as a "metric") is a function used to measure (or "score") the distance between two points in a space (actual or logical).

Distance functions have two inputs corresponding to the two points in the space for which the distance is being measured. To determine the "distance," ***two input values are required*** so that the distance between them can be computed using a metric or distance function. Thus, a metric or distance function is a two-input function. *See* McLaughlin Dec. at ¶ 27.

Correspondingly, branch metric functions (BMFs), like those disclosed in the CMU patents, are a species of distance functions that assign a value (i.e., a "branch metric value") to a branch of a trellis. At its most basic, the "distance" being measured is the distance between the digital value of a given signal sample at a sampling time instance (*r*) on the one hand and the "target value" (ideal, noise free value) for each given branch of the trellis at the sampling time instance on the other hand. *See* McLaughlin Dec. at ¶ 28.

To explain further, the two inputs for a BMF are (i) the signal sample *r* (or set of signal samples if multiple signal samples are used) and (ii) the target value (or set of target values) for the signal sample(s) for a given branch of the trellis (which corresponds to a sequence of symbols). Thus, a BMF scores the difference between the *r* value(s) and the target value(s) for that branch. In the context of Viterbi detection for HDDs, when the "target" (or ideal, noise free) values for a branch of the trellis are subtracted from the signal samples, the BMF is applied to that difference (e.g., squaring the difference), and the resulting branch metric value can be thought of as a measure (score) of the noise for that postulated branch. Seen this way, a path metric is the cumulative score of the noise structure for that given path, and the path through the trellis with the lowest aggregate score (the lowest path metric value) is determined by the detector to correspond to the correct (most likely) sequence of symbols (a's) actually written on the disk. *See* McLaughlin Dec. at ¶ 29.

Prior to the '839 patent, a common branch metric function used by Viterbi detectors in magnetic recording channels was the Euclidean branch metric function. This function is

$$M_{i,t} = (r_t - m_i)^2$$

where $M_{i,t}$ is the branch metric value at time t for branch i, r_t is the signal sample value for time t, and m_i is the target value for branch i. *See* '839 patent at col. 5:59-col. 6:14 (including Equation 8); Moura Dec. at ¶ 32.

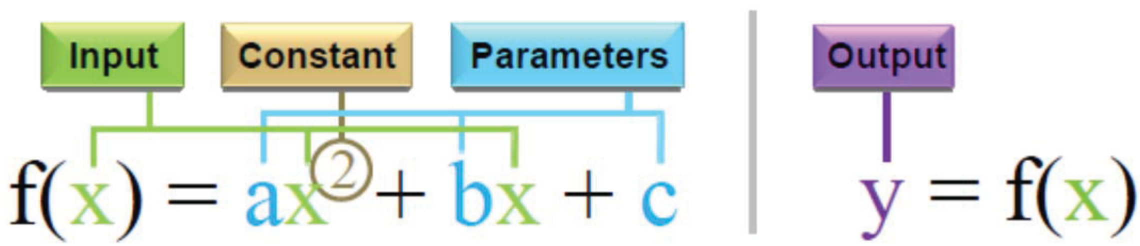
It is well known and universally accepted in mathematics that a single equation can represent a single "function." It is also well-known and universally accepted in the context of mathematical functions that changing an "input" to a function does not change the function itself (i.e., does not convert the single function into another function). *See* McLaughlin Dec. at ¶ 32. For example, with the function $f(x) = 2x$, the "input" is x and the function remains the same regardless of whether x takes on different values. Similarly, for a two-input function like the Euclidean metric, the function does not change when one or both of the inputs changes (e.g., either the signal sample r_t or the target value m_i). The Euclidean branch metric at time t for branch 1 is $M_{1,t} = (r_t - m_1)^2$ and for branch 2 it is $M_{2,t} = (r_t - m_2)^2$. These two mathematical expressions represent the ***same function*** because the target values m_1 and m_2 are "inputs" to the single function. *See* McLaughlin Dec. at ¶ 30.

A single equation, however, can represent a ***set*** of different functions where a "parameter(s)" of the equation can assume different values. For example, the equation $f(x) = Ax$ represents a set of parameterized functions, one for each value of the "parameter" A. In other words, $f(x) = 2x$ is a different function than $f(x) = 3x$. *See* McLaughlin Dec. at ¶ 31.

In the CMU case, Prof. Gilbert Strang, a professor at the Massachusetts Institute of Technology (MIT), filed a declaration elaborating on how to determine if a particular equation represents a single function or a set of functions. *See* Ex. 6 (Dkt. 325, Ex. 1). Prof. Strang[13] explained that there are four different symbol types that can be used in an equation to describe a mathematical function: ***inputs*** (sometimes called arguments); ***outputs***; ***constants***; and ***parameters***. *See* Ex. 6 at ¶ 11. Prof. Strang explained that "parameters are not constants, because parameters can change; that is their purpose. A change in the parameter causes the

[13] Dr. Strang's declaration lists his qualifications, which include teaching linear algebra (among other things) at MIT for more than 50 years, and having his lectures taped for MIT's OpenCourseWare project and viewed more than 2,000,000 times. *See* Ex. 6 at ¶ 3.

function to change." *Id.* To illustrate, Prof. Strang identified the input, output, constant, and parameters in the familiar quadratic equation (see below). *See id.* at ¶¶ 16-18. Dr. Strang explained that in this equation, "[t]he symbols 'a', 'b', and 'c' denote parameters in the quadratic equation …. Here, because 'a', 'b', and 'c' can vary, it is clear they are parameters." *Id.* at ¶ 18. Consequently, the quadratic equation is a "parameterized" family or "set" of multiple functions. *See id.* at ¶¶ 7-21.[14]



4. Complications from Noise

The read operation, however, is not as simple as detecting the presence of transitions from the magnetic fields emanating from the polarized regions on the disk as the disk's tracks move beneath the read head. Noise in the system, which manifests itself in the readback signal, complicates the read process. *See* Bajorek Dec. at ¶ 52. The Euclidean branch metric function, which was used in the early Viterbi detectors, assumes that the only noise in the system was additive white Gaussian noise. *See* '839 patent at col. 1:34-37; col. 5:59 - col. 6:13; Moura Dec. at ¶ 33. However, as data density increases, media noise in the system increases and becomes dominant. *See* Moura Dec. at ¶¶ 16 and 33; Bajorek Dec. ¶ 53; Ex. 8 (P-607) (email from Marvell engineer explaining that "now days the drives are dominated by media noise").

[14] In the CMU case, Marvell agreed that a single equation can represent one function while a different equation can represent a "set" of functions. *See* Ex. 7 (Dkt. 333) at slide 3. Marvell also agreed that when an equation contains "parameters," a change in the value of a parameter results in there being a new function arising from the same equation. *See* Ex. 7 (Dkt. 333) at slides 6-10. In fact, in the CMU case, Judge Fischer ruled that changing an "input" of a branch metric function (e.g., a "target value") does not change the underlying function into a new, second function. *See* Ex. 10 (Dkt. 337) at 10-17.

Media noise is noise in the readback signal that arises from fluctuations in the medium magnetization due to (1) the underlying polarity of the media at the time data are written to the disk, (2) the differing polarities of transitions written to the disk, and (3) the specific number and density of transitions in a given region of the disk. *See* Bajorek Dec. at ¶ 53. Two important, interrelated concepts about media noise are (1) the noise in the signal samples is ***correlated*** and (2) the correlated noise is ***signal-dependent***, such that it has a different noise structure for each specific sequence of symbols written to the disk. *See* Moura Dec. at ¶ 16; Bajorek Dec. at ¶¶ 71-73.

a. Correlated Noise

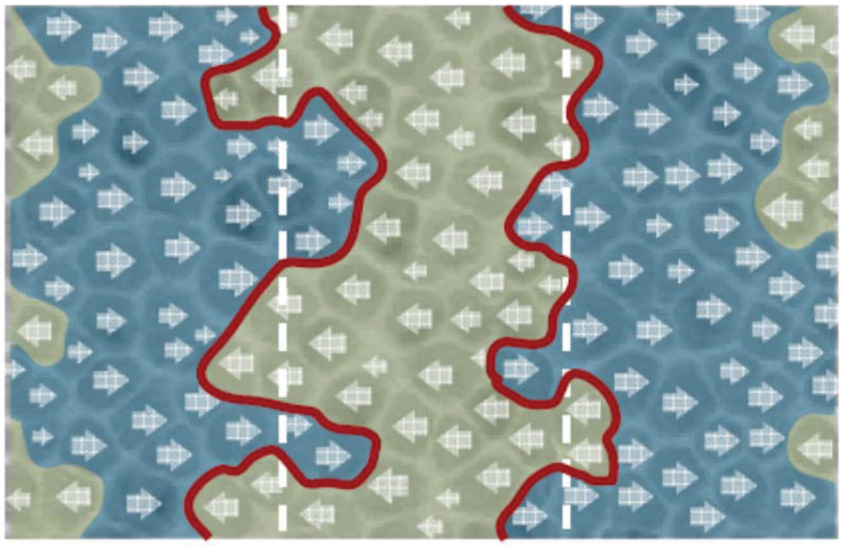
The '839 patent explains that there is "correlation between noise samples in the readback signal" and that these "correlations arise due to noise coloring by front-end equalizers, media noise, media nonlinearities, and magnetoresistive (MR) head nonlinearities." '839 patent at col. 1:57-61; *see also* Bajorek Dec. at ¶ 73; Moura Dec. at ¶ 33.

b. Signal-Dependent Noise

The amount of correlation between noise samples also depends on the specific pattern of symbols written to the disk. *See* Bajorek Dec. at ¶ 73. That is, the noise in the readback signal samples is attributable to the specific sequence of recorded symbols. The '839 patent explains that the branch metric function for a branch of the Viterbi trellis is "dependent on the postulated ***sequence of written symbols*** $a_{i-K_l}, \ldots, a_{i+L+K_t}$, which ensures the ***signal-dependence*** of the detector." '839 patent at col. 5:47-52 (emphasis added). In other words, each specific postulated symbol sequence, e.g., each trellis branch, has a different signal-dependent noise structure. *See* Moura Dec. at ¶ 19; Bajorek Dec. at ¶ 71.

Dr. Bajorek's declaration describes the predominant causes and effects of signal-dependent media noise, including that:

- ***First***, the transition between oppositely polarized regions on a track of the disk is jagged, as shown in the diagram below, due to the varying locations, shapes and sizes of the magnetic grains that make up the polarized regions. *See* Bajorek Dec. at ¶ 56. These jagged transitions affect the location and shape of the positive and negative pulses in the readback signal. *Id.* at ¶ 57.



TRANSITION TYPE: + −

- ***Second,*** the shift in the location of the pulses, referred to as "peak shift," is always influenced by the specific pattern or neighborhood of symbols written to the disk. *Id.* at ¶ 59. That is, the particular neighborhood of polarized regions written to the disk (including the polarity and number (density) of any transitions within the neighborhood) affects both the extent and direction of peak shift. *Id.* In other words, the degree and extent of the jaggedness (and hence the peak shift and shape) depends directly upon the polarity of the transition in question and the neighborhood of symbols. *Id.* at ¶¶ 61-65.

- ***Third***, the MR read head responds asymmetrically to positive and negative fields from the disk and, therefore, magnifies the signal-dependent distortions in the readback signal because of the read head's nonlinear response. This asymmetry causes two distortions in the readback signal: (1) a difference in amplitude between positive and negative fields; and (2) a difference in resolution -- that are both dependent on signal polarity (e.g., a positive transition versus a negative transition). These distortions further affect the positions (shift) and shapes (amplitude and sharpness) of the peaks in the readback signal. *See* Bajorek Dec. at ¶ 66.

- ***Fourth***, there are two demagnetizing fields emanating from the disk itself that contribute to the polarity- and neighborhood-dependence of the peak shifts. One is the underlying sequence of polarized regions on the disk existing at the time new data are written to the disk, which each produce a demagnetizing field. The other is the demagnetizing fields from the regions just written (as part of writing the new data). The demagnetizing field from a previously written transition (or even a neighborhood of previously written transitions) can interfere with the magnetic field from the write head during the write process. These demagnetizing fields result in so-called "hard" and "easy" transitions. *See* Bajorek Dec. at ¶ 67. A "hard" transition occurs when, for example, a region is written that is against the pre-existing magnetic orientation of the medium. Conversely, an "easy" transition occurs when a region to be written coincides with the direction of the pre-existing magnetization of medium. "Hard" transitions require more head field to saturate the magnetic orientation in the gap region of the head due to increased demagnetization fields, and these hard and easy transitions are further affected by the just-written data. *See* Bajorek Dec. at ¶¶ 68-69.

- ***Fifth***, there is even signal-dependent noise associated with sequences of non-transitions due to DC erase noise, baseline drift, and the demagnetizing fields (which lead to hard/easy transitions). Moreover, the noise structure for a stretch of positively polarized regions (e.g., a sequence of ⊕ regions) is different from a stretch of negatively polarized regions (e.g., a stretch of ⊖ regions). *See* Bajorek Dec. at ¶ 70.[15]

Below are a series of figures that illustrate the various signal-dependent noise effects. These two figures, Figures 1 and 2, each use the same symbol sequence but because the placement of the transitions on the disk depends on multiple factors (*see* Bajorek Dec. at ¶ 60), the readback signal resulting from the same sequence is not always the same. Figure 1 shows the peak shift and peak amplitude variation (rises for positive peaks and drops for negative peaks) in

[15] Zeng concedes this point. Zeng Thesis at 10 ("***more*** media noise are observed near transition regions than saturation regions," acknowledging that there is media noise in the saturation regions).

the modeled readback signal because of hard and easy transitions, as well as the polarity of the transitions. *See* Bajorek Dec. at ¶ 82.

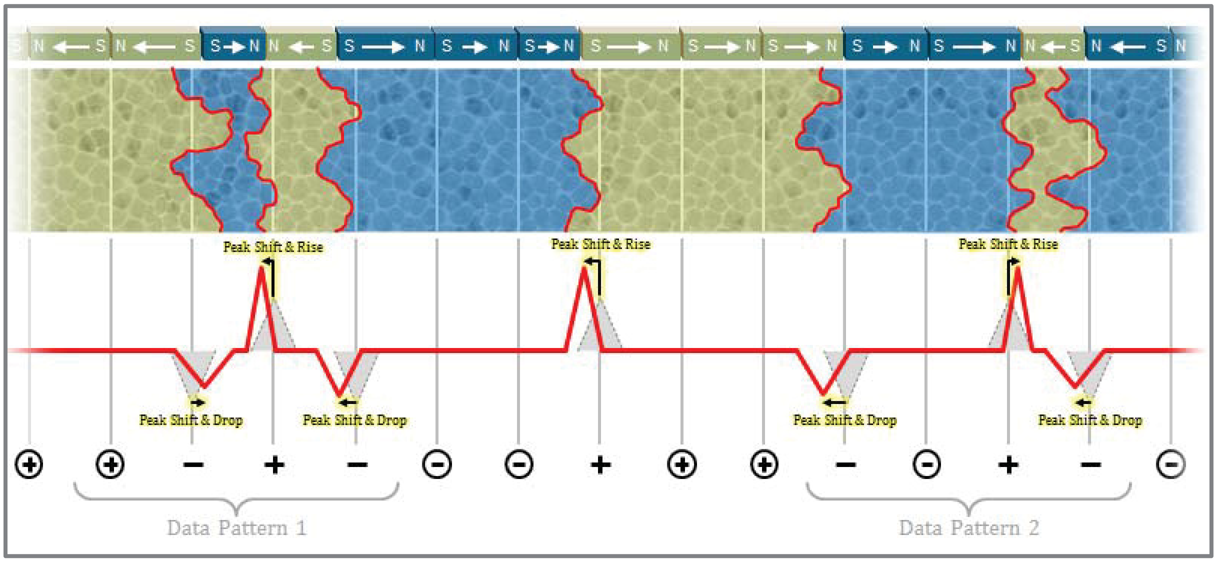


Figure 1

Figure 2 below shows an example of the modeled readback signal for the same symbol sequence, but this time with more pronounced transition interaction (e.g., percolation) and other neighborhood effects. *See* Bajorek Dec. ¶ 83.

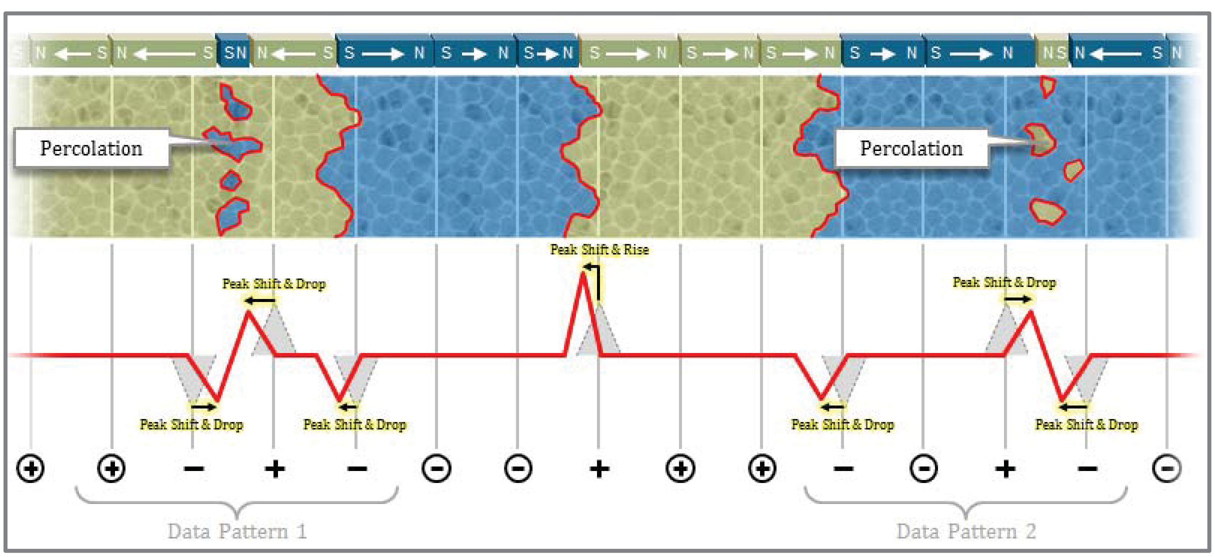


Figure 2

Figure 3 below illustrates these two readback signals placed on top of each other, along with several other possible readback signals for the same data sequence, each represented by a different color. These waveforms show that there is variation in the peak positions and size due to both the neighborhood effect and the polarities of the transitions. *See* Bajorek Dec. at ¶ 84.

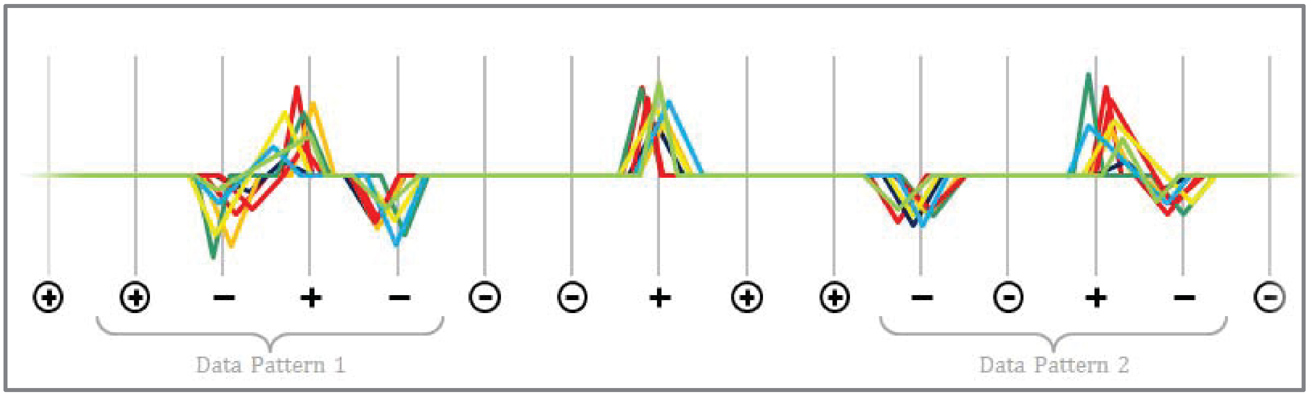


Figure 3

In fact, these waveforms illustrate the neighborhood and transition polarity dependence. The symbol sequence illustrated in all of these figures includes a first data pattern (labeled "Data Pattern 1") that is ⊕,-,+,- and a second data pattern (labeled "Data Pattern 2") that is -,⊖ ,+,-. These two sequences each involve three transitions (-,+,-) but the patterns are different because in Data Pattern 1 the three transitions are consecutive, whereas in Data Pattern 2 there is a non-transition region between the first and second transitions. *See* Bajorek Dec. at ¶ 81. The neighborhood dependence is shown in Figure 4 below. The +,- sequence in Data Pattern 1 has a different shape than the +,- sequence in Data Pattern 2 because of their different neighborhoods. The +,- sequence in Data Pattern 1 is preceded by ⊕,- , whereas the +,- sequence in Data Pattern 2 is preceded by -,⊖ , resulting in different pulse structures between the two. *See* Bajorek Dec. at ¶ 85.

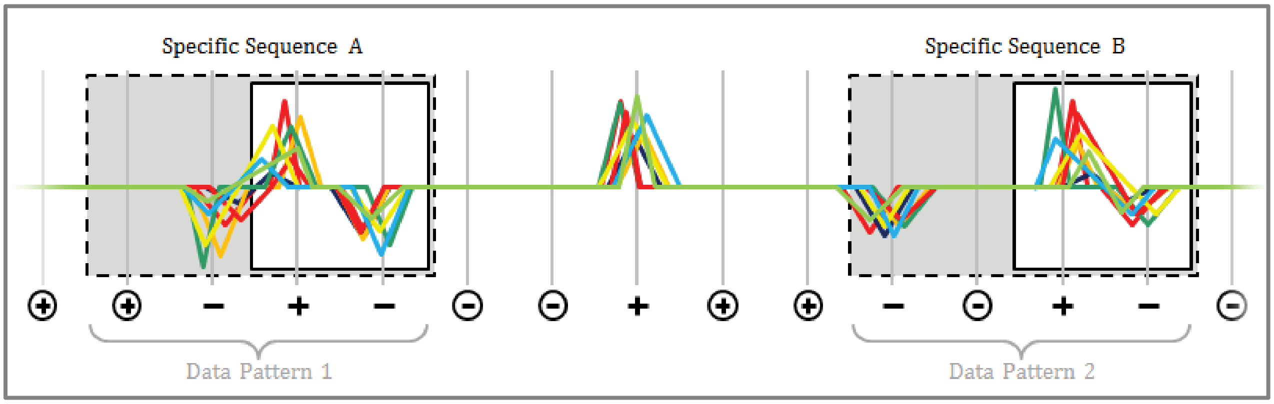


Figure 4

The transition polarity dependence is shown in Figure 5 below. The + and - transitions in the middle of the sequence lead to pulses having a different amplitudes and shape (even in absolute terms) because of their difference in polarity. *See* Bajorek Dec. at ¶ 86.

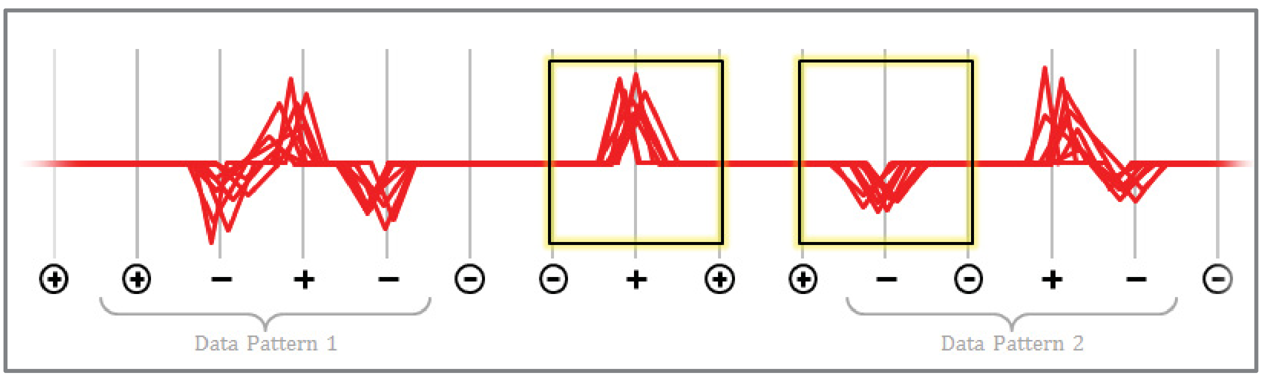


Figure 5

In summary, each specific sequence of polarized regions on the disk (written symbols) will have a different noise structure because each sequence will have different numbers (density) and polarities of transitions, with different numbers (and polarities) of non-transition regions therebetween. The different noise structures result from the fact that: (1) transitions of different polarity ("positive" and "negative" transitions) in and of themselves have different noise structures; (2) each specific sequence of adjacent and nearby symbols written to the disk (i.e., a "neighborhood) has a noise structure that differs from the noise structure embedded in the recording media by writing a different sequence of symbols; and (3) the magnetoresistive (MR) read head used to read the symbols written to the disk has an asymmetric response to positive and negative transitions written to the media, thereby reading these two differently polarized transitions differently. This signal-dependent noise greatly complicates the process of detecting the data written to the disk from the readback signal. *See* Bajorek Dec. at ¶ 71.

5. Important Factors in Developing the Invention of Claim 4

As explained in the background section of the '839 patent, in the early 1990s the HDD industry migrated from analog peak detectors to digital Viterbi detectors that used the Euclidean branch metric function. *See* '839 patent at col. 1:25-37. The Euclidean branch metric function assumes that the noise in the system is "additive white Gaussian noise (AWGN)." *Id.* at col. 1:34-36. As data density increased, however, this assumption became less valid because media noise became more dominant, and this media noise is signal-dependent and correlated -- not additive white Gaussian. The invention of claim 4 solves this problem. The inventors developed a Viterbi-type detector that computes branch metric values that account for the correlated, signal-dependent media noise. To do this, as noted in the '839 patent:

- "… the usual simplifying assumption that the noise samples are independent random variables" (e.g., white) should not be used ('839 patent at col. 1:65-67). This is the assumption made with detectors that use the Euclidean branch metric (*see id.* at col. 5:59-63) and, as shown below, by Zeng in Sections 4.4 and 5.2 of his Thesis for his "random jitter";

- Because different polarity transitions have different signal-dependent noise structures as described above, the detector's BMFs should not be "blind to signal asymmetries" because that is "inappropriate for the present problem" (*id.* at col. 3:60-64); and

- Because different specific symbol sequences have different signal-dependent noise structures as described above, the detector's BMFs should be a function of multiple samples of the readback signal and be "dependent on the postulated sequence of written symbols…, which ensures the signal-dependence of the detector." '839 patent at col. 5:49-52.

These were the guiding precepts that led to the inventive method of claim 4. *See* Moura Dec. at ¶¶ 23-25.

6. <u>The Invention of the CMU Patents</u>

The process by which the inventors discovered their new detector is set out in the patent specification at equations 1 through 6 and the corresponding text found at col. 4:2 - col. 5:49 of the '839 patent. *See* Moura Dec. at ¶ 40.[16] These teachings in the '839 patent show that the detector of claim 4 uses a "set of signal dependent branch metric functions" to account for the fact that each sequence of written symbols has its own specific noise structure: "the metric is a function of the observed samples, r_i, r_{i+1}, . . . r_{i+L}. ***It is also dependent on the postulated sequence of written symbols*** a_{i-K1}***, . . ,*** a_{i+L+Kt}***, which ensures the signal dependence of the detector.*** As a consequence, the branch metrics for each branch in the tree/trellis is [*sic*] based on its corresponding signal/noise statistics." '839 patent at col. 5:49-55 (emphasis added).

The BMFs of the CMU patents use multiple signal samples to compute the branch metric values for each specific sequence (or "neighborhood") of symbols (e.g., corresponding to a branch of the trellis). Thus, the CMU patents disclose that a different BMF is used for each specific neighborhood of symbols (e.g., a trellis branch) at a given time index of the trellis. *See, e.g.*, '839 patent at Eq. 13 and Figures 3A-B. In particular, the CMU patents expressly teach that a different branch metric function can be used for "each branch of the tree/trellis due to the signal dependent structure of the media noise." '839 patent at col. 2:18-20 (Summary of the Invention). This "set" of different BMFs allows the detector to account for the specific signal-dependent noise associated with specific symbol sequences. *See* Moura Dec. at ¶ 41.

[16] These equations and accompanying disclosure also prove mathematically that Drs. Kavcic and Moura discovered the optimal detector, i.e., provably unbeatable. *See* Moura Dec. at ¶ 40.

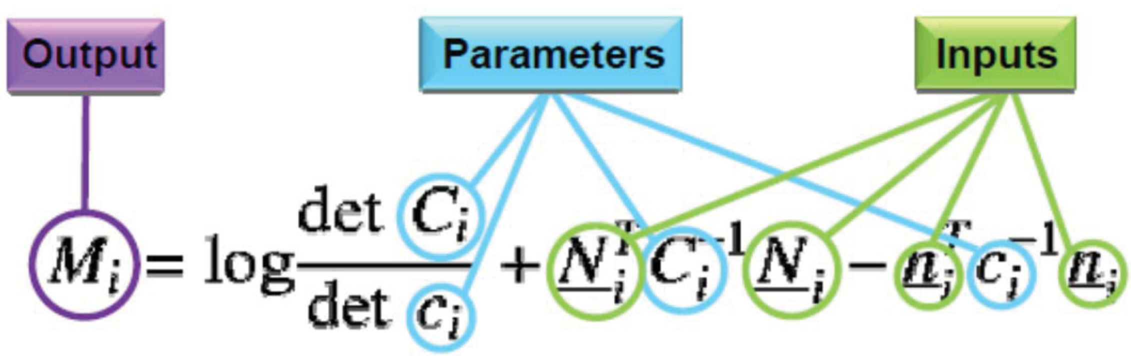
a. Equation 13

Equation 13 of the CMU patents discloses a general set of signal-dependent, correlation-sensitive BMFs:

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

where M_i is the branch metric value for the considered branch; C_i is a (L+1) x (L+1) noise covariance matrix for the considered branch; c_i is a LxL lower principal submatrix of C_i; $\underline{N}_i$ is a (L+1)-dimensional vector of differences between the signal samples (e.g., the *r* values) and their expected values (e.g., their "target" values) when a particular sequence of data is written; and $\underline{n}_i$ is a vector containing the last L elements of $\underline{N}_i$. *See* '839 patent at col. 6:56-65. In this formulation, the correlation length is L and L+1 signal samples are used to compute the branch metric value M_i. The noise covariance matrix C_i is signal-dependent, thereby making the BMF both signal-dependent and correlation-sensitive. *See* '839 patent at col. 2:15-20; 4:24-31; 6:36-55; Moura Dec. at ¶ 42.

Equation 13 of the CMU patents discloses a "***set of signal-dependent branch metric functions***" as required in claim 4 because each member of the set is a function with parameters that are different for each ***specific neighborhood*** of symbols, e.g., different for each trellis branch. In Equation 13, as shown below, these signal-dependent parameters are the noise covariance matrices C_i (including c_i, the lower principal submatrix of C_i). *See* McLaughlin Dec. at ¶ 37.



Equation 13 also discloses the application of the members of this set of signal-dependent BMFs to a "***plurality of signal samples***…, wherein each sample corresponds to a different sampling time instant,"[17] as required by claim 4 because as noted above, the $\underline{N}_i$ and $\underline{n}_i$ terms include multiple time-variant signal samples. *See* '839 patent at col. 6:57-65, equation 12 (collecting the multiple signal samples and their corresponding target values). The output of each member of this "set" of "signal dependent BMFs" is a "***branch metric value,***" denoted in the M_i in equation 13. *See* McLaughlin Dec. at ¶ 38.

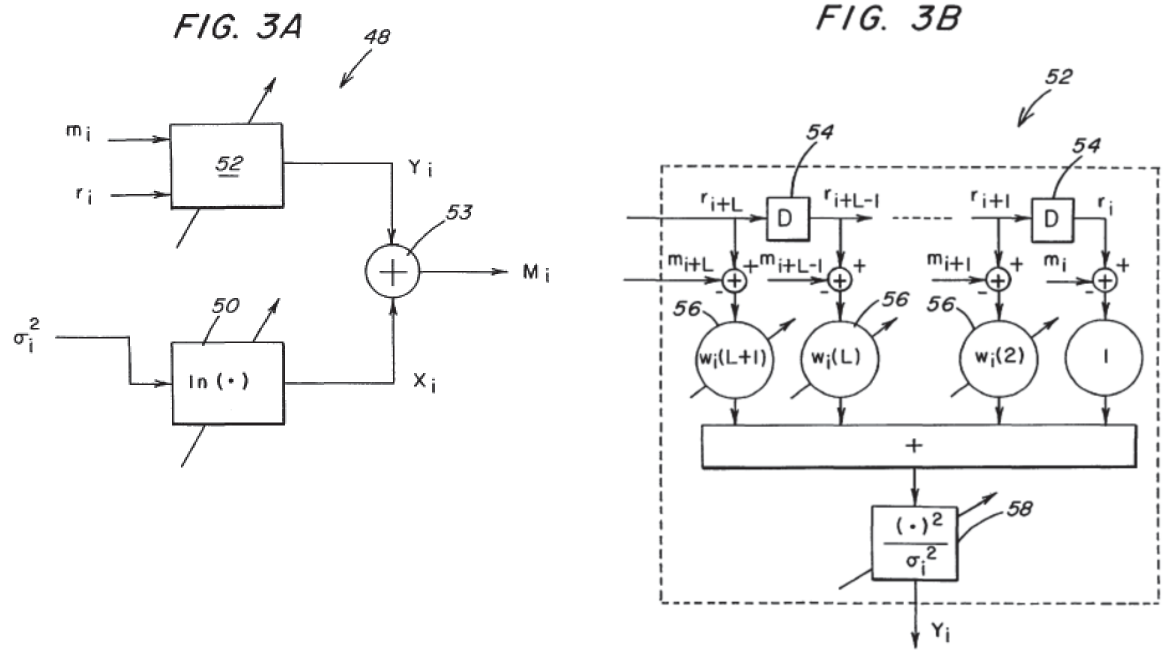
In Equation 13 the inputs for the BMFs are expressed as the vectors $\underline{N}_i$ and $\underline{n}_i$, which vectors comprise the difference of multiple time-variant signal samples (L+1 and L signal samples, respectively) and their associated target values (the *m* terms). *See* '839 patent at eq. 12, col. 6:61; McLaughlin Dec. at ¶ 39. Thus, the BMFs described by Equation 13 are two-input BMFs, the two inputs being (i) the set of signal samples and (ii) the target values. McLaughlin Dec. at ¶ 39. To compute a branch metric value for a given branch at a certain time index of the trellis, the signal-dependent BMF for the branch is selected by choosing the correct parameters for the BMF, e.g., the noise covariance matrix C_i as described by equation 13 or set of filter taps for a finite impulse response (FIR) filter embodiment (see following paragraph) that is tuned for the specific symbol sequence associated with that particular branch. The selected signal-dependent BMF is then applied to the plurality of time-variant signal samples in the vectors $\underline{N}_i$ and $\underline{n}_i$ for that certain time index. *See* McLaughlin Dec. at ¶ 40.

b. FIR Filter Embodiment

The CMU patents disclose that the signal-dependent, correlation-sensitive branch metric functions can be implemented with electronic circuits, referred to in the CMU patents as "branch metric computation circuits" and shown as element 48 in Figure 3A (see below). Each different branch of the trellis can use a different branch metric computation circuit to compute the branch metric value for the corresponding branch. *See* '839 patent at col. 7:10-13. For example, the CMU patents disclose that the each branch metric computation circuit 48 uses a separate FIR

[17] For convenience, these signal samples are sometimes referred to herein as "time-variant signal samples."

filter (shown in Figure 3B) to compute the branch metric value for its associated branch of the trellis based on Eq. 13. *See* Moura Dec. at ¶ 43; McLaughlin Dec. at ¶ 41.



The ‘839 patent shows mathematically how Eq. 13 can be implemented with a set of FIR filters. From Eq. 13 of the CMU patents, let $X_i = \log \frac{\det C_i}{\det c_i}$ and $Y_i = \underline{N}_i^T C_i^{-1} \underline{N}_i - \underline{n}_i^T c_i^{-1} \underline{n}_i$, in which case $M_i = X_i + Y_i$ according to Eq. 13, as shown in the CMU patents at equations 15-18. In Figure 3A, circuit 50 computes the X_i term and circuit 52, which is shown in more detail in Figure 3B, computes the Y_i term of Eq. 13. *See* Moura Dec. at ¶ 44. Equation 16 of the CMU patents is that $\sigma_i^2 = \frac{\det C_i}{\det c_i}$. Thus, circuit 50 computes the $\log \sigma_i^2 = \log \frac{\det C_i}{\det c_i}$, which is the X_i term of Equation 13. Equations 17 and 18 of the CMU patents (*see* ‘839 patent at col. 46-50) show that:

$$Y_i = \underline{N}_i^T C_i^{-1} \underline{N}_i - \underline{n}_i^T c_i^{-1} \underline{n}_i$$

$$= \frac{(\underline{w}_i^T \underline{N}_i)^2}{\sigma_i^2}$$

where $\underline{w}_i$ is a (L+1)-dimensional given by, according to Equations 19-20 (*see* '839 patent at col. 55-60):

$$\underline{w}_i^T = [1 \quad w_i(2) \quad w_i(3) \quad \ldots \quad w_i(L+1)]^T$$

$$= \begin{bmatrix} 1 \\ -c_i^{-1} \quad \underline{c}_i \end{bmatrix}.$$

See Moura Dec. at ¶ 45.

In this FIR filter embodiment, the weights $\underline{w}_i$ are the weights of the FIR filter in Figure 3B. The FIR filter comprises multipliers 56, which respectively multiply the mean-adjusted signal samples by the associated weight $\underline{w}_i$ for the multiplier. The products of each of the multipliers 56 are added by the adder (denoted in Figure 3B by the rectangle with a plus sign). The sum of the weighted signal samples (i.e., the output of the adder) is then squared and divided by σ_i^2 at block 58 to yield the Y_i term. Thus, the output Y_i of circuit 52 can be expressed as:

$$Y_i = \frac{\left(n_i + \sum_{j=1}^{L} w_i(j+1) n_{i+j}\right)^2}{\sigma_i^2}$$

where $n_i = r_i - m_i$. *See* Moura Dec. at ¶ 46.

In Figure 3A of the CMU patents, the adder 53 adds the X_i term from circuit 50 and the Y_i term from circuit 52 to produce the branch metric:

$$M_i = X_i + Y_i = \log \sigma_i^2 + \frac{\left(n_i + \sum_{j=1}^{L} w_i(j+1) n_{i+j}\right)^2}{\sigma_i^2}$$

See Moura Dec. at ¶ 47.

The weights $\underline{w}_i$ of the FIR filters and the σ_i^2 terms can be derived directly from the noise covariance matrix C_i, as explained in equations 16-20 of the CMU patents. Each specific sequence (neighborhood), i.e., each specific branch, has its own distinct weight vector $\underline{w}_i$ and its own distinct σ_i^2. To compute a branch metric value for a given branch at a time index of the trellis using the FIR filter approach, the BMF for each separate branch is selected by choosing the parameters for the BMF, e.g., weights $\underline{w}_i$ of the FIR filter, for that particular branch. The selected BMF is applied to a plurality of time variant signal samples , such that the signal samples are respectively multiplied by the weights $\underline{w}_i$ of the FIR filter for that particular branch. *See* Moura Dec. at ¶ 49.

B. Summary of the Detectors in Sections 4.4 and 5.2 of Zeng's Thesis

Claim 4 was rejected in the Office Action as anticipated by Sections 4.4 and 5.2 of Zeng's Thesis. Claim 4 was also rejected as being obvious based on two combinations of references, where for each combination Zeng's Thesis was the primary reference. Below is a summary of what Zeng's Thesis discloses.

1. Zeng's Channel Models and Branch Metric Functions

a. Zeng's General Channel Model

Zeng's Thesis provides its general channel model in Section 4.1 at equation 4.1 (p. 51), which is:

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

where *Z(t)* is the readback signal as a function of time *t*; a_k represents the transition response for a transition on the disk at position *k*; *N* represents the length of the entire data sequence; *T* is the symbol interval; *h(t)* is the channel response to a single transition; *w(t)* is the additive white noise process; and Δ_k "represents the random jitter in the position of the transition response." Zeng Thesis at 51. Zeng models the "random jitter" from transitions as "***independently and***

identically distributed random variables with zero mean" and variance σ_Δ^2. *See* Zeng Thesis at 52, 65; Kavcic Dec. at ¶ 17; McLaughlin Dec. at ¶ 44. That means that the random jitter Δ_k is ***independent*** of the written symbols, which is evident from eq. 4.1 itself, which shows that the value of the random jitter Δ_k is ***independent*** of the written symbols a_k. McLaughlin Dec. at ¶ 44. The subscript for the "random jitter" term Δ_k in equation 4.1 is k and not a sequence of a_k's (i.e., the written symbols), thereby demonstrating that the random jitter Δ_k is ***independent*** of the written symbols a_k. *See* McLaughlin Dec. at ¶ 44.

Because Zeng's "random jitter" Δ_k is "independently and identically distributed random variables with zero mean," Zeng's random jitter is ***white***, which is mutually exclusive of signal-dependent. McLaughlin Dec. at ¶ 59. A "***white*** Gaussian noise assumption" is one where "the noise samples are realizations of ***independent identically distributed*** Gaussian ***random variables*** with ***zero mean***…." '839 Patent at col. 5:59-62 (emphasis added). This explanation in the CMU patents is consistent with the ordinary understanding of a "white" random process in statistical signal processing, which is a process with a flat (or constant) power spectral density (PSD), which happens when random variables are independent and identically distributed with a zero mean. *See* McLaughlin Dec. at ¶ 59 (citing authorities). Consequently, Zeng's "random jitter" is merely a hypothesized form of peak shift (pulse shift) that assumes that such shift is entirely independent from one transition to another transition, and independent of the polarity and the specific sequence of transitions in the neighborhood of written symbols, which is not signal-dependent media noise. *See* McLaughlin Dec. at ¶ 61.[18] In stark contrast, the '839 patent explains that, "It has long been observed that the noise [which includes jitter] in magnetic recording systems is neither white nor stationary. The non-stationarity of the media noise results from its signal dependent nature." '839 patent col. 1:38-41.

Comparing the readback signal waveforms in Figure 3 on page 23 to ones generated using Zeng's "random jitter" model illustrates the differences between signal-dependent noise

---

[18] As explained below, Zeng's channel model also permits consecutive positive or consecutive negative transitions, which is physically impossible. A person having skill in the art could not make or use the detectors in Sections 4.4. and 5.2 of Zeng's Thesis. *See* Kavcic Dec. at ¶¶ 15-28.

and Zeng's "random jitter" model. Figure 6A shows a cumulation of possible readback signal waveforms for a symbol sequence in the presence of signal-dependent noise (including hard/easy transitions, transition interactions, polarity dependence, and neighborhood effects). *See* Bajorek Dec. at ¶¶ 82-84. The lower (purple) waveforms in Figure 6B show a cumulation of possible readback signals using Zeng's "random jitter" model (ignoring the d=1 RLL constraint that Zeng uses in Sections 4.4 and 5.2) for the exact same symbol sequence as Figure 6A. Zeng's model accounts only for random jitter, which are random shifts (left or right) in the peak positions with no variation in the peak shape, as shown in Figure 6B. *See* Bajorek Dec. ¶ 84.

Figure 6A (Signal-dependent model)

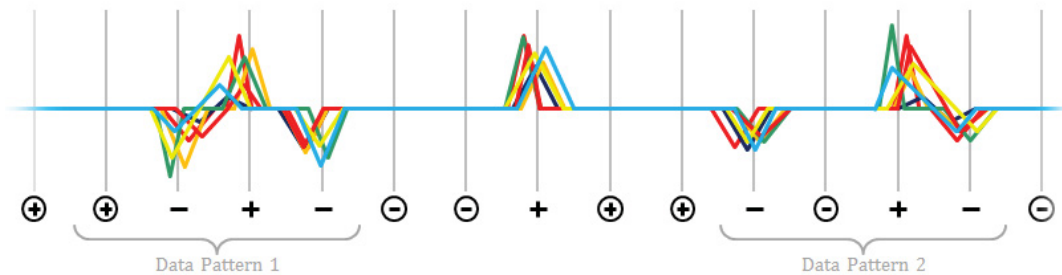
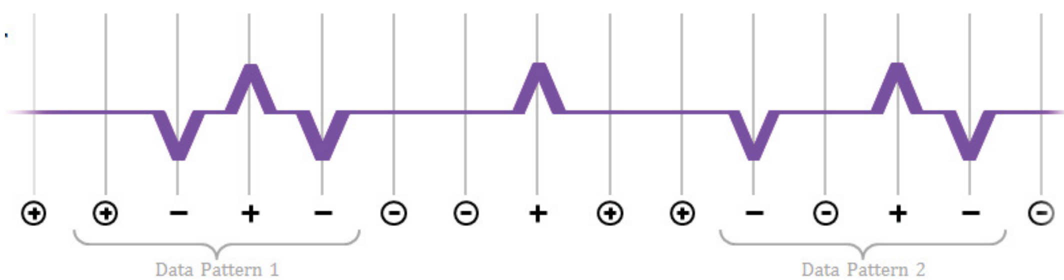


Figure 6B (Zeng's Random Jitter model)



Figures 7A and 8A below illustrate the neighborhood and transition polarity dependence for signal-dependent noise, which are absent in Zeng's random jitter model (Figures 7B and 8B). The neighborhood dependence is shown in Figures 7A-B below. In Figure 7A, the +,- sequence in Data Pattern 1 has a different readback signal shape than the +,- sequence in Data Pattern 2 because of their different neighborhoods (the +,- sequence in Data Pattern 1 is preceded by ⊕,- , whereas the +,- sequence in Data Pattern 2 is preceded by -,⊖), resulting in different pulse structures for the two +,- sequences. On the other hand, in Figure 7B, the readback signal shapes for the +,- sequences are identical for Data Patterns 1 and 2 with Zeng's random jitter model despite the differing "neighborhoods" because Zeng's random jitter model does not account for the neighborhood effects. *See* Bajorek Dec. at ¶ 85.

Neighborhood Dependent

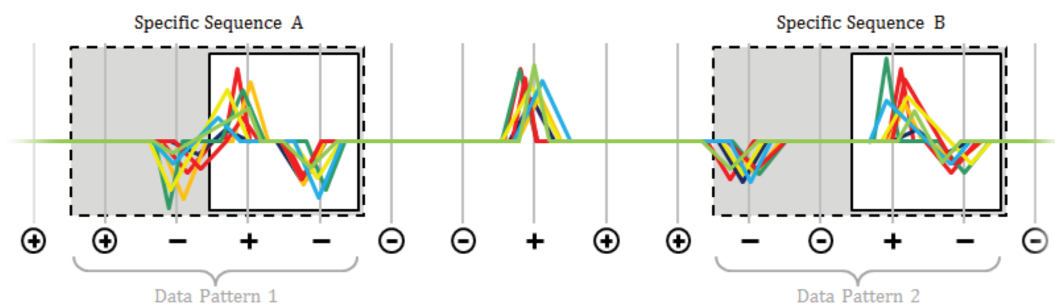


Figure 7A (Signal-dependent model)

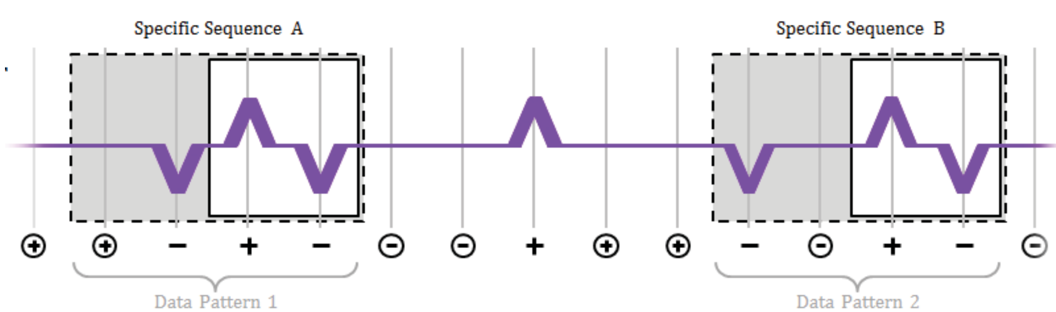


Figure 7B (Zeng's Random Jitter model)

The transition polarity dependence is shown in Figures 8A-B below. In Figure 8A, with signal-dependent noise, the + and - transitions in the middle of the sequence lead to readback signal pulses having different amplitudes and shapes because of their difference in polarity. On the other hand, in Figure 8B, the pulses from the two transitions are exactly the same (except for opposite direction) with Zeng's random jitter model since it does not account for the polarity of the transitions. *See* Bajorek Dec. at ¶ 86.

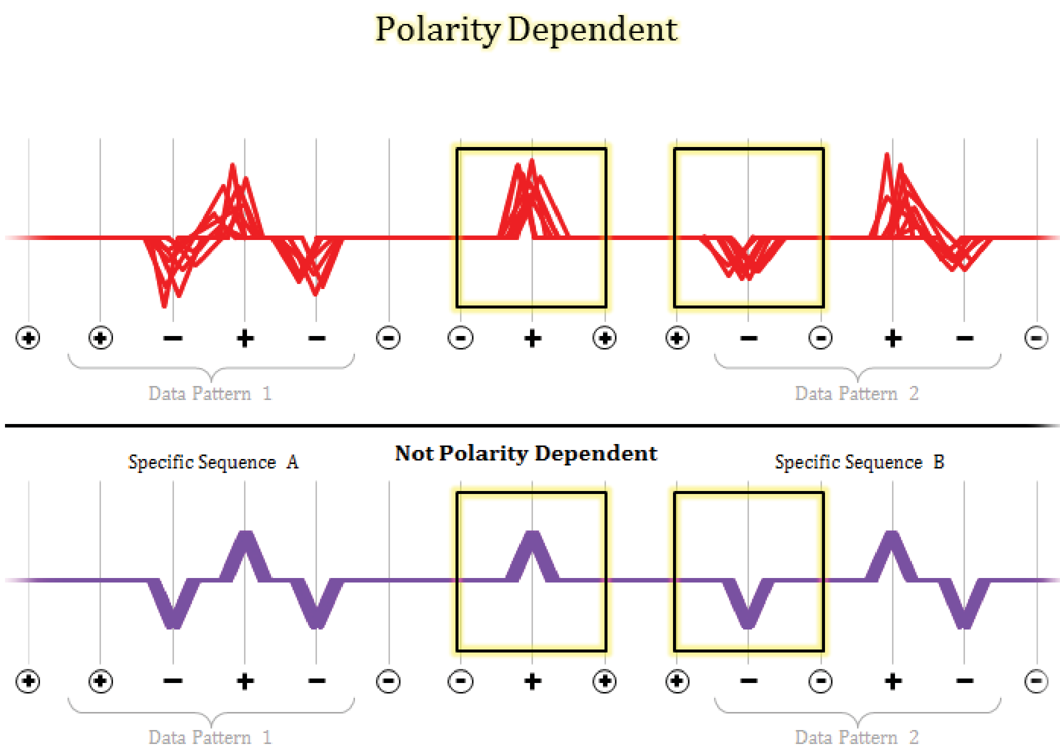


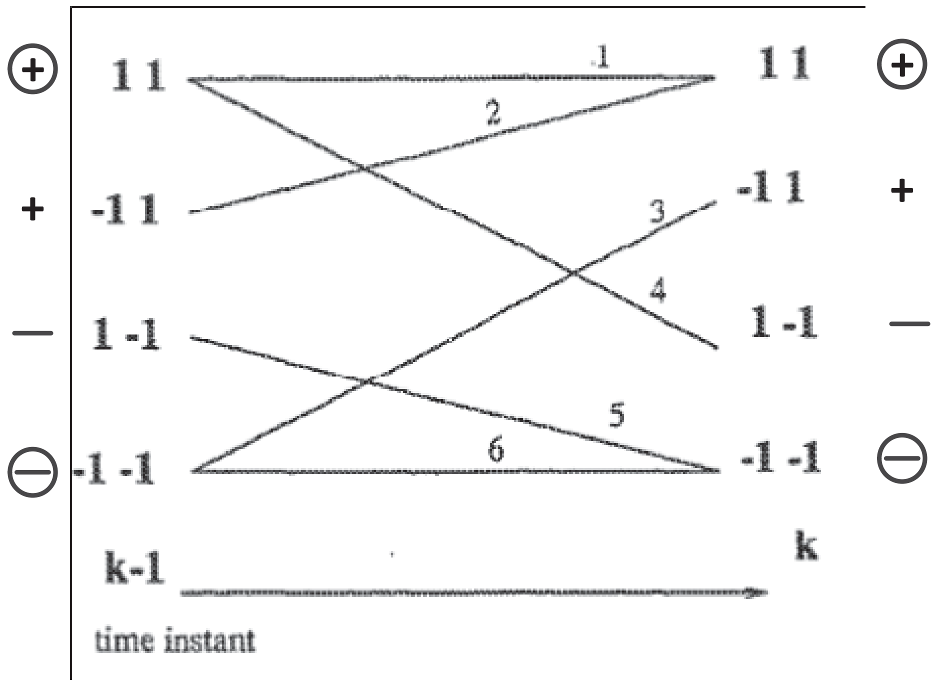
Figure 8A (Signal-dependent model)

Figure 8B (Zeng's Random Jitter model)

The '839 patent explains that to be signal-dependent, a detector needs to account for the neighborhood effect, including the polarity dependence. *See, e.g.,* '839 patent at col. 5:47-52; col. 3:54-64; Moura Dec. at ¶¶ 19-25. Zeng's BMFs do not account for either, let alone both.

b. Section 4.4

In Section 4.4 of his thesis, Zeng proposes a set of BMFs for a 4-state, 3-symbol sequence trellis with only six (6) branches. One section of this trellis is shown in Figure 4.5 of Zeng's Thesis, which is reproduced below. Zeng assigned each of the six branches a number in this diagram. This trellis is annotated below to also show, on the left- and right-hand sides of the diagram, the trellis notation used in the CMU patents.[19]



Zeng's trellis in Section 4.4 only has six (6) branches (instead of eight (8) as in Figure 4 of the CMU patents) because Zeng's device employs a d=1 Run Length Limited (RLL) constraint. *See* McLaughlin Dec. at ¶¶ 45-46. The d = 1 RLL constraint means that Zeng does not permit sequences with consecutive transitions; that is, transitions in the recorded symbol sequences must be separated by at least d = 1 non-transition regions. *See* Zeng Thesis at 4, 65. Zeng mandates the d=1 RLL code constraint because, according to Zeng, "it is ***impossible*** to implement the [Maximum Likelihood] detector for any reasonable length of data" without it. Zeng Thesis at p. 65 (emphasis added); Bajorek Dec. at ¶ 92.[20]

Zeng's BMFs in Section 4.4 are derived from a specific channel model for his 3-symbol sequence trellis. The channel model is Eq. 4.22 (shown below), and it represents the signal sample Z_k of the readback signal for position index (time instant) *k*.

[19] For example, branch 1 of Zeng' trellis corresponds to the ⊕⊕ branch of the trellis shown in Figure 4 of the CMU patents, and so on.

[20] Of course, Zeng is wrong. The CMU patents show how to implement a detector that accounts for signal-dependent noise without the d=1 RLL constraint. *See* Bajorek Dec. at ¶ 92.

$$Z_k = \underbrace{a_k + a_{k-1}}_{y_k} + \underbrace{w_k - a_k\Delta_k + a_{k-1}\Delta_{k-1}}_{n_k},$$

This equation shows that that the signal sample value Z_k is made up of a "signal component" y_k (target) and a "total noise contribution" n_k. *See* Zeng Thesis at 65. The a_k and a_{k-1} terms in the above equation represent the symbols on the disk at region *k* and *k*-1, respectively, and Zeng models them as being able to assume the values +2 (for a positive transition), 0 (for a non-transition), or -2 (for a negative transition). *See* Zeng Thesis at p. 51. The term w_k is the additive white noise and the Δ_k terms represent "the random jitter in the position of the transition response." *Id.*; *see also* McLaughlin Dec. at ¶ 47.

Because Zeng uses a d = 1 RLL code in Section 4.4, there cannot be two consecutive transitions. This means that in the channel model equation 4.22, at least one of a_k and a_{k-1} is zero at each time instance, and both are zero for a branch with no transitions (branches 1 and 6 in Zeng's trellis). *See* McLaughlin Dec. at ¶ 48.

Section 4.4 of Zeng's Thesis discloses three BMFs for the six branches in Zeng's trellis. The three BMFs are summarized in Table 1 below. In the functions, the z_k terms represent signal samples and the y_k terms represent targets.[21] σ_w^2 is the variance of the additive white noise and σ_Δ^2 (found in BMF No. 3 only) is the variance of Zeng's "random jitter" and it is a constant value across the various branches. *See* Zeng Thesis at 56 and 65-70.

[21] Zeng's Thesis does not disclose the targets for his BMFs and Dr. Lee did not identify them either. CMU's expert declarants determined them in order to generate this chart.

Function No.	Function	Target	Branch	Zeng Thesis Cite
1	$\ln \sigma_w^2 + \frac{[z_k - y_k]^2}{\sigma_w^2}$	$y_k = 0$	1	p. 67
		$y_k = 0$	6	
2	$\ln \sigma_w^2 + \frac{\left[\frac{1}{\sqrt{2}}(z_k - y_k) + \frac{1}{\sqrt{2}}(z_{k-1} - y_{k-1})\right]^2}{\sigma_w^2}$	$y_k = 2$ $y_{k-1} = 2$	2*	p. 68 Eq. 4.26
		$y_k = -2$ $y_{k-1} = -2$	5*	
3	$\ln[\sigma_w^2 + 8\sigma_\Delta^2] + \frac{\left[\frac{-1}{\sqrt{2}}(z_k - y_k) + \frac{1}{\sqrt{2}}(z_{k+1} - y_{k+1})\right]^2}{\sigma_w^2 + 8\sigma_\Delta^2}$	$y_k = 2$ $y_{k+1} = 2$	3*	p. 68 Eq. 4.28
		$y_k = -2$ $y_{k+1} = -2$	4*	

TABLE 1
(* indicates a branch with a transition in the symbol sequence)

See McLaughlin Dec. at ¶ 49.

A review of the functions above shows that none of them have any terms that represent signal-dependent noise. In fact, only one of them, BMF No. 3, has a term related to Zeng's ***white*** "random jitter" (σ_Δ^2). *See* McLaughlin Dec. at ¶ 49.

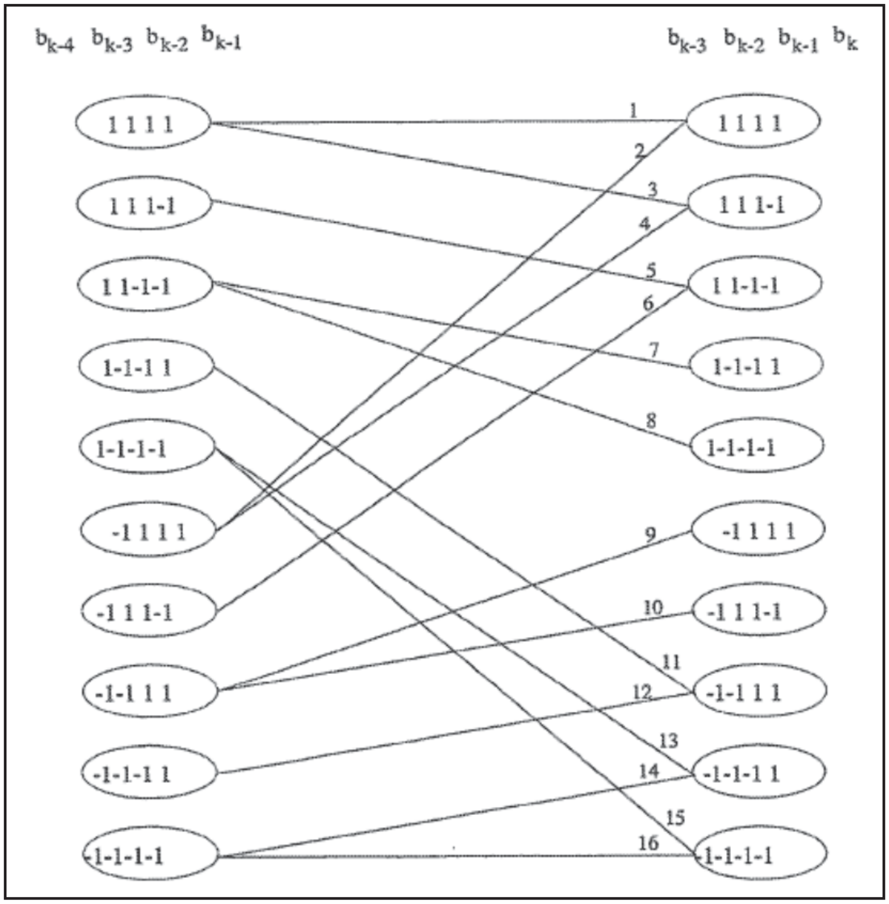
c. Section 5.2

Despite the numerous equations in Section 5.2, and despite Dr. Lee's erroneous statements to the contrary, Zeng's Section 5.2 only discloses ***three*** branch metric functions, in a manner very similar to what is disclosed in Zeng's Section 4.4. *See* McLaughlin Dec. at ¶ 50.[22]

Section 5.2 of Zeng's Thesis proposes three BMFs for a 5-symbol sequence trellis ($b_{k\text{-}4}$, … b_k) with sixteen (16) branches. Figure 5.3 of Zeng's Thesis, which is reproduced below,

[22] Also, Zeng's equations in Section 5.2 are riddled with errors, so much so that they are unusable in a detector. *See* Kavcic Dec. at ¶¶ 30-42. These errors are described below in Section VII.E.2 of this response.

shows one section of this trellis. Zeng assigned a number to each of the sixteen branches in this diagram.



Zeng's BMFs in Section 5.2 are derived from a channel model for his 5-symbol-sequence trellis. The channel model is Eq. 5.5 (shown below), and it models the signal sample ZD_k of the readback signal for polarized region k.

$$ZD_K = g_{-1}a_k + g_0a_{k-1} + g_1a_{k-2} + g_2a_{k-3} + w'_{k-1} + j_0a_{k-1}\Delta_{k-1} + j_1a_{k-2}\Delta_{k-2}$$

See Zeng Thesis at 76. Like with Eq. 4.22, channel model equation 5.5 shows that the signal sample value ZD_k is made up of a "signal component" (also known as the "target value") corresponding to $(g_{-1}a_k + g_0a_{k-1} + g_1a_{k-2} + g_2a_{k-3})$ and a "total noise contribution" corresponding to $(w'_{k-1} + j_0a_{k-1}\Delta_{k-1} + j_1a_{k-2}\Delta_{k-2})$. *See* Zeng Thesis at 76. As before:

a. The a_k terms are the responses to the non-return-to-zero (NRZ) bits b_k and can only assume the value +2, 0, or -2;

b. The w'_{k-1} term is the additive white noise;

c. The Δ_k terms represent "the random jitter in the position of the transition response"; and

d. The *g* and *j* terms according to Zeng are "some real numbers." *See* Zeng Thesis at pp. 51 and 75.

See McLaughlin Dec. at ¶ 52.

Zeng also mandates the d =1 RLL code constraint in Section 5.2. Consequently, the two "random jitter" terms Δ_{k-1} and Δ_{k-2} in Equation 5.5 (that contribute to the "noise component") cannot both appear at the same time in the "total noise contribution" of Equation 5.5 (because either a_{k-1} or a_{k-2} is zero). *See* McLaughlin Dec. at ¶ 53. Also, because of the $d = 1$ RLL code constraint, Zeng's trellis in Figure 5.3 has only sixteen (16) branches instead of the expected 32 for a 5-symbol-sequence trellis, and only 10 of the expected 16 nodes. *See* McLaughlin Dec. at ¶ 53.

The three BMFs in Section 5.2 are summarized in Table 2 below.[23] In the three functions, the ZD_k terms represent signal samples. The appropriate targets, t_a and t_b, are listed for the various branches in the third column of the chart (BMF No. 1 does not have a second target, t_b, since it only uses one signal sample).[24] Like before, σ_w^2 is the variance of the white noise and σ_Δ^2 (found in No. 3 only) is the variance of the "random jitter" and it is a constant value across the various branches. *See* Zeng Thesis at 56 and 65-70. The *g* and *j* terms are "real numbers" relating back to the channel model equation 5.5. *Id.* at 75-76.

[23] Zeng's derivation of these functions in Section 5.2 is erroneous as explained in Prof. Kavcic's declaration. This table uses the error-filled functions that Zeng actually discloses.

[24] Prof. Kavcic's declaration describes how Zeng's target values in his equations 5.13 to 5.21 are wrong. For the purposes of this table, Zeng's incorrect target values are used because that is what Zeng wrote and that is what Dr. Lee relied upon.

Function No.	Function	Target	Branch	Zeng Thesis Cite
1	$\ln\sigma_w^2 + \frac{[ZD_k - t_a]^2}{\sigma_w^2}$	$t_a = 0$	1	p. 76
		$t_a = 2g_2$	2*	Eq. 5.6
		$t_a = -2g_{-1}$	3*	Eq. 5.8
		$t_a = -2g_{-1} + 2g_2$	4*	Eq. 5.10
		$t_a = 2g_{-1} - 2g_2$	13*	Eq. 5.11
		$t_a = 2g_{-1}$	14*	Eq. 5.9
		$t_a = -2g_2$	15*	Eq. 5.7
		$t_a = 0$	16	p. 76
2	$\ln\sigma_w^2 + \frac{\frac{j_0}{j_o^2 + j_1^2}[ZD_k - t_a] - \frac{j_1}{j_o^2 + j_1^2}[ZD_{k-1} - t_b]^2}{\sigma_w^2}$	$t_a = 2g_{-1} - 2g_1$ $t_b = -2g_0$	7*	Eq. 5.18
		$t_a = -2g_1$ $t_b = 2g_0$	8*	Eq. 5.20
		$t_a = 2g_1$ $t_b = 2g_0$	9*	Eq. 5.21
		$t_a = -2g_{-1} + 2g_1$ $t_b = 2g_0$	10*	Eq. 5.19
3	$\ln[\sigma_w^2 + 4(j_o^2 + j_1^2)\sigma_\Delta^2] + \frac{\frac{j_1}{j_o^2 + j_1^2}[ZD_k - t_a] + \frac{j_0}{j_o^2 + j_1^2}[ZD_{k+1} - t_b]^2}{\sigma_w^2 + 4(j_o^2 + j_1^2)\sigma_\Delta^2}$	$t_a = -2g_0$ $t_b = -2g_1$	5*	Eq. 5.13
		$t_a = -2g_0 + 2g_2$ $t_b = -2g_1$	6*	Eq. 5.15
		$t_a = 2g_0 - 2g_2$ $t_b = 2g_1$	11*	Eq. 5.16
		$t_a = 2g_0$ $t_b = 2g_1$	12*	Eq. 5.14

TABLE 2
(* indicates a branch with a transition in the symbol sequence)

See McLaughlin Dec. at ¶ 54.

The reason why the numerous branch metric equations in Zeng's Section 5.2 constitute only three branch metric functions is because, as shown above, using differing target values, which are "inputs" to a branch metric function, does not create different functions. *See* McLaughlin Dec. at ¶¶ 55-57; Ex. 6 (Strang Dec) at ¶¶ 7-21. For example, Zeng's equations

5.18 to 5.21 are the same function, namely BMF No. 2 above. The only differences between the equations are the target values, which are inputs to the functions. *See* McLaughlin Dec. at ¶¶ 55-57.

Like with Section 4.4, a review of the functions in Table 2 above shows that none of them have any terms that represent signal-dependent noise. In fact, only one of them, BMF No. 3, has a term related to Zeng's ***white*** "random jitter" (σ_Δ^2). *See* McLaughlin Dec. at ¶ 54.

2. Zeng's Acknowledgement and Treatment of Signal-Dependent Noise

In his thesis, Zeng claims to be "interested in developing new detectors for high linear density recording channels, which experience significant amounts of nonlinear distortions and media noise." Zeng Thesis at 1. Zeng's Thesis identifies several sources of signal-dependent noise, but Zeng makes assumptions and simplifications such that his detectors ignore this signal-dependent noise in his BMFs. *See* Bajorek Dec. at ¶ 88. For example, Zeng identifies four sources of signal-dependent noise in Section 1.3 of the thesis: (1) "the overwrite effect" (which leads to hard and soft transitions); (2) peak shift where a transition is one bit interval earlier; (3) peak shift where "the separation between interacting transitions is twice the bit interval"; and (4) "transition broadening" (i.e., the shape and amplitude of the pulses). *See* Bajorek Dec. at ¶ 89. Zeng either ignores or mitigates each of these sources of signal-dependent noise in a manner wholly external to his BMFs, such that Zeng's BMFs do not need to account for these sources of signal dependent noise and, in fact, do not do so. *See* Bajorek Dec. at ¶¶ 88-95. In other words, Zeng assumed away the signal-dependent noise problem.

For the ***first*** signal-dependent effect identified by Zeng, the overwrite effect, Zeng acknowledges that for hard transitions, where a region is to be polarized in a direction opposite to its existing polarization, there is peak shift, i.e., "the resulting readback waveform is shifted later by an amount ε_0." Zeng Thesis at 9. Zeng, however, ***eliminates*** this effect "by removing the initial magnetization of the medium" with "an AC-erased medium," (*id.*), in other words degaussing the disk. That is, prior to writing any data, Zeng eliminates the "initial" (or pre-existing) magnetization on the disk by first writing tiny regions of alternating polarities in equal number over the whole disk (within what would otherwise be the boundaries of each individual symbol region) so that there is no net demagnetizing field (that is, an AC-erasure). By eliminating it, Zeng's BMFs do not need to account for it.[25] *See* Bajorek Dec. at ¶ 90.

Zeng's ***second*** signal-dependent effect is the peak shift where a transition is one bit interval earlier. Zeng acknowledges that this peak shift is related directly to the polarity of the transitions. Zeng Thesis at 9 ("The direction of the shift is determined by the polarities of the transitions"). In the first instance, Zeng's channel model and assumption regarding his "random jitter" ("independently and identically distributed random variables with zero mean…," *see* Zeng Thesis at 52), which is explained further below, ignore the polarities of the transitions (i.e., ignores an acknowledged cause of the peak shift) and ***his BMFs do not account for this signal-dependent noise***; nor does Zeng attempt to account for it. *See* Bajorek Dec. at ¶ 91. Zeng confirms this because he attempts to reduce this second signal-dependent effect not through his BMFs but "by precompensation, namely shifting the second transition later by ε_1 when we encounter two transitions in a row in the data sequence." Zeng Thesis at 9.

In fact, in Sections 4.4 and 5.2, Zeng further artificially eliminates this peak shift from his channel models by imposing the d=1 RLL code constraint that prevents the writing of two transitions in adjacent regions of the disk. *See* Zeng Thesis at 4, 65-67 and 75. Zeng mandates the d=1 RLL code constraint because, according to Zeng, "it is impossible to implement the [Maximum Likelihood] detector for any reasonable length of data" without it. Zeng Thesis at 65. As a consequence of using the d=1 RLL constraint in conjunction with his "random jitter" noise model, Zeng eliminates and ignores all interactions between neighboring transitions and he has no reason to (and does not) consider the signal-dependent media noise arising from interacting transitions (i.e., Zeng does not consider media noise that is attributable to a ***specific neighborhood*** (sequence) of interacting transitions/signals). *See* Bajorek Dec. at ¶ 92.

---

[25] As explained below, degaussing the disk prior to each write is not practical in a commercial HDD; it requires additional components (to buffer the data) and would take an unacceptably long time. *See* Bajorek Dec. at ¶ 113. This is one reason of many that a person of ordinary skill in the art would not be motivated to modify the detectors in Zeng's Thesis in view of any prior art. *See id.*

The ***third*** signal-dependent effect is peak shift where "the separation between interacting transitions is twice the bit interval," Zeng says this effect is "less important" than the second effect, so ***Zeng ignores it altogether***. Zeng states that these effects "are not discussed here." Zeng Thesis at 9; *see also* Bajorek Dec. at ¶ 93.

The ***fourth*** signal-dependent effect is transition broadening. Zeng says that it "is modeled as the second derivative of the readback signal." Zeng's Thesis, however, does not use or rely on the second derivative of the readback signal in his BMFs in Sections 4.4 and 5.2, so he ignores this effect as well. The fact that Zeng purports to model this fourth effect does not mean that his BMFs account for it; they do not. *See* Bajorek Dec. at ¶ 94.

Finally, Zeng also mentions but ignores other signal-dependent noise effects in his thesis.

- Zeng ignored the "asymmetry between the positive and negative pulses" in the MR head response. *See* Zeng Thesis at § 6.5, p. 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.***").

- Zeng ignored the signal-dependent noise associated with non-transition sequences (e.g., $\oplus \oplus \oplus$ or $\ominus \ominus \ominus$), even though Zeng acknowledges (as does the Requester) that sequences without transitions nevertheless have signal-dependent noise. *See* Zeng Thesis at 10 (acknowledging that "more media noise is observed near transition regions than saturation regions," i.e., non-transition sequences); Request at 15, 28, 33, 39, 44, 52.[26]

See Bajorek Dec. at ¶ 95.

---

[26] In both Sections 4.4 and 5.2, Zeng's BMF for non-transition sequences (*see* BMF Nos. 1 in both Tables 1 and 2 above) do not contain Zeng's random jitter term even though Zeng admits that such sequences contain signal-dependent media noise. Zeng Thesis at 10.

C. Peer-Reviewed and Other Commentary on Work by Zeng and Inventors of CMU Patent

Long before the CMU case and this reexamination, Dr. Zeng's peers acknowledged that the CMU invention (including claim 4) was a true breakthrough that had not been previously discovered by anyone, including Dr. Zeng. Indeed, on at least two occasions, Zeng's Thesis advisor at the University of Minnesota, Prof. Moon, recognized Profs. Kavcic and Moura as "first deriving" signal-dependent detectors ***after*** supervising the work of his own former student, Dr. Zeng. These acknowledgements confirm that Zeng's Thesis does not anticipate claim 4.

The first instance where Prof. Moon recognized Profs. Kavcic and Moura, and not his former student, is Prof. Moon's 2001 paper, "Pattern-Dependent Noise Prediction in Signal-Dependent Noise," IEEE Journal on Selected Areas of Communications, Vol. 19, No. 4, April 2001 (Ex. 2). This peer-reviewed paper[27] was first submitted in March 2000, more than five years after Dr. Zeng completed his thesis. The paper's topic is the same as the '839 patent -- Viterbi detectors for signal-dependent noise channels. In the introduction, Prof. Moon and his co-author describes signal-dependent noise channels as ones "where the noise characteristics depend highly on the local bit patterns," (i.e., on a specific sequence of symbols). Ex. 2 at 730 (see excerpt below).

---

[27] A peer-reviewed paper undergoes an "assessment of quality by impartial expert reviewers before publication in a scholarly journal. The peer reviewers check the manuscript for accuracy and assess the validity of the research methodology and procedures. If appropriate, they suggest revisions. If they find the article lacking in scholarly validity and rigor, they reject it." Moura Dec. at ¶ 5; *see also* Kavcic Dec. at ¶ 28. The IEEE journal in which Prof. Moon's paper appeared states: "All papers are reviewed by competent referees and are considered in the basis of their significance, novelty, and usefulness to the Journal readership." Ex. 51.

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].

Immediately after this, Prof. Moon says that the "maximum likelihood sequence detector (MLSD) for signal-dependent Gaussian noise has ***been first derived in*** " reference [5] (*see* excerpt above). Reference [5] is one of the IEEE papers by Profs. Kavcic and Moura describing the invention in the '839 patent that Marvell copied when designing its MNP detector, and this paper discloses the subject matter of claim 4. *See* pp. 26-31 above and footnote 42 below.

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

Pertinent to this reexamination, Prof. Moon did not identify Dr. Zeng as "first deriving" a detector for signal-dependent noise where the noise "depends highly on the local bit patterns" even though Prof. Moon was Dr. Zeng's thesis advisor; instead he identified Profs. Kavcic and Moura.

Prof. Moon again credited Profs. Kavcic and Moura, and not his former student, in his book chapter on Magnetic Storage (Chap. 101) in The Communications Handbook, J. Gibson, ed., CRC Press, 2002 (Ex. 3). In this book chapter (excerpted below), under the heading "Glimpse of the Future" (§ 101.5), Prof. Moon explained that even with powerful turbo codes, "it is difficult to overcome the severe rate penalty of the magnetic channel." Prof. Moon says one way to do this is "to improve detector capability." He added, "Especially when medium noise dominates, there is significant room for improvement in detector performance." A "viable strategy" to do this according to Prof. Moon is "to incorporate pattern [signal]-dependent noise prediction into the Viterbi algorithm." In support of this observation, Prof. Moon did not cite to Zeng, but instead cited to the 2000 IEEE article by the CMU inventors, which paper discloses the subject matter of claim 4. *See* pp. 26-31 above and footnote 42 below. Again, the fact that Prof. Moon did not cite the work of his prior student is evidence that Zeng's thesis does not anticipate claim 4.

> **Pattern-Dependent Noise Prediction**
>
> Unless the code is extremely powerful as in turbo and LDPC coded systems, it is difficult to overcome the severe rate penalty of the magnetic channel. In light of this observation, one effective strategy to achieve an SNR gain is to improve the detector capability. Especially when medium noise dominates, there is a significant room for improvement in detection performance. A viable strategy along this direction is to incorporate pattern-dependent noise prediction into the Viterbi algorithm [Caroselli et al., 1997; Kavcic and Moura, 2000; Moon and Park, 2001]. This approach is based on the observation that the noise power changes from one branch (state-transition) to the next in the Viterbi trellis because of the pattern-dependent nature of medium noise. It can be shown that if the noise can be predicted on each branch separately based on predetermined local noise statistics, the overall immunity against medium noise can be improved significantly. This method is particularly promising in the design of future high density storage systems where medium noise is likely to be a major limiting factor.

But that is not all. Nearly ***ten years before*** Prof. Moon's articles, Dr. Inkyu Lee (the Requester's expert declarant) published a peer-reviewed paper (the Lee '92 paper, Ex. 1) that he co-authored with Prof. Cioffi of Stanford University. In that paper, Dr. Lee identified the exact same Zeng channel model that is set out in the Zeng Thesis, and from which Zeng's Section 4.4

and 5.2 branch metric functions are derived.[28] Contrary to the Requester's statements and Dr. Lee's declaration, Drs. Lee and Cioffi stated affirmatively that Zeng's channel model was not "data-dependent."

Below is an excerpt from Section 4 of that paper. Ex. 1 at 963.

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

The "References" section of the Lee '92 paper (*see* Ex. 1 at 964) shows that reference [6] is a 1992 paper authored by Zeng and his thesis advisor, Prof. Moon (the "Zeng-Moon 1992 paper," Ex. 58[29]), the content of which Zeng generally included in his Thesis in Sections 4.1 to 4.3. *See* McLaughlin Dec. at ¶ 66. Also, equation (10) in the excerpt above is ***exactly*** the same channel model equation 4.1 in Zeng's Thesis. *See* Zeng Thesis at 51. In fact, the three (unnumbered) equations after equation 10 in Lee's '92 paper are substantively the same as equations 4.4, 4.2 and 4.5 respectively of Zeng's Thesis, as shown in the table below. *See* McLaughlin Dec. at ¶ 66. Thus, it is clear that in Section 4 of his 1992 paper, Dr. Lee is describing the channel model that Zeng's uses in his thesis.

---

[28] Dr. Lee's CV attached to his reexamination declaration (Ex. I of the Request) did not include a citation to this paper.

[29] The Zeng-Moon 1992 paper is W. Zeng and J. Moon, "Modified Viterbi Algorithm for a Jitter-dominant 1-D^2 Channel," IEEE Trans. on Magnetics, vol. 28, No. 5, pp. 2895-97, Sept. 1992.

Zeng-Moon 1992 paper

LEE '92 PAPER	ZENG THESIS
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 $z(t) = \sum_{k=1}^{M} a_k h(t - kT - \Delta_k T) + n(t)$ (10)	**Zeng's Eq. 4.1:** $Z(t) = \sum_{k=1}^{N} a_k h(t - kT - \Delta_k T) + w(t).$
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 $\pm$ 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's Eq. 4.4:** $h(t) = \frac{\sin \pi t/T}{\pi t/T} + \frac{\sin \pi(t-T)/T}{\pi(t-T)/T}$
For every k, $h(t)$ can be expanded in a Taylor series in terms of Δ_k. Approximate the series by its first two terms. Then, $h(t - kT - \Delta_k T) = h(t - kT) - \Delta_k T h'(t - kT)$	**Zeng's Eq. 4.2:** $h(t - kT - \Delta_k T) = h(t - kT) - \Delta_k T h'(t - kT)$
Sample $z(t)$ at $t = kT$, then approximate it by leaving the most important two terms Δ_k and Δ_{k-1}. We obtain $z_k = y_k - 2a_k\Delta_k + 2a_{k-1}\Delta_{k-1} + n_k$	**Zeng's Eq. 4.5:** $Z_k = y_k + w_k - a_k\Delta_k + a_{k-1}\Delta_{k-1}$

In his 1992 paper, Dr. Lee described that Zeng's "jitter term Δ_k is assumed to be a ***white*** Gaussian random variable with variance σ_Δ^2 …." Ex. 1 at 963 (see below).

> The jitter term Δ_k is assumed to be a white Gaussian random variable with variance σ^2_Δ 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 $\pm$ 1. Assume $h(t)$ to be the response of the channel to a single transtion. Then,

"White" random variables are, as Dr. Lee admits in his papers, ***independent*** of other variables. *See* McLaughlin Dec. at ¶¶ 59 and 68. Confirming this point, Dr. Lee wrote, "[w]e assume that Δ_k and n_k are uncorrelated with each other and have zero means."

> where y_k is the noiseless output of the channel $1 - D^2$ equal to $b_k - b_{k-2}$. We assume that Δ_k and n_k are uncorrelated with each other and have zero means. Since both Δ_k and n_k are assumed to be Gaussian, the overall noise also becomes Gaussian. Then, we can apply the alogrithm developed in section 2 which is based on a Gaussian distribution.

A "***white*** Gaussian noise assumption," as Dr. Lee described Zeng's "random jitter," is one where "the noise samples are realizations of ***independent identically distributed*** Gaussian ***random variables*** with ***zero mean***…." '839 Patent at col. 5:59-62 (emphasis added). Indeed, the ordinary understanding of a "white" random process in statistical signal processing is a process with a flat (or constant) power spectral density (PSD), which happens when random variables are independent and identically distributed with a zero mean. *See* McLaughlin Dec. at ¶ 59 (citing authorities). Consequently, by calling Zeng's "random jitter" ***white***, Lee confirms CMU's assertion that Zeng's "random jitter" is ***not signal-dependent***.

In fact, Dr. Lee reiterated his opinion that Zeng's "random jitter" is not signal-dependent in the next paragraph of his paper where he criticized Zeng's model because Zeng's random jitter is not "data-dependent," i.e., not dependent on "past data history"; in other words, not signal-dependent. *See* Ex. 1 at 963 (see below).

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 repeated this exact same analysis and statement in his 1995 thesis that the Requester submitted with the Request. *See* Exhibit E of the Request (Lee's Thesis) at Section 6.2.3 at 99-100 (see below; reference [67] is the 1992 Zeng-Moon paper).

6.2.3 Channel model

In this section, the same channel model as in [66, 67] has been defined for performance analysis. The output of class-IV partial response (PR4) channel corrupted by jitter and additive noise is expressed as

$$z(t) = \sum_{k=1}^{M} a_k h(t - kT - \Delta_k T) + n(t) \qquad (6.12)$$

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 formula $a_k = (b_k - b_{k-1})/2$ where b_k is i.i.d. with value $\pm$ 1. Assume $h(t)$ to be the response of the channel to a single transition. Then,

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

For every k, $h(t)$ can be expanded in a Taylor series in terms of Δ_k. Approximate the series by its first two terms. Then,

$$h(t - kT - \Delta_k T) = h(t - kT) - \Delta_k T h'(t - kT)$$

Sample $z(t)$ at $t = kT$, then approximate it by leaving the most important two terms Δ_k and Δ_{k-1}. We obtain

$$z_k = y_k - 2a_k\Delta_k + 2a_{k-1}\Delta_{k-1} + n_k$$

where y_k is the noiseless output of the channel $1 - D^2$ equal to $b_k - b_{k-2}$. We assume that Δ_k and n_k are uncorrelated with each other and have zero means. Since both Δ_k and n_k are assumed to be Gaussian, the overall noise also becomes Gaussian. Then,

Chapter 6. A Nonlinear Channel Model for Digital Storage System 100

we can apply the algorithm developed in Section 6.2.1 which is based on a Gaussian distribution.

The assumption that Δ_k are 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 $\Delta(i_k)$, *i.e.* a function of i_k where i_k indicates past data history. But in this section, a simple white

Zeng Thesis p. 99 Zeng Thesis p. 100

V. <u>Summary of the Request and the Office Action</u>

The Requester requested reexamination of only claim 4 of the '839 patent and the Request raised three substantial new questions of patentability (SNQs).

1. Claim 4 is anticipated by Zeng's Thesis, and in particular by the "Four-State Trellis" embodiment in Section 4.4 of Zeng's Thesis (*see* Request at pp. 17-21)

and by the "16-State Trellis" embodiment in Section 5.2 of Zeng's Thesis (*see* Request at pp. 21-25);

2. Claim 4 would have been obvious based on the combination of Zeng's Thesis and Lee's Thesis (*see* Request at pp. 31-41); and

3. Claim 4 would have been obvious based on the combination of Zeng's Thesis and the Coker patent (*see* Request at pp. 42-57).

The Office Action accepted the SNQs proposed in the Request and incorporated them by reference. *See* Office Action at pp. 4-5.

CMU traverses the rejections for the reasons set forth below.

VI. <u>APPLICABLE LEGAL PRINCIPLES</u>

A. <u>Claim Construction</u>

In reexamination of an unexpired patent, "claims are given their broadest reasonable interpretation consistent with the specification and limitations in the specification are not read into the claims…." MPEP § 2258.G (citing *In re Yamamoto*, 740 F.2d 1569, 222 USPQ 934 (Fed. Cir. 1984)). Construing claims consistent with the specification means that "the problem the inventor was attempting to solve, as discerned from the specification and the prosecution history, is a relevant consideration." *CVI/Beta Ventures, Inc. v. Tura LP*, 112 F.3d 1146, 1161 (Fed. Cir. 1997). This is because claims terms should be construed to "remain faithful to the invention actually described in the prosecution history." *St. Clair Intellectual Property Consultants, Inc. v. Canon Inc.*, 412 Fed. Appx. 270, 275 (Fed. Cir. 2005); *ResQNet.com, Inc. v. Lansa, Inc.*, 346 F.3d 1374, 1380 (Fed. Cir. 2003) (the portions of the specification "relating to extant problems in prior art, properly confirms the meaning of claim language"); *Ex Parte Dolan*, 2012 WL 889728 (BPAI 2012) (reversing examiner's construction because not consistent with problems solved by inventor); *Ex parte Sosalla*, 2010 WL 4262205 (BPAI 2010) (same).

Here, for example, "signal-dependent branch metric function" must be construed in a way that solves the problem the inventors addressed with their invention as discerned from the specification. That includes using different branch metric functions for each specific symbol sequence (e.g., trellis branch) to account for both the neighborhood effect and the polarity of the transitions. *See, e.g.*, '839 patent at col. 2:18-20 (neighborhood effect); col. 5:49-52 (neighborhood effect); col. 3:60-64 (transition polarity)*; see also* Moura Dec. at ¶¶ 23-25.

B. Anticipation

"A claim is anticipated only if each and every element as set forth in the claim is found, either expressly or inherently described, in a single prior art reference." MPEP § 2131 (quoting *Verdegaal Bros. v. Union Oil Co. of California*, 814 F.2d 628, 631, 2 USPQ2d 1051, 1053 (Fed. Cir. 1987)). The identical invention must be shown in as complete detail as is contained in the claim, and the elements must be arranged as required by the claim. *See id*. Here, Zeng's Thesis does not anticipate claim 4 at least because it does not disclose any, let alone a "set of signal-dependent branch metric functions" that are "applied to a plurality of signal samples" from different sampling time instances.

Further, "[t]he disclosure in an assertedly anticipating reference must provide an enabling disclosure of the desired subject matter; mere naming or description of the subject matter is insufficient, if it cannot be produced without undue experimentation." MPEP § 2121.01 (quoting *Elan Pharm., Inc. v. Mayo Found. For Med. Educ. & Research,* 346 F.3d 1051, 1054, 68 USPQ2d 1373, 1376 (Fed. Cir. 2003)). Here, Zeng's Thesis is not enabling for the reasons set forth in Prof. Kavcic's declaration. *See* Kavcic Dec. at ¶¶ 15-42.

C. Obviousness

"A patent may not be obtained though the invention is not identically disclosed or described as set forth in [35 U.S.C. § 102], if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains." 35 U.S.C. § 103(a); *see also* MPEP ¶ 2141.

"Obviousness is a question of law based on underlying factual inquiries." MPEP ¶ 2141. The factual inquiries enunciated by the Supreme Court in *Graham v. John Deere Co.*, 383 U.S. 1, 148 USPQ 459 (1966) and reiterated in *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 82 USPQ2d 1385 (2007) are as follows:

(A) Determining the scope and content of the prior art;

(B) Ascertaining the differences between the claimed invention and the prior art; and

(C) Resolving the level of ordinary skill in the pertinent art.

MPEP ¶ 2141. Further, "[o]bjective evidence relevant to the issue of obviousness must be evaluated by Office personnel." *Id.* (citing *Graham*, 383 U.S. at 17-18, 148 USPQ at 467). "Such evidence, sometimes referred to as 'secondary considerations,' may include evidence of commercial success, long-felt but unsolved needs, failure of others, and unexpected results." *Id.*

"The question of obviousness must be resolved on the basis of the factual inquiries set forth above. While each case is different and must be decided on its own facts, the Graham factors, including secondary considerations when present, are the controlling inquiries in any obviousness analysis." *Id.* Further, when a reference is unworkable, inoperative or riddled with errors, like Zeng's Thesis, it should be accorded little weight in the obviousness analysis because there is little motivation to combine it with other prior art. *See Dennison Mfg. Co. v. Ben Clements & Sons, Inc.*, 467 F. Supp. 391, 415 (S.D.N.Y. 1979) ("[The reference] is vague and indefinite as to its exact teachings. These factors, in conjunction with the unworkability of the disclosed invention, might well have caused one skilled in the art who was considering the idea of connecting the paddles to avoid tangling to discard the notion and look for other solutions."); *Azoplate Corp. v. Silverlith, Inc.*, 367 F. Supp. 711, 718 (D. Del. 1973), *aff'd*, 506 F.2d 1050 (3d Cir. 1974) ("This is not an instance where teachings of positive results can be readily combined to achieve the invention. Instead, one would have to assume that despite the negative results inherent in the references, a skilled lithographer would [modify and combine]. The Court is unconvinced that there would be good reason for him to do so."); *see also United States v. Adams,* 383 U.S. 39, 52 (1966) ("We do say, however, that known disadvantages in old devices which would naturally discourage the search for new inventions may be taken into account in determining obviousness.”).

VII. ZENG DOES NOT ANTICIPATE CLAIM 4

A. Zeng Does Not Disclose Any, Much Less a “Set,” of Signal-Dependent Branch Metric Functions

Zeng’s branch metric functions in Sections 4.4 and 5.2 are not signal-dependent, and in any event these sections of Zeng’s Thesis do not disclose a “set” of (two or more) signal-dependent branch metric functions.[30] This is clear from at least the following reasons.

1. Zeng’s “Random Jitter” is White and Not Signal-Dependent

The only noise terms in Zeng’s branch metric functions in Sections 4.4 and 5.2 are the variance of the additive white noise (σ_w^2) and the variance of Zeng’s “random jitter” (σ_Δ^2). *See* Zeng Thesis at 65-68 and 75-79; McLaughlin Dec. at ¶¶ 49, 54; Bajorek Dec. at ¶¶ 99 and 105. Since “white” and “signal-dependent” are mutually exclusive (noise cannot be both white and signal-dependent) (*see* McLaughlin Dec. at ¶ 59), the variance of the additive white noise is clearly a white, non-signal-dependent term (not even the Requester argues that this term is signal-dependent). That leaves the variance of Zeng’s “random jitter” (σ_Δ^2) as the only possible term in Zeng’s branch metric functions in Sections 4.4 and 5.2 that could possibly be signal-dependent and make Zeng’s BMFs signal-dependent, but it too is ***white and not signal-dependent.***[31]

In his thesis, Zeng assumes that the random jitter variables Δ_k’s “are independently and identically distributed random variables with zero mean” and “independent of additive noise.”

---

[30] As noted previously, both the Requester and Dr. Lee agree that a “set” of signal-dependent branch metric functions must have “more than one” such function. *See* Request at 20; Lee Dec. at ¶ 13.

[31] As described above (*see* Section IV.B.2, pp. 42-44), Zeng took several steps to eliminate or reduce several signal-dependent noise sources in his channel and ignored the others. By eliminating or ignoring all sources of signal-dependent noise, Zeng’s Sections 4.4 and 5.2 BMFs do not need to account for signal-dependent noise, and they don’t.

Zeng Thesis at 52. This is the definition of a "white" random variable. *See* McLaughlin Dec. at ¶ 59. Indeed, even Dr. Lee, the Requester's expert, previously confirmed that Zeng's "random jitter" is white and not signal-dependent. *See* Ex. 1 (Lee '92 paper) at 963; Exhibit E of the Request at 99-100; Section IV.C of this response. Thus, Dr. McLaughlin, Dr. Lee, and the CMU patent itself ('839 patent at col. 1:38-41; 63-67) all confirm that Zeng's "random jitter" is ***white*** and ***not signal-dependent***.

In fact, in his declaration, Prof. Kavcic proves mathematically that the noise in Zeng's channel models in Sections 4.4 and 5.2, from which Zeng's BMFs are derived, is not signal-dependent noise by analyzing the correlation structure (or covariance structure) of Zeng's modeled noise. *See* Kavcic Dec. at ¶¶ 42-46.

Zeng's BMFs in Sections 4.4. and 5.2 are derived from his white random jitter channel model using eigenvalue decompositions (*see* Zeng Thesis at 65-68), and none of those BMFs are signal dependent. *See* McLaughlin Dec. at ¶¶ 4 and 59-61; Bajorek Dec. at ¶¶ 107-110; *see also* '839 patent at col. 1:63-67 ("There is a need for an adaptive correlation-sensitive maximum likelihood sequence detector (MLSD) without making the usual simplifying assumption that the noise samples are ***independent random variables***.").[32] Since there are no signal-dependent BMFs in Sections 4.4 and 5.2, there cannot be a "set" of (two or more) signal-dependent BMFs, as required by claim 4. Therefore, Sections 4.4 and 5.2 do not anticipate claim 4.

2. Zeng's Mathematical Notation Confirms that his Channel Model and BMFs are Not Signal-Dependent

Zeng's channel model equation 4.1 (shown below) demonstrates that Zeng's random jitter Δ_k is unrelated to -- or ***independent*** of -- the written symbols a_k. The value of the random jitter Δ_k does not depend on the written symbols a_k. *See* McLaughlin Dec. at ¶ 44.

[32] The '839 patent teaches that the "MLSD" of the invention is both "signal dependent and sensitive to the correlations between signal samples." '839 patent at col. 2:9-14 (Summary of the Invention).

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

Zeng's discrete channel model equations 4.22 and 5.5, from which his BMFs in Sections 4.4 and 5.2 are derived using eigenvector decomposition, further demonstrate the independence between Zeng's random jitter Δ_k and the written symbols a_k. The fact that the random jitter term Δ_k in equations 4.1, 4.4 and 5.2 is not dependent on a specific sequence of symbols -- or the polarity of the transitions -- is apparent from the fact that the subscript for the "random jitter" term Δ_k in Zeng's channel model is k and not a sequence of a_k's (i.e., the written symbols). *See* Bajorek Dec. at ¶ 109; McLaughlin Dec. at ¶ 44. If Zeng accounted for the fact that the noise in the readback signal is dependent on the sequence of written symbols (including the polarity of any transitions in the sequence and the neighborhood of written symbols), his random jitter term would have been indexed in terms of a past neighborhood of a_k's, such as $\Delta(a_k, a_{k-1}, a_{k-2})$ instead of as Δ_k, as explained in the CMU patents. *See, e.g.*, '839 patent at col. 4:24-27 ("Due to the signal dependent nature of media noise in magnetic recording, the functional form of joint conditional pdf $f(r_1, .., r_N|a_1,\ldots, a_N)$ in [equation] (1) is different for ***different symbol sequences*** $\boldsymbol{a_1, \ldots, a_N}$.") (emphasis added); col. 6:15-35 ("the variance σ^{2i} [sic] where the index i denotes the dependence on the written symbol sequence"); col. 6:36 - col. 7:5; *see also* Bajorek Dec. at ¶ 109; McLaughlin Dec. at ¶¶ 44, 72.

This is further proof that Zeng's BMFs in Sections 4.4 and 5.2 are not signal-dependent, and therefore do not anticipate claim 4.

3. <u>Zeng's BMFs in Sections 4.4 and 5.2 are Blind to the Polarities of the Transitions</u>

The '839 patent explains that signal-dependent branch metric functions cannot be blind to the polarities of the transitions in the recorded symbols. *See* '839 patent at col. 3:51-64. This is because differently polarized transitions have different signal-dependent noise structures. *See* Moura Dec. at ¶ 24; Bajorek Dec. at ¶¶ 61-71.

Zeng's BMFs, however, are blind to the polarity of the transitions as shown in Figures 8A-B above (repeated below), which makes Zeng's BMFs "inappropriate for the present problem" of accounting for the signal-dependent, correlated noise in the readback signal that the CMU invention solves. *See* McLaughlin Dec. at ¶¶ 61, 71-72; Bajorek Dec. at ¶¶ 77-78, 82, 85-87, 105, 109. Zeng's discrete channel models (equations 4.22 and 5.5) confirm that his random jitter Δ_k is not dependent on a specific sequence of symbols and, consequently, has the same random distribution (independent Gaussian with zero mean) for both positive and negative transitions. *See* Bajorek Dec. at ¶ 109; McLaughlin Dec. at ¶ 44 (the subscript for the "random jitter" term Δ_k in Zeng's channel model is k and not a sequence of a_k's).

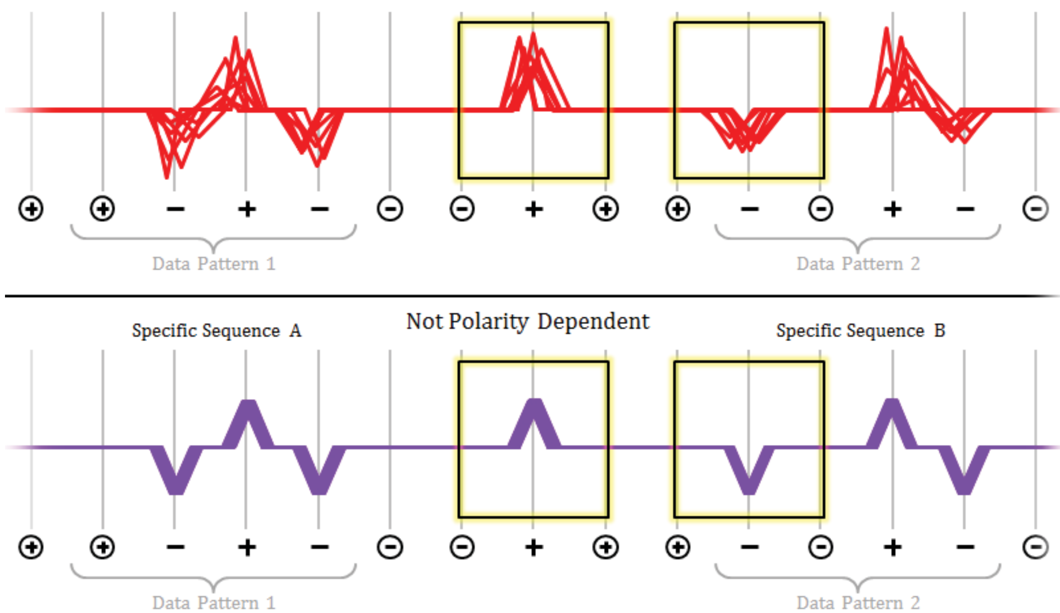


Figure 8A (Signal-dependent model)

Figure 8B (Zeng's Random Jitter model)

In other words, Zeng's model treats all transitions as 1's and all non-transitions as 0's (i.e., is blind to transition polarity). Further, the BMFs in Sections 4.4 and 5.2 based on those channel models are also blind to the transition polarity and not "signal-dependent." *See* McLaughlin Dec. at ¶ 70. Therefore, Zeng's BMFs in Sections 4.4 and 5.2 do not anticipate claim 4.

4. Zeng Uses the d=1 RLL Constraint in Sections 4.4 and 5.2, Which Eliminates the Signal-Dependent Neighborhood Effect from Zeng's BMFs

In addition to ignoring the polarity of the transitions, Zeng also uses a d=1 RLL constraint in Sections 4.4 and 5.2, which means that there must be at least d=1 nontransition regions between transitions because, according to Zeng, it is otherwise "impossible to implement" a detector. *See* Zeng Thesis at 4, 65-67, 75; Bajorek Dec. at ¶ 92. By using the d=1 RLL constraint, however, Zeng's BMFs cannot account for one of the "neighborhood effects" that he identified in his thesis, which is the peak shift that occurs where a transition is one bit interval earlier. *See* Zeng Thesis at 9; Bajorek Dec. ¶¶ 91-93 (Zeng also ignores the neighborhood effect when the transitions are "twice the bit interval"). By using the d=1 RLL constraint in conjunction with his "random jitter" noise model, Zeng eliminates and ignores all non-polarity effects, such as interactions (e.g., percolations) between neighboring transitions. *See* Bajorek Dec. ¶ 92. In essence, the d=1 RLL constraint allows Zeng to assume that there is no signal-dependent noise in his readback model. *See* Bajorek Dec. at ¶ 108. Consequently, Zeng has no reason to (and does not) consider the signal-dependent media noise arising from interacting transitions in his BMFs (i.e., Zeng does not consider media noise that is attributable to a *specific neighborhood* of interacting transitions/signals). *See* Bajorek Dec. ¶ 92. Therefore, for this independent reason, Zeng's BMFs in Sections 4.4 and 5.2 are not signal-dependent and therefore Sections 4.4 and 5.2 do not anticipate claim 4.

B. Even if Zeng's "Random Jitter" were Signal-Dependent, Zeng Still Does Not Disclose a Set of Signal-Dependent Branch Metric Functions

The Requester's argument that Zeng discloses a set of signal-dependent BMFs is premised on Dr. Lee's incorrect assertion that the variance of Zeng's random jitter (σ_{Δ}^2) is signal-dependent. *See* Request at 16; Lee Dec. at ¶ 34. As shown above, Zeng's random jitter (and its variance) is not signal-dependent. Requester's argument further fails because Zeng discloses only one BMF in Section 4.4 that includes the variance of the random jitter term and thus there is no set of signal-dependent BMFs. The same fatal flaw exists in the BMFs Zeng discloses in Section 5.2. *See* McLaughlin Dec. at ¶¶ 75-77. In the three BMFs for Section 4.4, only one equation, No. 3 in Table 3 below (*see* McLaughlin Dec. at ¶ 78), corresponding to Zeng's equation 4.28, includes the variance of Zeng's random jitter term (σ_Δ^2). The other two BMFs (BMF No. 1 and BMF No. 2 in Table 3) do not have such a term and, as a result, cannot be signal-dependent based on Dr. Lee's own rationale that the random jitter term σ_Δ^2 is what makes Zeng's BMFs signal-dependent. In other words, even under the incorrect interpretation that Zeng's "random jitter" is synonymous with or a form of "signal-dependent noise," (*see* Lee Dec. at ¶ 34), Zeng has at most only a ***single*** signal-dependent branch metric function in Section 4.4, and that cannot constitute a "set" of such functions as required by claim 4. *See* McLaughlin Dec. at ¶ 76.

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

Table 3
(* indicates a branch with a transition in the symbol sequence)

The same is true for Zeng's Section 5.2. Only one BMF in Section 5.2, No. 3 in Table 4 below (*see* McLaughlin Dec. at ¶ 79), has the random jitter term (σ_Δ^2) in it, and that lone function cannot constitute the claimed "set" of signal-dependent branch metric functions. *See* McLaughlin Dec. at ¶ 77.

Function 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	$\ln\sigma_w^2 + \frac{\frac{j_0}{j_o^2+j_1^2}[ZD_k - t_a] - \frac{j_1}{j_o^2+j_1^2}[ZD_{k-1} - t_b]^2}{\sigma_w^2}$	$t_a = 2g_{-1}-2g_1$ $t_b = -2g_0$	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	$\ln[\sigma_w^2 + 4(j_o^2+j_1^2)\sigma_\Delta^2] + \frac{\frac{j_1}{j_o^2+j_1^2}[ZD_k - t_a] + \frac{j_0}{j_o^2+j_1^2}[ZD_{k+1} - t_b]^2}{\sigma_w^2 + 4(j_o^2+j_1^2)\sigma_\Delta^2}$	$t_a = -2g_0$ $t_b = -2g_1$	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*		

Table 4
(* indicates a branch with a transition in the symbol sequence)

C. <u>Drs. Lee and Moon Confirm that Zeng's Thesis Does Not Anticipate</u>

There also is abundant additional evidence that confirms the technical analysis described in the preceding sections and demonstrates that Sections 4.4 and 5.2 of Zeng's Thesis do not anticipate claim 4. *See Dayco Prods. Inc. v. Total Containment, Inc.*, 329 F.3d 1358, 1368-69 (Fed. Cir. 2003) (the views of persons of ordinary skill are probative regarding the content of a reference); *see also Ciba-Geigy Corp. v. Alza Corp.*, 68 F.3d 487 (Fed. Cir. 1995) ("[E]xtrinsic evidence may be considered when it is used to explain … the meaning of a reference."); *AstraZeneca LP v. Apotex, Inc.*, 633 F.3d 1042, 1055-56 (Fed. Cir. 2010) (considering expert testimony regarding how someone of ordinary skill in the art would have understood the reference).

As explained above, Dr. Lee, the Requester's expert declarant, acknowledged twenty years ago that Zeng's "random jitter" is "***white***." *See* Ex. 1 (Lee '92 paper) at 963; Exhibit E of the Request (Lee's Thesis) at 99-100; McLaughlin Dec. at ¶¶ 64-69. "White" and "signal-dependent" are mutually exclusive; noise cannot be both white and signal-dependent. *See* McLaughlin at ¶ 59. So Zeng's "random jitter" cannot be signal-dependent, even according to Dr. Lee. In fact, Dr. Lee confirmed his opinion when he criticized Zeng's model because the random jitter term was not "data-dependent," i.e., it was not a function of "past data history." *See* Ex. 1 (Lee '92 paper) at 963; Exhibit E of the Request (Lee's Thesis) at 99-100; McLaughlin Dec. at ¶¶ 64-69. By concluding that Zeng's "random jitter" is not a function of "past data history," Lee admits that Zeng's "random jitter" is independent of the symbols written to the disk and, therefore, not signal-dependent.

Further, Zeng's own thesis advisor credited Profs. Kavcic and Moura with "first deriving" a detector that accounts for "***signal-dependent*** Gaussian noise…" six years after overseeing Zeng's work. *See* Ex. 2 at 730. In fact, Prof. Moon ignored his prior student's work a second time in favor of Prof. Kavcic's and Moura's work. *See* Ex. 3 at 101-12. The fact that Prof. Moon twice credited Profs. Kavcic and Moura, and not his former student, confirms that Zeng's Section 4.4 and 5.2 branch metric functions do not anticipate claim 4.

D. Marvell's Actions Both Before and During the CMU Case Confirm that Zeng's Thesis Does Not Anticipate

In all likelihood, Marvell requested this reexamination. But regardless of who requested it, both Marvell's actions in the CMU case and the actions of Marvell's engineers prior to being sued demonstrate the Zeng's Thesis does not anticipate claim 4.

1. Marvell is the Requester

More than a year after the verdict in the CMU case, J. Steven Baughman of the law firm of Ropes & Gray filed two Requests for Reexamination seeking to invalidate the claims of the CMU patents that CMU asserted against Marvell. Although Mr. Baughman identified himself as the "Requester," the following objective facts indicate that Marvell is the real-party-in-interest behind the Requests:

- The Request sought to invalidate only the claims that CMU asserted against Marvell in the litigation, namely claim 4 of the '839 patent and claim 2 of the '180 patent.

- Ropes & Gray's website lists Marvell as a client (*see* Ex. 9), and the PTO's public records confirm that Ropes & Gray presently represents Marvell before the PTO in connection with numerous pending patent applications, *see, e.g.*, U.S. Patent Serial. No. 14/166,428. In fact, two of the inventors for this application (Zining Wu and Gregory Burd) are Marvell employees that Marvell called as witnesses in the CMU case.

- Several arguments made in the Requests track arguments that Marvell made in the litigation. For example, Dr. Lee in his reexamination declaration discussion of Section 5.2 of Zeng's Thesis, the Requester, and Marvell in the CMU case, all take the position that changing the input to a function (here, the target value, which is one of the two inputs of a distance function) creates a new function. This assertion violates well-known principles of mathematics and the court in the CMU case denied Marvell's motions premised on this ill-conceived notion. *See* Ex. 10.

- Any interested party other than Marvell could have filed a petition for an *inter partes* review instead of an *ex parte* reexamination request.[33]

- Marvell has significant connections with Dr. Zeng (the author of the primary reference) and Dr. Lee (the Requester's declarant).

  - Dr. Zeng has been an employee at Marvell since April 2004. *See* Ex. 11 and Ex. 12.

  - As to Dr. Lee, Prof. John M. Cioffi at Stanford University was Dr. Lee's principal thesis advisor. *See* Exhibit E of Request at p. iii. From 1999-2006, Prof. Cioffi was a member of Marvell's board of directors, chaired its compensation committee and helped guide the company when it went public in 2000. *See* Ex. 14. Additionally, Dr. Zining Wu, Marvell's Chief Technology Officer, also was a Ph.D. student of Dr. Cioffi at the same time as Dr. Lee (1994-1995). *See* Ex. 15 (Tr. 12/13/12) at 7-9.

Regardless of whether Marvell is the "man behind the curtain," however, the facts and circumstances established during the CMU case confirm CMU's assertion that the Zeng Thesis does not invalidate Claim 4 of the '839 Patent.

2. <u>Marvell's Actions in the CMU Case Confirm CMU's Assertion that Zeng Does Not Invalidate Claim 4</u>

Just as the prior statements of Drs. Lee and Moon confirm CMU's position, the actions of Marvell and its expert witness in the CMU case, Dr. John Proakis, demonstrate that Sections 4.4 and 5.2 do not anticipate claim 4.

---

[33] CMU asked Marvell's litigation counsel whether Marvell was involved in any way with the Requests. Marvell's counsel refused to admit or deny Marvell's involvement. *See* Ex. 13.

a. Summary of the Litigation

In the CMU case, CMU asserted that Marvell infringed claim 4 of the '839 patent and claim 2 of the '180 patent when it and its customers used Marvell's "KavcicViterbi" simulator and its MNP and NLD-type chips and simulators during Marvell's so-called "sales cycle" which includes research, development, chip design, qualification, evaluation and sales. As a result of those infringing uses, Marvell sold more than 2.34 billion MNP and NLD-type HDD read channel chips from March 2003 to July 2012. CMU sought damages for Marvell's infringement in the form of a $0.50 royalty for the MNP and NLD-type chip Marvell sold. Marvell counterclaimed alleging, among other things, that the Asserted Claims were invalid.

The litigation up to trial consumed approximately three and a half years and included numerous hearings, a court-appointed technical special master, and numerous expert reports and expert depositions that concerned the merits of the CMU Patents.[34] A team of attorneys that included lawyers with Ph.D.'s in electrical engineering and physics represented Marvell over the course of the litigation. *See e.g.*, Ex. 16 (Dkt. 900) at 8, n.8; Ex. 17 (Dkt. 901) at 91-92, nn. 93-94.[35] Of course, Marvell also had access to its own vast array of technical personnel, including (i) its CEO, Dr. Sehat Sutardja, who has a Ph.D. in electrical engineering and is an IEEE Fellow, (ii) its present CTO, Dr. Zining Wu, whom Dr. Sutardja described as "the most brilliant scientist that I have ever known," and (iii) Mr. Gregory Burd, another Marvell engineer that Dr. Sutardja described as "another brilliant scientist that happens to work for Zining Wu." *See* Ex. 19 (Tr. 12/11/12) at 35-57. The CMU case culminated with a four-week trial from November through

[34] The district court engaged Prof. Thomas Costello, the Bettex Chair Professor Emeritus in the Department of Electrical Engineering at Notre Dame University, as a technical expert to provide the court with consultation on the technology described in the CMU patents. Professor Costello is a digital signal processing expert with experience in the field of the CMU patents. Ex. 18 (Dkt. 145-46). CMU and Marvell stipulated that Prof. Costello was qualified to be a technical expert for the court.

[35] Marvell's lead counsel for most of the litigation was Dr. David Radulescu who received a Ph.D. in Electrical Engineering from Cornell University. Marvell's team of lawyers also included Dr. Mark Tung, who received a Ph.D. in Physics from the University of California, Berkley, and Dr. Anna Ison, who received a Ph.D. in Electrical Engineering from the University of California, Berkley.

December 2012[36] and included testimony from each side's highly-qualified experts.[37] Marvell's invalidity expert was Dr. John Proakis, an adjunct professor at the University of California at San Diego, Professor Emeritus at Northeastern University in Boston, Massachusetts and an author or co-author of five textbooks in in the field of digital signal processing and communications.[38]

Despite Marvell's team of lawyers and the testimony of Dr. Proakis, a nine-member jury unanimously found that Marvell directly and indirectly infringed the Asserted Claims. The jury also found that Marvell failed to prove that the Asserted Claims were invalid under 35 U.S.C. §§ 102 and 103. In light of Marvell's pervasive and profitable infringement, the jury awarded CMU approximately $1.17 billion as damages (a $0.50 royalty on the MNP and NLD-type chips Marvell sold), which is one of the largest verdicts ever in a patent case.

Both parties filed post-trial motions. The court denied Marvell's motions for judgment as a matter of law and/or a new trial on infringement, validity and damages. Ex. 17 (Dkt. 901). The court granted CMU's motion for a finding of willful infringement and enhanced the jury's damages award by $287,198,828.60. Ex. 17 (Dkt. 901); Ex. 7 (Dkt. 933). In deciding to enhance the damages award, the court held that the amount "was sufficient to penalize Marvell

---

[36] *See* Ex. 16 (Dkt. 900) at 1, n. 1 ("The trial ran from 9:00 a.m. to 5:30 p.m. Monday through Friday, with counsel and the Court arguing objections and motions most days starting at 7:30 a.m. and after trial until 7:00 p.m., sometimes 9:00 p.m. The trial included 3 hours of openings, 3 hours of closing arguments, 171 exhibits, 20 witnesses, 1,100 slides, over 130 sidebars, and nearly 4,000 pages of trial transcript. (Docket Nos. 666-765, 770, 771). There were 9 jurors, each of whom were given a binder containing a copy of the patents, the Court's claim construction, initial instructions, a glossary of terms, and pages for notes. (Docket No. 671). Each juror also received a notepad to take notes, which all of them used, with one juror actually filling 3 such notepads by the end of the trial.").

[37] CMU's expert on infringement and validity was Prof. McLaughlin. Dr. Bajorek also testified as an expert on behalf of CMU.

[38] *See* www.jacobsschool.ucsd.edu/faculty/faculty_bios/index.sfe?fmp_recid=94. Marvell also relied on technical expert testimony from Dr. Jack Wolf (who passed away during the pendency of the litigation) and Dr. Richard Blahut, both of whom are well-credentialed in the field of digital signal processing. *See* cmrr.ucsd.edu/people/wolf/ and www.ece.illinois.edu/directory/profile.asp?blahut.

for its egregious behavior" which included "deliberate and extensive copying of [CMU's] patented methods." Ex. 7 (Dkt. 933) at 42, 45.

Marvell has appealed the verdict to the Court of Appeals for the Federal Circuit. The appeal is currently pending and the Federal Circuit has not made any rulings on the merits.

b. Despite Facing Substantial Damages and Raising the Zeng Thesis as an Anticipatory Reference During the Litigation, Marvell Decided Not To Use Zeng or His Thesis at Trial

As it faced a claim of damages in excess of $1 billion, Marvell initially asserted nearly every possible defense, including multiple invalidity defenses. For example, on August 17, ***2009***, Marvell identified, Dr. Zeng, an employee of Marvell since 2004, as a prospective fact witness with "discoverable information concerning prior art relevant to CMU's asserted patents." *See* Ex. 12 at 3-4 (see below).

> **A. Witnesses**
>
> Marvell identifies the following individuals who are likely to have discoverable information that Marvell may use to support its defenses and that MSI may use to support its Counterclaims. Unless otherwise indicated, these individuals are Marvell employees and may only be contacted through Marvell's counsel.
>
> * * *
>
> 7. Weining Zeng
>
> Mr. Zeng has discoverable information concerning prior art relevant to CMU's asserted patents.

Three months later, Marvell provided CMU with invalidity contentions in which Marvell asserted that sixteen purported prior art references, including Zeng's Thesis, either anticipated the claims of the CMU patents or rendered them obvious. Ex. 20 at 003-004. Marvell charted an anticipation argument based on Zeng's Thesis, ***relying on the identical sections that the Requester used here***. *Id.* at 161-166.[39] Marvell also contended that the Asserted Claims were invalid for failure to meet the written description, enablement, and definiteness requirements of 35 U.S.C. § 112. *Id.* at 054-060.

As the case progressed, Marvell narrowed its defenses -- presumably focusing on what it believed were its strongest arguments. *See* Ex. 17 (Dkt. 901) at 80 (Judge Fischer noting that Marvell "trotted out a number of different non-infringement and invalidity defenses throughout its four years litigating" the case and concluding that if Marvell thought any of those defenses were "reasonable," it would have presented them to the jury). After Marvell identified Zeng's Thesis in its invalidity contentions in 2009 (*see* Ex. 20 at 161-166), Marvell never again included it as part of its invalidity arguments leading up to and at trial in December 2012.[40] Even though in the interim Marvell's invalidity expert, Dr. Proakis, identified Zeng's Thesis as part of the materials he reviewed and considered in connection with his one-hundred page invalidity report, he did not even bother to copy into his report the anticipation chart that Marvell had prepared previously in its invalidity contentions. *See* Ex. 21 (Dr. Proakis's Jan. 17, 2012 report). And he said nothing about Zeng's Thesis at trial.[41] Additionally, during the four years that Marvell and

[39] Marvell produced Zeng's Thesis in the first one hundred fifty pages of the more than six million pages that Marvell produced in total during the litigation.

[40] In its pretrial motions and at trial, Marvell relied almost exclusively on U.S. Patent No. 6,282,251 ("Worstell") which the Court and the jury soundly rejected as invalidating prior art. *See, e.g.*, Ex. 17 (Dkt. 901) at 53-66, 78-80.

[41] Although Marvell did not rely on Zeng's Thesis at trial, Marvell pursued arguments based on a branch metric equation in a 1992 paper by Zeng and his Ph.D. advisor, Prof. Moon, Ex. 58 ("the Zeng-Moon 1992 paper"), during the course of the CMU case (including the trial and thereafter). The subject matter of the Zeng-Moon 1992 paper (which the Office considered during the original prosecution of the '839 patent) is included generally in Sections 4.1 to 4.3 of Zeng's Thesis. The branch metric equation and associated explanatory material in Sections 4.1-4.3 and in the Zeng-Moon 1992 paper differ from the equations and materials in Sections 4.4 and 5.2 of Zeng's Thesis upon which the Requester relies here (even though they start with the same channel model). The Requester acknowledges the overlap. *See* Request at footnote 7 (stating that the Zeng-Moon 1992 paper has "limited overlapping subject matter" with Zeng's Thesis but confirming that there are "multiple different branch metric functions" in the thesis). The Requester, however, confined its Request to the branch metric functions in Sections 4.4 and 5.2 of Zeng's Thesis. This strategic decision is not surprising because among other things: (1) the

its experts studied the validity of the CMU patents and the possible effect Zeng's Thesis had on the issue, Marvell never called Dr. Zeng as a witness at a deposition or at trial to testify even generally about his work or specifically that his thesis invalidates any of the claims of the CMU patents. The fact that Marvell would not call Zeng as a witness and prioritized other defenses over Zeng's Thesis, while facing (and ultimately losing) a judgment greater than $1 billion, demonstrates that Zeng's Thesis does not anticipate claim 4. If Zeng's Thesis truly was anticipatory, Marvell and its experts would have asserted it. The fact that they did not speaks volumes. Indeed, the court expressly found that throughout the course of the entire litigation, ***Marvell failed to proffer any objectively reasonable invalidity (or non-infringement) defenses***, which included its original assertion that the Zeng Thesis invalidates claim 4. Ex. 17 (Dkt. 901) at 80

Given the stakes of the litigation, the only reasonable conclusion from the foregoing is that the army of experts, lawyers and internal technical experts that Marvell brought to bear on the litigation agree with CMU's assessment that Zeng's Thesis does not invalidate claim 4 of the '839 patent. By asserting in this reexamination proceeding instead of the trial that Zeng's Thesis anticipates, the Requester chose a forum that shields Drs. Lee and Zeng from cross-examination -- presumably because cross-examination would promptly expose the flaws in the Request.

---

branch metric equation in Sections 4.1 to 4.3 and/or the branch metric equation in the Zeng-Moon 1992 paper uses only a single signal sample; (2) Marvell's invalidity expert, Dr. Proakis, did not opine that the Sections 4.1 to 4.3 and/or branch metric equation in the Zeng-Moon 1992 paper anticipate claim 4 (*see* Ex. 21)*;* and (3) the Court and jury determined that claim 4 is nonobvious in view of the Zeng-Moon 1992 paper. *See* Ex. 17 (Dkt. 901) at 62-66. Given the limitations of the Request and the prior rejection of the invalidity arguments based upon the branch metric equation in the 1992 Zeng-Moon paper, CMU does not address those arguments here.

3. Prior to Being Sued, Marvell's Engineers Copied Claim 4, Not Zeng's Thesis

Marvell's pre-litigation actions further confirm that Zeng's Thesis does not invalidate claim 4 of the '839 Patent. Marvell knew about Dr. Zeng's work in at least 2001 before it began developing its infringing MNP and NLD technologies. Dr. Zining Wu, Marvell's current Chief Technology Officer, stated under oath in a declaration filed with the district court that ***in early 2001***, Marvell was "researching various options for addressing media noise" and "reviewed literature and published papers by individuals in the field, such as by Dr. Cioffi, ***Dr. Zeng***, Dr. Moon and Dr. Kavcic." *See* Ex. 22 (Dkt. 802-2) at ¶ 11 (emphasis added). After performing that research, Marvell did ***not*** turn to Dr. Zeng's work, but instead copied CMU's patented method -- knowing at the time that it was patented. Marvell engineers admitted to reading and copying two IEEE papers by Profs. Kavcic and Moura that describe the invention in the '839 patent. Ex. 23 (Tr. 12/3/12) at 77-79; Ex. 24 (P-Demo 7) at 28-33, 73; Ex. 30 (P-366). Those papers are A. Kavcic and J. Moura, "Correlation-Sensitive Adaptive Sequence Detection," IEEE Trans. on Magnetics, vol. 34, pp. 763-771 (Ex. 25) and A. Kavcic and J. Moura, "The Viterbi Algorithm and Markov Noise Memory," IEEE Trans. on Information Theory, vol. 46, pp. 291-301 (Ex. 26). Indeed, the Marvell engineer responsible for designing the MNP admitted that Prof. Kavcic's work was the "launching pad" for Marvell's research. Ex. 24 (P-Demo 7) at 107.

After hearing all of the evidence and the parties' arguments over the four-week trial, the district court unequivocally found that Marvell knew about the CMU patents and consciously copied the Asserted Claims:

> The evidence at trial clearly and convincingly shows that Marvell had knowledge of the patents-in-suit at the time of infringement by 2002 and that the very people who designed the Accused [MNP and NLD] Technology knew of the patents. Ex. 17 (Dkt. 901) at 67.
>
> * * *
>
> Marvell engineer Gregory Burd, the developer of the Accused [MNP and NLD] Technology, stated that he read Dr. Kavcic's published papers and learned about his Viterbi detector. He told

> his supervisor, Toai Doan, about his work on "Kavcic's model" in 2001, and stated that he was able to develop a sub-optimal media noise detector based on the Kavcic model from Kavcic's IEEE Paper. Ex. 17 at 68 (citations omitted).
>
> * * *
>
> Mr. Burd testified that he used Dr. Kavcic's model to create a simulation program at Marvell. Mr. Burd named his model KavcicPP, and he named his optimal simulator KavcicViterbi. Ex. 17 at 68 (citations omitted).
>
> * * *
>
> "Although Mr. Burd stated that he was 'generally following the papers,' not the patents, and that he 'left it at that,' [citation omitted], Dr. McLaughlin testified that the papers are virtually identical to what is described in the patents" Ex. 17 at 82-83 (citations omitted).
>
> * * *
>
> "Marvell… comes before the Court with unclean hands ***after having engaged in deliberate and sustained copying of the patented method*** throughout the entire laches period and up to the present…." Ex. 54 (Dkt. 920) at 71 (emphasis added).
>
> * * *
>
> The "Court holds that the credible evidence presented at trial sufficiently establishes that ***Marvell deliberately copied CMU's Patents***…." Ex. 7 (Dkt. 933) at 19-20 (emphasis added).
>
> * * *
>
> The "Court believes that a penalty of enhanced damages should be assessed against Marvell given its… known willful infringement through its ***deliberate and extensive copying of the patented methods***…" Ex. 7 (Dkt. 933) at 42 (emphasis added).

Thus, despite express knowledge of Zeng's work, Marvell copied the Kavcic and Moura papers, knowing them to be the same as the CMU patents (which they are).[42] Marvell even named its

[42] CMU's witnesses established that the "Professor Kavcic papers" Mr. Burd used disclose the invention set out in claim 4 of the '839 patent. Ex. 23 (12/3/12) Tr. at 66-67; Ex. 31 (11/29/12

first generation infringing product after Dr. Kavcic, calling it the "KavcicPP." *See* Ex. 27 (P-279) (Burd writing he "developed sub-optimal media noise detector based on Kavcic model"); Ex. 28 (P-280) (Burd names the write up of that media noise detector "Kavcic PP"); Ex. 29 (P-196) (Burd's notebook write up of Kavcic PP); Ex. 23 (12/3/12 Tr.) at 65-67, 71-73.

Even though Marvell was aware of the '839 Patent and consciously copied it into its "Kavcic PP" detector (later named the MNP), Marvell still did not turn to Dr. Zeng's work when it developed its second generation read channel, the NLD-type chips. Instead, in 2003, Marvell doubled-down on its infringement. As Dr. Wu wrote, the enhancement to the "Kavcic PP" detector that ultimately became Marvell's NLD was "the original structure that Kavcic proposed in his paper." Ex. 30 (P-366) (shown below); Ex. 17 (Dkt. 901) at 82-83.

Tr.) at 68-69 (Dr. Moura testifying that the FIR filter implementation of the 2000 paper is in the patents); Ex. 32 (11/30/12 Tr.) at 154 (Dr. Kavcic testified that "[w]hat is described in this article is exactly the methods of the patents."); id. at 155-58 (Equation 19 in the article is equation 13 in the '180 patent, "[s]o what is described in this paper is exactly the same content that was described in the patent."); Ex. 23 (12/3/12 Tr.) at 77-79 (Dr. McLaughlin explaining that "'180 Equation 13, is actually the exact same Equation 13 in the 1998 papers," and Fig. 3B from the'180 patent shows the FIR filter described in P-169); Ex. 24 (P-Demo 7) at 29-33; Ex. 31 (11/29/12 Tr.) at 229-32; and Ex. 33 (P-Demo 3) at 60-66 (Dr. Kavcic mapping claim 4 of the '839 Patent into the specification which is the same text as the IEEE paper Dr. McLaughlin evaluated).

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.

Even after (i) the jury's $1.17 billion verdict, (ii) being found to have willfully infringed claim 4 of the '839 patent in the CMU case, and (iii) being ordered by Judge Fischer in the CMU case to pay $0.50 for each infringing chip it sells through the life of the '839 patent, ***Marvell still sells chips specifically designed to practice the method of claim 4***. This ongoing royalty amounts to ***over $15 million per month***. *See* Ex. 52 (Dkt. 943) at 2 (over 100 million chips sold in three-month time period from Nov. 2013 through Jan. 2014, at $0.50 per chip, amounts to over $15 million per month) and Ex. 59 (Dkt. 961) at 2 (over 185 million chips sold in the six-month time period from February through July 2014). Marvell continues to use the Asserted Claims in its MNP and NLD products since it is an industry standard. *See* Section VII.D.4 (pp. 74-79) below. And Marvell continues to use its Kavcic Viterbi simulator as the "yardstick" and "gold standard" by which to measure the performance of its read channels, *see* Ex. 23 (12/3/12 Tr.) at 172-73; Ex. 24 (P-Demo 7) at 110; Ex. 34 (11/28/12 Tr.) at 156, 166, because, according to Mr. Burd, Kavcic is "considered to be, you know, on a leading edge, or on the cutting edge of a field" and a "kind of VIP which everybody tries to cite." Ex. 24 (P-Demo 7) at 110.[43]

[43] In July 2013, more than seven months after the verdict, the parties in the CMU case filed a joint status report where Marvell updated the district court on the "status of 'next generation chips without NLD functionality.'" Ex. 53 (Dkt. 889) at 3. In that joint status report, Marvell

4. The Evidence in the CMU Case Showed that Claim 4 is Very Valuable

The method of claim 4 permits branch metric values to be computed in a Viterbi detector in a manner that accounts for the correlated, signal-dependent media noise. Consequently, as HDD manufacturers pack more and more data into smaller and smaller areas (i.e., increasing the data density), which increases the already dominant media noise in the channel, the data can still be accurately read, allowing for smaller HDDs with more storage capacity -- exactly what consumers demand. Indeed, the evidence in the CMU case showed that the branch metric value computation methods of the Asserted Claims were, ***and continue to be***, more than ten years after their introduction, "must have" for Marvell and a standard in the HDD industry.

a. CMU's Invention was "Must Have" for Marvell

Before it began infringing, Marvell's customers perceived that Marvell's read channel products were "a year behind" its then-biggest competitor, "Lucent." Ex. 35 (P-208) at 4; Ex. 36 (Tr. 12/4/12) at 117-118. Falling behind causes companies to lose sales and risk going out of business. Ex. 36 (Tr. 12/4/12) at 117-118. Marvell's product development efforts were focused on its "iterative" detector, a move which its customers called "risky," and which turned out to be a major mistake. The iterative chip became a "lost cause" and a "disaster" for Marvell. Ex. 36 (Tr. 12/4/12) at 118-23; Ex. 37 (P-209) at 12; Ex. 38 (P-240); Ex. 39 (P-285). Its failure "affect[ed] [Marvell's] ability to remain competitive in signal to noise ratio against other chip suppliers." Ex. 36 (Tr. 12/4/12) at 123. Dr. Bajorek and two of Marvell's own executives testified that betting on the wrong technology can drive companies out of business. Ex. 36 (Tr. 12/4/12) at 123; Ex. 40 (JX-C) at 360. Also, in 2001-02 Marvell's customers were specifically

claimed that it was redesigning its infringing NLD chips by "permanently and irreversibly disabl[ing] the accused NLD functionality such that there can be no argument that branch metric functions is applied a plurality of signal samples." *Id.* at 4. Marvell's timetable for this purported design-around "slipped by six months" and "there is no indication that Marvell's new allegedly noninfringing read channel will be acceptable to Marvell's customers or that Marvell will ever mass produce the [redesigned] circuits…." *Id.* at 4-5. Since making this claim in July 2013, Marvell has not provided the district court or CMU with any evidence that it has implemented this purported design-around.

demanding “a noise processing capability on these chips.” Ex. 36 (Tr. 12/4/12) at 123-24. Marvell therefore developed “the Kavcic postprocessor” (later renamed the “MNP”) and identified it as “critical” to Marvell’s success. Ex. 36 (Tr. 12/4/12) at 126, 130. For example, Marvell sent urgent emails indicating it aggressively planned to “pull in” (i.e., accelerate) the schedule for the MNP to meet customer demands, which was a “very serious objective for Marvell.” Ex. 36 (Tr. 12/4/12)at 130; Ex. 41 (P-304). Marvell also believed that its competitors had or were developing a media noise processor. Ex. 36 (Tr. 12/4/12) at 131; Ex. 42 (P-320). Marvell employees reported to “senior management” that “we must have MNP,” and “no one disagreed.” Ex. 42 (P-320). At the same time, Marvell employees reported that MNP “is [a] critical requirement” for customers such as Hitachi and Fujitsu. Ex. 36 (Tr. 12/4/12) at 132; Ex. 43 (P-328). Marvell’s efforts to “pull in” the MNP to chips in the “mature” part of the sales cycle was unusual and showed that Marvell was “in crisis mode.” Ex. 36 (Tr. 12/4/12) at 134; *see also id.* at 130; Ex. 44 (Tr. 12/7/12) at 118 (testimony from CMU’s damages expert, Ms. Catherine Lawton, that she “found that Marvell accelerated the delivery of the MNP-type chips by about one year.”).

b. Before Introducing the MNP, Marvell’s Sales Were Declining, But Afterwards Sales Spiked and Sales of Noninfringing Chips Dropped Almost Immediately to Zero

Before Marvell offered chips with an MNP, Marvell’s sales of successive chip families were decreasing, even though the market was growing. Ex. 44 (Tr. 12/7/12) at 115-17; Ex. 45 (P-Demo 13) at Chart 22. Also, this was a time of intense industry consolidation and various Marvell competitors were exiting the desktop market, but Marvell’s remaining competitors—not Marvell—were “getting this business.” Ex. 44 (Tr. 12/7/12) at 127, 203-04; Ex. 45 (P-Demo 13) at Chart 11. Marvell recognized that the industry would consolidate further such that there would be huge winners and losers in the marketplace and correspondingly huge risks and potential rewards for Marvell. Ex. 44 (Tr. 12/7/12) at 125-32; Ex. 46 (P-935). When Marvell introduced the MNP, its sales spiked. Ex. 44 (Tr. 12/7/12) at 119. Moreover, Marvell offered versions of its older, noninfringing chips without the MNP, but no customer elected to go to volume production with such chips. Ex. 44 (Tr. 12/7/12) at 108-110, 124-25; Ex. 45 (P-Demo 13) at Charts 5, 22-23. Instead, “Marvell’s entire business had converted by that point in time, that's 2005, to be entirely MNP and NLD-type chips." Ex. 47 (Tr. 12/10/12) at 84; Ex. 48 (P-Demo 16) at Chart 27. CMU's damages expert, Ms. Catherine Lawton, illustrated her testimony on this point, in part, with demonstrative exhibits for Marvell's stand-alone read channel products and its SOCs.[44] *See* Ex. 45 (P-Demo 13) at Charts 5 and 22.

Ms. Lawton's Chart 22 (shown below) shows the sales over time for Marvell's stand-alone read channel products. The blue, black, green and yellow bars indicate the declining sales of the noninfringing read channels prior to a spike in 2004 when the infringing MNP and later the NLD read channels (the red bars) were introduced to the market.

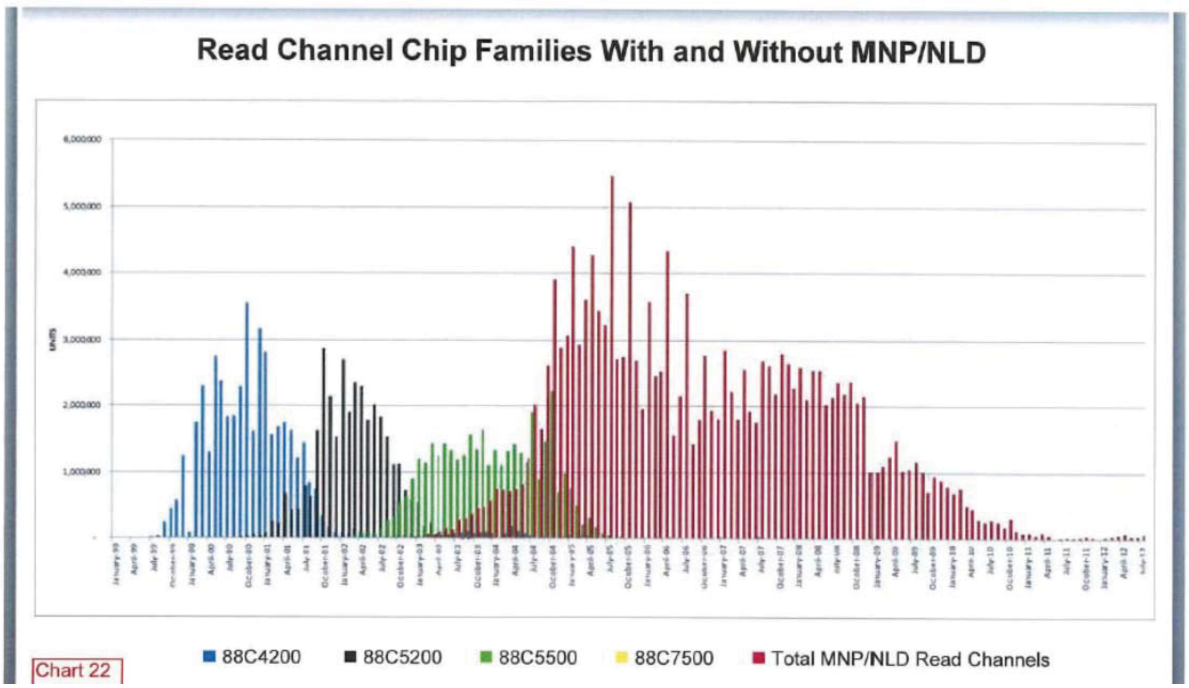


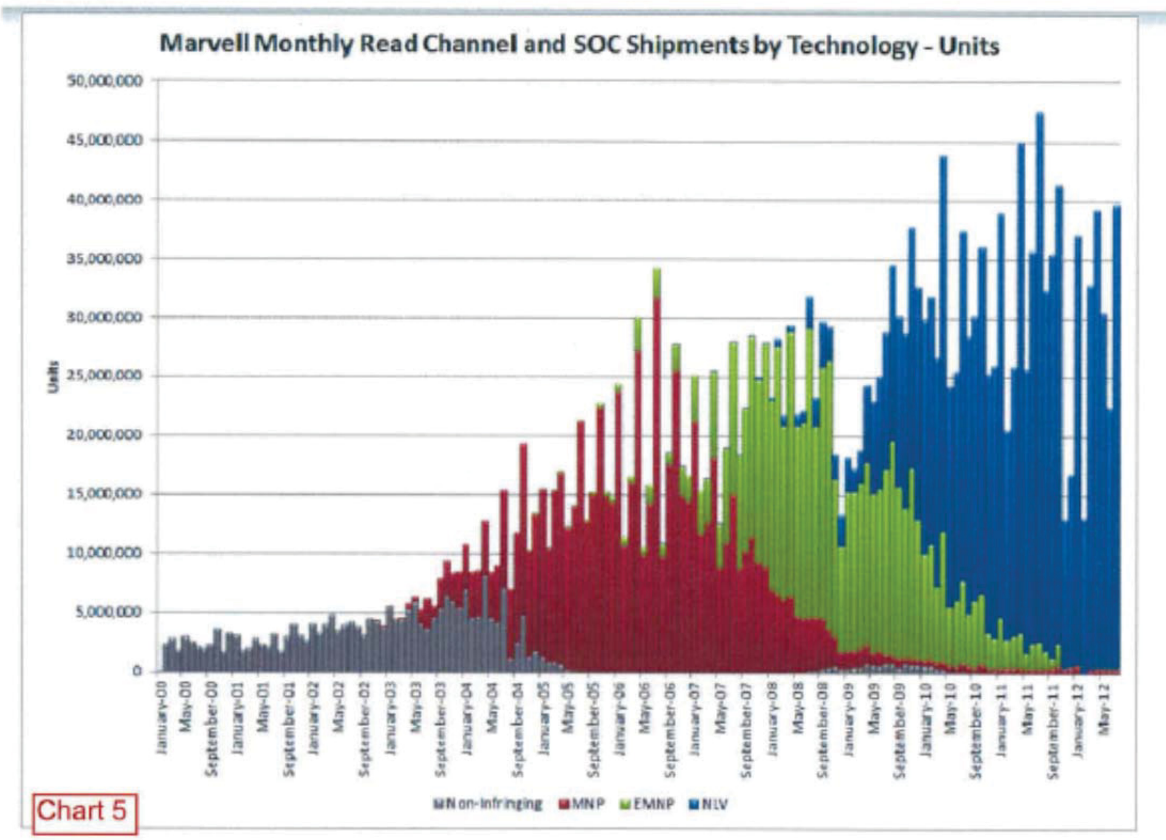
Chart 22

Ms. Lawton's Chart 5 (shown below) shows the monthly SOC sales. In this chart, the noninfringing products are indicated by the grey bars and the various generations of infringing MNP and NLD SOCs are indicated by the red, green and blue bars. Again, sales of the noninfringing SOCs began to decline almost immediately after the infringing SOCs were

---

44 Marvell's stand-alone read channel products include the detector. *See* Ex. 17 (Dkt. 901) at 15-16. Marvell's SOCs include the read channel (and hence the detector) as well as other components. *Id.* at 16-17.

introduced, fell to near zero soon thereafter, and the sales volume of the infringing SOCs dwarfed the sales volume of the noninfringing SOCs.



Marvell's customers' purchases showed their need and desire for CMU's technology. As Dr. Bajorek testified, "actions speak louder than words," and Marvell's customers "are not dummies. They wouldn't have bought the chips [with the MNP and NLD circuits] if they didn't plan to use them." Ex. 36 (Tr. 12/4/12) at 243; *see also id*. at 116 and Ex. 49 (P-Demo 8) at 44. Indeed, Marvell internal documents show that one of its HDD customers, Western Digital ultimately admitted the MNP delivered significant gains in SNR. *See* Ex. 36 (Tr. 12/4/12) at 135; Ex. 50 (P-506).

c. CMU's Technology Became an Industry Standard

Marvell's customers, the HDD manufacturers—with Marvell's full knowledge—adopted the technology of the CMU patents by purchasing large quantities of Marvell's infringing chips and using them, and Marvell was aware of and induced such use. *See* Ex. 36 (Tr. 12/4/12) at 111, 113-15.

d. The Nexus Between Marvell's Copying and its Commercial Success

The above facts demonstrate the nexus between Marvell's commercial success and its use of the method of claim 4. Marvell's chips that were designed to copy claim 4 were a "must have" for Marvell, so those chips became a commercial success for Marvell, and are now an industry standard. Marvell confirmed this nexus in writing. In a 2008 email announcing the promotion of Zining Wu, Marvell identified the MNP as one of two technologies that "helped firmly establish Marvell as a the market leader in the HDD IC business." *See* Ex. 56 (P-703) (excerpts below).

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

PLAINTIFF'S TRIAL EXHIBIT P-703

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.

Thus, seven years after it began copying claim 4, Marvell identified the copycat product as instrumental to its commercial success, thereby admitting the nexus between the commercial success and its use of the method of claim 4.

e. The Method of Claim 4 Continues to be "Must Have" for Marvell

Marvell's next generation detector after the MNP was its so-called "Non-linear Viterbi Detector" or "NLD." With the NLD, Marvell went all in on copying the CMU patents, noting that its new "enhancement" to the MNP "turns out to be the original structure that Kavcic proposed in his paper." Ex. 24 (P-Demo 7) at 73. Marvell continues to sell chips with its NLD design that "turns out to be the original structure that Kavcic proposed in his paper," more than 10 years after it first copied claim 4, and even though the NLD was found to infringe the CMU patents (including claim 4), and even though the court imposed an ongoing royalty of 50 cents for every chip that Marvell sells with the NLD through the life of the CMU patents. Ex. 4 (Dkt. 933). The fact that Marvell refuses to take the technology out of its products that infringe claim 4 despite the fact that it has to pay 50 cents for each such chip sold, shows that claim 4 has withstood the test of time as an important, "breakthrough," invention.

E. <u>Zeng is Not Enabling and Therefore Does Not Qualify as Prior Art</u>

"Prior art under § 102(b) must sufficiently describe a claimed invention to have placed the public in possession of that invention. … In particular, one must be able to make the claimed invention without undue experimentation." *In re Elnser*, 381 F.3d 1125 (Fed. Cir. 2004); *see also Elan Pharma., Inc. v. Mayo Foundation*, 346 F.3d 1051 (Fed. Cir. 2003) ("The disclosure in an assertedly anticipating reference must be adequate to enable possession of the desired subject matter. It is insufficient to name or describe the desired subject matter, if it cannot be produced without undue experimentation.").

Here, Zeng's Thesis is not enabling because the detectors described in Sections 4.4 and 5.2 of Zeng's Thesis could not be made or used by a person skilled in the art to which the '839 patent pertains regardless of the amount of experimentation (i.e., even undue experimentation would be insufficient) because Zeng derives his detector equations from a channel model that violates the laws of physics because it permits consecutive positive or consecutive negative transitions. The parameters of the BMFs of Zeng's detectors in Section 4.4 and 5.2 cannot be set since they depend on an eigenvalue decomposition of a physics-defying channel model. *See* Kavcic Dec. at ¶ 1, 25. Accordingly, such a person skilled in the art would be unable to set the parameters for the detector equations to detect data written to the disk in a real disk drive. In fact, it would be impossible to determine function parameter settings for the detector based on the disclosure in Zeng's Thesis. *See* Kavcic Dec. at ¶¶ 1 and 24-25. Apart from the problems associated with Zeng's physics-defying channel model, Zeng's equations in Section 5.2 have numerous -- and different kinds of -- mathematical errors, so much so that the Zeng's BMFs in Section 5.2 are unusable and non-operative in a detector. *See* Kavcic Dec. at ¶ 2.

1. <u>Zeng's Detectors are Derived from a Channel Model that Violates Laws of Physics</u>

As explained above, in magnetic recording the regions of the disk can be polarized in only one of two directions to record binary data, such as left or right for longitudinal magnetic recording (LMR), and those polarized regions can be conceptualized as bar magnets. *See* Bajorek Dec. at ¶¶ 39-40. Adjacent polarized regions can have the same polarization or opposing polarization, and when adjacent polarized regions are magnetized in opposing directions, there is a "transition" in the polarity of the regions that can be detected by the read head. *See* Bajorek Dec. at ¶¶ 40-41. As a matter of basic physics, there are only two types of transitions. In LMR, for example, one type of transition is when the N ends of the bar magnets are abutting, and in the other type of transition, the S ends of the magnets are abutting. *See* Kavcic Dec. at ¶ 15. Since bar magnets have opposite poles (N and S), a second positive transition cannot immediately follow a first positive transition; nor can a second negative transition immediately follow a first negative transition. Instead, a negative transition must follow a positive transition (with any number of non-transitions therebetween) and a positive transition must follow a negative transition (with any number of non-transitions therebetween). *See* Kavcic Dec. at ¶ 16.[45]

As mentioned above, Zeng's general channel model in Section 4.1 is at equation 4.1 (p. 51), which is:

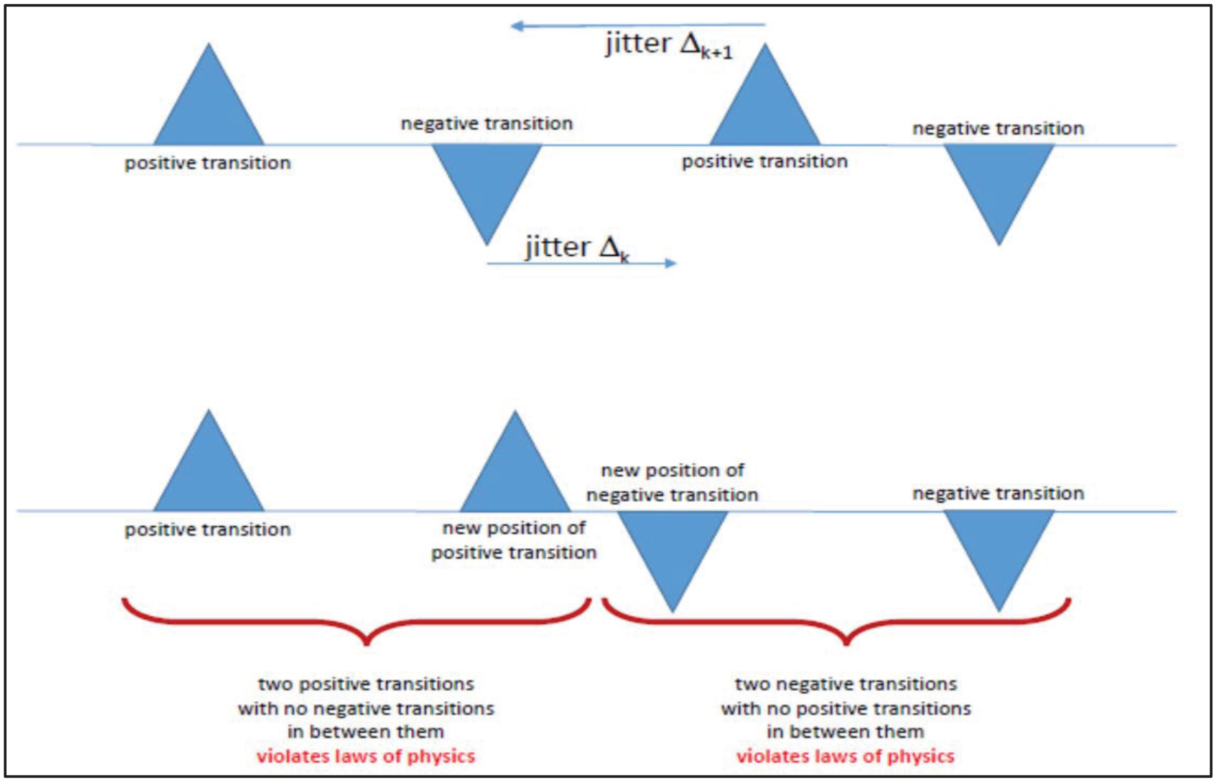
$$Z(t) = \sum_{k=1}^{N} a_k \, h(t - kT - \Delta_k T) + w(t)$$

[45] Lee concedes this point. Lee Dec. at ¶ 25.

Δ_k "represents the random jitter in the position of the transition response." Zeng Thesis at 51. Zeng models the "random jitter" from transitions as "independently and identically distributed random variables with zero mean" and variance σ_Δ^2. *See* Zeng Thesis at 52, 65; *see also* Kavcic Dec. at ¶ 17.

Zeng's "random jitter" is random shift in the position of the transition response. *See* Zeng Thesis at 51. This means, in LMR for example, that the position of the physical transition on the recording media will shift to the left or to the right from its ideal (jitterless) position by a random amount. In magnetic recording, however, consistent with the laws of physics, transitions cannot and do not shift by arbitrary random amounts. For example, jitter cannot take a transition (say a negative transition) farther than the next transition (a positive transition). If that were physically possible, then a positive and a negative transition would change places without cancelling each other, and there would be two consecutive positive transitions or two consecutive negative transitions. This cannot happen under the laws of physics because, as explained above, in magnetic recording there cannot be two consecutive positive transitions or two consecutive negative transitions. *See* Kavcic Dec. at ¶ 18.

Nevertheless, Zeng's "random jitter" channel model allows transitions to switch places such that there can be consecutive positive or consecutive negative transitions, which violates the laws of physics. The diagram below illustrates the physical impossibility.



See Kavcic Dec. at ¶ 19.

The upper part of the diagram shows that the transition at position *k* (a negative transition) could shift to the right randomly by random jitter Δ_k and that the transition at position *k*+1 (a positive transition) could shift to the left randomly by random jitter Δ_{k+1}. In Zeng's model, Δ_k and Δ_{k+1} are independent random variables, *see* Zeng Thesis at 52, so Δ_k could be a positive value (causing a shift to the right) and Δ_{k+1} could be a negative value (causing a shift to the left). If the random shifts were large enough, the positive transition at position *k*+1 could move before the negative transition at position *k*, so that there would be consecutive positive transitions and consecutive negative transitions, as shown in the lower part of the diagram. *See* Kavcic Dec. at ¶ 19.

The laws of physics dictate that two opposite transitions cancel each other (or otherwise destructively interfere) if one transition were ever moved to appear before (or close to) the previous, opposite transition. But Zeng's channel model does not permit transition cancellations at all. *See* Kavcic Dec. at ¶ 20. Further, because of the granular structure of the medium, even if the random jitter shifts two adjacent transitions as close as one grain from each other, the two transitions still have to cancel each other. That is, if the average grain diameter is denoted by τ, then two adjacent transitions cannot appear closer than τ to each other, or else they would cancel each other. However, Zeng's model in Section 4.1 never allows two adjacent transitions to cancel, even if (i) they appear closer than distance τ to each other and thereby violate the laws of physics, or (ii) they exchange order and also thereby violate the laws of physics. *See* Kavcic Dec. at ¶ 21. To the contrary, in Zeng's model the transitions can switch places and "a model that violates the laws of physics for any non-zero percentage of times is inherently flawed." Kavcic Dec. at ¶¶ 22-23.

Zeng's physically impossible channel model prevents a PSITA from making a detector based on Sections 4.4 or 5.2 of Zeng's Thesis that could work in a disk drive. Since the channel models in these sections violate the laws of physics, a PSITA would encounter fatal problems that would prevent the PSITA from making a useable detector, regardless of the amount of experimentation. *See* Kavcic Dec. at ¶ 24.

In attempting to make the detectors in Sections 4.4 and 5.2 of Zeng's Thesis to read data in a predictable and repeatable manner, a PSITA must set the parameters of Zeng's detector the BMFs to run in the disk drive. A disk drive, however, would never produce signals that violate the laws of physics. In a disk drive, opposing peaks that crossed would cancel each other, not switch places as in Zeng's model. Yet, short of models that violate the laws of physics, the parameters of Zeng's detectors of Section 4.4 and 5.2 simply cannot be set, since they rely on both a physically impossible random jitter model ***and*** depend on the eigenvalue decomposition of a physically non-existent model. *See* Kavcic Dec. at ¶ 25; Zeng Thesis at 65-68 (describing Zeng's eigenvalue decomposition). Detectors built according to either Section 4.4 or 5.2 would be inoperable in a disk drive.

The futility of the PSITA's efforts would be borne out if the PSITA tried to replicate the simulation results that Zeng disclosed in his thesis. For example, the results in Figure 4.6 of Zeng's Thesis (p. 69, shown below) are physically impossible and demonstrate that Zeng's detector cannot be used in a disk drive.

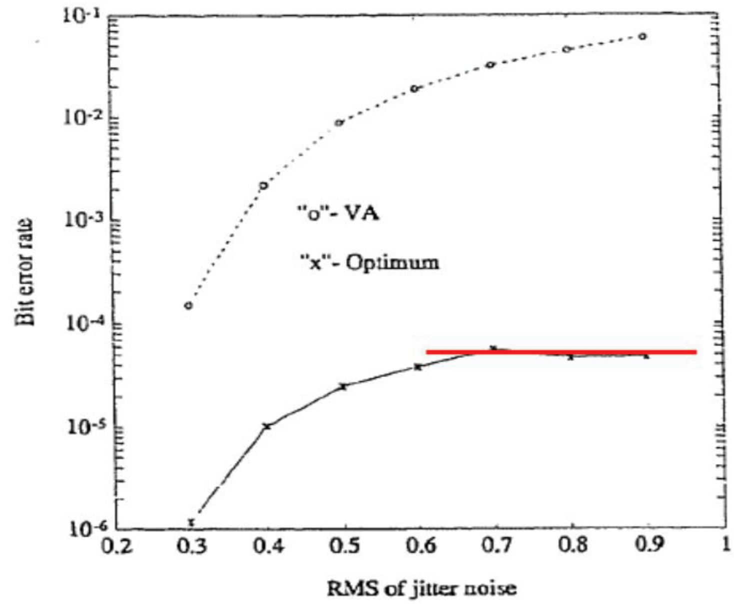


Figure 4.6: Error rate vs. σ_Δ for $\sigma_w = 0.001$.

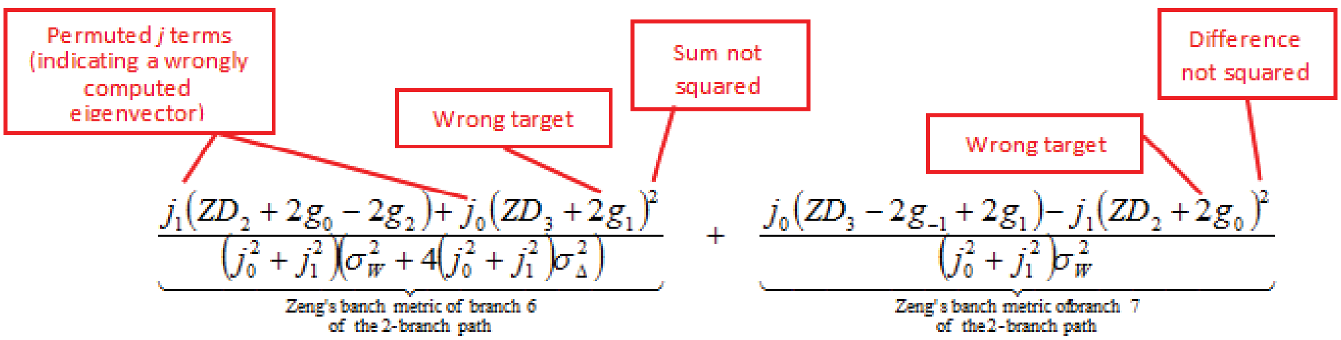
See Kavcic Dec. at ¶ 27; *see also* McLaughlin Dec. at ¶¶ 89-91 (Zeng's reported simulation results defy laws of physics).

This figure has an obvious flaw, which is that it shows that the detector's performance does not deteriorate (and, in fact, seems to improve) as the RMS of jitter (i.e., the variance of jitter) increases (as shown by the red line). That is simply a physical impossibility. The performance of any detector must deteriorate as the noise level in the system increases. Yet according to Zeng, his detector does not lose performance as the noise level increases beyond 0.7. The outcome of the experimentation and attempts to reproduce Zeng's results just described would cause the PSITA to conclude that Zeng's detector does not work rather than teach the PSITA that is could be made to work. *See* Kavcic Dec. at ¶ 27. Accordingly, it is not surprising that Zeng's supposedly ground-breaking work does not appear to have ever been published in a peer-reviewed journal. *See* Kavcic Dec. at ¶ 28.

Thus, Zeng's detectors in Sections 4.4 and 5.2 are not enabling, and therefore cannot anticipate claim 4.

2. Section 5.2 of Zeng's Thesis Has Pervasive Mathematical Errors that Make the Detector of Section 5.2 Inoperative and Unusable

Beyond the fact that the detector equations in Zeng's Section 5.2 are based on a physics-defying channel model and eigenvalue decomposition, Zeng's equations in Section 5.2 have numerous -- and different kinds of -- errors, so much so that Zeng's BMFs in Section 5.2 are unusable and non-operative in a detector. *See* Kavcic Dec. at ¶ 29.[46] Zeng's Thesis has errors in equations (5.13) to (5.16) and (5.18) to (5.21) that include (i) missing squares, (ii) wrongly computed eigenvectors, and (iii) wrongly computed targets. *See* Kavcic Dec. at ¶¶ 30-36. The graphic below illustrates Zeng's incorrect path metric for a path that follows branches 6 and 7 of Zeng's trellis (*see* Zeng Thesis at 77). *See* Kavcic Dec. at ¶ 36.



Instead of detecting the data accurately, Zeng's detector in Section 5.2 actually ***introduces systematic errors*** that make the detector useless for its intended purpose. *See* Kavcic Dec. at ¶ 37. For example, in Zeng's erroneous equations, if j_0 is negative, the numerator will be a lesser value than it should be because of the mathematical errors, thereby making the resulting branch metric value a lesser value than it should be (including having the theoretical possibility of taking a negatively infinite branch metric value if j_0 is negative and ZD_{k+1} is very large), which will tend to make any path with that branch in it appear vastly more likely than it otherwise should. Consequently, if a detector uses Zeng's erroneously written equations, the detector will tend to systematically favor paths (of branches) with transitions because the improper numerator is used in Zeng's BMFs for these branches. By systematically favoring

[46] Dr. Lee failed to describe any of these errors and instead copied Zeng's erroneous equations verbatim in his declaration.

paths (of branches) with transitions, Zeng's approach contains a built-in bias that runs directly contrary to the fundamental purpose of Viterbi detectors, i.e., to determine the most likely path through the trellis with both predictability and repeatability. *See* Kavcic Dec. at ¶ 37.

Furthermore, even an opposite effect may take place depending on the exact numerical values of j_0 and j_1. Namely, in simulations (such as those whose results are displayed in Zeng's Thesis), the detector may actually appear to be vastly better than it normally would be if employed in a real drive (because in simulations the detector would favor branches that have transitions, where Zeng's hypothetical "random jitter" is the dominant noise). Thus, in simulations, the detector would appear to be good when it really is not (i.e., false positive results). In fact, in a real drive, depending on the values of j_0 and j_1, Zeng's detector may make catastrophic errors, which can remain undetectable and uncorrectable through simulations for a long time. *See* Kavcic Dec. at ¶ 38.

Finally, the simulation results in Zeng's Section 5 appear on their face to be "too good," far better than detection performance results ever observed by even the best (provably optimal[47]) Kavcic-Moura detectors on real drives. This likely is a consequence of using the wrong channel model and/or incorrectly computed branch metric values. *See* Kavcic Dec. at ¶ 39.

All of the aforementioned errors in Zeng's Thesis are significant for a PSITA. Noticing and correcting the errors requires mastery of eigen-decompositions and their applications to a mesh of "synchronous" and "asynchronous" Viterbi algorithms,[48] and this goes beyond the skill level of a PSITA. Faced with Zeng's detector in Section 5.2, a PSITA would (sooner or later) notice that the detector operates wrongly and unpredictably, but likely would never figure out how to detect and correct the errors because of the high level of knowledge and skill needed to detect the exact locations of errors and equally high level of knowledge and skill needed to correct them. *See* Kavcic Dec. at ¶ 40. Indicative of the complexity of detecting and correcting the errors is the fact that neither Dr. Lee nor the Requester addressed or disclosed these errors.

---

[47] *See* footnote 16 above.

[48] Zeng uses an asynchronous trellis in Sections 4.4 and 5.2, as described below in Section VIII.B.2.

A PSITA could not make or use the detector described in Section 5.2 of Zeng's Thesis to predictably and repeatably detect data written to a disk, with or without undue experimentation, because of the errors in the equations. The BMFs disclosed by Zeng in Section 5.2 will yield unusable, undetectable, unpredictable, erratic, and systematic mistakes in the Viterbi detection process that makes detecting the data on the disk impossible. *See* Kavcic Dec. at ¶ 41. Therefore, Zeng's Section 5.2 is not enabling and cannot anticipate claim 4.

F. <u>There is Insufficient Evidence that Zeng's Thesis is a Printed Prior Art Publication</u>

"[U]ncorroborated third party oral testimony … is entitled to little, if any weight." *Ex Parte Haydon*, 2013 WL 5397786 at *5 (PTAB January 25, 2013, Appeal 2010-011645). Here, Ms. Heather Milliken's declaration as to the alleged publication of the Zeng Thesis (Ex. G of the Request) should be given "little, if any weight" since her testimony is uncorroborated and without sufficient foundation. According to Ms. Millken's declaration, she reviewed the "ProQuests's records regarding" Zeng's Thesis (Ex. G of Request at ¶ 6), but she did not include the records as part of her declaration. Therefore, her statements about what the records show (*see* Ex. G of Request at ¶ 6) are uncorroborated.

There are important policy reasons why uncorroborated testimony is given "little, if any weight" in patent cases. "The law has long looked with disfavor upon invalidating patents on the basis of mere testimonial evidence absent other evidence that corroborates that testimony." *Finnigan Corp. v. Int'l Trade Com'n*, 180 F.3d 1354, 1366 (Fed. Cir. 1999). This practice stems from a ruling by the United States Supreme Court that uncorroborated testimony alone is "unsatisfactory" to invalidate a patent. *The Barbed-Wire Patent*, 143 U.S. 275, 284 (1892) ("Witnesses whose memories are prodded by the eagerness of interested parties to elicit testimony favorable to themselves are not usually to be depended upon for accurate information."). The need for corroboration exists even when the testifying party is uninterested but testifying on behalf of an interested party. *See Finnigan Corp.*, 180 F.3d at 1367. Thus, the necessity of corroboration is defined not with reference to the level of interest of the witness, but rather by the inherent inability of testimonial evidence to meet the standard necessary to invalidate a patent. *Id.* at 1368.

Here, because Ms. Milliken's uncorroborated declaration testimony can be given "little, if any weight," there is insufficient evidence that Zeng's Thesis is a "printed publication" under 35 U.S.C. § 102(b) (pre-AIA). Therefore, Zeng's Thesis cannot qualify as prior art under § 102(b) (pre-AIA). This is yet another, independent reason that the patentability of claim 4 should be confirmed.

VIII. CLAIM 4 WOULD NOT HAVE BEEN OBVIOUS IN VIEW OF THE CITED REFERENCES

The Office Action includes two separate obviousness rejections of claim 4: (i) the combination of Zeng's Thesis and Lee's Thesis (*see* Request at 31-41); and (ii) the combination of Zeng's Thesis and the Coker patent (*see* Request at 42-57). There are multiple reasons that claim 4 is not obvious in view of these combinations.

A. None of the References Teach or Suggest a Set of Signal-Dependent Branch Metric Functions that Are Applied to a Plurality of Time Variant Signal Samples

Sections 4.4 and 5.2 of Zeng's Thesis do not disclose or suggest any signal-dependent branch metric functions, let alone a "set" of such functions, as described above. Lee's Thesis and the Coker patent likewise do not disclose any signal-dependent branch metric functions that are applicable to a plurality of signal samples, as required by claim 4. *See* McLaughlin Dec. at ¶¶ 86-87. In fact, neither the Requester nor Dr. Lee assert that Lee's Thesis or the Coker patent disclose any such signal-dependent branch metric functions. Instead, the Requester and Dr. Lee rely on Lee's Thesis and the Coker patent only in the event that CMU argues that "a portion of a time-dependent branch metric function must also in effect be [regenerated] (by recomputing a parameter) at each time index to constitute part of the claimed 'selecting.'" Request at p. 31 (brackets and parenthetical in original); *see also* Request at 34, 42 and 45; Lee Dec. at ¶¶ 57, 64. But the Requester's proposed obviousness rejections fail for the simple reason that none of the references teach or suggest a "set of signal-dependent branch functions" that are applied a "plurality of [time-variant] signal samples," as recited in claim 4. *See, e.g.*, *SynQor, Inc. v. Artesyn Techs.*, 709 F3.d 1365, 1375 (Fed. Cir. 2013) (claim not obvious where neither prior art reference "alone nor any combination of the asserted prior art teaches or suggests" an element of the claim).

B. A Person Having Ordinary Skill in the Art Would Not Have Been Motivated to Modify Zeng's Detectors in Sections 4.4 and 5.2 Based on Either Lee's Thesis or the Coker Patent

When an obviousness rejection is based on a combination of references, as is the case here, there must be some teaching, suggestion or motivation in the prior art that would have led a person having ordinary skill in the art (PHOSITA)[49] to combine the reference teaching to arrive at the claimed invention. *See* MPEP ¶ 2143; *see also Cheese Systems, Inc. v. Tetra Pak Cheese and Powder Systems, Inc.*, 725 F.3d 1341, 1352 (Fed. Cir. 2013) ("fact-finder must determine what the prior art teaches, whether prior art teaches away from the claimed invention, and whether there was motivation to combine teachings from separate references").

Here, even if the Zeng Thesis as modified based on Lee's Thesis or the Coker patent taught or suggested all of the elements of claim 4, there still would be no motivation to combine the references as suggested in the Request for at least the following reasons.

1. Zeng's Thesis Should be Accorded Little Weight Because of its Technical Flaws

Zeng's Thesis suffers from a physics-defying channel model, reports simulation results that defy the laws of physics, and is riddled with math errors. *See* Kavcic Dec. at ¶¶ 15-41. Because of these flaws, a PHOSITA would not be motivated to modify it. Consequently, Zeng's Thesis should be accorded little weight in the obviousness analysis. *See Dennison Mfg. Co. v. Ben Clements & Sons, Inc.*, 467 F. Supp. 391, 415 (S.D.N.Y. 1979) ("[The reference] is vague and indefinite as to its exact teachings. These factors, in conjunction with the unworkability of the disclosed invention, might well have caused one skilled in the art who was considering the idea of connecting the paddles to avoid tangling to discard the notion and look

[49] At ¶ 84 of his declaration, Prof. McLaughlin used the same standard for a PHOSITA espoused by Dr. Lee, which is "someone with at least a Master's degree in electrical engineering specializing in signal processing and digital communications with at least two years [sic] experience in that field or a related industry." Lee Dec. at ¶ 17.

for other solutions."); *Azoplate Corp. v. Silverlith, Inc.*, 367 F. Supp. 711, 718 (D. Del. 1973), *aff'd*, 506 F.2d 1050 (3d Cir. 1974) ("This is not an instance where teachings of positive results can be readily combined to achieve the invention. Instead, one would have to assume that despite the negative results inherent in the references, a skilled lithographer would [modify and combine]. The Court is unconvinced that there would be good reason for him to do so."); *see also United States v. Adams,* 383 U.S. 39, 52 (1966) ("We do say, however, that known disadvantages in old devices which would naturally discourage the search for new inventions may be taken into account in determining obviousness.")

2. Modifying a Detector Like Zeng's That Uses a d=1 RLL Constraint Defies Logic in Light of the Demand for Increased Data Density

Zeng's detectors in Sections 4.4 and 5.2 employ the d=1 RLL constraint, which prevents consecutive transitions from being written to the disk. *See* Zeng Thesis at 4, 65, 75; McLaughlin Dec. at ¶ 93; Bajorek Dec. at ¶¶ 100, 106. Neither Lee's Thesis nor the Coker patent use a d=1 RLL constraint, and for good reason. Zeng correctly notes that using the d=1 RLL constraint comes at the "cost of losing data rate." Zeng Thesis at 4. This is an understatement. By using the d=1 RLL code constraint, at least 30.6% of the disk is wasted with coding, making roughly ***one-third of the disk unusable for recording data***. In an industry highly focused on ***increasing data density***, no PHOSITA would use a system, like Zeng's, that wastes 30.6% of the disk. *See* McLaughlin Dec. at ¶ 93.

In fact, Dr. Bajorek explains that "[t]his amount of loss is now and has always been ***unacceptable*** in the HDD industry" and that based on his "industry experience, ***no HDD company would survive*** with a 30% disadvantage in data density/data capacity per disk." Bajorek Dec. at ¶ 112 (emphasis added). According to Dr. Bajorek:

> The HDD industry is extremely competitive and operates with modest profit margins. It typically increases the data capacity of HDDs by 50% every nine months. The customers of HDD producers demand HDDs of equal data capacity, from multiple producers simultaneously, every nine months. The 30% disadvantage, all else equal, would be equivalent to being six months behind competitors in HDD offerings, requiring extreme

> price discounts to sell any of such HDDs. Alternatively, compensating for the 30% disadvantage would require a 50% increase in disk and head components to achieve a specific HDD capacity with the concomitant cost increase. ***No HDD supplier in the history of the HDD industry would survive*** with such disadvantages in product timing or cost on a sustained basis.

Bajorek Dec. at ¶ 112 (emphasis added). This is why a PHOSITA would not be motivated to modify the detectors in Zeng's Thesis in view of either Lee's Thesis or the Coker patent. *See* Bajorek Dec. at ¶ 112. A PHOSITA, familiar with the cited references, recognizing that neither Lee's Thesis nor the Coker patent use the wasteful d=1 RLL constraint because disk space is at a premium, would not be motivated modify detectors like Zeng's that require the d=1 RLL constraint. *See* McLaughlin Dec. at ¶ 93. Indeed, it is not surprising that Marvell never commercialized the detectors in Zeng's Thesis even though Zeng is an employee of Marvell (and has been since 2004). *See* Bajorek Dec. at ¶ 112.

In this sense, Lee's Thesis and the Coker patent ***teach away*** from the modifications of Zeng's Thesis proposed in the Request. Both Lee's Thesis and the Coker patent teach that the goal of the HDD is ***increased data density.*** *See* Lee Thesis at p. iv ("Digital storage systems try to achieve the ***maximum data density***…."); Coker patent at col. 2:15-17 ("It is an object of the present invention to provide a method and apparatus to achieve ***higher linear storage density*** in direct access storage devices…."). In view of these teachings, a PHOSITA would be taught away from modifying Zeng's detectors in Sections 4.4 and 5.2 because these detectors waste at least 30.6% of the disk through the d=1 RLL coding. *See* McLaughlin Dec. at ¶ 94.

3. <u>A PHOSITA Would Not Be Motivated to Modify Zeng's Asynchronous Trellis Detector Based on the Synchronous Trellis Detectors in Lee's Thesis and the Coker Patent</u>

Zeng's detectors in Sections 4.4 and 5.2 use an asynchronous trellis. That means, that at one time instance of the trellis, Zeng uses different signal samples to compute the branch metric values for the various branches at that time instance, ***and*** uses the same signal samples to compute the branch metric values for the different branches at different time instances. *See* McLaughlin Dec. at ¶ 95. This is apparent from Zeng's BMFs. For example, in Section 4.4, Zeng uses signal samples z_k and z_{k-1} for branches 2 and 5, and samples z_k and z_{k+1} for branches 3 and 4. *See* Zeng Thesis at 68. So at time index 2 of the trellis (i.e., $k = 2$), Zeng uses z_2 and z_1 for branches 2 and 5, and samples z_2 and z_3 for branches 3 and 4. Thus, at one time index of the trellis, Zeng uses different signal samples to compute the branch metric values for branches 2 and 5 than he uses for branches 3 and 4. Further, at time index 3 of the trellis (i.e., $k = 3$), Zeng uses z_3 and z_2 for branches 2 and 5, and samples z_3 and z_4 for branches 3 and 4. Thus, Zeng uses the signal samples z_2 and z_3 to compute the branch metric values at two different time instances of the trellis (e.g., branches 3 and 4 at time index 2 and branches 2 and 5 at time index 3). *See* McLaughlin Dec. at ¶¶ 95-96. Zeng also uses an asynchronous trellis in Section 5.2. *See* McLaughlin Dec. at ¶ 95.

In contrast, neither Lee's Thesis nor the Coker patent uses an asynchronous trellis; they both use a synchronous trellis (as do the CMU patents, *see e.g.*, Eq. 6). *See* McLaughlin Dec. at ¶ 97.

In light of the fact that Zeng's signal sample inputs do not vary with time strictly (i.e., Zeng uses an asynchronous trellis so that the same sets of signal samples are used to compute branch metric value at different time indexes of the trellis), the Requester's and Dr. Lee's suggestion to modify Zeng to use time-varying parameters instead, such as a time-varying σ_{w}^2 (*see* Request at 34-42, Lee Dec. at ¶¶ 57-61) is nonsensical. Because Zeng's function inputs do not necessarily vary with time, a PHOSITA would not be motivated to modify Zeng as suggested in the Request and Dr. Lee to use time-varying parameters. *See* McLaughlin Dec. at ¶ 98.

In a related manner, it is not obvious for a PHOSITA to modify branch metric function components for an asynchronous trellis, like Zeng's, with function components from a synchronous trellis, like Lee's Thesis and the Coker patent. *See* McLaughlin Dec. at ¶ 99. For example, determining the target values for an asynchronous trellis based on a synchronous trellis is very complicated. *See* McLaughlin Dec. at ¶ 99. Indeed, Zeng incorrectly identified the targets for his BMFs in Section 5.2 precisely because he uses a synchronous trellis model to determine his targets and then erroneously applies them to the signal samples in his asynchronous trellis. *See* Kavcic Dec. at ¶¶ 34-35. Neither Dr. Lee nor the Requester disclosed these errors.

4. A PHOSITA Would Realize That Zeng's Reported Simulation Results are Physically Impossible and Would Not Be Motivated to Modify Teachings of a Physically Impossible Device

Zeng's reported simulation results in Section 4.4.1 of his thesis are physically impossible because they show the detector performance does not deteriorate as noise increases. In fact, some of the results show that the detector performance improves as noise increases. *See* McLaughlin Dec. at ¶¶ 89-90. For example, Figure 4.6 (reproduced below) on p. 69 of Zeng's Thesis is a plot of the bit error rate (BER) on the y-axis versus the root mean square (RMS) of the jitter noise on the x-axis over a range of RMS jitter noise from 0.3 to 0.9. This plot shows that the detector performance improves (that is, BER decreases) for jitter RMS values greater than 0.7 as indicated in the area of the red circle in the diagram below. That means Zeng's detector purportedly makes fewer errors when there is more noise. That is not physically possible and a PHOSITA would not be motivated to modify a device that yields physically impossible results. *See* McLaughlin Dec. at ¶ 90.

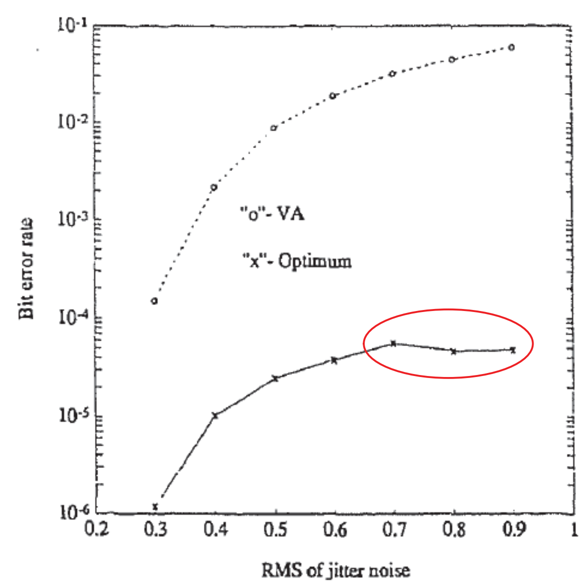


Figure 4.6: Error rate vs. σ_Δ for $\sigma_w = 0.001$.

Other plots in Section 4.4.1 show similar physically impossible results. For example, Figure 4.8 also shows that detector performance does not deteriorate as the noise increases. *See* McLaughlin Dec. at ¶ 91.

Despite allegedly achieving impossibly good results, Zeng inexplicably concludes that his detector in Section 4.4 "does not improve significantly upon the conventional Viterbi algorithm." Zeng Thesis at 71. Notwithstanding these alleged results, Zeng then proposes as an improvement "study[ing] the more realistic model in [the] next chapter." *Id*. Chapter 5 is the "next chapter" and it is filled with mathematical errors. *See* McLaughlin Dec. at ¶ 92; Kavcic Dec. at ¶¶ 29-41. A PHOSITA would be troubled by Zeng's denigration of a detector that allegedly shows impossibly good performance and proposal of an improved detector (in Section 5.2) that is based on numerous math errors. For these reasons, a PHOSITA would not be motivated to modify Zeng's detectors in either Sections 4.4 or 5.2 or even to rely upon his work in any way. *See* McLaughlin Dec. at ¶¶ 89-92. Instead, the "outcome of the experimentation and attempts to reproduce Zeng's results" would cause a person skilled in the art "to conclude that Zeng's detector ***does not work rather*** than teach" the person skilled in the art "that it could be made to work." Kavcic Dec. at ¶ 27 (emphasis added).

5. A PHOSITA Would Not Modify Zeng's Parameters Based on Either Lee's Thesis or the Coker Patent Because Zeng's Model is Mathematically Incompatible with Both Lee's Thesis and the Coker Patent

Dr. Lee and the Requester both assert that because Lee's Thesis has a time-dependent parameter σ_k^2, it would have been obvious to make σ_Δ^2 (the variance of Zeng's random jitter) in Zeng depend on time. *See* Lee Dec. at ¶¶ 57-61; Request at 34-35. Similarly, Dr. Lee and the Requester both assert that because the Coker patent has a time-dependent parameter p_i, it would have been obvious to make σ_Δ^2 in Zeng depend on time. *See* Lee Dec. at ¶¶ 64-68; Request at 45-47. These assertions by Dr. Lee and the Requester are wrong because Zeng cannot be modified based on either Lee's Thesis or the Coker patent because Zeng's channel model is mathematically incompatible with both Lee's Thesis and the Coker patent. *See* McLaughlin Dec. at ¶ 100.

Assuming Zeng's jitter were signal-dependent (it is not), then Zeng's random jitter variables no longer would be statistically independent, and Zeng's assumption that the noise covariance matrices break into 1x1 and 2x2 block diagonal forms no longer holds. *See* Zeng Thesis at 65 ("We then decompose V into a series of matrices with dimensions of 1 or 2 only."). Similarly, the noise in the Coker patent is correlated and does not adhere to block-diagonal covariance matrices whose blocks are of sizes 1x1 and 2x2 only, like Zeng's. Hence, not even Zeng's d=1 RLL constraint would preserve the 1x1 and 2x2 block diagonal covariance matrix forms in Zeng's BMF derivations. If the 1x1 and 2x2 block-diagonal forms are not preserved, Zeng's BMFs cannot be derived because they require the eigenvalue decomposition that Zeng can do only if he has block-diagonal forms of sizes 1x1 and 2x2. Thus, a PHOSITA would not be motivated to modify Zeng's equations to make them time varying as in Lee's Thesis or the Coker patent because such a modification would destroy Zeng's eigenvalue decomposition on which Zeng's BMFs are based. In fact, a PHOSITA would be ***led away*** from such a modification of Zeng for this reason. *See* McLaughlin Dec. at ¶ 101.

6. It is Commercially Unacceptable and Impractical to Degauss the Disk Prior to Each Write as Zeng Does

One of the techniques Zeng uses to compensate for the fact that his BMFs in Sections 4.4 and 5.2 do not account for signal-dependent noise is to degauss the disk prior to each write. *See* Zeng Thesis at 9 and 20. While an AC-erase will substantially eliminate the overwrite effect, it is ***utterly impractical*** for a commercial HDD. *See* Bajorek Dec. at ¶ 113. With such a system, each time some portion of the disk is to be updated with new data, (i) the existing data on the disk would have to be read into a memory, (ii) a significant portion of the entire disk would then have to be AC-erased, and then (iii) this portion of the disk would have to be rewritten with the data in the memory but updated with the new data. In addition to requiring another memory device to buffer the data, such a process would make write times unacceptably and impractically long, so Zeng's device (whether or not modified as suggested in the Request) would not be commercially practical or acceptable. This is another reason that no PHOSITA would modify Zeng's device in Sections 4.4 or 5.2 in view of either Lee's Thesis or the Coker patent. Zeng's devices so modified would still be unacceptable in the HDD market for several reasons, including (i) that the write times would be unacceptably and impractically long and (ii) the device would be more expensive because of the additional memory needed to buffer the data during the AC-erase. In fact, a PHOSITA would be led to ***not*** use Zeng's device or the suggested modifications thereto in the Request (and the Office Action) at all. *See* Bajorek Dec. at 113.

C. Overwhelming Secondary Considerations Show That Claim 4 Would Not Have Been Obvious

Secondary considerations of nonobviousness are relevant to the question of obviousness. *See e.g.*, *KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 415 (2007). They must always be considered when present because they can "serve as an important check against hindsight bias." *Bristol-Myers Squibb Co. v. Teva Pharm. USA, Inc.*, 752 F.3d 967, 977 (Fed. Cir. 2014). Here, the secondary considerations overwhelmingly demonstrate that claim 4 is nonobvious. The secondary considerations that demonstrate the nonobviousness of claim 4 include: copying; commercial success; praise and acclaim in the field; failure by others; and satisfaction of a long-felt need. *See, e.g.*, MPEP ¶ 2141 (listing secondary considerations).

1. Copying

Evidence that that a claimed invention was copied can be an indication that the invention is nonobvious. *See Iron Grip Barbell Co. v. USA Sports, Inc.*, 392 F.3d 1317, 1325 (Fed. Cir. 2004). Here, after presiding over the litigation, including the four-week trial in late 2012, Judge Fischer concluded unequivocally that Marvell copied the Asserted Claims (including claim 4 of the '839 patent):

> "Marvell… comes before the Court with unclean hands ***after having engaged in deliberate and sustained copying of the patented method*** throughout the entire laches period and up to the present…." Ex. 54 (Dkt. 920) at 71 (emphasis added).
>
> * * *
>
> The "Court holds that the credible evidence presented at trial sufficiently establishes that ***Marvell deliberately copied CMU's Patents***…." Ex. 4 (Dkt. 933) at 19-20 (emphasis added).

* * *

> The "Court believes that a penalty of enhanced damages should be assessed against Marvell given its… known willful infringement through its ***deliberate and extensive copying of the patented methods***…" Ex. 4 (Dkt. 933) at 42 (emphasis added).

The bases for Judge Fischer's determination are clear and Prof. McLaughlin summarizes the evidence at ¶ 107(a)-(j) of his declaration. The salient points include:

- The Marvell engineer, Mr. Gregory Burd, responsible for designing Marvell's first infringing product, the MNP (which was originally called the "KavcicPP" by Marvell), had knowledge of the CMU patents at the Marvell commenced its infringement in 2001- 2002. *See* Ex. 17 (Dkt. 901) at 67.
- Mr. Burd reviewed the IEEE papers by Profs. Kavcic and Moura, and used them as the "launching pad" for his research, even though he knew those papers were the same as the CMU patents. *See* Ex. 24 (P-Demo 7) at 107-108.
- Marvell named its first infringing product, the MNP, after Prof. Kavcic, originally calling it the "KavcicPP." *See* Ex. 24 (P-Demo 7) at 23 and 27.
- When Marvell designed its second generation infringing product, the NLD, Zining Wu, Marvell's current CTO, wrote an email to Mr. Burd, describing the new design as, "the original structure that Kavcic proposed in his paper." Ex. 30 (P-366); Ex. 24 (P-Demo 7) at 73. The reference to the Kavcic "paper" refers to either of the Kavcic-Moura IEEE papers (Ex. 25 and Ex. 26 hereto), which are essentially the same as the CMU patents (*see* McLaughlin Dec. at ¶ 43).

In fact, Prof. McLaughlin reviewed documents produced by Marvell in the CMU case relating to its development of the MNP, and he concluded that he "saw no evidence that Marvell's engineers developed its signal-dependent detectors independent of the work of Profs. Kavcic and Moura" and that "Marvell's development of its signal-dependent detectors (the MNP and NLD) flowed directly from its access to the Kavcic-Moura IEEE papers." *See* McLaughlin Dec. at ¶ 108.

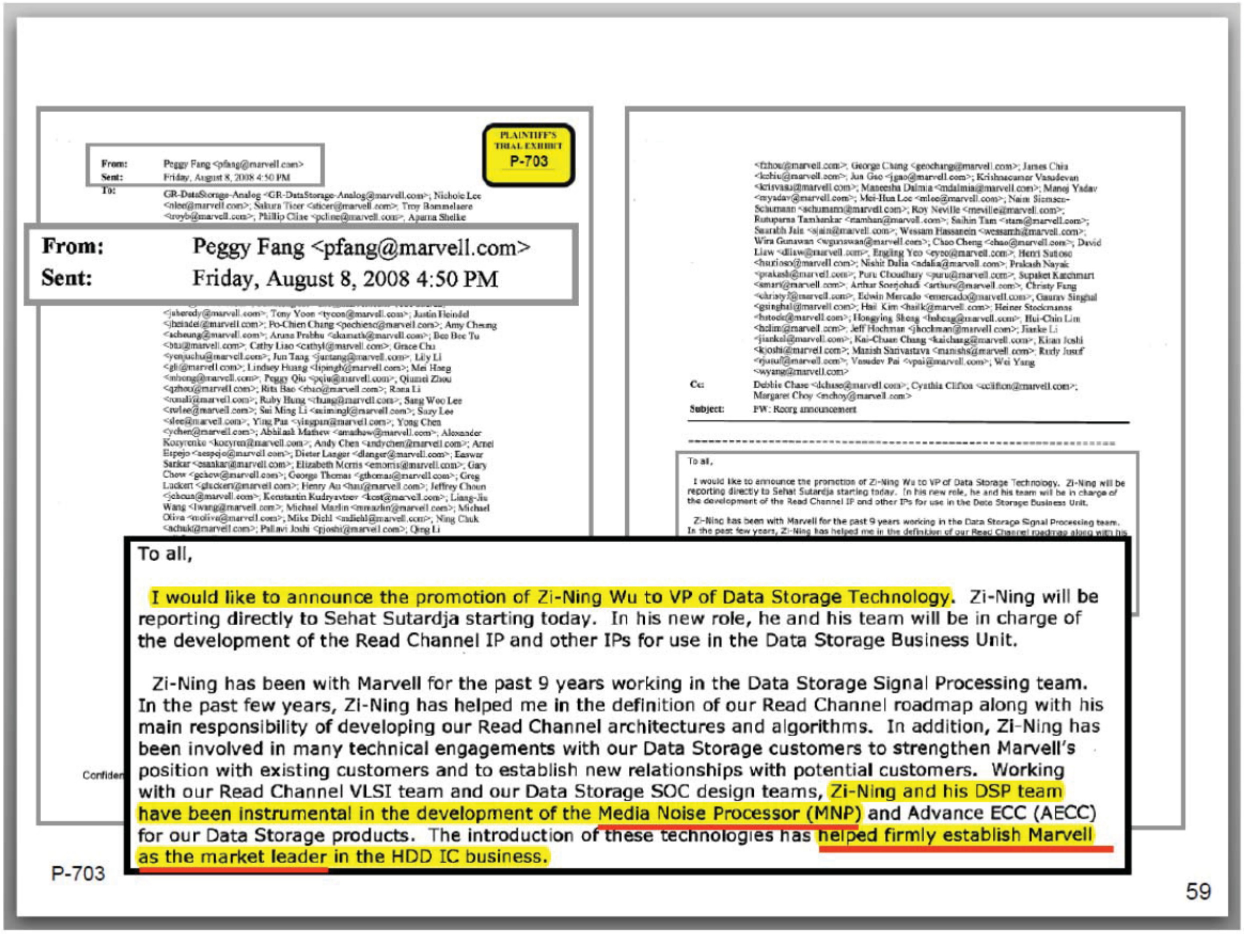
The fact that Marvell did, in fact, copy the claim 4, as determined by a federal judge, shows that claim 4 is nonobvious. *See Spectralytics, Inc. v. Cordis Corp.*, 649 F.3d 1336, 1344 (Fed. Cir. 2011).

2. Commercial Success

Commercial success is another secondary consideration that "might be utilized to give light to the circumstances surrounding" whether a claim is obvious or not. *See KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). Here, the CMU case demonstrated that Marvell experienced enormous commercial success from selling products (read channel chips) that practiced the methods of claim 4 of the '839 patent and claim 2 of the '180 patent. The jury found and the judge affirmed that two types of Marvell read channel chips practiced the methods of claim 4 of the '839 patent and claim 2 of the '180 patent: (i) the MNP-type read channel chips and (ii) the NLD-type read channel chips. Sales data through July 28, 2012 for these chips was available for the trial. From March 6, 2003[50] to July 28, 2012, Marvell had sold ***2.34 billion*** MNP- and NLD-type chips, with an average revenue of $4.42 per chip and an average operating profit of $2.16 per chip. *See* Ex. 17 (Dkt. 901) at 18. That means, in this time period Marvell earned approximately ***$10.34 billion*** in revenue and ***$5.05 billion in operating profit***. *See* McLaughlin Dec. at ¶ 103.

There is a direct nexus between Marvell's use of the CMU invention and its commercial success. Marvell itself credited the MNP, which practices claim 4 of the '839 patent, as a technology that "helped firmly establish Marvell as the market leader in the HDD IC business." *See* McLaughlin Dec. at ¶ 104. Below is a trial demonstrative from the CMU case showing an internal Marvell email dated August 8, 2008 (Ex. 49 (P-Demo 8) at p. 59) announcing the promotion of a Marvell engineer, Zining Wu to be the Vice President of Data Storage Technology (and reporting directly to Marvell's CEO, Sehat Sutardja). In the email Marvell boasts that Dr. Wu was "instrumental in the development of the [infringing] Media Noise Processor (MNP)" and that the MNP, which practices claim 4 of the '839 patent and claim 2 of the '180 patent, "helped firmly establish Marvell as the market leader in the HDD IC business."

[50] March 6, 2003 is six year prior to the date that the complaint in the CMU case was filed.



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.

P-703

59

Additional evidence in the CMU case confirms the nexus. That evidence showed that the CMU patents "had become industry standard" in the HDD industry since it was "adopted by the majority of drive makers for two or more generations of drives" (*see* Ex. 17 (Dkt. 901) at 23-24) and was "must have" for Marvell in order to remain competitive with other read channel chipmakers. *See* Ex. 17 (Dkt. 901) at 109 (summarizing evidence that CMU's technology was "must have" for Marvell); *see also* Ex. 36 (Tr. 12/4/12) at 109-140 (Dr. Bajorek's testimony in the CMU case pertaining the "must have" nature of the invention and that is became an "industry standard"); *see also* Ex. 49 (P-Demo 8) at 45-61 (Dr. Bajorek's trial demonstratives); McLaughlin Dec. at ¶ 105.

The commercial success of Marvell from using the CMU patents is further demonstrated by the fact that Marvell's MNP- and NLD-type chips displaced earlier Marvell products that are not capable of performing the method of the Asserted Claims. *See* McLaughlin Dec. at ¶ 106. When Marvell was describing its first generation of infringing chips to its customers it told them in writing that the "only" or "key" difference between the infringing chips and the then latest Marvell chips on the market was the addition of the infringing technology. *See* Ex. 55 (P-Demo 10) at 161 ("the only difference between 5575 and 5575P is the MNP addition to the post processor…." "The key difference is in the detector only."). Marvell's sales of non-accused products (i.e., products without the functionality of claim 4) declined rapidly after the first MNP-type products were introduced in September 2002 and practically ceased by 2005, as shown in chart below used in the trial in the CMU case. *See* Ex. 45 (P-Demo 13) at 1 (Chart 5). In this chart, the noninfringing predecessor products are shown in gray, and the various generations of Marvell products that practice claim 4 of the '839 patent and claim 2 of the '180 patent are shown in red, green, and blue. This chart shows that not only do the sales of the infringing product dwarf in volume the sales of the noninfringing products, the sales of the noninfringing products ceased shortly after Marvell introduced the infringing products to the market. *See* McLaughlin Dec. at ¶ 106

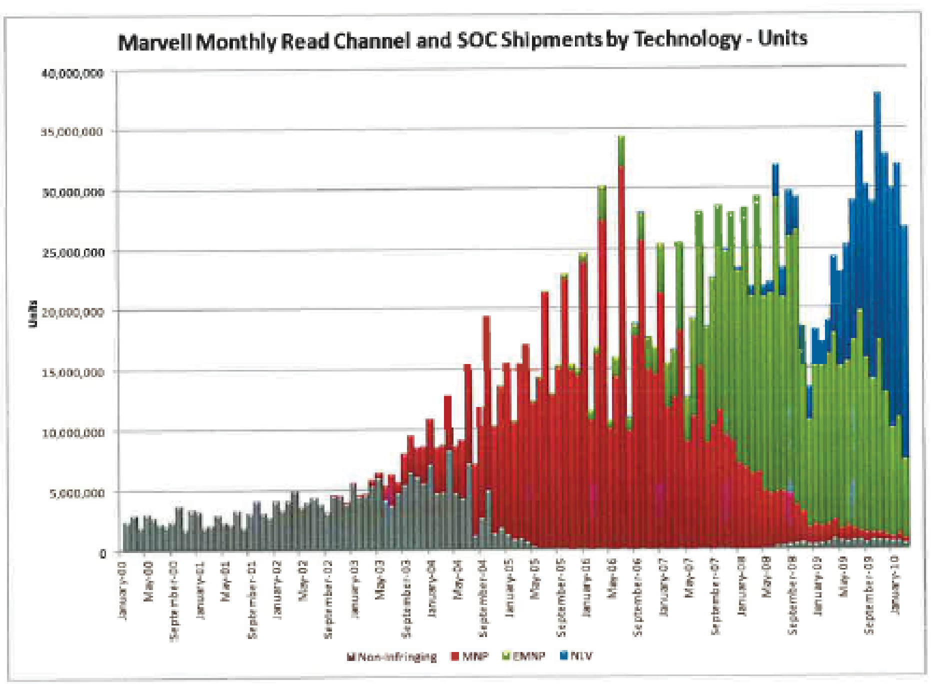
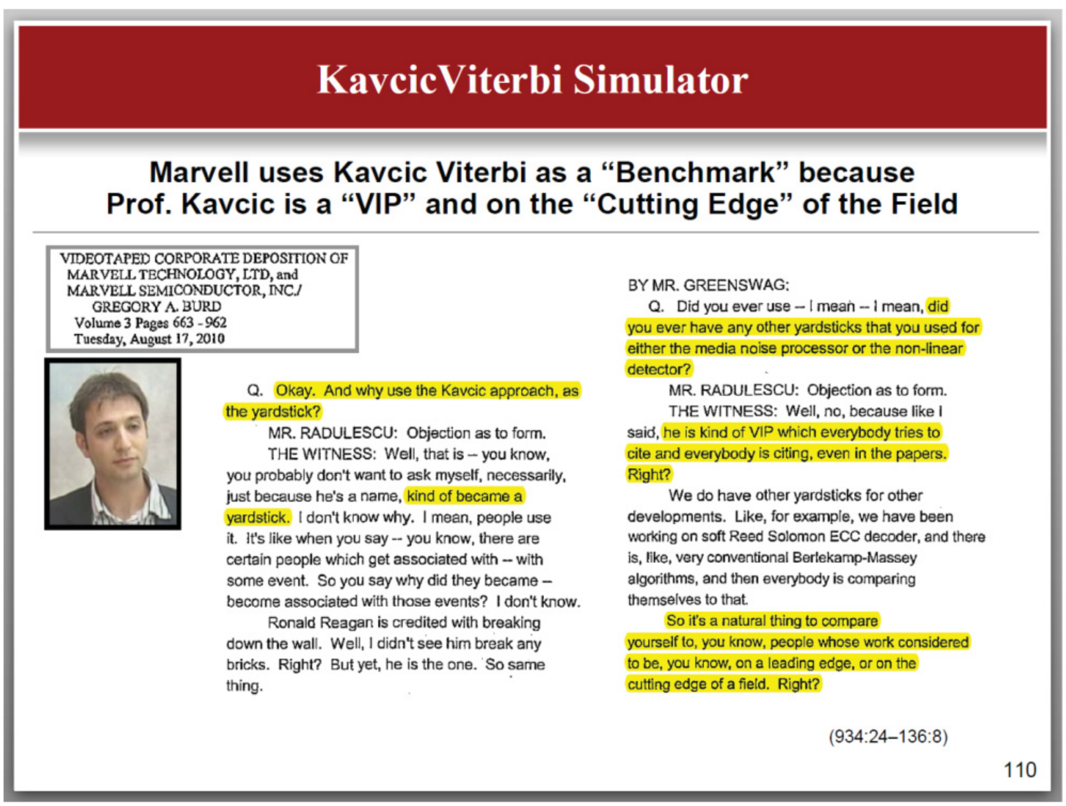


CHART 5

Thus, Marvell enjoyed immense commercial success from practicing (and copying) claim 4 of the '839 patent, and that commercial success is directly attributable to it use (and copying) of claim 4, which demonstrates conclusively the nonobviousness of claim 4.

3. Praise and Acclaim in the Industry

Praise for the invention in the relevant industry is another secondary consideration of nonobviousness. *See Brown & Williamson Tobacco Corp. v. Philip Morris Inc.*, 229 F.3d 1120, 1129 (Fed.Cir.2000). Here, despite the fact that Zeng works at Marvell (*see* Ex. 11, Dr. Zeng's Linked-In page), Marvell's engineers heap praise on the CMU patents, not Zeng's Thesis. The evidence in the CMU case was that Marvell's engineers "continuously run Kavcic algorithm to benchmark any subsequent algorithm that we develop at Marvell," use the CMU invention as a "***yardstick***" because Prof. Kavcic is "kind of ***VIP*** which everybody tries to cites and everybody is citing" and "it's a natural think to compare yourself to, you know, people whose work considered to be, you know, or a ***leading edge or on the cutting edge of a field***." *See* Ex. 24 (P-Demo 7) at 109 and 110 (slide 110 is reproduced below); *see also* McLaughlin Dec. at ¶ 110.

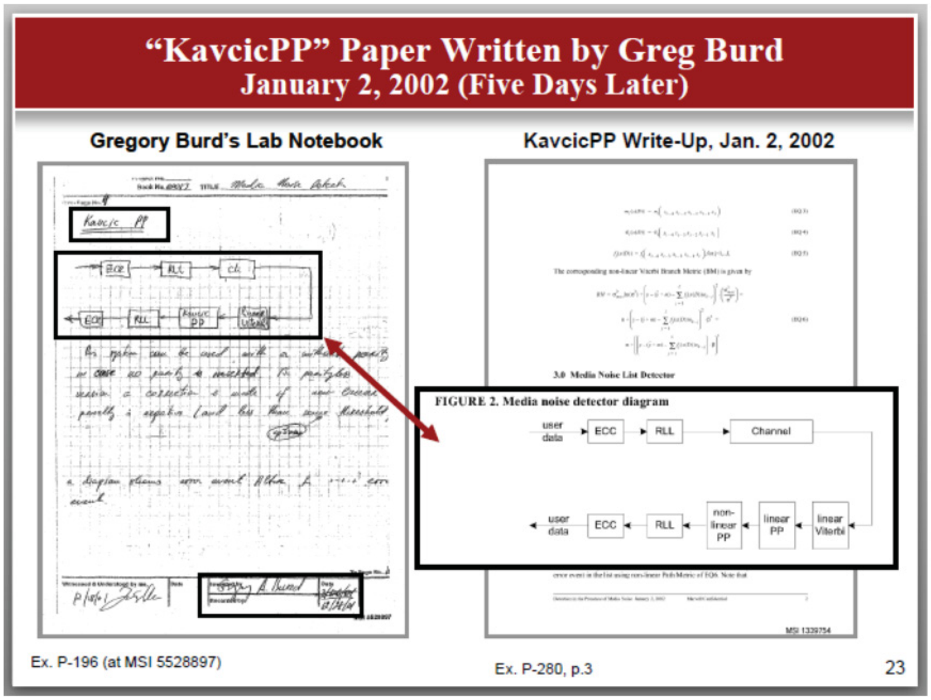


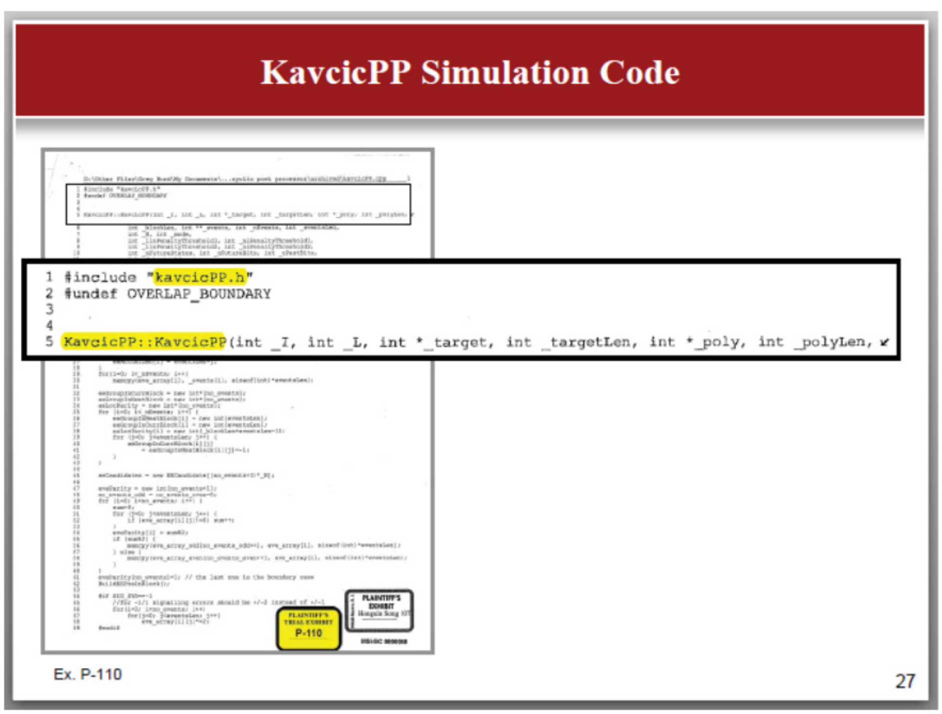
According to the testimony at trial in the CMU case, Marvell's current Chief Technology Officer, Dr. Zining Wu, Marvell named its "Kavcic Simulator" after Dr. Kavcic, because it is "common practice" in engineering to name something after the "author" of the solution, just as the Viterbi detector is named after Dr. Viterbi. Below is an excerpt from Dr. Wu's testimony of December 11, 2012 in the CMU case (Ex. 19 at p. 302):

Q. Why did Marvell name its simulator, Kavcic Viterbi?

A. Kavcic Viterbi is a simulator that simulates the method proposed by Dr. Kavcic in his paper. All right. So we use that simulator to benchmark our own development. It is actually a very common practice to name a simulator after the author. We, at Marvell we have simulator, a Viterbi simulator that is named after Dr. Viterbi, who developed the Viterbi algorithm. We also have a simulator called PCHR. And that's named after the scientist who developed that algorithm. They are initials of P, C, H and R. So we name that after those authors. Its a common practice. I've beep doing that since I was a student.

As mentioned before, Marvell also initially named its first infringing chip design, the MNP, after Prof. Kavcic, by initially calling it the "KavcicPP." *See* Ex. 24 (P-Demo 7) at 23 and 27.





The fact that Marvell named its own products after Prof. Kavcic, the first named inventor of the '839 patent, even though Dr. Zeng works at Marvell and even though Marvel's engineers reviewed Zeng's work before deciding the copy claim 4 of the '839 patent, shows that Zeng's Thesis does not invalidate claim 4. It also is consistent with the view of Prof. Moon, who acknowledged in a peer-reviewed paper in 2001 (six years after Zeng's Thesis) that Profs. Kavcic and Moura, and not his own student, "first derived" a detector for "signal-dependent Gaussian noise" where "the noise characteristics depend highly on the local bit patterns," i.e., on the specific symbol sequence. Ex. 2; *see also* McLaughlin Dec. at ¶¶ 80 and 111. Prof. Moon again acknowledged the CMU inventors and their breakthrough in his 2002 book chapter without a mention of Zeng's work. *See* Ex. 3. If Marvell or Prof. Moon considered claim 4 to be obvious, they would have directed their praise at their colleague or former student. Instead, their repeated and consistent industry praise for Profs. Kavcic and Moura shows that claim 4 is not obvious.

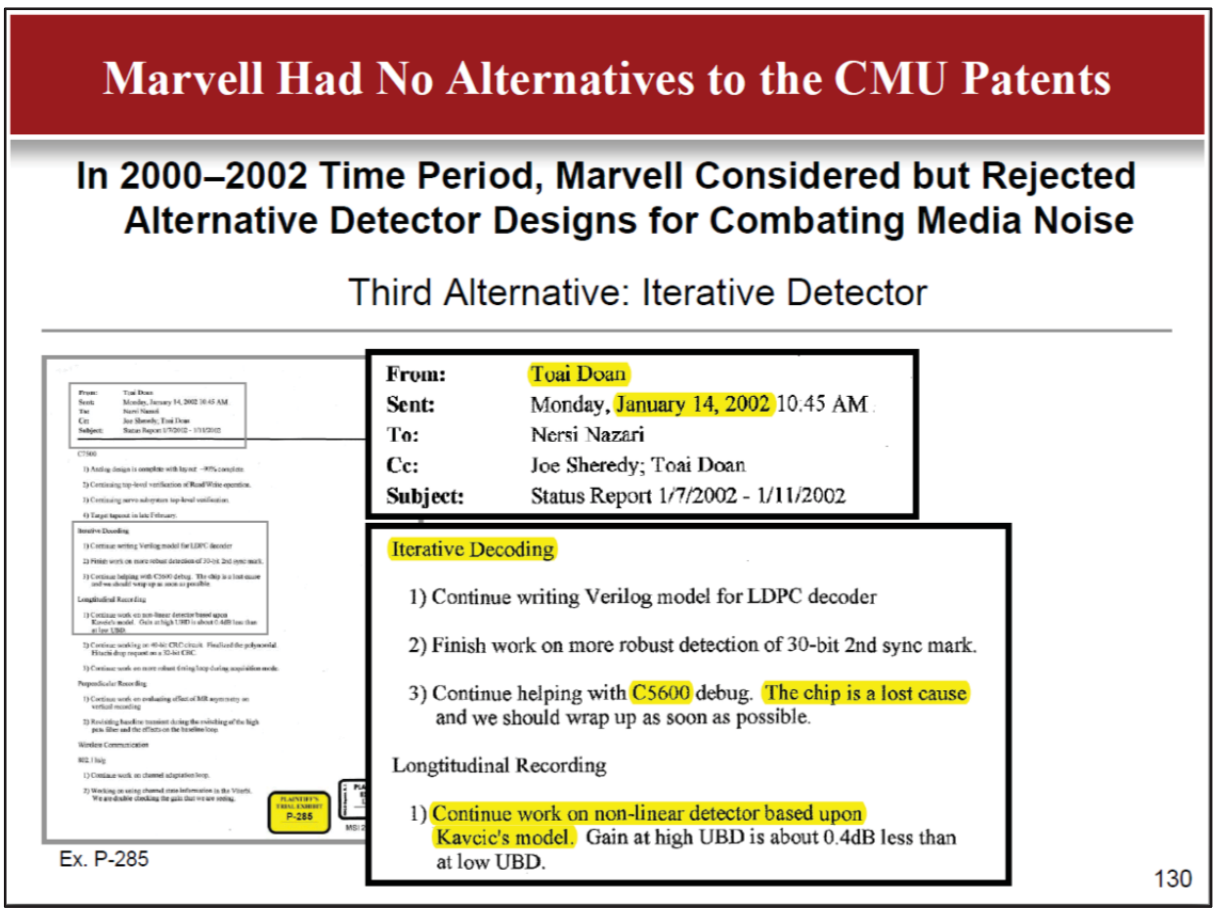
4. <u>Satisfaction of a Long-felt, Unresolved Need</u>

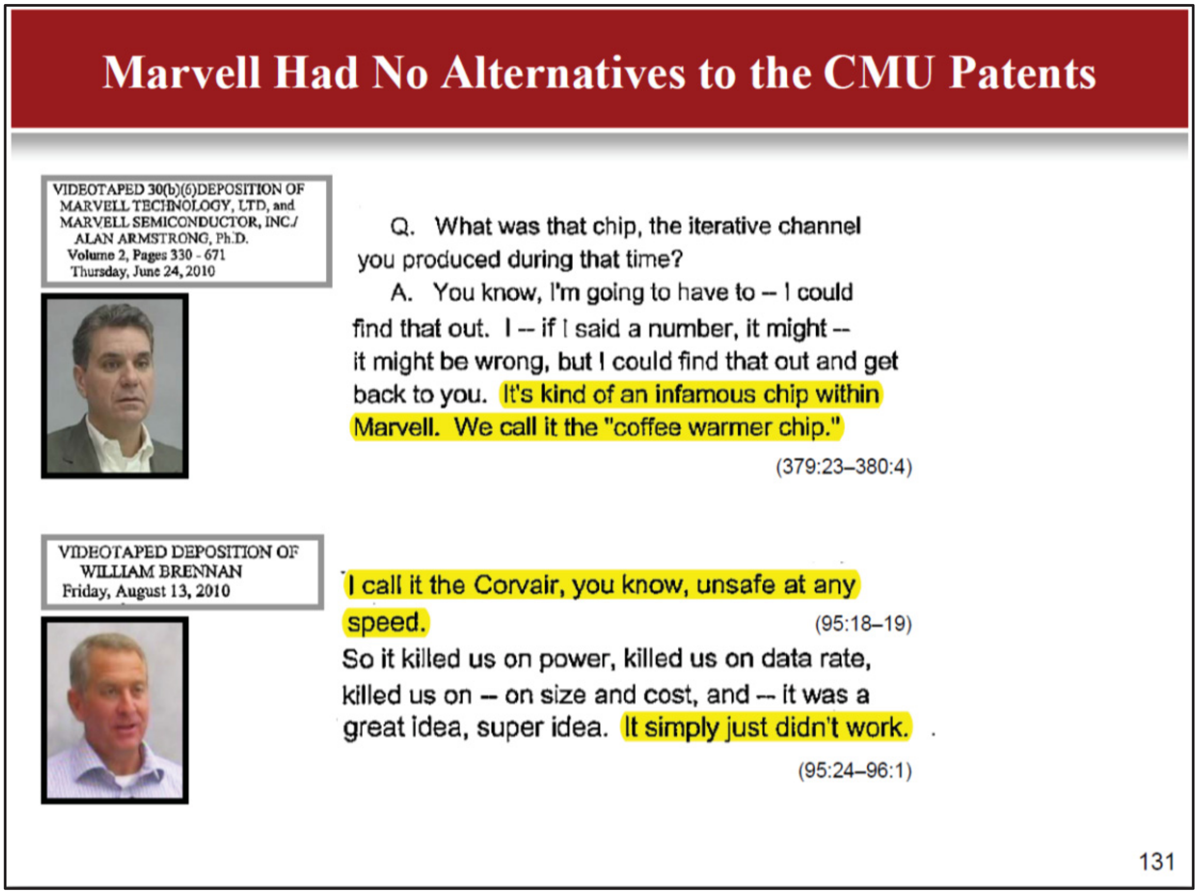
Another indication that an invention is nonobvious is when it satisfies a long felt but unresolved need. *See KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007). Again, the evidence in the CMU case shows that there was a long felt, unresolved need for a method of determining branch metric values for branches of a trellis that accounted for the correlated, signal-dependent noise in the channel. The preceding section shows that Marvell's attempts at combating this increasing problem failed. To overcome these failures, Marvell adopted and copied the methods of the CMU patents, and now Marvell engineers refer to its MNP and NLD products that practice the CMU patents as "a must" to combat the media noise in a HDD. *See e.g.,* Ex. 8 (P-607) (2007 email by a Marvell engineer the CMU technology in Marvell's infringing MNP and NLD chips "a must" because "now days the drives are dominated by media noise"); McLaughlin Dec. at ¶ 116. This evidence too shows that claim 4 is nonobvious.

5. <u>Failure by Others</u>

Failure by others is another secondary consideration that, when present, shows the nonobviousness of the invention. *See KSR Int'l Co. v. Teleflex Inc.*, 550 U.S. 398, 406 (2007).

Here, the evidence in the CMU case showed that prior to copying the methods of claim 4 of the '839 patent and claim 2 of the '180 patent in late 2001 - early 2002, Marvell considered other noninfringing alternatives to combat the problem of ever increasing media noise, but none of those alternative worked. One alternative Marvell considered in this time frame to combat the problem of increasing media noise was a noninfringing "iterative detector." Marvell's engineering team not only failed to get the iterative detector to work but it also became a corporate joke as its executives called it "a lost cause," a "coffee warmer chip," and a "Corvair, unsafe at any speed." *See* Ex. 24 (P-Demo 7) at 130-131 (shown below); *see also* McLaughlin Dec. at ¶ 113.





Judge Fischer, in her post-trial opinions, confirmed the failure of Marvell's prior attempts to address the media noise problem, stating:

> "One of [Marvell's] earlier projects from around 1999 to 2001 was implementing iterative coding, a different method of improving [signal-to-noise ratio] on chips…. However, iterative coding was not initially successful for Marvell.... Marvell was not able to install iterative coding on chips until the 2007-2008 time period…. In fact, Mr. Doan called these chips a 'lost cause' and Mr. Brennan said many people referred to them as 'coffee warmer' chips because they used so much power." Ex. 17 (Dkt. 901) at 15.

Other alternatives that Marvell considered in this time frame (circa 2002) were (i) a modification to a then-existing product, the c5500, and (ii) a so-called "non-linear signal-bit post processor." Marvell never commercialized either of these designs. *See* McLaughlin Dec. at 115; Ex. 24 (P-Demo 7) at 126-129.

In addition to Marvell's failures, Prof. McLaughlin identified numerous others who tried and failed to solve the media noise problem addressed in claim 4, as well as evidence of the long-felt need for the invention, in his rebuttal expert report on validity in the CMU case (Ex. 57) at ¶¶ 94-107. This persistent failure by Marvell and others shows that claim 4 is not obvious.

IX. DR. LEE'S DECLARATION SHOULD BE GIVEN LITTLE WEIGHT

Declarations that lack factual support offer "little probative value in a validity determination." *Ashland Oil, Inc. v. Delta Resins & Refractories, Inc.*, 776 F.2d 281, 294 (Fed. Cir. 1985). Here, Dr. Lee's declaration is of "little probative value" and should be given little if any weight, because it lacks factual support for the following reasons:

- Dr. Lee contradicted his own prior technical writings. Twenty years ago, in a peer-review paper (Ex. 2) and his Ph.D. thesis (Ex. E of the Request), Dr. Lee said Zeng's "random jitter" is white, which is mutually exclusive of signal-dependence. He even criticized Zeng's model in those prior writing because Zeng's "random jitter" is not "data-dependent," i.e., not based on "past data history." Now, in this reexamination, he says the exact opposite, stating that Zeng's "random jitter" "reflects signal-dependent noise." Lee Dec. at ¶ 34. Lee's reexamination declaration provides no explanation for this reversal in positions. Thus, his expert declaration is not factually supported by his own writings.

- In his discussion of Section 5.2 of Zeng's Thesis (*see* Lee Dec. at ¶¶ 43-53), Dr. Lee ignored the well-known, universally-accepted mathematical principle that changing the ***input*** to a function does not change the function into a different function. *See* McLaughlin Dec. at ¶¶ 27-32. Dr. Lee's reexamination declaration ignored this principle when he concluded that Section 5.2 of Zeng's Thesis has 16 branch metric functions, when it only has three. *See* McLaughlin Dec. at ¶¶ 54-57. As explained in

Prof. McLaughlin's declaration, Dr. Lee's error was assuming that changing the target value, which is an input for a branch metric function, changes the function. *See* McLaughlin Dec. at ¶¶ 129-133. Thus, Dr. Lee's expert declaration is not factually supported because it is inconsistent with well-known, universally-accepted mathematical principles.

- Dr. Lee's declaration is also internally inconsistent regarding whether the target values are inputs to the branch metric functions because for Section 4.4 of Zeng's Thesis, Dr. Lee took the correct view that changing the target values does not change the function. For example, at ¶ 38 of his declaration, Dr. Lee states that for Zeng's 4-state trellis embodiment in Section 4.4, branches 2 and 5 use the same branch metric function (eq. 4.26) even though they have different targets. *See* McLaughlin Dec. at ¶ 132. The fact that Dr. Lee takes internally inconsistent positions in his declaration undercuts the reliability of all of his conclusions.

- Dr. Lee did not describe the pervasive errors in Zeng's Thesis. Dr. Lee failed to disclose that Zeng's channel model violates the laws of physics (*see* Kavcic Dec. at ¶¶ 15-23), that Zeng's simulation results defy the laws of physics (*see* Kavcic Dec. at ¶ 28; McLaughlin Dec. at ¶¶ 89-91), and the pervasive math errors in Section 5.2 of Zeng's Thesis (*see* Kavcic Dec. at ¶¶ 30-41).

- Dr. Lee's declaration also relies on flawed logic. At ¶¶ 52-53 of his declaration, Dr. Lee takes the illogical view that when the ***signal samples*** of the readback signal have signal-dependent noise, then the branch metric functions where those signal samples are inputs "are confirmed to be signal-dependent metric functions." *See* McLaughlin Dec. at ¶¶ 125-126. That is illogical because under that view, the Euclidean branch metric function would be "confirmed to be signal-dependent metric functions" merely if the ***signal samples*** have signal-dependent noise. *See* McLaughlin Dec. at ¶ 128. The '839 patent explains, however, that the Euclidean branch metric function is not signal-dependent. *See e.g.*,'839 patent at col. 5:59-64. And even Dr. Lee agrees that the Euclidean branch metric function is not signal dependent. *See* Ex. E of the

Request at 96 (stating that "when the noise is assumed to be white and stationary, … the new error metric is reduced to $(z_k - y_k)^2$, the original [i.e., Euclidean] error metric of the Viterbi detector."); *see also* McLaughlin Dec. at ¶ 128.[51]

- Many of Dr. Lee's assertions in his declaration lack factual support because he relies on quoted passages from Zeng's Thesis that do not support his assertions.
  - For example, at ¶ 34 of his declaration, Dr. Lee relies upon the following quote from Zeng's Thesis to support his erroneous assertion that "Zeng discloses a Viterbi-like sequence detector that expressly takes into consideration signal dependent noise":

    > It is also found that media noise in thin film media is associated with transitions, i.e., more media noise are observed near transition regions that saturation regions [3]. This kind of data pattern dependent noise causes severe degradation in detection performance for most detectors; see [19, 20, 21], for example. The objective of the works summarized in Chapter 4, 5, and 6 are to find new detectors and equalizers to combat this kind of noise and, thus, enhance the data reliability of magnetic storage. [Zeng Thesis at 1].

    That Zeng ***tried*** to "combat" signal-dependent noise does not mean or imply that his branch metric functions in Sections 4.4 and 5.2, in fact, actually account for the signal-dependent noise. As explained in ¶¶ 88-95 of Dr. Bajorek's declaration, to the extent Zeng "combats" signal-dependent noise at all, Zeng does so through means ***other than his BMFs in Sections 4.4 and 5.2,*** such as by: (i) degaussing (AC erasing) the disk prior to writing data to address the transition polarity-driven overwrite effect; (ii) using precompensation to account for the mean peak shift, while ignoring the variance of the peak shift; (iii) ignoring peak shift where transitions are separated by two or more non-transition regions;

51 This assertion by Dr. Lee in his reexamination declaration, therefore, is not only logically flawed, but also represents yet another instance of Dr. Lee contradicting his own prior writings. *See* McLaughlin Dec. at ¶ 120.

(iv) ignoring transition broadening in the readback signal pulses; (v) using the d = 1 RLL code to ignore interacting transitions; and (vi) ignoring the "asymmetry between positive and negative pulses" caused by the interaction between the MR heads and the signal-dependent noise associated with each sequence of symbols written on the media. *See* Zeng Thesis at pp. 9, 65-67, 92. Dr. Lee is wrong to suggest that Zeng combats signal-dependent noise through his BMFs when none of Zeng's techniques for combatting signal-dependent noise utilizes BMFs at all, let alone signal-dependent BMFs. *See* McLaughlin Dec. at ¶ 123.

- Another quote that Dr. Lee relies upon for his assertion that "Zeng discloses a Viterbi-like sequence detector that expressly takes into consideration signal dependent noise," is:

  > Thus, for the Viterbi-like detector, the decoding process is similar to the conventional Viterbi algorithm, except that the recursive cycle is now from 0 to N+1 (with readback samples 0, Z_0, Z_1, …, Z_N), and only the first N decoded bits are stored as the detected user bits.

  Zeng Thesis at 75-76 (quoted at ¶ 34 of Lee's declaration). This quote in no way supports Dr. Lee's assertion that "Zeng discloses a Viterbi-like sequence detector that expressly takes into consideration signal dependent noise." Nothing in this quote refers to or even implicates signal-dependent noise. Nor does Dr. Lee explain why this quote supports his assertion. The fact that the Viterbi algorithm might use a greater recursive algorithm has nothing to do with whether the detector accounts for signal-dependent noise. *See* McLaughlin Dec. at ¶ 124.

For at least these reasons, Dr. Lee's declaration should be given no or little weight.

X. WRITTEN STATEMENT OF INTERVIEW

CMU's Written Statement of Interview is filed concurrently with this response.

XI. SERVICE ON REQUESTER

A certificate of service is submitted herewith showing proof of service of this response and all declarations (including their exhibits and appendices) on the Requester at the address set forth below in accordance with 37 C.F.R. §§ 1.248 and 1.550(f).

J. Steven Baughman
Ropes & Gray LLP
Prudential Tower.
Patent Department
800 Boylston Street
Boston, MA 02199-3600

XII. CONCLUSION

The patentability of claim 4 should be confirmed for the reasons stated above. A representative of the Central Reexamination Unit is invited to contact the undersigned with any questions regarding this reexamination.

Respectfully submitted,

Date: September 3, 2014

/Mark G. Knedeisen/
Mark G. Knedeisen
Reg. No. 42,747

K&L GATES LLP
K&L Gates Center
210 Sixth Ave.
Pittsburgh, Pennsylvania 15222

Ph. (412) 355-6342
Fax (412) 355-6501
email: mark.knedeisen@klgates.com