Choir Codes: Coding for Full Duplex Interference Management

Lorne Applebaum*, Waheed U. Bajwa†, Robert Calderbank‡, and Stephen Howard§
* lappleba@princeton.edu Princeton University
† waheed.bajwa@rutgers.edu Rutgers University
‡ robert.calderbank@duke.edu Duke University
§ stephen.howard@dsto.defence.gov.au Defence Science & Technology Organisation

***Abstract*—Communication networks conventionally operate with half-duplex methods and interference avoiding schemes to manage multiple transceivers. Here we consider a method in which nodes transmit and receive in concert to achieve full duplex communication without transmitter coordination. We build on a recent framework for full-duplex communication in ad-hoc wireless networks recently proposed by Zhang, Luo and Guo. An individual node in the wireless network either transmits or it listens to transmissions from other nodes but it cannot do both at the same time. There might be as many nodes as there are MAC addresses but we assume that only a small subset of nodes contribute to the superposition received at any given node in the network. We develop deterministic algebraic coding methods that allow simultaneous communication across the entire network. We call such codes *choir codes*. Users are assigned subspaces of $\mathbb{F}_{2^m}$ to define their transmit and listen times. Codewords on these subspaces are designed and proven to adhere to bounds on worst-case coherence and the associated matrix spectral norm. This in turn provides guarantees for multi-user detection using convex optimization. Further, we show that matrices for each receiver's listening times can be related by permutations, thus guaranteeing fairness between receivers. Compared with earlier work using random codes, our methods have significant improvements including reduced decoding/detection error and non-asymptotic results. Simulation results verify that, as a method to manage interference, our scheme has significant advantages over seeking to eliminate or align interference through extensive exchange of fine-grained channel state information.**

***Index Terms*—Duplex Codes, coding theory, wireless, full duplex, sparse recovery, random access**

I. Introduction

In many wireless networks, nodes communicate in an uncoordinated fashion. Users are not allocated channels or time-slots and they independently choose when to transmit their data. We call such networks *random access* networks. Examples of wireless random access networks include control channels of cellular systems as well as certain ad hoc and sensor networks. Conventionally, nodes in these scenarios compete for channel resources using contention resolution schemes such as ALOHA or CSMA. In these schemes, nodes either wait for acknowledgements and retransmit collided data or preemptively detect activity on the channel to avoid collisions. However, such schemes can introduce significant delays and waste channel resources. Addressing this problem, a new approach has developed which uses the fact that typically only a few users simultaneously compete for the channel.

Since the set of active users is small, if they simultaneously transmit their codewords, the signal at a receiver is, in a sense, sparse. This was recognized in [1] where recent advances in sparse signal recovery were applied to random access uplink communication. It was shown that data from multiple users could be detected simultaneously, without interference avoidance or coordination. This was strengthened and extended to asynchronous uplink communication in [2]. Recently, [3] introduced a novel scheme in which similar ideas were applied beyond the uplink. It was shown that network-wide virtual full-duplex communication could be achieved with half-duplex hardware. By switching radios between transmitting and receiving on user specific intervals, nodes could simultaneously recover data from neighbors while transmitting data themselves. This was further developed in [4] and [5]. It is upon the scheme in [3]–[5] which the work in this paper is based. In this paper, we consider a code design for the virtual full-duplex framework of [4] for random access wireless networks.

In [4] and [5], randomly generated codewords, defining both the receive periods and transmitted symbols, were considered. It was proved that, using a group testing or message-passing algorithm with the random codewords, data from transmitting nodes can be recovered with high probability. Further, in [4] they simulated the use of deterministic second-order Reed-Muller codewords with random erasures defining receive periods. Using the chirp decoding algorithm of [6] they empirically showed successful recovery. In this paper, we develop a fully deterministic code with several advantageous properties. Firstly, as a fully deterministic code, storage and generation is relatively simple. Second, by proving bounds on metrics of the codebook, we are able to give data recovery guarantees when a variety of algorithms are used. Finally, we show that the recovery problems exposed to each receiver are equivalent in a manner that ensures fairness between them.

The remainder of the paper is organized as follows. In Section II we describe the virtual full-duplex system and its model along with our underlying assumptions. Section III is the bulk of the paper and contains the description of our deterministic codes. In the subsections we analyze the code's properties in the context of three metrics; the worst-case coherence, the average coherence, and the spectral norm. In Section IV we apply our analysis to results from the literature to provide recovery guarantees from the code. Results from simulations are described in Section V and we conclude our discussion in Section VI.

II. System Model

We consider a framework like that of [4] in which nodes in a wireless network simultaneously transmit codewords. While the nodes are assumed to be half-duplex devices on the symbol time-scale, we exploit the fact that hardware can rapidly switch between transmission and reception. Codewords are designed to describe not only the transmitted signal but also time periods during which the node is set to receive. In effect, full-duplex is achieved on the codeword time-scale. In this section we describe our model of code transmission and reception while assuming a set of known codewords. In Section III, we make our code explicit.

Let $\mathcal{U}$ be an indexing set for the codewords/users. For each $a \in \mathcal{U}$, we associate a codeword $\mathbf{x}_a$ as a vector with elements in $\{-1, +1, 0\}$. We collect these vectors together and define the *full codebook* matrix as

$$\tilde{\mathbf{X}} = \begin{bmatrix}\mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{M+1}\end{bmatrix} \tag{1}$$

where, for clarity, we have assumed $\mathcal{U} = \{1, \dots, M+1\}$. Using a simple random access model, for each $a \in \mathcal{U}$, we define an independent and identically distributed (iid) Binomial random variable I_a with $\mathbb{P}[I_a = 1] = p_t$. The probability p_t is assumed to be small. The set of active users is defined as $\mathcal{I} = \{a \in \mathcal{U} : I_a = 1\}$ which is a small subset of $\mathcal{U}$. While the codebook is known to users, the set $\mathcal{I}$ is not.

We model communication as follows. Each user $a \in \mathcal{I}$ simultaneously transmits their codeword $\mathbf{x}_a$ by modulating the non-zero elements of $\mathbf{x}_a$ on the channel while switching their radio to receive during the 0 elements. Restricted to the time slots of the 0 elements, users receive a truncated version of other users' codewords. For a receiver $a \in \mathcal{U}$, we define the *collapsed codewords* of the remaining users as $\{\mathbf{x}_b^{(a)}\}_{b \in \mathcal{U}/\{a\}}$ where $\mathbf{x}_b^{(a)}$ is the vector $\mathbf{x}_b$ with the rows corresponding to the non-zero elements of $\mathbf{x}_a$ removed. Correspondingly we define the *collapsed measurement matrix* as

$$\mathbf{X}_a = \begin{bmatrix}\mathbf{x}_{b_1}^{(a)} & \mathbf{x}_{b_2}^{(a)} & \cdots & \mathbf{x}_{b_M}^{(a)}\end{bmatrix} \tag{2}$$

where we have indexed $\mathcal{U}/a$ by $b_1, \dots, b_M$. $\mathbf{X}_a$ is a sub-matrix of $\tilde{\mathbf{X}}$ with rows and one column removed.

We can model the signal received by user a as

$$\mathbf{y}_a = \mathbf{X}_a \beta + \mathbf{n} \tag{3}$$

where $\mathbf{y}_a$ is the vector of samples received during the user's receiving time slots and $\mathbf{n}$ is vector of noise distributed as $\mathcal{N}(0, \sigma\mathbf{I})$. The vector β has non-zero elements corresponding to $\mathcal{I}$ with values determined by the fading and power modulation of transmitting users. Since $\mathcal{I}$ is a small portion of $\mathcal{U}$, β has few non-zero components and is a sparse vector.

The goal of user a is to recover $\mathcal{I}$ from the support of the vector β to decode the data. Devoid of the communication network context, this problem is known as *model selection* and, since β is sparse, recent work in sparse recovery and compressed sensing suggest recovery is possible [7], [8]. Indeed, there is a large body work providing recovery conditions and methods for formulations similar to (3) when considering a single user. However, unique to this problem is that codewords jointly generate a family of recovery problems (one for each user). In the original paper by Zhang, Luo and Guo [4], vectors $\mathbf{x}_a$ generated at random or partially at random are shown to work in this framework with high probability. In this paper, we develop a fully deterministic construction.

III. A Small Choir Code

In this section we provide a code designed to operate in the framework described in Section II. We construct a code deterministically by operating in the finite field $\mathbb{F}_{2^m}$ with m odd. In particular we take $\mathcal{U} = \mathbb{F}_{2^m}^*$ and enumerate codeword symbols by $\mathbb{F}_{2^m}^*$, where we use $\mathbb{F}_{2^m}^*$ to denote the multiplicative group of the field.

The code makes extensive use of the field trace operator denoted as $\mathrm{Tr}(\cdot)$. As a review, the field trace has the following relevant properties for $a, b \in \mathbb{F}_{2^m}$.

(i) $\mathrm{Tr} : \mathbb{F}_{2^m} \to \mathbb{F}_2$
(ii) $\mathrm{Tr}(a^2) = \mathrm{Tr}(a)$
(iiia) $\mathrm{Tr}(a + b) = \mathrm{Tr}(a) + \mathrm{Tr}(b)$
(iiib) $(a, b) \mapsto \mathrm{Tr}(ab)$ is a bilinear form

Letting m be odd and $x \in \mathbb{F}_{2^m}^*$ enumerate the elements of a codeword, we define the *choir code* as

$$[\mathbf{x}_a]_x = (-1)^{\mathrm{Tr}(a^3x^3)} \delta_{\mathrm{Tr}(ax),0} \tag{4}$$

where δ denotes the Kronecker delta so that elements are only non-zero when $\mathrm{Tr}(ax) = 0$. These elements correspond to the transmission symbols of the codewords. When $\mathrm{Tr}(ax) = 1$, the user a switches its radio to receive signals. To begin, we investigate some properties of the sets of x during which $\mathrm{Tr}(ax) = 0$.

Considering $\mathbb{F}_{2^m}$ as a vector space and using property (iiib) of the trace, each user a is associated with a subspace we denote $\mathcal{N}_a = \{x \in \mathbb{F}_{2^m} : \mathrm{Tr}(ax) = 0\}$. These subspaces correspond to the transmission times for each user and the complementary sets, denoted $\mathcal{N}_a^c$, define the receiving times. The subspaces have the following useful property.

Fact 1: For $a_1, \dots, a_l$ as linearly independent elements in $\mathbb{F}_{2^m}^*$ the cardinality of the set $\mathcal{N}_{a_1} \cap \mathcal{N}_{a_2} \cap \cdots \cap \mathcal{N}_{a_l}$ is 2^{m-l}. This fact can be proved using dual bases, for example see [2, Proposition 5].

Since each subspace is of size $|\mathcal{N}_a| = 2^{m-1}$, users transmit during approximately half of their length $2^m - 1$ codeword. Further, intersections of subspaces are of size 2^{m-2} meaning users are able to receive approximately half of every other users' transmission. Assigning subspaces of $\mathbb{F}_{2^m}$ to users this way ensures that no user's transmissions completely mask any other user. While these non-overlapping transmission supports are necessary to allow recovery, they are not sufficient to ensure recovery. Below we consider properties of the matrices $\mathbf{X}_a$ which we later apply to recovery guarantees in Section IV.

A. Worst-Case Coherence

Worst-case coherence is a common metric found in sparse recovery literature. The worst-case coherence is the largest magnitude of inner-products between columns and is defined as

$$\mu(\mathbf{X}_a) = \frac{1}{2^{m-2}} \max_{\substack{b,c\in\mathcal{U}/\{a\}\\ b\neq c}} |\langle \mathbf{x}_b^{(a)}, \mathbf{x}_c^{(a)} \rangle| \tag{5}$$

where we have added a normalization factor of $2^{-(m-2)}$ to account for the fact that worst-case coherence is customarily applied to unit-normed columns. Using (4), we can write the inner-product of columns as the sum

$$\langle \mathbf{x}_b^{(a)}, \mathbf{x}_c^{(a)} \rangle = \sum_{\substack{x\in\mathbb{F}_{2^m}\\ x\in\mathcal{N}_a^c\cap\mathcal{N}_b\cap\mathcal{N}_c}} (-1)^{\mathrm{Tr}((b^3+c^3)x^3)} \tag{6}$$

where the sum is over all of $\mathbb{F}_{2^m}$ since 0 is already excluded from $\mathcal{N}_a^c$. First, note that if $a = b+c$ the sum is 0 since $\mathcal{N}_a^c \subset (\mathcal{N}_b \cap \mathcal{N}_c)^c$. For the non-trivial case, we have the following lemma.

Lemma 1: For the linearly independent elements $a, b, c \in \mathbb{F}^*_{2^m}$, $\langle \mathbf{x}_b^{(a)}, \mathbf{x}_c^{(a)} \rangle^2 \leq 2^{m+1}$.

Proof: Defining $g = (b^3 + c^3)$ for brevity, we can write the squared sum of (6) as

$$\begin{aligned}\langle \mathbf{x}_b^{(a)}, \mathbf{x}_c^{(a)} \rangle^2 &= \sum_{\substack{x,y\in\mathbb{F}_{2^m}\\ x,y\in\mathcal{N}_a^c\cap\mathcal{N}_b\cap\mathcal{N}_c}} (-1)^{\mathrm{Tr}(g(x^3+y^3))}\\ &= \sum_{\substack{x,y\in\mathbb{F}_{2^m}\\ x,y\in\mathcal{N}_a^c\cap\mathcal{N}_b\cap\mathcal{N}_c}} (-1)^{\mathrm{Tr}(g((x+y)^3+xy(x+y)))}\\ &= \sum_{\substack{z,y\in\mathbb{F}_{2^m}\\ y\in\mathcal{N}_a^c\cap\mathcal{N}_b\cap\mathcal{N}_c\\ z\in\mathcal{N}_a^c\cap\mathcal{N}_b\cap\mathcal{N}_c+y}} (-1)^{\mathrm{Tr}(g(z^3+zy(z+y)))}\end{aligned} \tag{7}$$

where in the last equality we have used the change of variables $z = x + y$. Using the linearity of the trace, we know that $\mathrm{Tr}(ax) = 1$ with $\mathrm{Tr}(ay) = 1$ implies $\mathrm{Tr}(a(x+y)) = 0$. Thus, for all $y \in \mathcal{N}_a^c \cap \mathcal{N}_b \cap \mathcal{N}_c$, we have $\mathcal{N}_a^c \cap \mathcal{N}_b \cap \mathcal{N}_c + y = \mathcal{N}_a \cap \mathcal{N}_b \cap \mathcal{N}_c$. We can therefore factor the above sum as

$$\begin{aligned}\langle \mathbf{x}_b^{(a)}, \mathbf{x}_c^{(a)} \rangle^2 &= \sum_{\substack{z\in\mathbb{F}_{2^m}\\ z\in\mathcal{N}_a\cap\mathcal{N}_b\cap\mathcal{N}_c}} (-1)^{\mathrm{Tr}(gz^3)} \times \\ &\qquad \sum_{\substack{y\in\mathbb{F}_{2^m}\\ y\in\mathcal{N}_a^c\cap\mathcal{N}_b\cap\mathcal{N}_c}} (-1)^{\mathrm{Tr}(g(z^2y+y^2z))}\\ &= \sum_{\substack{z\in\mathbb{F}_{2^m}\\ z\in\mathcal{N}_a\cap\mathcal{N}_b\cap\mathcal{N}_c}} (-1)^{\mathrm{Tr}(gz^3)} \times \\ &\qquad \sum_{\substack{y\in\mathbb{F}_{2^m}\\ y\in\mathcal{N}_a^c\cap\mathcal{N}_b\cap\mathcal{N}_c}} (-1)^{\mathrm{Tr}((gz^2+\sqrt{gz})y)}\end{aligned} \tag{8}$$

where for the final equality we have used the properties (ii) and (iiia) of the trace in the inner sum. Focusing on the inner sum, in the exponent we have a linear function of y with the null space $\mathcal{N}_{gz^2+\sqrt{gz}}$. For most values of z, the inner sum is 0 since, by Fact 1, precisely half the summands are (-1). However, if $\mathcal{N}_{gz^2+\sqrt{gz}} \supset \mathcal{N}_a \cap \mathcal{N}_b \cap \mathcal{N}_c$, all the summands are equal and the inner sum evaluates to $\pm|\mathcal{N}_a^c \cap \mathcal{N}_b \cap \mathcal{N}_c| = \pm 2^{m-3}$. In what follows, we bound the number of z in the outer sum for which this occurs.

The condition $\mathcal{N}_{gz^2+\sqrt{gz}} \supset \mathcal{N}_a \cap \mathcal{N}_b \cap \mathcal{N}_c$ is equivalent to $gz^2 + \sqrt{gz} = s$ for $s \in \mathrm{Span}\{a, b, c\}$. By Proposition 1 in the Appendix, this equation is linear with two solutions to the homogeneous equation. Therefore, there are at most $2 \times |\,\mathrm{Span}\{a, b, c\}| = 2^4$ values of z for which the inner sum of (8) evaluates to $\pm 2^{m-3}$. Applying the triangle inequality on the outer sum yields the result. ■

An application of the above lemma to the definition in (5) yields the following theorem.

Theorem 1: For any $a \in \mathbb{F}_{2^m}$, $\mu(\mathbf{X}_a) \leq 2^{-\left(\frac{m+5}{2}\right)}$.

B. Average Coherence

Where the worst-case coherence considers the magnitude of inner-products pairwise, the average coherence considers the inner-product with the average received codeword. As a metric, it is useful for guaranteeing support recovery using one-step thresholding [7]. The average coherence is defined as

$$\nu(\mathbf{X}_a) = \frac{1}{2^{m-2}} \frac{1}{|\mathcal{U}| - 2} \max_{b\in\mathcal{U}/\{a\}} \Big| \sum_{c\in\mathcal{U}/\{a,b\}} \langle \mathbf{x}_b^{(a)}, \mathbf{x}_c^{(a)} \rangle \Big| \tag{9}$$

where, once again, we have added an additional factor of $2^{-(m-2)}$ to account for codeword normalization.

To bound the average coherence, we begin with a simple fact about the full code matrix $\tilde{\mathbf{X}}$.

Lemma 2: With appropriate enumeration of the users and codeword elements, the matrix $\tilde{\mathbf{X}}$ is circulant.

Proof: Let z be a generator for the multiplicative group $\mathbb{F}^*_{2^m}$ and consider the following enumeration of the codeword elements and users. Let the ith element be indexed by $x_i = z^i$ and the jth user be indexed by $a_j = z^{-j}$. Thus, elements of the matrix $\tilde{\mathbf{X}}$ are given by

$$[\tilde{\mathbf{X}}]_{i,j} = (-1)^{\mathrm{Tr}(a_j^3 x_i^3)} = (-1)^{\mathrm{Tr}(z^{3(i-j)})} \tag{10}$$

which is a function of $(i - j) \mod 2^m - 1$. ■

Since $\tilde{\mathbf{X}}$ is circulant, the row sums of $\tilde{\mathbf{X}}$, (and $\mathbf{X}_a$) are constant. We denote this value R and can say the following.

Lemma 3: The row sum of $\tilde{\mathbf{X}}$ is $R = \pm 2^{\frac{m-1}{2}} - 1$.

Proof: Completing the row with a element of value 1 to sum over $\mathbb{F}_{2^m}$ rather than $\mathbb{F}^*_{2^m}$ and taking our arbitrary row to be $x = 1$, we have

$$R + 1 = \sum_{\substack{a\in\mathbb{F}_{2^m}\\ \mathrm{Tr}(a)=0}} (-1)^{\mathrm{Tr}(a^3)}. \tag{11}$$

Squaring the sum gives

$$\begin{aligned}(R+1)^2 &= \sum_{a,b\in\mathcal{N}_1} (-1)^{\mathrm{Tr}(a^3+b^3)} \\ &= \sum_{a,b\in\mathcal{N}_1} (-1)^{\mathrm{Tr}((a+b)^3+ab(a+b))} \\ &= \sum_{w\in\mathcal{N}_1} (-1)^{\mathrm{Tr}(w^3)} \sum_{b\in\mathcal{N}_1} (-1)^{\mathrm{Tr}(w^2b+b^2w)} \\ &= \sum_{w\in\mathcal{N}_1} (-1)^{\mathrm{Tr}(w^3)} \sum_{b\in\mathcal{N}_1} (-1)^{\mathrm{Tr}((w^2+\sqrt{w})b)}\end{aligned} \tag{12}$$

where in the final two equalities we first made the substitution $w = a+b$ and second, as we did in (8), used properties of the trace to produce a linear function of b in the exponent. The inner sum is precisely zero unless $w^2+\sqrt{w} = 0$ or 1 and we consider these two cases in turn.

Using the fact that the field has characteristic 2, solutions to $w^2+\sqrt{w} = 1$ are also solutions to $w^4+w+1 = 0$. This is an irreducible polynomial and therefore its solutions are in $\mathbb{F}_4$. However, since we take m to be odd, $\mathbb{F}_4$ is not a sub-field. Thus, no solutions exists to $w^2+\sqrt{w} = 1$ in $\mathbb{F}_{2^m}$ for m odd.

By Proposition 1 in the Appendix, the solutions to $w^2+\sqrt{w} = 0$ are 0 and 1. However, only $w = 0$ is an element of $\mathcal{N}_1$. Thus, the outer sum has only one non-trivial term with the value $(R+1)^2 = |\mathcal{N}_1| = 2^{m-1}$. ■

In addition to the row sum, we also require a bound on sums of arbitrary columns of the collapsed matrix $\mathbf{X}_a$. The following lemma provides such a bound.

Lemma 4: For any $b \neq a \in \mathbb{F}^*_{2^m}$, the sum of a collapsed codeword obeys $\langle \mathbf{x}_b^{(a)}, \mathbf{1}\rangle \leq 2^{\frac{m+1}{2}}$, where $\mathbf{1}$ is the vector of all 1.

Proof: The proof is similar to that of Lemma 1 since the column sum has the form

$$\langle \mathbf{x}_b^{(a)}, \mathbf{1}\rangle = \sum_{x\in\mathcal{N}_a^c\cap\mathcal{N}_b} (-1)^{\mathrm{Tr}(b^3x^3)} \tag{13}$$

which differs from (6) only in the subspaces of the index and the coefficient of x^3. Letting $g = b^3$ and following similar manipulations yields an equation identical to (8), though with the outer and inner sum indices as $\mathcal{N}_a\cap\mathcal{N}_b$ and $\mathcal{N}_a^c\cap\mathcal{N}_b$, respectively. In this case, $|\mathcal{N}_a^c\cap\mathcal{N}_b| = 2^{m-2}$. Further $gz^2+\sqrt{gz} \in \mathrm{Span}\{a,b\}$ has at most 2^3 solutions for z. As a result, applying the triangle inequality yields

$$\langle \mathbf{x}_b^{(a)}, \mathbf{1}\rangle^2 \leq 2^3 \times 2^{m-2}. \tag{14}$$

■

Using the results of Lemmas 3 and 4 we can derive the following bound on average coherence.

Theorem 2: For any $a \in \mathbb{F}^*_{2^m}$, $\nu(\mathbf{X}_a) \leq 5/(2^m-3)$.

Proof: Using the fact established in Lemma 2 that the row sums of $\mathbf{X}_a$ are constant, we can write

$$\sum_{c\in\mathcal{U}/\{a,b\}} \mathbf{x}_c^{(a)} = R\mathbf{1} - \mathbf{x}_b^{(a)}. \tag{15}$$

Applying this to the inner-product of (9) gives

$$\begin{aligned}\sum_{c\in\mathcal{U}/\{a,b\}} \langle \mathbf{x}_b^{(a)}, \mathbf{x}_c^{(a)}\rangle &= R\langle \mathbf{x}_b^{(a)}, \mathbf{1}\rangle - \langle \mathbf{x}_b^{(a)}, \mathbf{x}_b^{(a)}\rangle \\ &\leq 2^m(3/4+2^{\frac{1-m}{2}})\end{aligned} \tag{16}$$

where we have used Lemmas 3 and 4 with the fact that $\langle \mathbf{x}_b^{(a)}, \mathbf{x}_b^{(a)}\rangle = 2^{m-2}$. Taking $m \geq 3$ to bound the second term in the brackets by $1/2$ and including the additional multiplicative factors in (9) gives the result. ■

C. Spectral Norm

Here, we investigate the spectral norm (or induced ℓ_2 norm) $\|\mathbf{X}_a\|_2$. This norm can be used to give recovery guarantees for the lasso algorithm. We begin with the simple fact that, since $\mathbf{X}_a$ is a sub-matrix of $\tilde{\mathbf{X}}$, $\|\mathbf{X}_a\|_2 \leq \|\tilde{\mathbf{X}}\|_2$. Thus, the bulk of this subsection is devoted to bounding the spectral norm of the full codebook matrix.

The spectral norm of $\tilde{\mathbf{X}}$ is determined by the eigenvalues of the Gram matrix $\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}$. We first investigate the off diagonal entries of the Gram matrix.

Lemma 5: For $x \neq y \in \mathbb{F}_{2^m}$ indexing a row and column, the element $[\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}]_{x,y}$ is either -1 or $\pm 2^{\frac{m-1}{2}} - 1$.

Proof: Let S be the element at location x, y. Its value is given by the sum

$$\begin{aligned}(S+1) &= \sum_{a\in\mathcal{N}_x\cap\mathcal{N}_y} (-1)^{\mathrm{Tr}((ax)^3+(ay)^3)} \\ &= \sum_{a\in\mathcal{N}_x\cap\mathcal{N}_y} (-1)^{\mathrm{Tr}(ga^3)}\end{aligned} \tag{17}$$

where we have added 1 to allow a sum over $\mathbb{F}_{2^m}$ rather than $\mathbb{F}^*_{2^m}$ and defined $g = x^3+y^3$.

Following similar steps to (7) and (8) we have

$$\begin{aligned}(S+1)^2 &= \sum_{a,b\in\mathcal{N}_x\cap\mathcal{N}_y} (-1)^{\mathrm{Tr}(g((a+b)^3+ab(a+b)))} \\ &= \sum_{z\in\mathcal{N}_x\cap\mathcal{N}_y} (-1)^{\mathrm{Tr}(gz^3)} \times \\ &\qquad \sum_{b\in\mathcal{N}_x\cap\mathcal{N}_y} (-1)^{\mathrm{Tr}((gz^2+\sqrt{gz})b)}\end{aligned} \tag{18}$$

where we have used the change of variables $z = a+b$. As we have found in earlier sections, unless $gz^2+\sqrt{gz} \in \mathrm{Span}\{x,y\}$, the inner sum vanishes which motivates the definition of the set $\mathcal{Z} = \{z \in \mathbb{F}_{2^m} : gz^2+\sqrt{gz} \in \mathrm{Span}\{x,y\}\}$. Unlike in earlier sections, we can take advantage of not indexing over a complementary set such as $\mathcal{N}_a^c$. Here, the inner sum is identically $|\mathcal{N}_x\cap\mathcal{N}_y| = 2^{m-2}$ for $z \in \mathcal{Z}$ (i.e., we know the sign is not negative). As result, we have

$$\frac{(S+1)^2}{2^{m-2}} = \sum_{z\in\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}} (-1)^{\mathrm{Tr}(gz^3)} \tag{19}$$

where we have included $\mathcal{Z}$ in the index to reflect that they are the only terms that remain in the outer sum. By a simple application of Proposition 1 in the Appendix, $\mathcal{Z}$ and consequently $\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}$ is a subspace of $\mathbb{F}_{2^m}$. Therefore, seeing that (19) is similar to (17), we can apply the same transformations in (18) and find

$$\begin{aligned}\left(\frac{(S+1)^2}{2^{m-2}}\right)^2 &= \sum_{z\in\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}} (-1)^{\mathrm{Tr}(gz^3)}\times \\ &\qquad \sum_{b\in\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}} (-1)^{\mathrm{Tr}((gz^2+\sqrt{gz})b)} \\ &= \sum_{z\in\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}} (-1)^{\mathrm{Tr}(gz^3)} \times |\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}| \\ &= |\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}|\frac{(S+1)^2}{2^{m-2}}\end{aligned} \tag{20}$$

where, in the second equality we use the fact that outer sum is restricted $z\in\mathcal{Z}$ making the inner sum constant for every z. In the last equality, we use that the sum in (19) has reemerged. Thus, we have a quadratic equation for $(S+1)^2$ which has two solutions, $(S+1)^2 = 0$ or

$$(S+1)^2 = |\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}|\times 2^{m-2}. \tag{21}$$

Focusing on the second case, we first note this can occur only if all the terms in (19) are 1. Considering that $z = g^{-1/3}\in\mathcal{Z}$ has the violating property $\mathrm{Tr}(gz^3) = \mathrm{Tr}(1) = 1$, we find $\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}$ is a strict subset of $\mathcal{Z}$. Next, since $|\operatorname{Span}\{x,y\}| = 4$, using Proposition 1, $|\mathcal{Z}|$ is at most 8. However, by definition, $(S+1)$ must be an integer. Therefore, to make (21) a perfect square, $|\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}|$ must be an odd power of 2. The only possibility is $|\mathcal{N}_x\cap\mathcal{N}_y\cap\mathcal{Z}| = 2$ which gives the result. ■

Having characterized the entries of $[\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}]_{x,y}$ we can bound the norm of $\mathbf{X}_a$ with the following theorem. We state the result in terms of $\frac{1}{2^{m-2}}\|\mathbf{X}_a\|_2^2$, since conventionally the spectral norm is calculated with normalized columns.

Theorem 3: For arbitrary $a\in\mathbb{F}_{2^m}^*$, $\frac{1}{2^{m-2}}\|\mathbf{X}_a\|_2^2 \le 2^{\frac{m+5}{2}}$

Proof: The diagonal entries of $\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}$ are all $2^m - 1$ while the off-diagonal entries are, by Lemma 5, bounded in magnitude by $2^{\frac{m-1}{2}}+1$. Thus, by Gershgorin's circle theorem [9],

$$\begin{aligned}\|\tilde{\mathbf{X}}\|_2^2 &= \lambda_{\text{max}}(\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}) \\ &\le (2^{m-1}-1) + (2^{\frac{m-1}{2}}+1)(2^m-2) \\ &\le 2^{\frac{3m+1}{2}}\end{aligned} \tag{22}$$

Since $\mathbf{X}_a$ is a sub-matrix of $\tilde{\mathbf{X}}$, it also obeys this bound. Normalizing by $\frac{1}{2^{m-2}}$ gives the result. ■

We concede that the bound in Theorem 3 is not tight due to the crude application of Gershgorin circle theorem. However, it is possible to efficiently calculate a tighter bound on the spectral norm. In particular, from Lemma 2 we know that $\tilde{\mathbf{X}}$ is diagonalized by the Fourier matrix. Therefore, $\|\tilde{\mathbf{X}}\|_2$ can be calculated with the computationally efficient Fast Fourier Transform (FFT) applied on any codeword $\mathbf{x}_a$. This in turn bounds $\|\mathbf{X}_a\|_2$. In Figure 1 we display computations of this bound. For comparison, we include direct computations of $\frac{1}{\sqrt{2^{m-2}}}\|\mathbf{X}_a\|_2$ for small values of m. We also include a the lower bound $\sqrt{\frac{2^m-2}{2^{m-1}}}$ provided by a tight frame of unit normed columns with dimensions of $\mathbf{X}_a$. Figure 1 shows that the Theorem 3 is loose but a tighter bound is computable. Further, we conjecture that $\frac{1}{\sqrt{2^{m-2}}}\|\mathbf{X}_a\|_2$ scales like a tight frame. Methods extending those of Lemma 5 to give this result are in progress.

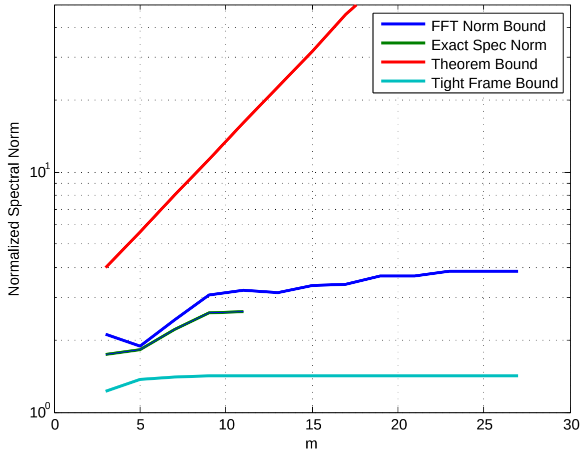

Fig. 1. Comparison of spectral norm bounds for $\mathbf{X}_a$ for various values of m in a semi-log plot. Normalizations are applied for comparison with Theorem 3.

D. Receiver Fairness

While the above results for $\mu(\mathbf{X}_a)$, $\nu(\mathbf{X}_a)$ and $\|\mathbf{X}_a\|_2$ apply to every user, they are upper bounds and apply to sufficiency conditions in Section IV. As such, it may be the case that some collapsed measurement matrices perform better than others. In this section we show that this is not the case. We do so by showing that the collapsed measurement matrices are related by row and column permutations.

We define the following matrices to aid in our discussion. For $a\in\mathbb{F}_{2^m}^*$ let $\mathbf{C}_a$ be the $(2^m-1)\times(2^m-1)$ matrix which, when multiplied on the left, "zeros" rows indexed by $x\in\mathbb{F}_{2^m}^*$ that satisfy $\mathrm{Tr}(ax) = 0$. That is, $\mathbf{C}_a$ is the identity matrix with the defined rows/columns set to zero. Further, for $\alpha\in\mathbb{F}_{2^m}^*$, define $\mathbf{T}_\alpha$ as the $(2^m-1)\times(2^m-1)$ permutation matrix which, when multiplied on the right, permutes columns such that column a is moved to column $a\alpha$. From the basic facts of permutation matrices $\mathbf{T}_\alpha$ is also a row permutation matrix which, when multiplied on the left, permutes rows such that row x is moved to row $x\alpha^{-1}$.

$\mathbf{C}_a\tilde{\mathbf{X}}$ is very closely related to $\mathbf{X}_a$. $\mathbf{X}_a$ is a sub-matrix of $\mathbf{C}_a\tilde{\mathbf{X}}$ since the latter merely contains extra rows of zeros and one extra column of zeros. We will show that $\mathbf{C}_b\tilde{\mathbf{X}}$ is formed from a permutation of rows and columns of $\mathbf{C}_a\tilde{\mathbf{X}}$. As sub-matrices, $\mathbf{X}_a$ and $\mathbf{X}_b$ are like-wise related by permutations. We begin with the following lemmas.

Lemma 6:

$$\mathbf{C}_b\mathbf{T}_{ba^{-1}} = \mathbf{T}_{ba^{-1}}\mathbf{C}_a \tag{23}$$

Proof: The following compound operations are equivalent:

- moving row x to $b^{-1}ax$ then setting it to zero if $\mathrm{Tr}(bb^{-1}ax) = \mathrm{Tr}(ax) = 0$
- setting row x to zero if $\mathrm{Tr}(ax) = 0$ then moving to $b^{-1}ax$

■

Lemma 7:

$$\mathbf{T}_\alpha \tilde{\mathbf{X}} = \tilde{\mathbf{X}} \mathbf{T}_{\alpha^{-1}} \tag{24}$$

Proof: By a fact of permutation matrices, $\mathbf{T}_{\tilde{\alpha}}^{-1} = \mathbf{T}_{\alpha^{-1}}$. Thus, here we equivalently show $\mathbf{T}_\alpha \tilde{\mathbf{X}} \mathbf{T}_\alpha = \tilde{\mathbf{X}}$. The element at location (x, a) is moved to $(a\alpha, x\alpha^{-1})$ by the pre and post multiplication of T_α. The value at $(a\alpha, x\alpha^{-1})$ is $(-1)^{\mathrm{Tr}(a^3\alpha^{-3}\alpha^3 x)}\delta(\mathrm{Tr}(a\alpha\alpha^{-1}x) = 0)$ which is equal to the value at (x, a). ■

Theorem 4: For $a \neq b \in \mathbb{F}_{2^m}$, $\mathbf{X}_b$ is a permutation of the rows and columns of $\mathbf{X}_a$.

Proof: As noted above, since $\mathbf{X}_b$ and $\mathbf{X}_a$ are sub-matrices, it is enough to show that $\mathbf{C}_b\tilde{\mathbf{X}}$ is a permutation of $\mathbf{C}_a\tilde{\mathbf{X}}$.

$$\mathbf{C}_b\tilde{\mathbf{X}} = \mathbf{T}_{ba^{-1}}\mathbf{C}_a\mathbf{T}_{b^{-1}a}\tilde{\mathbf{X}} = \mathbf{T}_{ba^{-1}}\mathbf{C}_a\tilde{\mathbf{X}}\mathbf{T}_{ba^{-1}} \tag{25}$$

where the first equality is due to Lemma 6 and the second equality is due to Lemma 7. ■

IV. Performance with Recovery Methods

In this section, we take the Theorems 1–3 of Section III and apply them to known results in the literature which can guarantee the recovery of β or $\mathcal{I}$ for the problem posed in Section II.

A. Restricted Isometry Property

Perhaps the best known recovery guarantees for sparse signals are those based on the restricted isometry property (RIP). For example, in [10], the RIP is used to provide sparse signal recovery guarantees using a linear program. We say a matrix $\mathbf{A}$ with unit-normed columns satisfies the RIP of order S with parameter δ_S is if

$$(1 - \delta_S)\|\mathbf{v}\|_2^2 \leq \|\mathbf{A}_S\mathbf{v}\|_2^2 \leq (1 + \delta_S)\|\mathbf{v}\|_2^2 \tag{26}$$

for all $\mathbf{v} \in \mathbb{R}^S$ and for all sub-matrices $\mathbf{A}_S$ of $\mathbf{A}$ constructed by selecting S columns. While the choir codes were not designed with the RIP in mind, due to its prolific nature we characterize the RIP of the matrix $\mathbf{X}_a$ using $\mu(\mathbf{X}_a)$.

The condition in (26) is equivalently a bound on the eigenvalues of $\mathbf{A}_S^T\mathbf{A}_S$. Using the methods of [11] any eigenvalue λ of $\mathbf{A}_S^T\mathbf{A}_S$ satisfies $|\lambda - 1| \leq (S-1)\mu(\mathbf{A})$. Applied to $\mathbf{X}_a$ we have that $\mathbf{X}_a$ satisfies the RIP for all δ_S and S satisfying

$$\delta_S \geq 2^{-\frac{m+5}{2}}(S - 1). \tag{27}$$

This RIP result can be applied to a wide variety of recovery methods. For example, applied to the Dantzig selector [12], gives a guarantee that β is estimated accurately when

$$|\mathcal{I}| \leq C_d 2^{\frac{m}{2}} \tag{28}$$

for a known constant C_d. This result is somewhat weak due to the reliance on $\mu(\mathbf{X}_a)$ to prove the RIP.

Recent advances, however, have provided recovery guarantees that depend directly on the metrics proved in Section III rather than the RIP. Further, they consider the support recovery problem directly and address estimates of $\mathcal{I}$ rather than estimates of β. We consider two of these results in the subsections below.

B. One-Step Thresholding

The model selection problem of estimating $\mathcal{I}$ from $\mathbf{y}_a$ using one-step thresholding is studied in [7]. One-step thresholding is the simple algorithm of back-projecting $\mathbf{y}_a$ onto $\mathbf{X}_a^T$ and thresholding the resulting vector. The two conditions

$$\mu(\mathbf{X}_a) \leq \frac{0.1}{\sqrt{2\log(2^m - 1)}} \quad \text{and} \tag{29}$$

$$\nu(\mathbf{X}_a) \leq \frac{\mu(\mathbf{X}_a)}{\sqrt{2^{m-1} - 1}} \tag{30}$$

are proven to allow one-step thresholding to recover $\mathcal{I}$ with high probability. Using Theorems 1 and 2, choir codes satisfy both conditions. As a result, [7] gives the following guarantee. With an appropriately chosen threshold, one-step threshold recovers $\mathcal{I}$ with high probability when

$$|\mathcal{I}| \leq C_O \frac{2^{m-1} - 1}{m \log 2} \tag{31}$$

for a known constant C_O dependent on the noise [7, Theorem 1]. Compared with (28), we find that with one-step thresholding a large set $\mathcal{I}$ can be guaranteed to be recovered.

C. The Lasso

Known as the lasso, the minimization

$$\hat{\beta} = \mathrm{argmin}_b \frac{1}{2}\|\mathbf{y}_a - \mathbf{X}_a b\|_2^2 - \lambda\sigma\|b\|_1 \tag{32}$$

is an estimation technique for the sparse signal β. It is studied in [8] in the context of model selection, whereby an estimate of $\mathcal{I}$ is formed from the support of $\hat{\beta}$. In [8], the following two conditions are given.

$$\mu(\mathbf{X}_a) \leq \frac{C_{L_0}}{\log(2^m - 1)} \tag{33}$$

$$[\beta]_a > 8\sigma\sqrt{2\log(2^m - 1)} \quad \forall a \in \mathcal{I} \tag{34}$$

where C_{L_0} is a known constant. For the choir code, (33) is satisfied by Theorem 1, while (34) is a mild requirement on the received signal power. When satisfied, $\mathcal{I}$ is successfully recovered with high probability as long as $|\mathcal{I}| \leq C_{L_1}\frac{2^m - 1}{(2^{m-2})^{-1}\|\mathbf{X}_a\|_2^2\log(2^m - 1)}$ for a known constant C_{L_1} [8, Theorem 1.3]. Applying Theorem 3 and assuming received powers satisfy (34), recovery of $\mathcal{I}$ is assured with high probability if

$$|\mathcal{I}| \leq C_{L_2}\frac{2^{\frac{m}{2}}}{m \log 2} \tag{35}$$

for a known constant C_{L_2}. This scales slightly worse than (28). However, the bound on $\|\mathbf{X}_a\|_2$ calculable using the FFT can be used. As shown in Figure 1 and discussed in Section III-C,

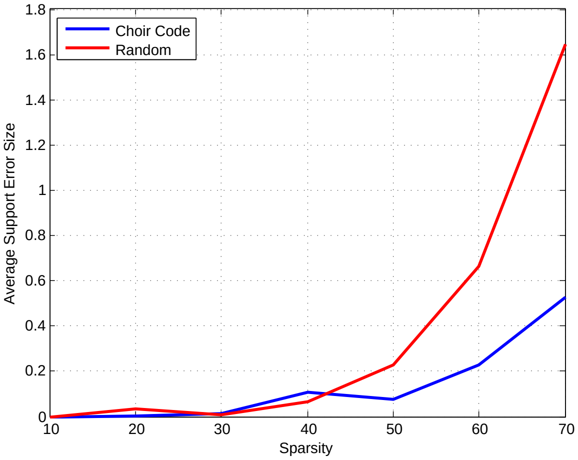

Fig. 2. Monte Carlo experiments illustrating the quality of recovery for choir codes and random codes as a function of expected size of $\mathcal{I}$.

evidence shows that $\|\mathbf{X}_a\|_2$ scales as a tight frame. In this case,

$$|\mathcal{I}| \leq C_{L_3} \frac{2^m}{m \log 2}. \tag{36}$$

guarantees recovery. This scales as (31).

In Section V we simulate the use of the choir code with lasso recovery.

V. Receiver Simulations

To verify the results presented in this paper, we use simulations of a receiver in a network using choir codes. In our experiments, we take $m = 11$ and select an arbitrary user a as a receiver. The active user set $\mathcal{I}$ and Gaussian noise $\mathbf{n}$ is generated at random in Monte Carlo iterations. We populate the vector β on the support corresponding to $\mathcal{I}$ with the value 1 and choose the noise parameter σ such that the SNR is 20dB. 500 Monte Carlo trials are used for each experiment. We choose the lasso as our recovery method and use the SpaRSA [13] package as a solver.

For comparison, we also simulate a receiver in a network using randomly generated codewords. The random codewords are generated with iid symbols. In expectation, half the symbols are 0. The transmitted symbols are ±1 with equal probability. The random codebook is generated once per experiment.

Results of these simulations are shown in Figure 2. The experiments show the average quality of recovery of $\mathcal{I}$ as a function of the sparsity $\mathbb{E}[|\mathcal{I}|]$. The sparsity level is adjusted via activation probability p_t. The quality of recovery is measured as the average size of the error set $(\mathcal{I} \cap \hat{\mathcal{I}}^c) \cup (\mathcal{I}^c \cap \hat{\mathcal{I}})$ (i.e., the average number of missed detections and false positives). We see that we are able to recover the active user set with few errors when the average number of users is less than 70. By comparison, we find that the randomly generated code begins to show significant errors with a smaller active user set sizes.

VI. Conclusion

In this paper, we introduce choir codes for use in random access wireless networks. The code's intent is to allow full duplex communication in the network by including 0 symbols indicating sampling periods during which a user's radio is set to receive. We allocate the 0 symbols to users by assigning subspaces of $\mathbb{F}_{2^m}$, the field upon which the code is defined. This ensures each user can receive a sufficient portion of other users' transmitted codewords when restricted to the listening symbols. The set of codewords creates a family of estimation problems where each user must recover data from sets of collapsed codewords of other users. On the sets of these collapsed codewords, we calculate bounds on three important metrics: worst-case coherence, average coherence and spectral norm. We draw these metrics from literature on sparse signal recovery and model selection. By bounding them, we provide guarantees that users can recover transmitted data when receivers use one-step thresholding, the lasso or various other algorithms. Further, we show that no user is at a disadvantage to another by proving that the recovery problems are equivalent via permutations.

Compared to past work, choir codes have several advantages. Firstly, the code is a purely deterministic construction. As such, allocation, storage and retrieval of codewords and the matrices $\mathbf{X}_a$ is relatively simple. The code also exhibits performance benefits. As shown in Figure 2, the performance of the code exceeds that of the randomly generated codewords of [4] when using the lasso. We expect that this extends to other recovery methods as well. Further, the results in Section III and their subsequent application in Section IV are non-asymptotic which gives designers more insight into parameter selection.

These codes represent a movement away from avoiding interference to managing interference. Conventional peer-to-peer random access wireless networks operate with orthogonal signaling or collision detection and avoidance mechanisms. This can be costly in delays or pre-communication coordination. Choir codewords, on the other hand, work in harmony to provide simultaneous network-wide communication.

Appendix

Proposition 1: For m odd and $g \in \mathbb{F}_{2^m}$, let $f : x \mapsto gx^2 + \sqrt{gx}$. Then f is linear and $f = 0$ has two solutions given by $x = 0$ and $x = g^{-1/3}$.

Proof: Since $\mathbb{F}_{2^m}$ has characteristic 2, $(x+y)^2 = x^2+y^2$. Further, $\sqrt{x} = x^{2^{(m-1)}}$. The linearity of f comes from these two facts.

Solutions to $f = 0$ also satisfy $f^2 = 0$ which factors as $gx(gx^3+1) = 0$. Thus, the two solutions are $x = 0$ and $x = g^{-1/3}$. The cubed root is well defined since, by Proposition 2, $\left(\frac{a}{b}\right)^3 = 1$ has the unique solution $a = b$. Therefore, $a \mapsto a^3$ is a bijection. ■

Proposition 2: For m odd, there are no non-trivial cubed roots of unity in $\mathbb{F}_{2^m}$.

Proof: Assume a non-trivial cubed root exists. Since the order of an element must divide the order of $\mathbb{F}_{2^m}^*$, we must have $3 \mid 2^m - 1$. However, by [14, Theorem 2.3], $\gcd(3, 2^m - 1) = 2^{\gcd(2,m)} - 1 = 1$, where for the last equality we use the fact that m is odd. ■

References

[1] A. Fletcher, S. Rangan, and V. Goyal, "On-off random access channels: A compressed sensing framework," submitted [arXiv:0903.1022v2].

[2] L. Applebaum, W. U. Bajwa, M. F. Duarte, and R. Calderbank, "Asynchronous code-division random access using convex optimization," *Physical Communication*, 2011, accepted for publication.

[3] D. Guo and L. Zhang, "Virtual full-duplex wireless communication via rapid on-off-division duplex," in *Forty-Eighth Annual Allerton Conference on Communication, Control, and Computing*, 2010.

[4] L. Zhang, J. Luo, and D. Guo, "Compressed neighbor discovery for wireless networks," *preprint*, vol. abs/1012.1007, 2010.

[5] L. Zhang and D. Guo, "Wireless peer-to-peer mutual broadcast via sparse recovery," *preprint*, vol. abs/1101.0294, 2011.

[6] R. Calderbank, S. D. Howard, and S. Jafarpour, "Sparse reconstruction via the Reed-Muller sieve," *IEEE Trans. Info. Theory*, 2010, submitted.

[7] W. U. Bajwa, R. Calderbank, and S. Jafarpour, "Why Gabor frames? two fundamental measures of coherence and their role in model selection," *J. Commun. Netw.*, pp. 289–307, Aug. 2010.

[8] E. J. Candès and Y. Plan, "Near-ideal model selection by ℓ_1 minimization," *Ann. Statist.*, vol. 37, no. 5A, pp. 2145–2177, 2009.

[9] R. S. Varga, *Geršgorin and His Circles*. Berlin, Germany: Springer-Verlag, 2004.

[10] E. Candès and T. Tao, "Decoding by linear programming," *Information Theory, IEEE Transactions on*, vol. 51, no. 12, pp. 4203 – 4215, dec. 2005.

[11] L. Applebaum, S. D. Howard, S. Searle, and R. Calderbank, "Chirp sensing codes: Deterministic compressed sensing measurements for fast recovery," *Applied and Computational Harmonic Analysis*, vol. 26, no. 2, pp. 283 – 290, 2009.

[12] E. Candès and T. Tao, "The Dantzig selector: Statistical estimation when p is much larger than n," *Ann. Statist.*, vol. 35, no. 6, pp. 2313–2351, Dec. 2007.

[13] S. Wright, R. Nowak, and M. Figueiredo, "Sparse reconstruction by separable approximation," *IEEE Trans. Signal Processing*, pp. 2479–2493, Jul. 2009.

[14] R. J. McEliece, *Finite Fields for Computer Scientists and Engineers*. Norwell, MA, USA: Kluwer Academic Publishers, 1987.