

Deep Neural Network Based Media Noise Predictors for Use in High-Density Magnetic Recording Turbo-Detectors

Amirhossein SAYYAFAN¹, Benjamin J. BELZER¹, Krishnamoorthy SIVAKUMAR¹, Jinlu SHEN¹, Kheong Sann CHAN² and Ashish JAMES³

1) Washington State University, Pullman, WA, USA, (a.sayyafan, belzer, siva, jinlu.shen)@wsu.edu

2) Nanjing Institute of Technology, Nanjing, China, kheongsann@ieee.org

3) Institute for Infocomm Research (I2R), A*STAR, Singapore, ashish_james@i2r.a-star.edu.sg

I. INTRODUCTION

Trellis based detection with pattern dependent noise prediction (PDNP) [1] has become standard practice in the HDD industry. In a typical single-track signal processing scheme, the received samples from the read head are first filtered by a linear equalizer with a 1D partial response (PR). The linear filter output flows into a trellis-based (e.g. BCJR [2]) detector that employs a super-trellis based on the PR mask ISI channel and a 1D pattern dependent noise prediction (1D PDNP) algorithm. The effective channel model has a media noise term which models signal dependent noise due to, e.g., magnetic grains intersected by bit boundaries. The media noise can influence two or more bit readback values. The trellis detector sends soft estimates of the coded bits to a channel decoder, which outputs estimates of the user information bits. There are two problems with traditional PDNP. First, when the number of tracks N_t simultaneously processed is greater than one, e.g. in two-dimensional magnetic recording (TDMR), the number of trellis states can become impractically large; this is the state explosion problem. Second, the relatively simple autoregressive noise model and linear prediction used in PDNP is somewhat restrictive and may not accurately represent the media noise, especially at high storage densities; this is the modeling problem. To address the state explosion problem, we separate the ISI detection and media noise prediction into two separate detectors and use the turbo-principle to exchange information between them, thus avoiding use of a super-trellis. To address the modeling problem, we design and train deep neural network (DNN) based media noise predictors. As DNN [3] models are much more general than autoregressive models, they give a more accurate model of magnetic media noise than PDNP; this more accurate modeling results in reduced detector BERs compared to PDNP.

II. SYSTEM MODEL

The BCJR-DNN turbo detector assumes a channel model for the k th linear equalizer filter output $y(k)$:

$$y(k) = (\mathbf{h} * \mathbf{u}) + n_m(k) + n_e(k) \quad (1)$$

where $\mathbf{h}$ is the PR target, $\mathbf{u}$ are the coded bits on the track, $*$ indicates 1D convolution, $n_m(k)$ is media noise, $n_e(k)$ is reader electronics AWGN, and the ISI channel length $L = \text{length}(\mathbf{h}) - 1$. The media noise $n_m(k)$ is not modeled as an AR process; instead a more general model for $n_m(k)$ is learned by the DNN through offline training. We use grain flipping probabilistic (GFP) model data to train and evaluate our system. The GFP waveforms are generated based on micro-magnetic simulations [4]. The simulated media has grain density of 11.4 Teragrains/in². The GFP waveforms correspond to five tracks of coded bits (± 1), denoted as tracks 0 through 4. They are written using shingled writing technology. Fig. 1 shows a cartoon representation of the GFP model output data. The blue and red stripes represent -1 and $+1$ coded bits. In our GFP simulations, the track pitch (TP) is 48 nm, the bit length (BL) is 11 nm, and there are 9.33 grains per coded bit (GPB). Each track in a GFP data set corresponds to 41206 coded bits, which is close to the sector size of 32768 bits (4K bytes) in a typical HDD. The readings from Track #2 are used as input to the BCJR-DNN turbo detector, and to a comparison baseline 1D PDNP detector.

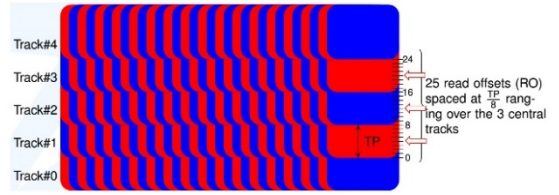


Fig. 1 Cartoon representation for GFP model output data.

III. BCJR-DNN DETECTOR

Fig. 2 shows the system block diagram for the proposed BCJR-DNN turbo detector. This system is a turbo-equalization structure that separates the ISI detection and media-noise prediction functions into two detectors that iteratively exchange LLRs estimates of coded bits and noise samples until convergence to a low BER occurs. In Fig. 2, the GFP simulated HDD read-head output vector $\mathbf{r}$ contains two samples per coded bit. The

Benjamin J. BELZER

E-mail: belzer@wsu.edu

tel: +1-509-335-4970

Work funded by NSF grant CCF-1817083

odd samples $\mathbf{r}^{(1)}$ (the “first samples” per bit) are first filtered by a linear equalizer designed to minimize the mean squared error (MMSE) between the filter output $\mathbf{y}^{(1)}$ and the convolution of the coded bits $\mathbf{u}$ with the 1D PR mask $\mathbf{h}$. The filter output $\mathbf{y}^{(1)}$ is input to the BCJR detector, which handles only ISI equalization based on the PR target $\mathbf{h}$ and outputs a block $\mathbf{LLR}_b$ of 41206 coded bit LLRs. In this work, we designed the PR target $\mathbf{h}$ with three taps, so that the BCJR detector has four states and eight total branches, and the ISI channel length $I = 2$. The BCJR’s coded bit LLRs $\mathbf{LLR}_b$ are sent to the DNN, which provides an estimate $\hat{\mathbf{n}}_m$ of the media noise to the next iteration of the BCJR detector in order to improve the BCJR’s estimate $\mathbf{LLR}_b$. The dotted lines and box in Fig. 2 indicate future work beyond the present paper; they show how the detector can interface to an LDPC channel decoder.

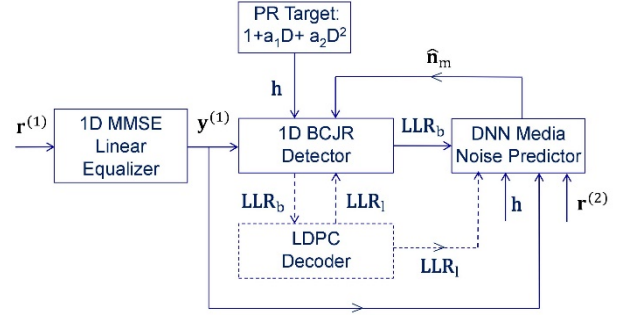


Fig. 2 Block diagram for the BCJR-DNN turbo detector

We employ two methods for interfacing the BCJR detector to the DNN. In the first method, labeled as “1 DNN” in Table I, one DNN estimates the media noise for the k th BCJR trellis stage based on $\mathbf{LLR}_b$ (and on $\mathbf{y}^{(1)}$ and even samples $\mathbf{r}^{(2)}$) and then passes this estimate $\hat{\mathbf{n}}_{m_k}$ to all eight BCJR branches. In the second method, labeled as “8 DNNs” in Table I, a media noise estimate $\hat{\mathbf{n}}_{m_{kj}}$, $0 \leq j \leq 7$, for the j th branch of the k th BCJR trellis stage is provided by a DNN, denoted DNN_j , dedicated to (and trained for) the j th branch. We investigate two neural network architectures for the noise predictor. The first architecture is the traditional fully connected deep neural network (FCDNN). The second architecture is the convolutional neural network (CNN), wherein each CNN layer has a bank of trained finite length filters connected to an output layer.

IV. SIMULATION RESULTS

Table I summarizes the simulation results for three-tap unit energy (UE) and monic PR masks designed using the method described in [5]. One turbo-loop between the BCJR and the DNN is performed. The disjoint training and test data sets have 16 blocks each. We consider four input scenarios involving some combination of the sign of the LLRs or their corresponding probabilities, the filtered first sample sequence $\mathbf{y}^{(1)}$, the PR target $\mathbf{h}$, and even samples $\mathbf{r}^{(2)}$ of GFP readings. For 1 DNN, the BCJR-CNN detector achieves $0.371 \times$ the BER and $0.116 \times$ the per bit run time (PBRT) of a 1D PDNP detector with ISI channel length $I = 2$ and 128 trellis states; for 8 DNNs, it achieves $0.343 \times$ the BER and $0.731 \times$ PBRT of a 1D PDNP detector.

REFERENCES

- 1) J. Moon and J. Park, “Pattern-dependent noise prediction in signal dependent noise,” *IEEE Jour. Sel. Areas Commun.*, vol. 19, no. 4, pp. 730–743, Apr 2001.
- 2) L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, “Optimal decoding of linear codes for minimizing symbol error rate,” *IEEE Trans. Inform. Theory*, vol. 20, pp. 284–287, Mar. 1974.
- 3) Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, pp. 436–444, May 2015.
- 4) K. S. Chan, R. Radhakrishnan, K. Eason, M. R. Elidrisi, J. J. Miles, B. Vasic, and A. R. Krishnan, “Channel models and detectors for two-dimensional magnetic recording,” *IEEE Trans. Mag.*, vol. 46, no. 3, pp. 804–811, March 2010.
- 5) C. K. Matcha and S. G. Srinivasa, “Generalized partial response equalization and data-dependent noise predictive signal detection over media models for TDMR,” *IEEE Trans. Mag.*, vol. 51, pp. 1–15, Oct. 2015.

Table I Simulation results for PDNP and BCJR detectors

Method		PR Mask	BER		MSE	
PDNP		UE	4.37e-4		-	
		Monic	5.39e-4		-	
DNN	Input		8 DNNs	1 DNN	8 DNNs	1 DNN
FC	Sign[LLR], h, y ⁽¹⁾	UE	1.08e-3	9.06e-4	2.12e-3	1.87e-3
		Monic	2.70e-4	2.79e-4	9.89e-4	9.88e-4
	Pr[LLR], h, y ⁽¹⁾	UE	7.24e-4	6.87e-4	3.33e-3	3.74e-3
		Monic	2.34e-4	2.15e-4	9.47e-4	9.98e-4
	Sign[LLR], h, y ⁽¹⁾ , r ⁽²⁾	UE	2.96e-4	5.76e-4	1.38e-3	1.40e-3
		Monic	2.25e-4	2.78e-4	9.73e-4	9.85e-4
CNN	Sign[LLR], h, y ⁽¹⁾ , r ⁽²⁾	UE	2.11e-4	2.23e-4	1.29e-3	1.32e-3
		Monic	2.17e-4	2.06e-4	9.28e-4	9.22e-4
	Pr[LLR], h, y ⁽¹⁾ , r ⁽²⁾	UE	1.77e-4	2.38e-4	5.74e-4	7.65e-4
		Monic	1.58e-4	1.67e-4	6.22e-4	6.69e-4
	Sign[LLR], h, y ⁽¹⁾ , r ⁽²⁾	UE	1.76e-4	1.68e-4	5.67e-4	5.54e-4
		Monic	1.50e-4	1.62e-4	5.90e-4	5.78e-4