Deep Neural Network Media Noise Predictor Turbo-detection System for One and Two Dimensional High-Density Magnetic Recording

Amirhossein SAYYAFAN[1], Benjamin J. BELZER[1], Krishnamoorthy SIVAKUMAR[1], Kheong Sann CHAN[2] and Ashish JAMES[3]

1) Washington State University, Pullman, WA, USA, (a.sayyafan, belzer, siva)@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. SYSTEM MODEL

The hard disk drive (HDD) industry is facing a physical limit on the areal density (AD) of one-dimensional magnetic recording (1DMR) on traditional magnetic media. To increase capacity without media redesign, two-dimensional magnetic recording (TDMR) has been introduced. The effective channel model has a media noise term which models signal dependent noise due to, e.g., magnetic grains intersected by bit boundaries. Trellis based detection with pattern dependent noise prediction (PDNP) [1] is standard practice in HDDs. The trellis detector sends soft coded bit estimates to a channel decoder, which outputs user information bit estimates. PDNP uses a relatively simple autoregressive noise model and linear prediction; this model is somewhat restrictive and may not accurately represent the media noise, especially at high storage densities. To address this modeling problem, we design and train deep neural network (DNN) based media noise predictors. As DNN [2] models are more general than autoregressive models, they more accurately model media noise compared to PDNP. The proposed turbo detector assumes a channel model for the kth linear equalizer filter output $y(k)$:

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

where $\mathbf{h}$ is the PR target, $\mathbf{u}$ are the coded bits on the track, $*$ indicates 1D/2D convolution, $n_m(k)$ is media noise, and $n_e(k)$ is AWGN. We use a grain flipping probabilistic (GFP) model data (based on micro-magnetic simulations [3]) to train and evaluate our system. The simulated media grain density is 11.4 Teragrains/in^2. The GFP waveforms correspond to five tracks of coded bits (±1), denoted as tracks 0 through 4, written using shingled writing with a bit length (BL) of 11 nm. For 1DMR, the track pitch (TP) is 48 nm, and there are 9.33 grains per coded bit (GPB). For TDMR, TP =18 nm, and GPB = 3.50. There are 41206 and 41202 coded bits per track for TP 48 nm and 18 nm respectively, which are close to the sector size of 32768 bits in a typical HDD. For 1DMR, we use the readings from Track #2 as inputs, and for TDMR, the readings are from Track #1 through #3. The proposed methods are evaluated in comparison with baseline PDNP detectors.

II. BCJR-LDPC-DNN TURBO DETECTOR

In [4], we proposed a BCJR-DNN turbo detector for 1DMR, without LDPC decoding. Here, we generalize to three-track TDMR and include decoding. Fig. 1 shows the BCJR-LDPC-DNN turbo detector for TDMR with separate trellis-based ISI/ITI detection and DNN-based media-noise prediction. Log-likelihood-ratio (LLR) and media noise estimates are exchanged between these two detectors until the BER converges. In Fig. 1, the GFP simulated read-head output vector $\mathbf{r}$ contains two samples per coded bit. The odd samples $\mathbf{r}^{(1)}$ (the "first samples" per bit) for all five input tracks are passed to a 2D partial response (PR) equalizer designed to minimize the mean squared error (MMSE) between the three-track filtered output $\mathbf{y}^{(1)}$ and the convolution of the coded bits $\mathbf{u}$ with the 2D PR mask $\mathbf{h}$. The 2D-BCJR trellis detector performs ISI equalization on input $\mathbf{y}^{(1)}$, and generates LLR outputs. The PR target $\mathbf{h}$ is 3×3, hence, the 2D-BCJR state bits are 3×2, so the trellis has 64 states. In the 1st iteration, the 2D-BCJR LLRs $\mathbf{LLR}_{b_0}$, $\mathbf{y}^{(1)}$ and the

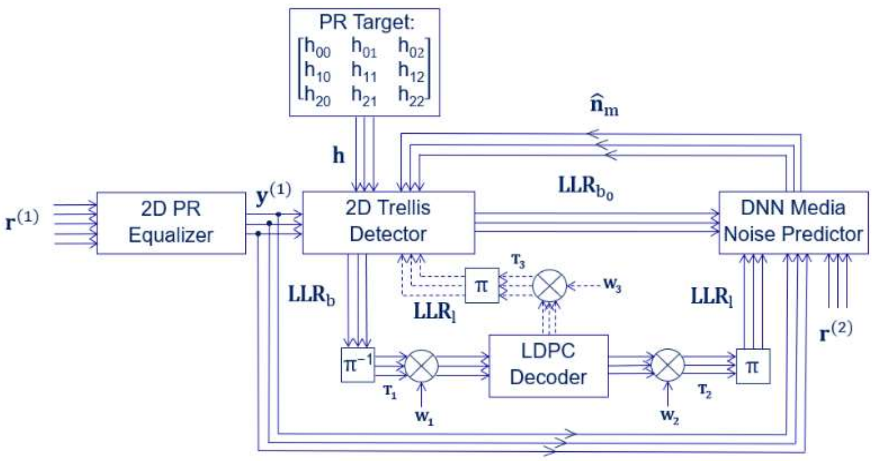

Fig. 1 BCJR-LDPC-DNN turbo detector for TDMR.

Benjamin J. BELZER
E-mail: belzer@wsu.edu
tel: +1-509-335-4970
Work funded by NSF grant CCF-1817083

unfiltered even samples $\mathbf{r}^{(2)}$ are passed to the DNN to estimate the media noise $\hat{\mathbf{n}}_m$. The noise $\hat{\mathbf{n}}_m$ is fed back to the 2D-BCJR to obtain a lower BER. Next, the 2D-BCJR passes LLRs $\mathbf{LLR}_b$ to a low density parity check (LDPC) decoder. The decoder generates the final LLRs $\mathbf{LLR}_l$ at the end of each turbo-iteration.

AD is determined by increasing the LDPC code rate until the decoded BER is $\leq 10^{-5}$. For the 2nd iteration, the decoder LLRs $\mathbf{LLR}_l$ are passed as inputs to the DNN instead of $\mathbf{LLR}_b$. The system exchanges LLRs between DNN, BCJR and decoder iteratively to achieve higher AD. The dotted lines in Fig. 1 indicate optional inner iterations between the BCJR and the LDPC decoder. We investigate two approaches to interface the BCJR and LDPC with the DNN. In the first approach labeled as “3 CNNs 1 pass” in Table II, we use one convolutional neural network (CNN) to estimate the media noise $\hat{n}_{m_k}$ for the kth BCJR trellis stage for the three tracks. The second approach labeled as “3 CNNs 1 pass” predicts $\hat{n}_{m_k}$ using a separate CNN for each track. By employing three CNNs, we expand the search space to improve the estimation of the media noise.

III. SIMULATION RESULTS

Table I presents simulation results for PDNP and BCJR-LDPC-DNN detectors for 1DMR. The 1D turbo detector processes only the middle track, and its PR target **h** has three taps, thus its trellis has 4 states. The 1D-PDNP uses the same PR target, but has 128 trellis states. LLRs are exchanged between the 1D-PDNP and the LDPC decoder iteratively in a turbo architecture. To train and test the 1D-PDNP parameters, 20 and 80 blocks are used respectively. The performance of the detectors is evaluated on the TP 48 nm dataset. To be consistent with [4] we use 16 blocks for the training data set, and 84 blocks for the test dataset. Two turbo-iterations are performed for both methods. For 1D-PDNP, the second iteration does not improve the AD, hence, we only report its first iteration. The first turbo-iteration of BCJR-LDPC-DNN detector labeled as “CNN 1 pass” acheives 1.60% density gain over 1D-PDNP. The second iteration labeled as “CNN 2 passes” has density gain of 2.01% over 1D-PDNP.

Table II presents simulation results for PDNP and BCJR-LDPC-DNN detectors for TDMR; the PDNP results are from [5]. The two-track 2D-PDNP looks at 2×3 bit patterns and has 64 states; 40 and 60 blocks are used for training and testing, respectively. The best PDNP performance belongs to 1D-PDNP with two turbo-iterations labeled as “1D-PDNP 2 passes”, which we consider as the baseline. For the BCJR-LDPC-DNN detector, we use 80 and 20 blocks as the training and test datasets respectivly. The one-CNN architecture (labeled “1 CNN 1 pass”) achieves a density gain of 27.52% over the baseline. For the three-CNNs architecture (labeled “3 CNNs 1 pass”), the BCJR-LDPC-DNN detector has a density gain of 27.75% compared to the baseline.

Table I Simulation of PDNP and BCJR-LDPC-DNN detectors for 1DMR at TP 48 nm

TDMR Detectors	Areal Density (Tb/in^2)	User Bits per Grain	Code Rate
1D-PDNP	1.186	0.1041	0.9710
CNN 1 pass	**1.205**	0.1057	0.9865
CNN 2 passes	**1.210**	0.1062	0.9905

Table II Simulation of PDNP and BCJR-LDPC-DNN detectors for TDMR at TP 18 nm

TDMR Detectors	Areal Density (Tb/in^2)	User Bits per Grain	Code Rate
1D-PDNP 1 pass	2.482	0.2177	0.7600
1D-PDNP 2 passes	2.531	0.2220	0.7750
2D-PDNP 1 pass	2.230	0.1957	0.6830
1 CNN 1 pass	**3.228**	0.2833	0.9883
3 CNNs 1 pass	**3.234**	0.2838	0.9901

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] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, pp. 436–444, May 2015.

[3] K. S. Chan, R. Radhakrishnan, K. Eason, M. R. Elidrissi, 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.

[4] A. Sayyafan, B. J. Belzer, K. Sivakumar, J. Shen, K. S. Chan, and A. James, “Deep neural network based media noise predictors for use in high-density magnetic recording turbo-detectors,” *IEEE Trans. Mag.*, vol. 55, no. 12, pp. 1–6, Dec. 2019.

[5] J. Shen, A. Aboutaleb, K. Sivakumar, B. J. Belzer, K. S. Chan, A. James, “Deep Neural Network a Posteriori Probability Detector for Two-dimensional Magnetic Recording”, *IEEE Trans. Mag.*, vol. 56, no. 6, pp. 1-12, June 2020.