Joint Equalization and Decoding for Two-Dimensional Intersymbol Interference Channels: A Comparative Study

Naveen Singla and Joseph A. O'Sullivan
Electronic Systems & Signals Research Laboratory
Department of Electrical & Systems Engineering
Washington University
St. Louis, MO 63130, USA
Email: [singla, jao]@wustl.edu

Abstract— **We study different joint equalization and decoding schemes for two-dimensional intersymbol interference (ISI) channels vis-à-vis their performance and computational complexity. Three detection schemes are considered; two of these three schemes perform sum-product message-passing on different factor graph representations of the channel ISI and the third one uses a minimum mean-square-error (MMSE) equalizer. One of the two factor graph representations considered is kernel-based and the other is state-based. Sum-product message-passing on the state-based representation is equivalent to a multi-track version of the BCJR algorithm. Low-density parity-check (LDPC) codes are used for error-correction. Simulations show that the two message-passing based schemes have almost identical performance. The MMSE-LDPC scheme, even though it has much lower complexity than the message-passing based schemes, performs nearly as well.**

I. Introduction

Two-dimensional intersymbol interference (ISI) channels have received a lot of attention lately. The primary reason for this is the emergence of two-dimensional paradigms for data storage. Current prevalent data storage technologies are either magnetic or optical. Where magnetic recording has made ultra-high density storage possible, optical storage has provided an affordable, portable medium for data storage and dissemination. However, it is well-known that current storage technologies, that are inherently one-dimensional, are restricted by physical limits which will prevent them from keeping up with the ever-increasing demands for data storage. Superparamagnetism is, and has been for a long time, a major concern for longitudinal magnetic recording. The recently announced blu-ray disk uses a blue-violet laser with a wavelength that is almost at the lower end of the visible light spectrum; storage capacity for optical recording has increased (partly) by switching from using infra-red (CD) to red (DVD) to blue-violet (blu-ray disc) lasers and sustaining this trend is going to become more and more challenging. Factors like this, and many others, have motivated research into alternate storage technologies. Perpendicular magnetic recording has shown a lot of promise in this regard. However, a promising development (and perhaps the long-term solution) is the shift in research focus in the storage industry towards developing a two-dimensional paradigm for storage.

The principle underlying two-dimensional storage is simple: since the storage medium is two-dimensional it makes sense to store data in a manner that utilizes both dimensions. Current storage technologies utilize the second dimension only loosely; data is stored in concentric tracks and adjacent tracks are separated by a guard band which reduces the area available for storing data. The two-dimensional storage paradigm is being applied to both magnetic and optical recording and the results, although preliminary, are quite encouraging. Patterned magnetic media, in which information is stored in isolated single grain islands, make areal densities on the order of $10^{12} bits/in^2$ feasible, which is far beyond the limits of conventional magnetic recording. The two-dimensional optical storage (TwoDOS) paradigm involves storing data in a "broad" spiral on the disc. Each revolution of the spiral (a track) contains multiple rows of bits (typically 11) stacked together. Doing this allows storage capacities up to twice that of the blu-ray disc while using the same optics.

However, the increase in storage capacity by using a two-dimensional paradigm comes at a price. Due to the two-dimensional nature of storage, the ISI during the readback process is also two-dimensional. Conventional recording technologies like magnetic hard disks and DVDs have one-dimensional ISI for which many low-complexity schemes exist for detection. Partial response maximum-likelihood (PRML) decoding is one such scheme that has proved highly successful, especially for magnetic recording. Extending PRML to two dimensions is not straightforward since the Viterbi algorithm, which is used for maximum-likelihood decoding, becomes computationally intractable in two dimensions, as does the BCJR algorithm. In fact, in a recent paper Ordentlich and Roth have shown that the problem of maximum-likelihood sequence detection for two-dimensional ISI channels is NP-complete [1]. This motivates the need for new methods to combat two-dimensional ISI. Besides advanced storage technologies, multi-user communication scenarios, like cellular communication or optical communication systems employing dense wavelength-division multiplexing, also have situations where two-dimensional ISI is prevalent.

The importance of two-dimensional ISI channels can also be gauged by the amount of research being carried out to develop decoding algorithms for such channels. Currently many schemes exist that have had varying degrees of success in dealing with two-dimensional ISI [2]-[15]. These schemes range from using two one-dimensional detectors, one for each dimension of the ISI to using simple two-dimensional minimum mean-square-error (MMSE) equalizers to the use of more sophisticated ideas, like message-passing. The schemes proposed by Singla *et al.* [10]-[13] consider using error-control coding in conjunction with the equalization schemes and show that significant improvement in performance can be obtained by doing so. In this paper we study the application of three different joint equalization and decoding schemes for two-dimensional ISI channels vis-à-vis their performance and computational complexity. Two of the schemes involve performing sum-product message-passing on two different joint code/channel factor graphs that represent the underlying system. The third scheme uses an MMSE equalizer and an error-control code in an iterative setting. The difference between the two factor graph based approaches is that one uses a kernel-based representation of the channel and the other uses a state-based [16] representation. Sum-product message-passing on the state-based representation is equivalent to a multi-track version of the BCJR algorithm. Low-density parity-check (LDPC) codes are used for error-correction and in all the cases their factor graph representation is kernel-based (commonly referred to as the Tanner graph). These schemes are chosen because they include both linear and nonlinear equalization techniques. Due to its pervasiveness, MMSE equalization is the logical choice for a linear equalization technique. The two graphical representations are chosen because they are the most commonly used graphical representations in error-control coding. All the schemes compute either a bit-by-bit MMSE or bit-by-bit MAP estimate of the data.

The rest of the paper is organized as follows. Section 2 describes the model that is used in this paper for the system with two-dimensional ISI. Section 3 describes the three joint equalization and decoding schemes. Simulation results and complexity comparison for the three schemes are provided in Section 4. Section 5 concludes the paper.

II. Channel Model

The system is modeled as a discrete-time communication system as shown in Fig. 1. The uncoded user data, $\mathbf{a}$, is encoded using the error-control code to obtain the encoded data. The encoded data is arranged in a matrix, $\mathbf{X}$, which is typically rectangular. The user data and the encoded data are assumed to be binary. For error correction we use LDPC coset codes [17] with the code graph chosen uniformly at random from the ensemble of regular graphs. Prior to transmission over the channel the encoded data is modulated so that $0\rightarrow 1$ and $1\rightarrow -1$. This corresponds to the two most popular forms of recording: saturation recording on magnetic media or absence and presence of pits on optical media. It also corresponds to a communication system employing BPSK modulation. A guard band of all ones is appended around the encoded data matrix. The guard band is used for the initialization and termination of our schemes. The guard band width is one less than the extent of the ISI.

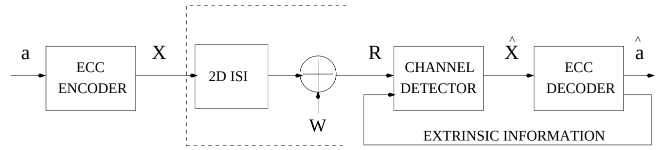

Fig. 1. Equivalent discrete-time model for the system with two-dimensional ISI.

The output of the channel is a blurred and noisy version of the encoded data. $\mathbf{W}$ represents additive white Gaussian noise (AWGN) with zero mean and variance σ_w^2. The ISI is of course two-dimensional and assumed to be linear, although this restriction can be relaxed. The channel output, $\mathbf{R}$, is a matrix with elements

$$r(i,j) = \sum_{k_1=0}^{L-1}\sum_{k_2=0}^{L-1} x(i-k_1, j-k_2)h(k_1,k_2) + w(i,j), \quad (1)$$

where L represents the number of elements over which the ISI extends in each dimension; $\mathbf{h} = \{h(k_1,k_2)_{k_1,k_2=0}^{L-1}\}$ is the $L{\times}L$ ISI matrix also referred to as the channel point spread function in recording; $x(i,j)$ are the encoded data; and $w(i,j)$ are instances of the AWGN. The coefficients of the point spread function are assumed to be real. Note that, without loss of generality, we are assuming a square point spread function since it is always possible to pad any point spread function with a sufficient number of zeros to make it square. After the data is received it is presented to the receiver where it first passes through an equalizer followed by the decoding of the error-control code. The receiver is iterative; that is, the equalizer and decoder exchange extrinsic information and use it as *a priori* information for their computations. This idea, first applied in the decoding of turbo codes, is the well-known turbo principle. In this paper we consider only soft information exchange between the receiver components. Typically, exchanging soft information improves performance over exchanging hard information, albeit at the price of increased communication cost between the equalizer and the decoder.

III. Joint Equalization and Decoding Algorithms

This section describes the three joint equalization and decoding schemes for two-dimensional ISI channels. The primary focus will be on the equalization part of the joint schemes since the decoding of LDPC codes is standard and well understood [16],[18],[19] and will be omitted here.

A. Full Graph Algorithm

It is well-known now that the sum-product algorithm is an algorithm that computes the marginals of a global function of several variables [16]. The algorithm derives its computational efficiency by exploiting the way in which the global function factors into a product of local functions each of the which depend only on a subset of the variables. For decoding, the sum-product algorithm computes the bit-by-bit maximum *a posteriori* (MAP) estimate of the codeword given the measured data. In this case the global function under consideration is the posterior probability of the encoded data given the measured data, $P(\mathbf{X}|\mathbf{R})$. For our problem this joint probability can be factored as follows;

$$\begin{aligned}\Pr(\mathbf{X}|\mathbf{R}) &\propto p(\mathbf{R}|\mathbf{X})\Pr(\mathbf{X})\\ &\propto \prod_{(i,j)} p(r(i,j)|\{x(k,l):(k,l)\in\mathcal{N}(i,j)\})\Pr(\mathbf{X}),\end{aligned}$$

where $\mathbf{X}$ and $\mathbf{R}$ are as defined in the previous section; and $\mathcal{N}(i,j)$ are the indices of all the bits that interfere with $x(i,j)$, including $x(i,j)$. This factorization can be expressed graphically by means of a factor graph. This graph has three levels corresponding to the measurements (measured data nodes), codeword bits (codeword bit nodes), and the parity-check equations of the LDPC code (check nodes). The codeword bit nodes are connected to the check nodes and the measured data nodes via the parity-check matrix of the LDPC code and the two-dimensional ISI, respectively. This is commonly termed as the "kernel" representation of the system. Performing sum-product message-passing on this "full graph" allows computation of the MAP estimate of the codeword bits given the measured data. The following message-passing schedule is used: messages are passed from the codeword bit nodes to the measured data nodes and back, then from the codeword bit nodes to the check nodes and back. This completes one iteration. The channel part of the full graph has many length 4 cycles, as illustrated by Fig. 2 for a 2×2 ISI, hence the computed posterior probabilities are only approximate [16].

B. Multi-Track BCJR

In certain cases it is seen that the introduction of hidden or auxiliary variables can transform a factor graph with cycles into one that is cycle-free. A well-known example of such a transformation is the trellis. For one-dimensional ISI channels this gives an alternate, cycle-free representation of the kernel-based channel graph which otherwise has many short cycles (when the ISI length is greater than 2). The channel in (1) also has a state-based representation that is cycle-free if the columns of $\mathbf{X}$ are considered to be state variables and the row direction is considered as "time." Denoting the ith column of $\mathbf{X}$ as $\mathbf{X}_i$, the state at time t for an $L\times L$ ISI is $\mathbf{S}_t = (\mathbf{X}_{t-L+1}, \mathbf{X}_{t-L+2}, \ldots, \mathbf{X}_t)$. Also let $\mathbf{S}_0$ be a known state (determined by the known guard band), typically the all zeros state. Then the joint posterior probability can be factored as follows,

$$\Pr(\mathbf{X}|\mathbf{R},\mathbf{S}_0) \propto \prod_{t=1}^{n} p(\mathbf{R}_t|\mathbf{X}_t,\mathbf{S}_{t-1})\Pr(\mathbf{S}_t|\mathbf{S}_{t-1}), \qquad (2)$$

where n denotes the number of columns in the matrix $\mathbf{R}$. This factorization can be represented conveniently as a trellis or as a Wiberg-type graph [20] which is just the factor graph representation of a trellis. Sum-product message-passing can then be performed on this graph to obtain the MAP estimate; this is what the BCJR algorithm does. Since the factor graph in this case is cycle-free the posterior probabilities will be exact provided that the encoded bits are independent. This would work well except that message-passing on this trellis is computationally infeasible since the size of the state space grows exponentially with the size of the data. For example, for an $n\times n$ encoded (binary) data matrix and an $L\times L$ ISI the number of states is 2^{nL}. However, if instead of treating an entire column as a state, we use only a fixed number of rows per column to define a state then we obtain a computationally tractable version of the BCJR algorithm for two-dimensional ISI channels. This is akin to the multi-track version of the Viterbi algorithm proposed by Krishnamoorthi [?] and Weeks [14]. Marrow and Wolf [7] have proposed an algorithm for detection on two-dimensional ISI channels which is very similar to the multi-track BCJR algorithm. However, none of the aforementioned papers consider the use of error-control coding. As mentioned earlier, we want to study

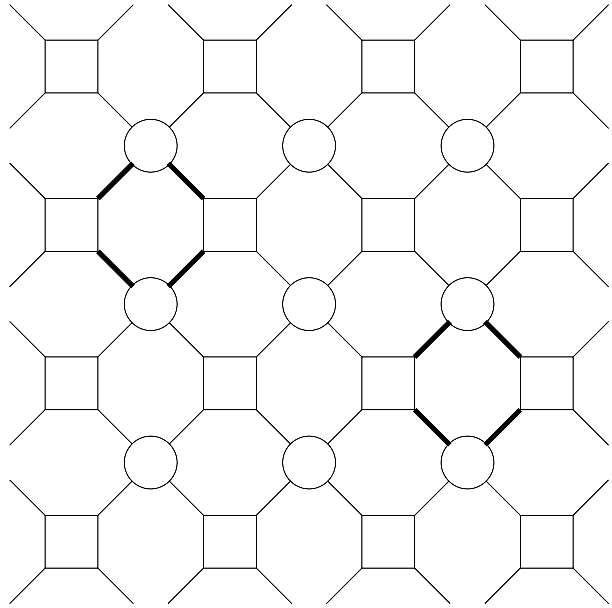

Fig. 2. The channel part of the full graph for a 2×2 ISI. The squares and circles represent the measured data nodes and the codeword bit nodes, respectively. The two-dimensional ISI introduces length 4 cycles in the graph as shown by the bold lines. The check nodes (not shown) are connected to the codeword bits nodes through the parity-check matrix of the LDPC code.

the performance of the multi-track BCJR algorithm when used with LDPC decoding.

Assuming that T tracks are processed simultaneously, the multi-track BCJR algorithm proceeds as follows: at the kth step the strip consisting of rows $k-T+1$ through k are selected and a forward/backward sweep is performed on it. This sweep uses the extrinsic information from the previously decoded rows $k-T-L+2$ through $k-T$ as-well-as the extrinsic information of the LDPC decoder for the rows in the strip. The extrinsic information from the previous rows is used for computation of the branch probabilities. Recall that the extrinsic information from the previous rows is soft, so while calculating the branch probabilities the contribution from the bits in the previous rows is averaged over all the possible combinations (which are $2^L(L-1)$ in this case).

After calculating the probabilities the algorithm moves to the strip consisting of rows $k+1-T+1$ through $k+1$ and performs a forward/backward sweep using extrinsic information from rows $k-T-L+3$ through $k-T+1$ and the extrinsic information of the LDPC decoder. This process is continued till all the rows have been processed and then extrinsic information is passed to the LDPC decoder. The LDPC decoder uses the extrinsic information to initialize its message-passing and after a fixed number of iterations passes its extrinsic information to the BCJR equalizer starting the next iteration. Fig. 3 shows the Wiberg-type graph for the multi-track BCJR algorithm that processes strips of 2 rows for a 2×2 ISI. The figure clearly shows the presence of length 4 cycles in the factor graph, thus the computed posteriors are still only approximate.

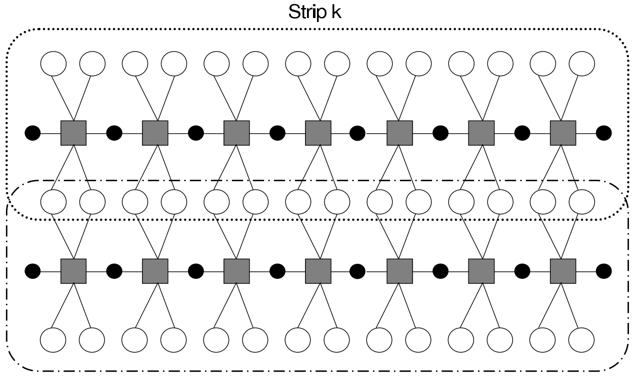

Fig. 3. The Wiberg-type graph for the multi-track BCJR algorithm that processes strips of 2 rows for a 2×2 ISI. The local functions (squares) correspond to the trellis sections and are connected to the state variables (filled circles), the codeword bits and the measured data (both represented as empty circles).

C. MMSE-LDPC Decoding

Tüchler *et al.* [21] proposed soft-in soft-out MMSE based equalizers for one-dimensional ISI channels and showed that when used in an iterative scheme with error-control coding the receiver performs almost as well as a receiver employing the much more complex MAP (BCJR algorithm based) equalizer. Singla and O'Sullivan [12] extended the work of Tüchler *et al.* to two-dimensional ISI channels and showed that the MMSE equalizer based scheme when used with LDPC decoding performs very close to the full graph algorithm.

The MMSE equalizer computes the linear MMSE estimate of the codeword bits using an $N\times N$ support which is usually a little bigger than the size of the ISI, $L\times L$. The estimate for codeword bit $x(i,j)$ is,

$$\hat{x}(i,j) = E[x(i,j)] + \\ + \sum_{l_1,l_2=-N}^{N} \Big(\hat{r}(i-l_1, j-l_2)\Big) c(i,j:l_1,l_2), \quad (3)$$

where $E[\cdot]$ denotes the expectation operator; $\hat{r}(\cdot,\cdot)$ denotes the measured data with its mean subtracted; and $\{c(i,j:l_1,l_2)\}_{l_1,l_2=-N}^{N}$ are the MMSE filter coefficients for bit $x(i,j)$. The filter coefficients and the MMSE estimate are computed using the channel model, the observed data, and the statistics of the codeword bits. These statistics are computed using the extrinsic information provided by the LDPC decoder. Hence the equalizer recalculates its coefficients and the MMSE estimate every iteration. The decoding schedule involves computation of the MMSE estimate using the measured data and the extrinsic information of the LDPC decoder. The equalizer then passes its extrinsic information to the LDPC decoder which uses it to initialize its message-passing and performs a fixed number of iterations before passing its extrinsic information to the equalizer. This completes one iteration of decoding.

IV. Results

In this section we present bit-error rate (BER) versus signal-to-noise ratio (SNR) curves for the three joint equalization and decoding schemes of the previous section. This is followed by a comparison of the computational complexity of the three schemes. For our simulations we use the following 3×3 point spread function:

$$h = \begin{pmatrix} 0.074 & 0.184 & 0.074 \\ 0.184 & 0.918 & 0.184 \\ 0.074 & 0.184 & 0.074 \end{pmatrix}. \quad (4)$$

The SNR is calculated as

$$SNR = 10\cdot\log_{10}\frac{\sum_{l_1,l_2=0}^{L-1} h^2(l_1,l_2)}{2c\sigma_w^2}, \quad (5)$$

where c is the rate of the LDPC code used.

The results after 5, 10, and 20 iterations are shown in Figs. 4, 5, and 6, respectively. The LDPC code used is a block length 10000, regular (3,6) code. The data are arranged in a 100×100 matrix prior to transmission over the channel. The results are compared to the performance of the LDPC code on an AWGN channel; this curve is labeled "LDPC No ISI" in the figures. The multi-track BCJR algorithm considers 3 rows simultaneously.

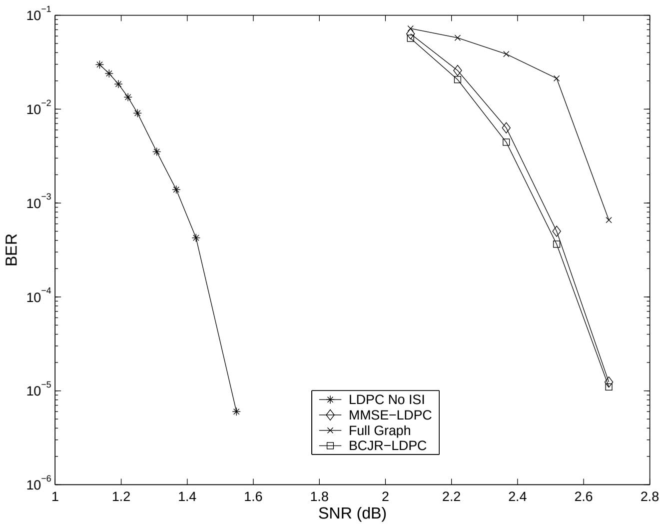

Fig. 4. Comparison of the three schemes at the fifth iteration.

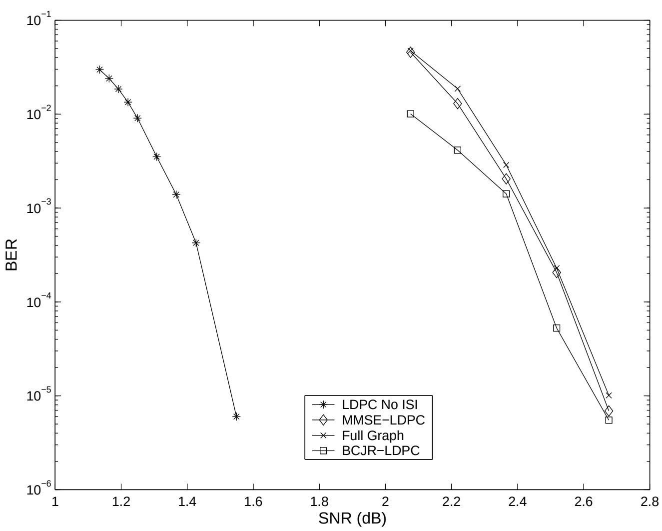

Fig. 5. Comparison of the three schemes at the tenth iteration.

Fig. 4 shows that the full graph algorithm has a poor starting behavior whereas the BCJR-LDPC and the MMSE-LDPC schemes have almost identical performance after 5 iterations. Fig. 5 shows that after 10 iterations the performance of the full graph algorithm improves vastly and is very close to that of the MMSE-LDPC decoder, although still worse. The performance of the BCJR-LDPC decoder is slightly better than that of the MMSE-LDPC decoder. One can also see the onset of the error-floor region in the BCJR-LDPC curve. Fig. 6 shows that after 20 iterations the full graph algorithm has surpassed the MMSE-LDPC algorithm and has almost the same performance as the BCJR-LDPC algorithm. The BCJR-LDPC and full graph schemes use different schedules to perform sum-product message-passing on very loopy graphs representing the same

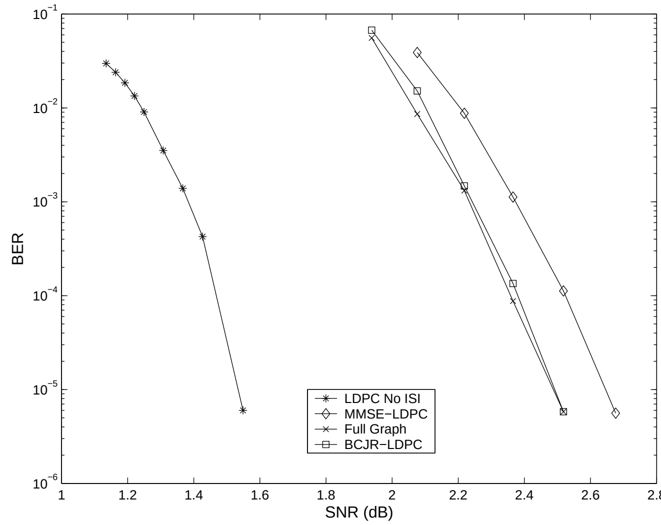

Fig. 6. Comparison of the three schemes at the twentieth iteration.

system hence the performance is expected to be the same. The performance of the full graph algorithm is about 1 dB worse than the performance of the LDPC code. Looking across the figures one can observe that the full graph algorithm performance improves most significantly whereas the performance of the other two schemes improves marginally. This shows that the full graph algorithm converges slower than the other two schemes.

The performance trend can be understood by looking at the complexity of the schemes. The complexity is determined by the equalizer complexity only since the LDPC decoder is common to all the schemes. For the full graph algorithm the complexity is determined by the computation of the messages from the measured data nodes to the codeword bit nodes. This computation involves a sum over all the values assumed by the set of neighboring bits of a particular codeword bit node. For an $L \times L$ ISI each measured data node is connected to L^2 codeword bit nodes and since the data is assumed to be binary, the number of values to be considered are 2^{L^2}. For each combination a product is taken over all the neighboring bits giving an overall complexity of the order of $L^2 2^{L^2}$. In the special case when the value of the measured data node depends only on the sum of the values of the connected codeword bits, for example memoryless channels and the channel model of TwoDOS, the computation of the messages from the measured data to codeword bit nodes can be done efficiently using the BCJR algorithm [18], [11]. In this case the complexity is reduced to L^4. For the multi-track BCJR algorithm that processes strips of T rows, the number of branches in the trellis is 2^{LT}, plus the algorithm also takes into account the extrinsic information from the previous $L-1$ rows, giving an overall complexity of the order of $2^{L(L-1+T)}$. For the MMSE equalizer the complexity is dominated by the computation of the inverse of an $N^2 \times N^2$ matrix (recall that the filter has an $N \times N$ support) thus the complexity is of the order of N^6 [12]. N is chosen so that the support of the filter is of the order of the extent of the ISI, hence the complexity is of the order of L^6.

Comparing the complexities we see that the multi-track BCJR algorithm is the most complex, followed by the full graph algorithm and then the MMSE-LDPC scheme. Preliminary simulations suggest that the performance of the BCJR-LDPC algorithm surpasses that of the full graph algorithm as we consider strips of more than 3 rows, however, the exponential increase in complexity makes the use of more rows very expensive. The MMSE-LDPC scheme has the lowest complexity among the three schemes yet its performance is very close to the message-passing based schemes.

V. Conclusion

We studied three different joint equalization and decoding schemes for two-dimensional ISI channels. Two of the schemes involve performing sum-product message-passing on two different factor graphs that represent the underlying system and the third scheme uses an MMSE equalizer. The factor graph representations considered are a kernel-based and a state-based representation of the channel ISI. Sum-product message-passing on the state-based representation is equivalent to a multi-track version of the BCJR algorithm. The schemes are compared with respect to their BER versus SNR performance and computational complexity. Simulations show that the schemes based on message-passing have nearly the same performance. They achieve the same BER as the LDPC code on an AWGN channel (no ISI) with less than a dB loss in SNR for the point spread function of (4). The MMSE-LDPC algorithm has a much lower complexity than the message-passing based algorithms yet it has comparable performance. Preliminary results suggest that using strips of more than 3 rows in the multi-track BCJR algorithm improves the performance even further. However, even then the performance is worse than that of the LDPC code. It is clear that message-passing algorithms are the way to go in order to ultimately achieve the capacity on ISI channels with bounded computational complexity. What is not clear, however, is whether this requires considering alternate graphical representations of the system or alternate message-passing algorithms or perhaps a combination of the two.

VI. Acknowledgment

The authors would like to thank Jiening Zhan for help with the simulations and the reviewers for very helpful suggestions and comments. This work was supported by the Office of Naval Research under Award N00014-03-1-0110.

References

[1] E. Ordentlich and R. M. Roth, "On the computational complexity of 2D maximum-likelihood sequence detection," *Info. Theory and Applications Workshop*, San Diego, California, Feb. 2006.

[2] K. M. Chugg, X. Chen, and M. A. Neifeld, "Two-dimensional equalization in coherent and incoherent page-oriented optical memory," *J. Opt. Soc. Amer. A*, vol. 16, pp. 549-562, Mar. 1999.

[3] K. M. Chugg, X. Chen, and M. A. Neifeld, "Two-dimensional linear MMSE for page-oriented optical memories," *31st Annual Asilomar Conf. Sig., Sys. and Comps.*, vol. 1, pp. 343-347, Nov. 1997.

[4] A. H. J. Immink *et al.*, "Signal processing and coding for two-dimensional optical storage," *GlobeComm.*, Piscataway, New Jersey, 2003.

[5] R. Krishnamoorthi, "Two-dimensional Viterbi like algorithms," M.S. thesis, University of Illinois at Urbana-Champaign, 1998.

[6] P. S. Kumar and S. Roy, "Two-dimensional equalization: theory and application to high density magnetic recording," *IEEE Trans. Comm.*, vol. 42, pp. 386-395, Feb. 1994.

[7] M. Marrow and J. K. Wolf, "Iterative detection of 2-dimensional channels," *IEEE Info. Theory Workshop*, Paris, France, Mar. 2003.

[8] A. Moinian, L. Fagoonee, and B. Honary, "Symbol detection for multilevel two-dimensional optical storage using a non-linear channel model," *Intl. Conf. Comm.*, Seoul, South Korea, May 2005.

[9] O. Shental, A. J. Weiss, N. Shental, and Y. Weiss, "Generalized belief propagation receiver for near-optimal detection of two-dimensional channels with memory," *IEEE Info. Theory Workshop*, San Antonio, Texas, Oct. 2004.

[10] N. Singla, J. A. O'Sullivan, R. S. Indeck, and Y. Wu, "Iterative decoding and equalization for 2-D recording channels," *IEEE Trans. Magn.*, vol. 38, pp. 2328-2330, Sept. 2002.

[11] N. Singla and J. A. O'Sullivan "Joint equalization and decoding for nonlinear two-dimensional intersymbol interference channels," *IEEE Intl. Symp. Info. Theory*, Adelaide, Australia, Sept. 2005.

[12] N. Singla and J. A. O'Sullivan, "Minimum mean square error equalization using priors for two-dimensional intersymbol interference," *IEEE Intl. Symp. Inform. Theory*, Chicago, USA, June-July 2004.

[13] Y. Wu, J. A. O'Sullivan, R. S. Indeck, and N. Singla. "Iterative detection and decoding for separable two-dimensional intersymbol interference," *IEEE Trans. Magn.*, vol. 39, pp. 2115-2120, July 2003.

[14] W. Weeks IV, "Full surface data storage," Ph.D. thesis, University of Illinois at Urbana-Champaign, 2000.

[15] Z. Zhao and R. Blahut, "The Richardson-Lucy algorithm based demodulation algorithms for the two-dimensional intersymbol interference channel," *Proc. Conf. on Info. Sciences and Systems*, Baltimore, Maryland, Mar. 2005.

[16] F. R. Kschischang, B. J. Frey, and H. A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, pp. 498-519, Feb. 2001.

[17] A. Kavčić, X. Ma, and M. Mitzenmacher, "Binary intersymbol interference channels: Gallager codes, density evolution and code performance bounds," *IEEE Trans. Inform. Theory*, vol. 49, pp. 1636-1652, July 2003.

[18] D. J. C. Mackay, "Good error-correcting codes based on very sparse matrices," *IEEE Trans. Inform. Theory*, vol. 45, pp. 399-431, March 1999.

[19] T. Richardson and R. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599-618, Feb. 2001.

[20] N. Wiberg, "Codes and decoding on general graphs," Ph.D. dissertation, Linköping Univ., Linköping, Sweden, 1996.

[21] M. Tüchler, R. Koetter, and A. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Comm.*, vol. 50, pp. 754-767, May 2002.