Abstract—We introduce a new ensemble of random bipartite graphs, which we term the ‘smearing ensemble’, where each left node is connected to some number of consecutive right nodes. Such graphs arise naturally in recovering sparse wavelet coefficients when signal acquisition is in the Fourier domain, such as in magnetic resonance imaging (MRI). Graphs from this ensemble exhibit small, structured cycles with high probability, rendering current techniques for determining iterative decoding thresholds inapplicable. In this paper, we develop a theoretical platform to analyze and evaluate the power of smearing-based structure. Despite the existence of these small cycles, we derive exact density evolution recurrences for iterative decoding on graphs with smear-length two. Furthermore, we give lower bounds on the performance of a much larger class from the smearing ensemble. Such graphs arise naturally in recovering sparse wavelet representations from Fourier domain samples asymptotically using $O(K)$ Fourier domain samples asymptotically using $O(K \log K)$ operations.

I. INTRODUCTION

We explain our problem through an intriguing balls-and-bins game. There are $n$ distinct colors, $d$ balls of each color and $M$ bins. You know beforehand that only $K \ll n$ of the colors, which are selected uniformly at random from the $n$ possible colors, will be ‘active’, but you do not know which ones they are. You have to throw all the $(dn)$ balls into the $(M)$ bins. The rules of the game are as follows:

R1) For each color $c$, you choose a subset (possibly using a randomized strategy) $B_c \subset \{0, \ldots, M - 1\}$ of size $d$. Then, the system throws the balls of that color $c$ into bins $\{b + b_c : b \in B_c\}$ modulo $M$ where $b_c$ is sampled uniformly at random from $\{0, \ldots, M - 1\}$.1.

R2) If a bin contains a single active ball, then all $d$ balls having the same color as that ball can be removed.

R3) The process continues iteratively until either (a) all active balls have been removed or (b) there is no bin having a single active ball.

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

1) What is the optimal design policy of where to dispatch the $d$ balls? (That is, what is the optimal strategy for designing the subsets in rule R1?)

2) Given $(n, K, d)$, what is the minimum number of bins $(M)$ necessary?

While this is an intriguing game in its own right, more importantly, it has connections to the design of Low Density Parity Check (LDPC) codes and peeling decoding. Surprisingly, and more relevant here, it is also intimately related to the recovery of sparse wavelet representations from Fourier domain samples (see Section III).

To illustrate, suppose that $d = 3$, then the best known strategy 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 sparse-graph codes. Now suppose that $d = 6$. The natural strategy is to again throw each ball at a bin selected uniformly at random. Surprisingly, this strategy is not optimal. To see this, we give a brief introduction to the smearing ensemble. Consider a $g$ dimensional vector $s = [s_1, s_2, \ldots, s_g]$ where $\sum_{i=1}^{g} s_i = d$ that describes a partitioning of $d$ balls into $g$ groups. The ensemble is such that $g$ bins are selected at random, and for the $i$th bin, the immediately following $s_i - 1$ bins are deterministically selected2; this is what we term the smearing ensemble. Fig. 1 shows an example illustration for $s = [2, 2]$, and we formally define the smearing ensemble in Section II. Examining Table 1, one can see that many simple smearing strategies outperform the fully random ensemble. In this paper, we do not claim to design an optimal strategy for this game; rather, we provide a theoretical platform to analyze these structure-exploiting policies.

Table 1: Thresholds $(M/K)$ for $d = 6$. Note that this table contains a strict subset of all possible strategies.

| Regime | 1,1,1,1,1,1 | 1,1,1,1,3 | 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 |

Our balls-and-bins game is motivated by an extension of the recently proposed FFAST (Fast Fourier Aliasing-based Sparse Transform) algorithm [2] to the case where sparsity is with

This work was supported in part by NSF grant 1527767.

1The only effect of this is to randomly offset the bins for each color.

2For example, if $s_1 = 2$ (we henceforth refer to this as the smear-length), then bins $b_i$ and $b_i + 1$ (modulo $M$) are selected, where $b_i$ is chosen uniformly at random on $\{0, 1, \ldots, M - 1\}$. See Fig. 1 for an illustration.
A. Related Works

Density evolution methods have proven powerful in analyzing the performance of LDPC codes and their extensions [1], [5]. 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 [6]. These techniques were rigorously analyzed by Bayati and Montanari in [7]. Although AMP has been successfully applied to many problem domains, e.g., [8], it imposes a dense structure on the factor graph. Additionally, Kudekar et. al have been able to show the benefit of structure in convolutional LDPC codes through the spatial coupling effect [9], [10]. However, if the bipartite graph is sparse, but contains small, structured cycles, it may not be necessary to invoke such methods.

B. Main Results and Organization of the Paper

The main results of this paper are the derivation of exact thresholds for random graphs with smear-length 2, and bounds for higher smear-length 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. In particular, given a signal with ambient dimension $n$ which is $K$-sparse with respect to the 1-stage Haar wavelet, our analysis shows that $2.63K$ Fourier domain samples are needed to recover the signal in $O(K \log K)$ time.

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. We outline the connection between the ball coloring game and the recovery of sparse wavelet representations in Section III. We conclude with Section IV by summarizing some interesting open problems and conjectures that have resulted from this work in.

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 with smear-length 2, and we give lower bounds for smear-length $L$.

We now formally define the smearing ensemble.

**Definition 1.** Let $G(K, M, s)$ denote the ‘smearing graph’ with $K$ left nodes and $M$ right nodes. The vector $s = [s_1, s_2, \ldots, s_g]$ characterizes the connections between left and right nodes according to the following the algorithm:

for each left node:

for $j$ in $[1, \ldots, g]$:

connect left node to a right node chosen uniformly at random and its $s_{j-1}$ following neighbors

We begin by deriving the density evolution recurrence for smearing ensembles with the maximum smearing length of 2. Then, we provide lower bounds for the general smearing case.

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 $[1, 1, 1, 1, 1]$. In this section, we give exact analysis for these thresholds.

First, we define our notation in Table 2 and illustrate it on the neighborhood of a reference node in Fig. 2.

**Table 2: Notation for density evolution with smear-length 2**

| Symbol | Description |
|--------|-------------|
| $x_t$ | 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 $t$ |

Unlike as in conventional density evolution methods in the LDPC literature which are based on an edge perspective, we

\[ ^3 \text{We note that the full version of this work, explicit derivation of thresholds for smearing ensembles with smear-length 3 is given, but for the sake of brevity it is omitted here.} \]

\[ ^4 \text{Although thresholds can be derived for the other elements in the table, we give the [2, 2, 2] derivation for clarity.} \]
take a node perspective here because dependencies between edges in the smearing setting complicate the analysis. Our goal is to derive recurrences for the quantities given in Table 2. In this case, the equations for \( x_t \) and \( q_t \) are clear:

\[
x_t = q_t^3, \\
q_t = 1 - d_t(1 - (1 - s_t)^2).
\]

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 the stream must be removed in another stage. Let \( D_1 \) denote the number of balls in this stream other than the reference ball we are looking at. It follows that \( D_1 \) is distributed as Binomial\( \left(3(K - 1), 1/M\right) \), which can be approximated well for large \( K \) and \( M \) by Poisson\( (\lambda) \), where \( \lambda = 3K/M \). It follows that:

\[
d_t = \sum_{i=0}^{\infty} P(D_1 = i)(1 - q_{t-1}^2)^i = e^{-\lambda} \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} (1 - q_{t-1}^2)^i = e^{-\lambda q_{t-1}^2}. \tag{3}
\]

Now, we tackle \( s_t \). Note that intuitively, \( s_t \geq d_t \) 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. 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:

\[
s_t = \sum_{i=0}^{\infty} P(D_2 = i)(1 - q_{t-1}^2)^i + is_{t-1}q_{t-1}^2(1 - q_{t-2}^2)^{t-1} = e^{-\lambda q_{t-1}^2} + \lambda s_{t-1}q_{t-1}^2 e^{-\lambda q_{t-2}^2}. \tag{4}
\]

Note that the first term in this recursion is exactly \( d_t \). The second term describes the help received from the shared bin. In order to properly characterize this term, it is necessary to introduce the notion of memory (term involving time \( t - 2 \)): the shared bin can help if it has 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.

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.** Follows from the density evolution derived above. \( \square \)

This style of thinking leads to the principles of generalization listed in below.

**Generalizing to \( L \)-smearing**

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

**B. Lower Bound on \( L \)-smearing**

For clarity, we will consider the ensembles with \( s = [L, L, L] \). Along with \( x_t, q_t, d_t \) as described in Table 2, we define quantity \( s_t^{(1)} \) in Table 3.

**Table 3: Notation for density evolution with smear-length \( L \)**

| \( s_t^{(1)} \): 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 \( t \) |

As described in the principles of generalization, there will be \( L \) recursions, and up to \( L - 1 \) memory. The probability \( d_t \)

\[\text{Although it is possible to derive exact recurrences in such a case, we pursue lower bounds for intellectual clarity.}\]
The critical quantity \( q_t \) only one step of memory. The following lemma characterizes the reference stream in \( L \) for \( i \) with respect to \( d_t \), which we may use to complete the bound.

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

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

**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].

We now bound \( d_t, s_t^{(1)}, \ldots, s_t^{(L-1)} \):

\[
d_t = e^{-\lambda q_t^{(i)}},
\]

\[
s_t^{(i)} \geq e^{\lambda q_t^{(i)} - 1} + \lambda q_t^{(2)} e^{-\lambda q_t^{(i-2)} - 1},
\]

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

\[
l_t^{(i)} \triangleq \sum_{j=1}^{i} \sum_{k=1}^{j} s_t^{(j)} s_t^{(k)} e^{-(L-k-1)\lambda q_t^{(i-2)}} e^{-(k-1)\lambda q_t^{(i)}}.
\]

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: \( e^{\lambda q_t^{(i)}} \).

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 implies that it is also contained in \( L - j \) bins that do not intersect the reference stream. It can be removed by any of these bins, as long as it is the only contribution. 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 experimentally find the thresholds for full recovery by sweeping \( \lambda \) and compare them to the thresholds implied by the bounds. Fig. 4 shows these for filters with different lengths corroborating that the bounds in Corollary 1 capture the actual thresholds well.

**III. CONNECTIONS TO RECOVERY OF SPARSE WAVELET REPRESENTATIONS**

In MRI, one acquires samples of the Fourier transform of an input signal of interest. MRI speed is directly related to the number of samples acquired. An inverse transform is then used to recover the original signal. Mathematically, let \( x \) be an \( n \)-length signal, and \( X \triangleq F_n x \) 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.

We now present how the game of balls-and-bins and its analysis as described in Sections I and II relates to MRI. For ease of illustration we confine ourselves to the noiseless setting and exact sparsity, but these assumptions can be
relaxed. 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 [2], [13]. However, the images of interest in MRI are generally not sparse, but they do have sparse wavelet representations [4]. That is, we can express \( x = W_n^{-1} \alpha \), where \( W_n^{-1} \) is an appropriate wavelet, and \( \alpha \) is sparse. Under this signal model, the problem of recovering \( \alpha \) can be transformed into the problem of decoding on an erasure channel using a sparse-graph code. In particular, the graph for the code is drawn from a smearing ensemble with smearing length \( a \) sampled with 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 \( \alpha \) 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 \( D_{m,n} \) be the regular downsampling matrix from length \( n \) to \( m \), so \( D_{m,n} = [I_m \cdots I_m] \) with \( I_m \) is repeated \( n/m \) times. Let \( y_{f_i} \) for \( i \in \{1, 2, 3\} \) be the inverse Fourier transform of the downsized \( X \), that is, \( y_{f_i} \equiv F^{-1}_{f_i} D_{f_i,n} X \). Using the properties of Fourier Transform, it follows that \( y_{f_i} = D_{f_i,n} x = D_{f_i,n} W_n^{-1} \alpha \).

Now, for simplicity, assume \( W \) is a block transform with block size \( L \) (e.g., for 1 stage Haar wavelets \( L = 2 \)), and the support of \( \alpha \) is chosen uniformly random over the subsets of size \( K \). Using the relations between \( y_{f_i}, y_{f_2} \) and \( y_{f_3} \) and \( \alpha \), recovering \( \alpha \) is equivalent to decoding on a random graph from \( G(K, M = f_1 + f_2 + f_3, s = \{L, L, L\}) \).

We can actually ‘improve’ the induced graph if a factor of the signal length has \( L \) as a factor. Say that \( L \) divides \( f_1 \), so \( A_{f_1,n} W_n^{-1} = [W_{f_1} \cdots W_{f_1}] \), where \( W_{f_1}^{-1} \) is repeated \( n/f_1 \) times. It can be verified that \( y_{f_1} = W_{f_1} y_{f_1} = A_{f_1,n} \alpha \), hence it aliases the wavelet coefficients regularly without smearing. The relation between \( y_{f_1}, y_{f_2} \) and \( y_{f_1} \) and \( \alpha \) then induces a graph from \( G(K, M = f_1 + f_2 + f_3, s = \{L, L, L\}) \), which gives raise to a better threshold.

To complete the equivalence to decoding a sparse-graph code on an erasure channel, we need a mechanism 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 \( x \) (incurring an additional factor of 2 of oversampling). Due to space limitations, we defer the details to the extended version [12]. We end this section with the following lemma.

**Lemma 4.** Consider a signal \( \alpha \) with ambient dimension \( n \) and sparsity \( K \), and access to samples from \( F_n W_n^{-1} \alpha \). Then, the subsampling scheme described above along with the peeling decoder is able to exactly recover the sparse signal \( \alpha \) using \( 2.63K \) samples and time complexity \( O(K \log K) \).

**Proof.** The threshold \( 2.63K \) follows from the density evolution obtained in Section II. We refer the reader to the extended version [12] for complexity proof.

IV. CONCLUSIONS AND FUTURE WORK

We have introduced a new random graph ensemble, termed the ‘smearing ensemble’ and developed a framework for deriving density evolution recurrences for random graphs from this ensemble. Recalling our balls-and-bins game, our results show that some amount of smearing can lead to a better strategy than the full random case. 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. 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 [10] 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, Feb. 2001.

[2] S. Pawar and K. Ramchandran, “Computing a \( k \)-sparse \( n \)-length discrete Fourier transform using at most \( 4k \) samples and \( O(k \log k) \) complexity,” in *IEEE Int. Symp. on Inf. Theory*. IEEE, Jul. 2013, pp. 464–468.

[3] M. Vetterli and J. Kovacevic, “Wavelets and subband coding,” 1995.

[4] 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, Mar. 2008.

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

[6] D. L. Donoho, A. Maleki, and A. Montanari, “Message-passing algorithms for compressed sensing,” *Proc. Natl. Acad. Sci. USA*, vol. 106, no. 45, pp. 18914–18919, Sep. 2009.

[7] 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, Jan. 2011.

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

[9] 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, Jan. 2011.

[10] 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, Jan. 2013.

[11] C. Hager, H. D. Pfister, F. Brannström et al., “Density evolution for deterministic generalized product codes on the binary erasure channel,” arXiv preprint arXiv:1512.00433, Dec. 2015.

[12] K. Chandrasekher, O. Ocal, and K. Ramchandran, “Density evolution on a class of smeared random graphs: A theoretical framework for fast mri,” arXiv preprint, May 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, Sep. 2015.

[14] D. L. Donoho, A. Maleki, and A. Montanari, “Message passing algorithms for compressed sensing: I. motivation and construction,” in *IEEE Int. Theory Workshop on Inf. Theory*. IEEE, Jan. 2010, pp. 1–5.