Density Evolution on a Class of Smeared Random Graphs

Kabir Chandrasekher, Orhan Ocal, and Kannan Ramchandran
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley
{kabirc, ocal, kannanr}@berkeley.edu

***Abstract*—We introduce a new ensemble of random bipartite graphs which we term as the 'smearing ensemble', where each left node is connected to some number of consecutive right nodes. These types of graphs arises in inverse problems related to magnetic resonance imaging (MRI), neighbor discovery in the presence of multipath, and in sparse spectrum recovery off-the-grid. By design, these graphs have short cycles of known nature with high probability. In this paper, we develop a theoretical platform to analyze and evaluate the power of smearing-based structure in a class of random graphs. In particular, we give exact density evolution recurrences for smearing of order 2, and give lower bounds on the performance of a much larger class from the smearing ensemble. For these, we give simulation results showing tight agreement between thresholds determined by our bounds and empirical thresholds. Finally, we detail the use of our algorithm in recovering the sparse wavelet coefficients when signal acquisition is in the Fourier domain, which is the case with MRI.**

I. Introduction

Consider the following problem where you have n distinct colors, K of which are active. Additionally, for each active color, you have d balls. There is also a set of M bins into which you throw the dK balls. The rules of the game are as follows:

R1) Each of the d balls of an 'active' color must go to a different bin[1],
R2) A bin which contains a single ball 'releases' all the other $d-1$ balls of the same color,
R3) Process continues iteratively until either (a) all balls have been released or (b) no further single-occupied bins exist.

The goal of the game is to release all active balls using the minimum number of bins. We focus on the regime in which $(n, K, M) \to \infty$, $d = \mathcal{O}(1)$ and ask the following questions:

1) What is the optimal design policy of where to dispatch the d balls?
2) Given (n, K, d), what is the minimum number of bins necessary[2]?

An interesting problem in its own right, it has connections to the design of Low Density Parity Check (LDPC) codes. Surprisingly, it is also intimately related to the recovery of sparse wavelet representations from Fourier domain samples (see Section III).

[1]i.e. each bin may contain only one ball of each color

[2]We note here the importance of the parameter n. If the set of K colors were known a priori, it would be possible to use a deterministic strategy using $K + d$ bins.

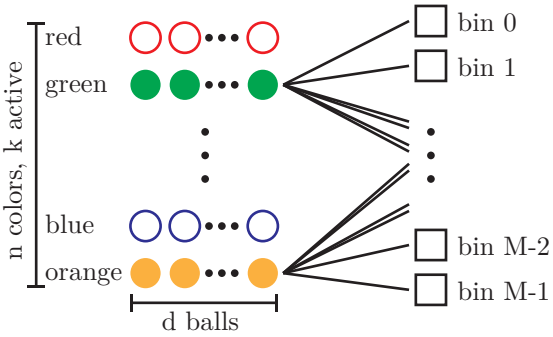

Fig. 1: We have n colors and d balls of each color. The balls are thrown into M bins, and k out of n colors become active (shown as filled circles). Using the rules R1-R3, we want to find the active colors.

To illustrate, say that $d = 3$ and the strategy that requires the smallest number of bins is to throw each ball at a bin selected uniformly at random. It has been demonstrated that we need asymptotically $M \simeq 1.222K$ bins as K grows. This can be shown through density evolution methods, introduced by Richardson and Urbanke in [1], which have proven powerful in analyzing the performance of LDPC codes.

Now, say that $d = 6$, and we throw each ball at a bin selected uniformly at random as before. Then, we require $M \simeq 1.570K$ asymptotically in K. However, interestingly, by changing how the balls of each color are thrown, one can do better than uniformly at random throws. For each color, throw three balls uniformly at random, and for the remaining 3 balls at the bins following those first three. This last method requires $M \simeq 1.489K$ outperforming the uniformly at random setting, showing that some 'structure' in throwing the balls can help reduce the required number of bins. In this paper, we provide a theoretical platform to analyze and evaluate such policies exploiting structure. The effects of structure in different regimes can be seen in Table 1.

Table 1: Pairs of regimes and thresholds (M/K) for $d = 6$. The regimes are s vectors as per Definition 1.

Regime	1,1,1,1,1,1	1,1,1,1,2	1,1,1,3	1,1,4	1,1,2,2	1,2,3	2,2,2
M/K	1.570	1.533	1.489	1.518	1.533	1.542	1.547

A. Related Works

Density evolution methods have proven powerful in analyzing the performance of LDPC codes and their extensions [1], [2]. Unfortunately, the well-known density evolution equations apply only for sparse random graphs that are locally tree-like. This is not the case for all ball-throwing strategies in the game we outlined above, e.g. the $[2, 2, 2]$ scheme. Recently Donoho et. al. have introduced approximate message passing (AMP) techniques to extend the message passing paradigm to the case when the underlying factor graph is dense [3]. These techniques were rigorously analyzed by Bayati and Montanari in [4]. Although AMP has been successfully applied to many problem domains, e.g., [5], [6], it imposes a dense structure on the factor graph. However, if the bipartite graph is sparse, but contains cycles, it may not be necessary to invoke such methods. Recently, Kudekar et. al have been able to show the benefit of structure in convolutional LDPC codes through the spatial coupling effect [7], [8]. In this paper, we show how to extend the density evolution methods of [1] to a certain class of smeared random graphs which contain structured cycles.

Our observations and motivation have come from an extension of the recently proposed FFAST (Fast Fourier Aliasing-based Sparse Transform) algorithm [9] to the case where sparsity is with respect to some wavelet basis. We note that in MRI, images are observed to be sparse with respect to a wavelet basis, but acquisition is in the Fourier domain [10]. The computational bottleneck has been observed to be the multiple FFT's taken. We extend the FFAST algorithm to address this problem.

B. Main Results and Paper Organization

The main result of this paper is the derivation of exact thresholds for random graphs with maximum smearing of order 2 and bounds for higher order smearing which are empirically shown to be very tight. We additionally detail an application to fast recovery of sparse wavelet representations when acquisition is in the Fourier domain.

The organization of this paper is as follows. In Section II-A, we derive sharp thresholds for the specific case of 2-smearing. In Section II-B, we derive lower bounds for the probability of recovery in the case of arbitrary smearing. In Section III, we outline the connection between the ball coloring game and the recovery of sparse wavelet representations. We summarize some interesting open problems and conjectures that have resulted from this work in IV.

II. Main Results

In this section, we introduce a new random graph ensemble, termed the 'smearing ensemble', and show how to derive density evolution recursions for graphs from this ensemble. We detail the derivation of density evolution for smearing of order 2. Additionally, we give lower bounds for smearing of order k.

We now formally define the smearing ensemble.

Definition 1. Let $\mathcal{G}(K, M, s)$ denote the 'smearing graph ensemble' with K left nodes and M right nodes. The vector $s = [s_1, s_2, \ldots, s_i]$ characterizes the connections between left and right nodes according to the following the algorithm:

```
for each left node:
    for j in [1, ..., i]:
        connect left node to a right node
        chosen uniformly at random and its
        s_j-1 following neighbors
```

Denote the ratio $\lambda := \frac{K}{M}$. Now, note that if K balls are thrown into the M bins randomly, the degree of each bin will be distributed like Poisson(λ). It is clear that under smearing, bins in stage i will have degree distributed like Poisson$(s_i\lambda)$, as the degree of each bin is the sum of s_i independent 'streams' of balls.

We begin by carefully deriving the density evolution recurrence for smearing ensembles with the maximum smearing of order 2.[3] Finally, we give lower bounds for the general smearing case.

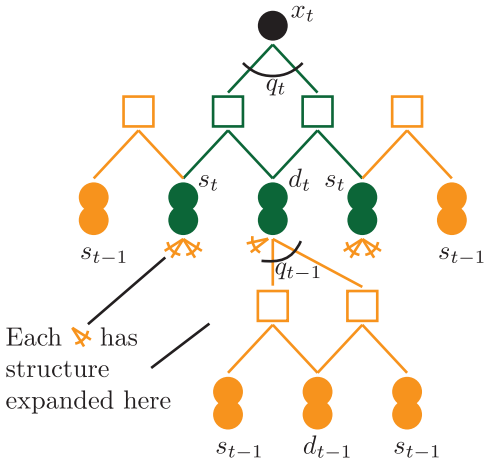

Fig. 2: Depth 2 neighborhood of a ball in a graph with $[2, 2, 2]$ smearing. The bins and balls are grouped by their colors with respect to distance from the root.

A. Density Evolution on 2-smear Graphs ($s = [2, 2, 2]$)

In our ball-coloring game, we noted that threshold for the setting $[2, 2, 2]$ outperforms that for the setting $[1, 1, 1, 1, 1, 1]$. In this section, we give exact analysis for these thresholds. First, we define our notation in Table 1.

Table 2: Notation for density evolution of order 2

x_t :	Probability that a random node is *not* removed at iteration t
q_t :	Probability that neither bin in a smeared pair is removed at iteration t
d_t :	Probability that *all* nodes in the same stream as the node of reference are removed at time t
s_t :	Probability that *all* nodes in a stream which intersects, but does not fully overlap, with the reference stream are removed at time t

[3] We note that the full version of this work, explicit derivation of thresholds for smearing ensembles of order 3 is given, but for the sake of brevity it is omitted here.

Unlike conventional density evolution, note that we are taking a node perspective. Our goal is to derive recurrences for each of these quantities. In this case, the equations for x_t and q_t are clear:

$$x_t = q_t^3 \tag{1}$$

$$q_t = 1 - d_t(1 - (1 - s_t)^2) \tag{2}$$

In order to see this, note that for a node to be removed, at least one of the bins to which it is connected must be removed. Thus, x_t only occurs when none of the pairs of smeared bins to which it is connected are removed. On the other hand, for a bin to be removed, *all* of its connected nodes must be removed. Thus, all of the balls in the same stream as the reference ball must be removed, which happens with probability d_t. Additionally, at least one of the two streams that do not fully overlap with the reference ball must be removed, which happens with probability $(1-(1-s_t)^2)$. Now, note that in Fig. 2, the recurring structures are those with probabilities s_t, d_t, q_t. We first calculate d_t. Each of the balls in this stream act independently of each other, and each of the balls in this stream must be removed in another stage. Letting D_1 denote the number of balls in this stream, we find:

$$\begin{aligned} d_t &= \sum_{i=0}^{\infty} P(D_1 = i)(1 - q_{t-1}^2)^i \\ &= \sum_{i=0}^{\infty} e^{-\lambda} \frac{\lambda^i}{i!}(1 - q_{t-1}^2)^i \\ &= e^{-\lambda} \sum_{i=0}^{\infty} \frac{(\lambda(1 - q_{t-1}^2))^i}{i!} \\ &= e^{-\lambda q_{t-1}^2} \end{aligned} \tag{3}$$

Now, we tackle s_t. Note that intuitively, $s_t \geq d_t$. This is because each ball in a stream tracked by s_t gets the same independent help from other stages. However, there is additionally a shared bin between all these balls. In many cases, this bin will not help; however, it can provide an extra boost. This shared bin is able to aid in the removal of the stream tracked by s_t when exactly one ball is left in the stream, and the bin has no contributions from elsewhere. Letting D_2 be the number of balls in this stream, we precisely characterize this as follows:

$$\begin{aligned} s_t &= \sum_{i=0}^{\infty} P(D_2 = i) \cdot \left[(1 - q_{t-1}^2)^i + i q_{t-1}^2 (1 - q_{t-2}^2)^{i-1}\right] \\ &= e^{-\lambda q_{t-1}^2} + \lambda s_{t-1} q_{t-1}^2 e^{-\lambda q_{t-2}^2} \end{aligned} \tag{4}$$

Note that the first term in this recursion is exactly d_t. The second term describes the positive effect of the shared bin. In order to properly characterize this term, it is necessary to introduce the notion of memory: the shared bin must have all contributions removed except for one by time t. Unlike when the neighborhood is tree-like and contains no cycles, branches of the tree become independent [11], when cycles are introduced, dependence between branches is introduced. The introduced memory captures exactly this dependence. This style of thinking leads to the principles of generalization listed in Table 3.

Table 3: Principles of generalization

1)	In a stage with L smearing, there will be $L-1$ steps of memory necessary in order to properly capture the smearing structure
2)	In a stage with L smearing, the number of recurring structures will be L
3)	Shared bins can be used through the introduction of memory in the recursion

In the following section, we will use these principles to derive simple, but effective, lower bounds for L-smearing.

We summarize the results of this section with the following lemma:

Lemma 1. *Consider a random graph from the ensemble $G(K, M, [2, 2, 2])$. Then, for $M \geq 1.547$, recovery using the peeling decoder will succeed with high probability.*

Proof. The threshold 1.315 follows from the density evolution derived above. □

We note that the threshold induced by these recursions are in tight agreement with simulations. In the extended version of this work [12], we derive explicit recursions for the ensembles with $s = [1, 1, 3]$ (smearing of order 3), but omit it here for the sake of clarity. Following the above principles of generalization, we now derive lower bounds on the density evolution for the general smearing ensemble.

B. Lower Bound on L-smearing

For simplicity, we will consider the ensemble $[L, L, L]$. Before we begin, we will clarify the notation to be used in this section (note that x_t, q_t, d_t remain as in the previous section):

Table 4: Notation for density evolution of order L

$s_t^{(i)}$:	Probability that *all* nodes in the streams which *does not* intersect with the reference stream in j bins where $1 \leq j \leq i$ bins are removed at time t

Following the principles of generalization outlined in the previous section, we know that there will be L recursions (this is shown to be true in Lemma 2), and up to $L-1$ memory involved in these recursions. It is clear that the probability d_t is unaffected by such memory, as it depends only on the other stages. All of the $s_t^{(i)}$, however, can have up to i memory (corresponding to the number of bins that do not overlap with the reference ball). Thus, we take an approach much like a first-order approximation of a Taylor series, and allow each $s_t^{(i)}$ to use only one step of memory. It is clear that approximating in this manner is a lower bound to using the full power of memory to the decoder's advantage. The following lemma characterizes the critical quantity q_t in terms of its component streams.

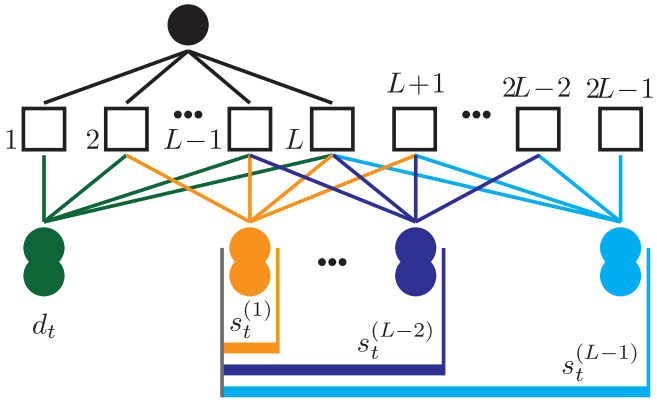

Fig. 3: Depth 1 neighborhood of a ball under L-smearing. The quantities d_t and $s_t^{(i)}$ are the recurrent structures, where $s_t^{(i)}$ tracks the joint probability that all the streams which intersect with the reference stream in j bins, where $i \leq j \leq L-1$, are removed at time t.

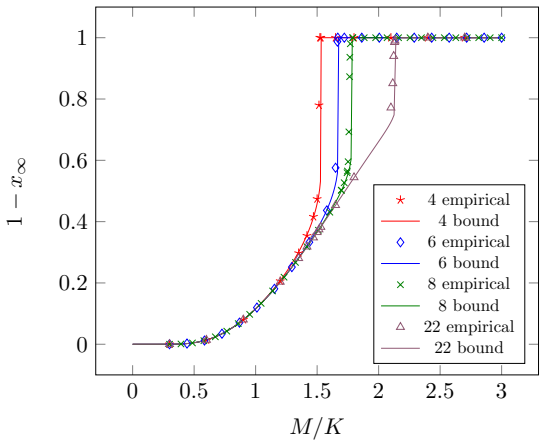

Fig. 4: We plot our lower bounds against empirical simulations with settings $s = [1, 1, L]$ for various L. The y-axis here denotes the probability that a random ball is removed (i.e. $1 - x_t$ as $t \to \infty$). We note that this ensemble is used as it is of the most relevant practical interest to recovery of sparse wavelet representations.

Lemma 2. *In a stage with L-smearing,*

$$1 - q_t = d_t \cdot \left[2s_t^{(L-1)} + \sum_{i=2}^{L-1} s_t^{(i-1)} s_t^{(L-i)} - \sum_{i=1}^{L-1} s_t^{(i)} s_t^{(L-1)} \right]$$

Proof. See the appendix in the extended version [12]. □

The following lemma then establishes monotonicity of $1-q_t$ with respect to $d_t, s_t^{(i)}$, which we may use to complete the bound.

Lemma 3. *Let $f(d_t, s_t^{(1)}, \ldots s_t^{(L-1)}) = 1 - q_t$, then $f(d_t, s_t^{(1)}, \ldots s_t^{(L-1)})$ is non-decreasing in $(d_t, s_t^{(1)}, \ldots, s_t^{(L-1)})$.*

Proof. See the appendix in the extended version [12]. □

Given the above characterization of q_t, it now remains to give lower bounds for $d_t, s_t^{(1)}, \ldots, s_t^{(L-1)}$. We can bound these as follows:

$$d_t = e^{-\lambda q_{t-1}^2} \tag{5}$$

$$s_t^{(i)} \geq e^{i\lambda q_{t-1}^2} + \lambda q_{t-1}^2 e^{-\lambda q_{t-2}^2} r_{t-1}^{(i)}, \tag{6}$$

for $i \in \{1, \ldots, L-1\}$, where

$$r_t^{(i)} \stackrel{\text{def}}{=} \sum_{j=1}^{i} \sum_{k=1}^{j} s_{t-1}^{(k)} e^{-(L-k-1)\lambda q_{t-2}^2} e^{-(k-1)\lambda q_{t-1}^2}. \tag{7}$$

There is a simple way to think about the problem so that these bounds appear. Consider the bound on $s_t^{(i)}$. This tracks the joint probability that all the streams which intersect the reference stream in $L-j$ bins where $1 \leq j \leq i$ are removed at time t. In order for all of these streams to be removed, there are two cases:

1) Each stream was removed from another stage. This probability is tracked by the first term: $e^{i\lambda q_{t-1}^2}$.
2) Exactly one ball remains among all the streams. This probability is tracked by the second term.

We focus on the second case. Suppose that the remaining ball is in the stream which intersects the reference stream in j bins. This means that it is also contained in $L-j$ shared bins that do not intersect the reference stream. It can be removed by any of these shared bins, as long as it is the only contribution to that bin. This help is tracked by the summation in the second term.

Corollary 1. *The lower bounds given in equations* (5) *and* (6) *imply a lower bound on x_t, the probability that a random node is removed at time t.*

Proof. This follows directly from Lemma 2 and Lemma 3. □

We have derived these density evolution recursions using the assumption of a locally pseudo-tree-like neighborhood, excepting for the high probability cycles created by the smearing ensemble. We now give numerical experiments to show that the bounds stated in Corollary 1 captures the actual thresholds well. Towards this end, we experimentally find the thresholds for full recovery by sweeping λ and compare it to the thresholds implied by the bounds. Fig. 4 shows the bound for 1-depth of some filters (relevant to practical regimes of interest).

III. Applications to Wavelet Transform

We now present how the game of balls and bins described in the introduction relates to MRI. In MRI, one acquires samples from the Fourier transform of an input signal, which is then inverted to recover the original signal. Mathematically, let x be an n length signal, and $X \stackrel{\text{def}}{=} F_n x$ be its Fourier transform, where F_n is the Fourier matrix of size $n \times n$. In MRI, the problem is to recover x from $\{X_i : i \in \mathcal{I}\}$ where the set $\mathcal{I}$ denotes the sampling locations, and this set is a design parameter. If the signal x is K-sparse, one can use

the FFAST algorithm to recover x from $O(k)$ samples with $O(k \log k)$ computations [9], [13]. However, the images of interest in MRI are generally not sparse, but they do have sparse wavelet representations [10]. That is, we can express $x = W_n^{-1}\alpha$, where W_n^{-1} is an appropriate wavelet, and α is sparse. Under this signal model, the problem of recovering α can be transformed into the problem decoding on an erasure channel a sparse-graph code. In particular, the graph for the code is drawn from a smearing ensemble with smearing length L a function of the length of the underlying wavelet filter.

Furthermore, assume α is K-sparse and the length of x is of the form $n = f_1 f_2 f_3$ where f_1, f_2 and f_3 are co-prime. For $m \in \mathbb{Z}$ that divides n, let $D_{m,n} \in \{0,1\}^{m \times n}$ be the regular downsampling matrix from length n to m, that is, $D_{i,j}$ is 1 if $j = ni/m$ or 0 otherwise. Let y_{f_ℓ} for $\ell \in \{1,2,3\}$ be the inverse Fourier transform of the downsampled X, i.e.,

$$y_{f_\ell} \stackrel{\text{def}}{=} F_{f_\ell}^{-1} D_{f_\ell,n} X.$$

Using the properties of Fourier Transform, it follows that

$$y_{f_\ell} = A_{f_\ell,n} x = A_{f_\ell,n} W_n^{-1} \alpha,$$

where $A_{m,n}$ is the aliasing matrix, i.e., $A_{m,n} = [I_m \cdots I_m]$ with I_m is repeated n/m times,

Now, for simplicity, assume that W is a block transform with block size b (eg., for 1 stage Haar wavelets $b = 2$), and the non-zero coefficients of α are distributed uniformly at random. Using the relations between y_{f_1}, y_{f_2} and y_{f_3} and α, recovering α is equivalent to decoding on a random graph from $G(K, M = f_1 + f_2 + f_2, s = [L, L, L])$.

We can actually 'improve' the induced graph if a factor of the signal length has b as a factor. In particular, say that b divides f_1. In this case, it follows that $A_{f_1,n} W_n^{-1} = [W_{f_1}^{-1} \cdots W_{f_1}^{-1}]$, where $W_{f_1}^{-1}$ is repeated n/f_1 times. It can be checked that $y'_{f_1} \stackrel{\text{def}}{=} W_{f_1} y_{f_1} = A_{f_1,n} a$, hence it only aliases the wavelet coefficients. The relation between y'_{f_1}, y_{f_2} and y_{f_3} and α then induces a graph from $G(K, M = f_1 + f_2 + f_2, s = [1, L, L])$, which gives raise to a better threshold.

To complete the equivalence to decoding a sparse-graph code on an erasure channel, we need to be able to check if there is a single component in a bin (a single color in a bin). This can be implemented by processing a shifted version of x. Due to space limitations we defer the detailed explanation to the extended version [12].

IV. Conclusions and Future Work

We have introduced the new random graph ensemble, termed the 'smearing ensemble' and derived density evolution recurrences for certain cases from this ensemble. Looking back to our ball coloring game, these recurrences allow us to explicitly analyze strategies involving smearing bounded by order 3. Intriguingly, these results show that some amount of smearing helps to achieve an optimal ball-throwing strategy. A fascinating open question arises here: what is the optimal ball-throwing strategy and what are the density evolution recurrences for such a strategy? In this paper, we have given the first steps in analyzing this problem rigorously and have shown bounds on the smearing ensemble. To do this, we have leveraged the existence of small, structured cycles and introduced the notion of memory into our density evolution. We believe there to be a deep connection between the introduction of memory in our recurrences and the introduction of the 'Onsager' term in the update equations of AMP [14]. We additionally believe the gains seen in spatially coupled ensembles [8] are intimately related to the structural gains of the smearing ensemble. An extremely interesting open problem is to determine the nature of these connections. We have additionally shown the practical connection between the smearing ensemble and the recovery of a sparse wavelet representation of a signal whose samples are taken in the Fourier domain.

References

[1] T. J. Richardson and R. L. Urbanke, "The capacity of low-density parity-check codes under message-passing decoding," *IEEE Trans. Inf. Theory*, vol. 47, no. 2, pp. 599–618, 2001.

[2] T. Richardson and R. Urbanke, *Modern coding theory*. Cambridge University Press, 2008.

[3] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Natl. Acad. Sci. USA*, vol. 106, no. 45, pp. 18 914–18 919, 2009.

[4] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, no. 2, pp. 764–785, 2011.

[5] J. Tan, Y. Ma, and D. Baron, "Compressive imaging via approximate message passing with image denoising," *IEEE Trans. Sig. Process.*, vol. 63, no. 8, pp. 2085–2092, 2015.

[6] A. Maleki, L. Anitori, Z. Yang, and R. G. Baraniuk, "Asymptotic analysis of complex lasso via complex approximate message passing (camp)," *IEEE Trans. Inf. Theory*, vol. 59, no. 7, pp. 4290–4308, 2013.

[7] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Threshold saturation via spatial coupling: Why convolutional ldpc ensembles perform so well over the bec," *IEEE Transactions on Information Theory*, vol. 57, no. 2, pp. 803–834, 2011.

[8] S. Kudekar, T. Richardson, and R. L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Transactions on Information Theory*, vol. 59, no. 12, pp. 7761–7813, 2013.

[9] S. Pawar and K. Ramchandran, "Computing a k-sparse n-length discrete fourier transform using at most 4k samples and O(klogk) complexity," in *IEEE Int. Symp. on Inf. Theory*, 2013.

[10] M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly, "Compressed sensing mri," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 72–82, 2008.

[11] C. Häger, H. D. Pfister, F. Brännström *et al.*, "Density evolution for deterministic generalized product codes on the binary erasure channel," *arXiv preprint arXiv:1512.00433*, 2015.

[12] K. Chandrasekher, O. Ocal, and K. Ramchandran, "Density evolution on a class of random graphs with cycles (extended)," 2017.

[13] F. Ong, S. Pawar, and K. Ramchandran, "Fast and efficient sparse 2d discrete fourier transform using sparse-graph codes," *arXiv preprint arXiv:1509.05849*, 2015.

[14] D. L. Donoho, A. Maleki, and A. Montanari, "Message passing algorithms for compressed sensing: I. motivation and construction," in *IEEE Inf. Theory Workshop on Inf. Theory*, 2010.