Virtual Full Duplex Wireless Broadcasting via Compressed Sensing

Lei Zhang and Dongning Guo, *Senior Member, IEEE*

Abstract—A novel solution is proposed to undertake a frequent task in wireless networks, which is to let all nodes broadcast information to and receive information from their respective one-hop neighboring nodes. The contribution in this paper is twofold. First, as each neighbor selects one message-bearing codeword from its unique codebook for transmission, it is shown that decoding their messages based on a superposition of those codewords through the multiaccess channel is fundamentally a problem of compressed sensing. In the case where each message is designed to consist of a small number of bits, an iterative algorithm based on belief propagation is developed for efficient decoding. Second, to satisfy the half-duplex constraint, each codeword consists of randomly distributed on-slots and off-slots. A node transmits during its on-slots and listens to its neighbors only through its own off-slots. Over one frame interval, each node broadcasts a message to its neighbors and simultaneously receives the superposition of neighbors' signals through its own off-slots and then decodes all messages. The proposed solution fully exploits the multiaccess nature of the wireless medium and addresses the half-duplex constraint at the fundamental level. In a network consisting of Poisson distributed nodes, numerical results demonstrate that the proposed scheme often achieves several times the rate of slotted ALOHA and CSMA with the same packet error rate.

Index Terms—Ad hoc network, belief propagation, full duplex, message passing, multiaccess channel, Poisson point process, rapid on–off-division duplex, sparse recovery.

I. INTRODUCTION

CONSIDER a frequent situation in wireless peer-to-peer networks, where every node wishes to broadcast messages to all nodes within its one-hop neighborhood, called its *neighbors*, and also wishes to receive messages from the same set of neighbors. We refer to this problem as *mutual broadcasting*. Such traffic can be dominant in many applications, such as messaging or video conferencing of multiple parties in a

L. Zhang was with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208 USA. He is now with Qualcomm Research, Bridgewater, NJ 08807 USA (e-mail: leizhang2011@u.northwestern.edu).

D. Guo is with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208 USA.

spontaneous social network or at an incident scene. Wireless mutual broadcasting is also critical to efficient network resource allocation, where messages are exchanged between nodes about their local states, such as queue length, channel quality, code and modulation format, and request for certain resources and services.

A major challenge in wireless networks is the half-duplex constraint, namely, currently affordable radio cannot receive useful signals at the same time over the same frequency band over which it is transmitting. This is largely due to the limited dynamic range and noise of the radio frequency circuits, which are likely to remain a physical restriction in the near future. An important consequence of the half-duplex constraint is that if two neighbors transmit their packets (or frames which are used interchangeably hereafter) at the same time, they do not hear each other. To achieve reliable mutual broadcasting using a usual packet-based scheme, nodes have to repeat their packets many times interleaved with random delays, so that all neighbors can hear each other after enough retransmissions. This is basically the ubiquitous random channel access solution.

A closer examination of the half-duplex constraint, however, reveals that a node does not need to transmit an entire packet before listening to the channel. An alternative solution is conceivable: Let a frame (typically of a few thousand symbols) be divided into some number of slots, where a node transmits over a subset of the slots and assumes silence over the remaining slots, then the node can receive useful signals over those nontransmission slots. If different nodes activate different sets of on-slots, then they can all transmit information during a frame and receive useful signals within the same frame and decode messages from neighbors, as long as sufficiently strong error-control codes are applied.

This on–off signaling, called *rapid on–off-division duplex (RODD)*, was originally proposed in [1]. Using RODD, reliable mutual broadcasting can be achieved using a single frame interval. Thus, RODD enables half-duplex radios to achieve virtual full-duplex communication. Despite the half-duplex physical layer, the radio appears to be full-duplex in the medium access control (MAC) layer and above. Not only is RODD signaling applicable to the mutual broadcasting problem, it can also be the basis of a clean-slate design of the physical and MAC layers of wireless peer-to-peer networks.

In this paper, we focus on a special use of RODD signaling and a special case of mutual broadcasting, where every node has a small number of bits to send to its neighbors. It is assumed that node transmissions are perfectly synchronized. The goal here is to provide a practical algorithm for encoding and decoding the short messages to achieve reliable and efficient

mutual broadcasting. Decoding is fundamentally a problem of *sparse recovery* (or *compressed sensing*) based on linear measurements since the received signal is basically a noisy superposition of neighbors' codewords selected from their respective codebooks. There are many algorithms developed in the compressed sensing literature to solve the problem, the complexity of which is often polynomial in the size of the codebook (see, e.g., [2]–[6]). In this paper, an iterative message-passing algorithm based on belief propagation (BP) with linear complexity is developed. Numerical results show that the proposed RODD scheme significantly outperforms slotted ALOHA with multipacket reception capability and CSMA in terms of data rate.

The excellent performance of the proposed scheme is because it departs from the usual solution where a highly reliable, highly redundant, capacity-achieving, point-to-point physical-layer code is paired with a rather unreliable MAC layer. In this paper, the physical and MAC layers are treated as a whole, and the proposed scheme achieves better overall reliability at much higher efficiency.

It is easier to implement RODD in a synchronized network. Regardless of whether RODD or any other physical- and MAC-layer technology is used, it is necessary to acquire their timing (or relative delay) in order to decode messages from neighbors. Timing acquisition and decoding are generally easier if the frames arriving at a receiver are synchronous locally within each neighborhood, although synchronization is not a necessity. In a wireless network, synchronization can be achieved using various distributed algorithms for reaching consensus [7]–[9] or using a common source of timing, such as the Global Positioning System (GPS). Whether spending the resource to synchronize nodes is worthwhile is a challenging question, which is not discussed further in this paper.

In contrast to virtual full duplex using RODD signaling, real full duplex becomes feasible if the interference a node's transmit chain causes to its own receive chain can be suppressed to weaker level than the desired received signal. Two self-interference cancellation techniques have recently received much attention: One employs a balanced/unbalanced transformer to negate the transmitted signal for analog cancellation [10]; the other separates the transmit and receive antennas and uses analog and optional digital cancellation [11]. However, those techniques do not always apply. For instance, self-interference cancellation may be insufficient if the signals have large dynamic range; space limitations may not allow adequate antenna separation; and, with multiple transmit antennas, canceling self interference in multiple chains may be hard. In those cases, RODD is a more viable solution. Whether full duplex is achieved using RODD or an alternative means, the proposed compressed sensing framework and technique provide a competitive solution to mutual broadcasting.

The remainder of the paper is organized as follows. The system model is presented in Section II. The proposed coding scheme for mutual broadcasting is described in Section III. The message-passing decoding algorithm is developed in Section IV. Section V studies the conventional random access schemes, namely slotted ALOHA with multipacket reception capability and CSMA. Numerical comparisons are presented in Section VI. Section VII concludes the paper.

II. SYSTEM MODEL

A. Linear Channel Model

Let $\Phi = \{Z_i\}_i$ denote the set of nodes on the plane. We refer to a node by its location Z_i. Suppose all transmissions use the same single carrier frequency.[1] Let time be slotted and all nodes be perfectly synchronized.[2] Suppose each node has l bits to broadcast to its neighbors. Let $d_i \in \{1, \ldots, 2^l\}$ denote the data or message that node Z_i wishes to broadcast. In discrete-time baseband, let $\boldsymbol{S}_i(d_i)$ denote the on–off signature (codeword) of length M_s transmitted by node Z_i, whose entries take values in $\{-1, 0, +1\}$. Here, the zero entries of $\boldsymbol{S}_i(d_i)$ correspond to the off-slots in which Z_i listens to the channel. The design of the on–off signatures will be discussed in Section III. Let U_{0i} denote the complex-valued coefficient of the wireless link from Z_i to Z_0. The signal received by node Z_0, if it could listen over the entire frame, is described by

$$\widetilde{\boldsymbol{Y}} = \sqrt{\gamma} \sum_{Z_i \in \Phi \backslash \{Z_0\}} U_{0i} \boldsymbol{S}_i(d_i) + \widetilde{\boldsymbol{W}} \qquad (1)$$

where the noise $\widetilde{\boldsymbol{W}}$ consists of independent identically distributed (i.i.d.) circularly symmetric complex Gaussian entries with zero mean and unit variance, and γ denotes the nominal signal-to-noise ratio (SNR). Denote the set of neighbors of Z_0 by $\mathcal{N}(Z_0)$. If we further assume that transmissions from non-neighbors, if any, are accounted for as part of the additive Gaussian noise, (1) can be rewritten as

$$\widetilde{\boldsymbol{Y}} = \sqrt{\gamma} \sum_{Z_i \in \mathcal{N}(Z_0)} U_{0i} \boldsymbol{S}_i(d_i) + \overline{\boldsymbol{W}} \qquad (2)$$

where each element in $\overline{\boldsymbol{W}}$ is assumed to be circularly symmetric complex Gaussian with variance σ^2. The variance, to be derived in Section III, accounts for interference from non-neighbors and depends on the network topology.

B. Network Model

Consider a network with nodes distributed across the plane according to a homogeneous Poisson point process (p.p.p.) with intensity λ. The number of nodes in any region of area A is a Poisson random variable with mean λA. Without loss of generality, we assume node Z_0 is located at the origin and focus on its performance, which should be representative of any node in the network.

Poisson point process is the most frequently used model to study wireless networks (see [13], [14], and references therein). The homogeneous p.p.p. model is assumed here to facilitate analysis and comparison of competing technologies. The proposed RODD signaling and the mutual broadcasting scheme can be applied to general networks.

[1]The frequency offset between different transmitters is assumed to be small, so as not to cause phase rotation over one frame.

[2]A discussion of synchronization issues is found in [1]. In [12], cyclic codes are proposed to accommodate different user delays under the same compressed sensing framework.

C. Propagation Model and Neighborhood

The large-scale signal attenuation over distance is assumed to follow the power law with some path-loss exponent $\alpha > 2$. The small-scale fading of a link is modeled by an independent Rayleigh random variable with mean equal to 1. The neighborhood of a node can be defined in many different ways. For concreteness, we say that nodes Z_i and Z_j are neighbors of each other if the channel gain between them exceeds a certain threshold, θ. Link reciprocity is regarded as given.

For any pair of nodes $Z_i, Z_j \in \Phi$, let $R_{ij} = |Z_i - Z_j|$ denote the distance and and G_{ij} the small-scale fading gain between them in a given frame, respectively. Then, the channel gain between Z_j and Z_i is $G_{ij}R_{ij}^{-\alpha}$. The neighborhood of a node depends on the instantaneous fading gains. Specifically, we denote the set of neighbors of node Z_i as

$$\mathcal{N}(Z_i) = \left\{ Z_j \in \Phi : G_{ij}R_{ij}^{-\alpha} \geq \theta, j \neq i \right\}. \qquad (3)$$

The channel coefficient U_{ij} should satisfy $|U_{ij}|^2 = G_{ij}R_{ij}^{-\alpha}$, where its phase is assumed to be uniformly distributed on $[0, 2\pi)$ independent of everything else. Assuming the Poisson point process network model introduced in Section II-B, the distribution of the amplitude of coefficient U_{0i} in (1) for an arbitrary neighbor $Z_i \in \mathcal{N}(Z_0)$ is derived in the following.

Without loss of generality, we drop the indices 0 and i, and use R and G to denote the distance and the fading gain, respectively. Since the two nodes are assumed to be neighbors, G and R satisfy $GR^{-\alpha} \geq \theta$, i.e., $R \leq (G/\theta)^{1/\alpha}$. Under the assumption that all nodes form a p.p.p., for given G, this arbitrary neighbor Z_i is uniformly distributed in a disc centered at node Z_0 with radius $(G/\theta)^{1/\alpha}$. Therefore, the conditional distribution of R given G can be expressed as

$$\mathsf{P}(R \leq r|G) = \min\left\{ 1, \left(\frac{\theta}{G}\right)^{\frac{2}{\alpha}} r^2 \right\}. \qquad (4)$$

Hence, for every $u \geq \sqrt{\theta}$

$$\mathsf{P}(GR^{-\alpha} \geq u^2) = \mathsf{E}_G\left\{ \mathsf{P}\left(R \leq \left(\frac{G}{u^2}\right)^{\frac{1}{\alpha}} \middle| G \right) \right\}$$
$$= \mathsf{E}_G\left\{ \left(\frac{G}{u^2}\right)^{\frac{2}{\alpha}} \left(\frac{\theta}{G}\right)^{\frac{2}{\alpha}} \right\}$$
$$= \frac{\theta^{\frac{2}{\alpha}}}{u^{\frac{4}{\alpha}}}. \qquad (5)$$

Therefore, the probability density function (pdf) of $|U_{0i}|$ is

$$p(u) = \begin{cases} \frac{4}{\alpha} \frac{\theta^{2/\alpha}}{u^{4/\alpha+1}}, & u \geq \sqrt{\theta} \\ 0, & \text{otherwise.} \end{cases} \qquad (6)$$

In fact, the coefficient vector $\mathcal{G}_i = (G_{ji})_j$ for all $j \neq i$ can be regarded as a mark of node Z_i, so that $\tilde{\Phi} = \{(Z_i, \mathcal{G}_i)\}_i$ is a marked p.p.p. [15]. Denote

$$\hat{\Phi} = \tilde{\Phi} \backslash (Z_0, \mathcal{G}_0) \qquad (7)$$

given that $(Z_0, \mathcal{G}_0)$ is at the origin. By the Slivnyak–Meche theorem [13], $\hat{\Phi}$ is also a marked p.p.p. with intensity λ. By the Campbell's theorem [13], the average number of neighbors of Z_0 can be obtained as

$$c = \mathsf{E}_{\hat{\Phi}}\left\{ \sum_{(Z_i, \mathcal{G}_i) \in \hat{\Phi}} 1\left(G_{0i}R_{0i}^{-\alpha} \geq \theta\right) \right\}$$
$$= 2\pi\lambda \int_0^\infty \int_0^\infty 1(gr^{-\alpha} \geq \theta)re^{-g}\mathrm{d}r\mathrm{d}g$$
$$= \frac{2}{\alpha}\pi\lambda\theta^{-2/\alpha}\Gamma\left(\frac{2}{\alpha}\right) \qquad (8)$$

where $1(\cdot)$ is the indicator function and $\Gamma(\cdot)$ is the Gamma function.

III. ENCODING FOR MUTUAL BROADCASTING

Recall that each node Z_i is assigned a codebook of 2^l on–off signatures (codewords) of length M_s, denoted by $\{\boldsymbol{S}_i(1), \ldots, \boldsymbol{S}_i(2^l)\}$. For simplicity, let each element of each signature be generated randomly and independently, which is 0 with probability $1 - q$, and 1 and -1 with probability $q/2$ each.[3] Node Z_i broadcasts its l-bit message (or information index) $d_i \in \{1, \ldots, 2^l\}$ by transmitting the codeword $\boldsymbol{S}_i(d_i)$. Transmissions of all nodes are synchronized. All nodes finish one message exchange after each frame of transmission.

In each symbol slot, those transmitting nodes in $\hat{\Phi}$ defined in (7) form an independent thinning of $\hat{\Phi}$ with retention probability q, denoted by $\hat{\Phi}_q$. $\hat{\Phi}_q$ is still an independent marked p.p.p., but with intensity λq. Thus, the sum power from all transmitting non-neighbors of node Z_0 in each time-slot is derived as

$$\mathsf{E}_{\hat{\Phi}_q}\left\{ \sum_{(Z_i, \mathcal{G}_i) \in \hat{\Phi}_q} \gamma G_{0i}R_{0i}^{-\alpha} 1\left(G_{0i}R_{0i}^{-\alpha} < \theta\right) \right\}$$
$$= 2\pi\lambda q\gamma \int_0^\infty \int_0^\infty gr^{-\alpha} 1\left(gr^{-\alpha} < \theta\right)re^{-g}\mathrm{d}r\mathrm{d}g$$
$$= 2\pi\lambda q\gamma \int_0^\infty r^{-\alpha+1}\left[1 - (\theta r^\alpha + 1)e^{-\theta r^\alpha}\right]\mathrm{d}r$$
$$= \frac{4\pi\lambda q\gamma\theta}{\alpha - 2} \int_0^\infty re^{-\theta r^\alpha}\mathrm{d}r$$
$$= \frac{4}{\alpha(\alpha-2)}\pi\lambda q\gamma\theta^{1-2/\alpha}\Gamma\left(\frac{2}{\alpha}\right). \qquad (9)$$

Therefore, the variance of each element of $\overline{\boldsymbol{W}}$ in (2) is

$$\sigma^2 = \frac{4}{\alpha(\alpha-2)}\pi\lambda q\gamma\theta^{1-2/\alpha}\Gamma\left(\frac{2}{\alpha}\right) + 1. \qquad (10)$$

The signal received by the typical node Z_0, if it could listen over the entire frame, is described by (2). Suppose $|\mathcal{N}(Z_0)| = K$ and the neighbors of Z_0 are indexed by $1, 2, \ldots, K$. The total number of signatures of all neighbors is $N = 2^l K$. Due to the half-duplex constraint, however, node Z_0 can only listen

---

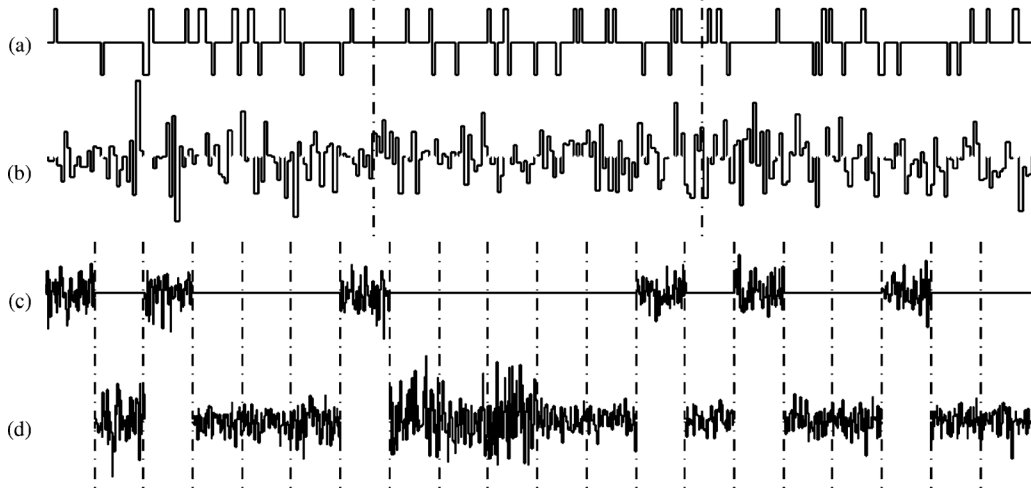[3]The optimal design of the signatures is out of the scope of this paper.

Fig. 1. Signals a node transmits and receives using RODD and slotted ALOHA, respectively: (a) RODD signaling, three consecutive frame transmissions by a node; (b) RODD signaling, three frames received by the same node (blanks represent erasures); (c) slotted ALOHA, transmissions of a node over 20 frame intervals; and (d) slotted ALOHA, signals received by the same node over 20 frame intervals. All signals are of the same symbol rate and bandwidth. The received samples are complex-valued due to the complex-valued channel gain. We only plot the in-phase component of the received signals.

during its off-slots, the number of which has binomial distribution, denoted by $M \sim \mathcal{B}(M_s, 1 - q)$, whose expected value is $\mathsf{E}\{M\} = M_s(1 - q)$. Let the matrix $\boldsymbol{S} \in \mathbb{R}^{M \times N}$ consist of columns of the signatures from all neighbors of node Z_0, observable during the M off-slots of node Z_0, and then normalized by $\sqrt{M_s q(1 - q)}$ so that the expected value of the Euclidean norm of each column in $\boldsymbol{S}$ is equal to 1. The number of rows in $\boldsymbol{S}$ for each node varies depending on the number of off-slots in that node's signature, where the standard deviation of the fluctuation in percentage is about $1/\sqrt{M_s}$, which is quite small in cases of practical interest. Based on (2), the M-vector observed through all off-slots of node Z_0 can be expressed as

$$\boldsymbol{Y} = \sqrt{\gamma_s} \boldsymbol{S} \boldsymbol{X} + \boldsymbol{W} \tag{11}$$

where

$$\gamma_s = \gamma M_s q(1 - q)/\sigma^2 \tag{12}$$

$\boldsymbol{W}$ consists of circularly symmetric complex Gaussian entries with unit variance, and $\boldsymbol{X}$ is an N-vector indicating which K signatures are selected to form the sum in (2) as well as the signal strength for each neighbor. Precisely

$$X_{(j-1)2^l + i} = U_{0j} \mathbb{1}(d_j = i)$$

for $1 \leq j \leq K$ and $1 \leq i \leq 2^l$. For example, for $K = 3$ neighbors with $l = 2$ bits of information each, where the messages are $d_1 = 3, d_2 = 2$, and $d_3 = 1$, we have

$$\boldsymbol{X} = [0 \ \ 0 \ \ U_{01} \ \ 0 \ \ 0 \ \ U_{02} \ \ 0 \ \ 0 \ \ U_{03} \ \ 0 \ \ 0 \ \ 0]^\top \tag{13}$$

where $[\cdot]^\top$ represents the transpose of a vector. The sparsity of $\boldsymbol{X}$ is exactly 2^{-l}, which is typically very small. The average system load is defined as

$$\beta = \frac{\mathsf{E}\{N\}}{M_s(1 - q)} = \frac{2^l c}{M_s(1 - q)} \tag{14}$$

where c is the average number of neighbors defined in (8).

An illustration of the RODD signaling scheme is given in Fig. 1(a) and (b). Fig. 1(a) plots the transmitted signal of a node in three consecutive frame intervals, separated by dash-dotted lines. Each frame carries one signature representing a message, which consists of mostly off-slots and a smaller number of on-slots, where the signal takes the values of ± 1. Fig. 1(b) plots the received signal of the same node, which is the superposition of all neighboring nodes' transmissions subject to fading, corrupted by noise and interference from non-neighbors. The received signal is erased whenever the node transmits (hence the blank segments in the waveform).

Given the received signal, the decoding problem node Z_0 faces is to identify, out of a total of $N = 2^l K$ signatures from all its neighbors, which K signatures were selected. This requires every node to know the codebooks of all neighbors. One solution is to let the codebook of each node be generated using a pseudo-random number generator using its network interface address (NIA) as the seed, so that it suffices to acquire all neighbors' NIAs. This, in turn, is a neighbor discovery problem, which has been studied in [16]–[18]. The discovery scheme proposed in [17] and [18] also uses RODD signaling and solves a compressed sensing problem.

IV. Sparse Recovery (Decoding) via Message Passing

The problem of recovering the support of the sparse input $\boldsymbol{X}$ based on the observation $\boldsymbol{Y}$ has been studied in the compressed sensing literature. In this section, we develop an iterative message-passing algorithm based on belief propagation. The reasons for the choice include: 1) it is one of the most competitive decoding schemes in terms of error performance; and 2) the complexity is only linear in the dimensionality of the signal to be estimated. BP belongs to a general class of message-passing algorithms for statistical inference on graphical models, which has demonstrated empirical success in many applications including error-control codes, neural networks, and multiuser detection in code-division multiple access (CDMA) systems.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

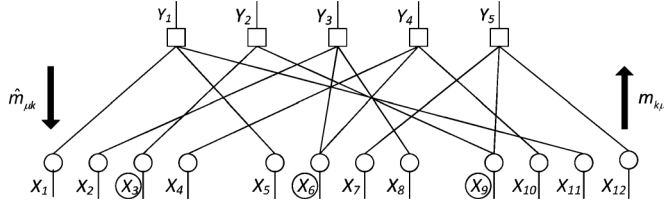ZHANG AND GUO: VIRTUAL FULL DUPLEX WIRELESS BROADCASTING VIA COMPRESSED SENSING

5

Fig. 2. Forney-style factor graph of coded mutual broadcasting. Three neighbors' signatures observable during the five measurements are as follows. Neighbor 1: [[±1 0 0 0 0]', [0 0 ±1 0 0]', [0 ±1 0 0 0]', [0 0 0±1 0]']; Neighbor 2: [[±1 0 0 0 0]', [0 0 ±1 ±1 0]', [0 0 0 0 ±1]', [0 0±1 0 0]']; Neighbor 3: [[±1 0 0 0 ±1]', [0 0 0 ±1 0]', [±1 0 0 0 0]', [0 0 0 0±1]'].

A. Factor Graph

We construct a Forney-style bipartite factor graph to represent the model (11). Here, we separate the real and imaginary parts in (11) as

$$\mathbf{Y}^{(1)} = \sqrt{\gamma_s}\mathbf{S}\mathbf{X}^{(1)} + \mathbf{W}^{(1)}, \mathbf{Y}^{(2)} = \sqrt{\gamma_s}\mathbf{S}\mathbf{X}^{(2)} + \mathbf{W}^{(2)}$$
(15)

where the superscripts (1) and (2) represent the real and imaginary parts, respectively, and $\mathbf{W}^{(i)}, i = 1, 2$ consists of i.i.d. Gaussian random variables with zero mean and variance $1/2$. The message-passing algorithm we shall develop based on (15) is not optimal, but such separation facilitates approximation and computation, which will be discussed in Section IV-B. Since two parts in (15) share the same factor graph, we treat one of them and omit the superscripts

$$y_\mu = \sqrt{\gamma_s}\sum_{k=1}^{N} s_{\mu k}x_k + w_\mu$$
(16)

where $\mu \in 1, 2, \ldots, M$ and $k \in 1, 2, \ldots, N$ index the measurements and the input "symbols," respectively. For simplicity, we ignore the dependence of the symbols $\{X_k\}$ for now, which shall be addressed toward the end of this section. Each X_k then corresponds to a symbol node, and each Y_μ corresponds to a measurement node, where the joint distribution of all $\{X_k\}$ and $\{Y_\mu\}$ are decomposed into a product of $M + N$ factors, one corresponding to each node. For every (μ, k), symbol node k and measurement node μ are connected by an edge if $s_{\mu k} \neq 0$. A simple example is shown in Fig. 2 for five measurements and three neighbors each with four messages, i.e., $M = 5, K = 3$, and $N = 3 \times 4 = 12$. The actual messages chosen by the three neighbors, $d_1 = 3, d_2 = 2$, and $d_3 = 1$, correspond to the circled variables X_3, X_6, and X_9, respectively.

B. Message-Passing Algorithm

In general, an iterative message-passing algorithm involves two steps in each iteration, where a message (or belief, which shall be distinguished from an information message) is first sent from each symbol node to every measurement node to which it is connected, and then a new set of messages are computed and sent in the reverse direction, and so forth. The algorithm performs exact inference within a finite number of iterations if there are no loops in the graph (the graph becomes a tree if it remains connected). In general, the algorithm attains a good

approximate solution for loopy graphs as the one in the current problem.

For convenience, let $\partial\mu$ (resp. ∂k) denote the subset of symbol nodes (resp. measurement nodes) connected directly to measurement node μ (resp. symbol node k), called its neighborhood.[4] Let $|\partial\mu|$ (resp. $|\partial k|$) represent the cardinality of the neighborhood of measurement node μ (resp. symbol node k). Also, let $\partial\mu\backslash k$ denote the neighborhood of measurement node μ excluding symbol node k and let $\partial k\backslash\mu$ be similarly defined.

Algorithm 1: Message-Passing Decoding Algorithm

1: *Input: $\mathbf{S}, \mathbf{Y}, \gamma_s, M_s, q$.*
2: *Initialization:*
3: $z_{\mu k}^{0,i} \leftarrow y_\mu^i/(\sqrt{\gamma_s}s_{\mu k})$ for all $s_{\mu k} \neq 0$ and $i = 1, 2$.
4: Initialize $\tau^{0,i}$ to a large positive number for $i = 1, 2$.
5: *Main iterations:*
6: **for** $t = 1$ to $T - 1$ **do**
7:     **for all** μ, k with $s_{\mu k} \neq 0$ and $i = 1, 2$ **do**
8:       $m_{k\mu}^{t,i} \leftarrow \mathsf{E}\left\{X \mid Y = \frac{\sum_{\nu\in\partial k\backslash\mu} z_{\nu k}^{t-1,i}}{|\partial k|-1}; \frac{\tau^{t-1,i}}{|\partial k|-1}\right\}$.
9:       $(\sigma_{k\mu}^{t,i})^2 \leftarrow \mathrm{var}\left\{X \mid Y = \frac{\sum_{\nu\in\partial k\backslash\mu} z_{\nu k}^{t-1,i}}{|\partial k|-1}; \frac{\tau^{t-1,i}}{|\partial k|-1}\right\}$.
10:       $z_{\mu k}^{t,i} \leftarrow \frac{1}{\sqrt{\gamma_s}s_{\mu k}}\left(y_\mu^i - \sqrt{\gamma_s}\sum_{j\in\partial\mu\backslash k} s_{\mu j}m_{j\mu}^{t,i}\right)$.
11:     **end for**
12:     $\tau^{t,i} \leftarrow \frac{1}{\sum_\mu |\partial\mu|}\sum_\mu |\partial\mu|\sum_{j\in\partial\mu}(\sigma_{j\mu}^{t,i})^2 + \frac{1}{2\gamma_s}M_s q(1-q)$ for $i = 1, 2$.
13: **end for**
14: $m_k^i \leftarrow \mathsf{E}\left\{X \mid Y = \frac{\sum_{\nu\in\partial k} z_{\nu k}^{T-1,i}}{|\partial k|-1}; \frac{\tau^{T-1,i}}{|\partial k|-1}\right\}$ for all k, $i = 1, 2$.
15: *Output: $\hat{w}_k = \arg\max_{j=1,\ldots,2^l} |m_{(k-1)2^l+j}^1 + \sqrt{-1}\,m_{(k-1)2^l+j}^2|, k = 1, \ldots, K$.*

The message-passing algorithm, given as Algorithm 1, decodes the information indexes $d_1, \ldots, d_K$ and is ready for implementation. It is based on the conventional belief propagation algorithm. Central limit theorem and some other approximation techniques are used to reduce the computational complexity, which will be discussed in detail shortly.

The superscripts $i = 1, 2$ in Algorithm 1 represent the real and imaginary parts, respectively. Here, $\mathsf{E}\{X \mid Y = y; \xi\}$ and $\mathrm{var}\{X \mid Y = y; \xi\}$ represent the conditional mean and variance of the input given the Gaussian channel output $Y = X + W$ with $W \sim \mathcal{N}(0, \xi)$ is equal to y. Mathematically, assume X has cumulative distribution function $P_X(x)$, then for $i = 1, 2, \ldots$

$$\mathsf{E}\{X^i \mid Y = y; \sigma^2\} = \frac{\int x^i e^{-\frac{(y-x)^2}{2\xi}}\,\mathrm{d}P_X(x)}{\int e^{-\frac{(y-x)^2}{2\xi}}\,\mathrm{d}P_X(x)}$$
(17)

and

$$\mathrm{var}\{X \mid Y = y; \xi\} = \mathsf{E}\{X^2 \mid Y = y; \xi\} - (\mathsf{E}\{X \mid Y = y; \xi\})^2$$
(18)

where $\int \cdot \,\mathrm{d}P_X(x)$ denotes the Riemann–Stieltjes integral.

---

[4]This is to be distinguished from the notion of neighborhood in the wireless network defined in Section II-C.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6   IEEE/ACM TRANSACTIONS ON NETWORKING

In the following, we derive Algorithm 1 starting from (16), which is valid for both real and imaginary parts in (15). It is a simplification of the original iterative BP algorithm, which iteratively computes the marginal *a posteriori* distribution of all symbols given the measurements, assuming that the graph is free of cycles. For each $k \in \partial\mu$ (hence $\mu \in \partial k$), let $\{V_{k\mu}^t(x)\}$ represent the message from symbol node k to measurement node μ at the tth iteration and $\{U_{\mu k}^t(x)\}$ represent the message in the reverse direction. Each message is basically the belief (in terms of a probability density or mass function) the algorithm has accumulated about the corresponding symbol based on the measurements on the subgraph traversed so far, assuming it is a tree. Let $p_X(x)$ denote the *a priori* probability density function of X. In the tth iteration, as in the belief propagation algorithm [19], [20], $V_{k\mu}^t(x)$ is computed by combining the prior information $p_X(x)$ and messages from measurement nodes in the previous iteration $U_{\nu k}^{t-1}(x)$, and $U_{\mu k}^t(x)$ is calculated based on the Gaussian channel model and the messages from symbol nodes in the current iteration $V_{j\mu}^t(x)$, which involves a integral of the product of Gaussian noise pdf and $V_{j\mu}^t(x)$'s for all $j \in \partial\mu\backslash k$. Therefore, we have

$$V_{k\mu}^t(x) \propto p_X(x) \prod_{\nu \in \partial k\backslash \mu} U_{\nu k}^{t-1}(x) \tag{19a}$$

for all (k, μ) with $s_{\mu k} \neq 0$, and then

$$U_{\mu k}^t(x) \propto \int_{(x_j)_{\partial\mu\backslash k}} \exp\left[-\left(y_\mu - \sqrt{\gamma_s}s_{\mu k}x\right. \right.$$
$$\left.\left. -\sqrt{\gamma_s}\sum_{j\in\partial\mu\backslash k} s_{\mu j}x_j\right)^2\right] \prod_{j\in\partial\mu\backslash k}\left(V_{j\mu}^t(x_j)\mathrm{d}x_j\right) \tag{19b}$$

where $\int_{(x_j)_{\partial\mu\backslash k}}$ denotes integral over all x_j with $j \in \partial\mu\backslash k$, and $V(x) \propto u(x)$ means that $V(x)$ is proportional to $u(x)$ with proper normalization such that $\int_{-\infty}^{\infty} V(x)\mathrm{d}x = 1$. In case X is a discrete random variable, the integral shall be replaced by a sum over the alphabet of X. In this problem, X follows a mixture of discrete and continuous distributions, so the expectation can be decomposed as an integral and a sum.

The complexity of computing the integral in (19b) is exponential in $|\partial\mu| = \mathcal{O}(qN)$, which is in general infeasible for the problem at hand. However, as $qN \gg 1$, the computation carried out at each measurement node admits a good approximation by using the central limit theorem. A similar technique has been used in the CDMA detection problem, for fully connected bipartite graph in [19], [21], and [22], and for a graph with large node degrees in [20].

To streamline (19a) and (19b), we introduce $m_{k\mu}^t$ and $(\sigma_{k\mu}^t)^2$ for all (μ, k) pairs with $s_{\mu k} \neq 0$ to represent the mean and variance of a random variable with distribution $V_{k\mu}^t(x)$. Using Gaussian approximation, one can reduce the message-passing algorithm to iteratively computing the following messages with the initial conditions that $z_{\mu k}^0 = y_\mu / \left(\sqrt{\gamma_s}s_{\mu k}\right)$ and τ^0 is a large positive number:

$$m_{k\mu}^t = \mathsf{E}\left\{X \,\middle|\, Y = \frac{\sum_{\nu\in\partial k\backslash\mu} z_{\nu k}^{t-1}}{|\partial k| - 1} \,;\, \frac{\tau^{t-1}}{|\partial k| - 1}\right\} \tag{20a}$$

$$(\sigma_{k\mu}^t)^2 = \mathsf{var}\left\{X \,\middle|\, Y = \frac{\sum_{\nu\in\partial k\backslash\mu} z_{\nu k}^{t-1}}{|\partial k| - 1} \,;\, \frac{\tau^{t-1}}{|\partial k| - 1}\right\} \tag{20b}$$

$$z_{\mu k}^t = \frac{1}{\sqrt{\gamma_s}s_{\mu k}}\left(y_\mu - \sqrt{\gamma_s}\sum_{j\in\partial\mu\backslash k} s_{\mu j}m_{j\mu}^t\right) \tag{20c}$$

$$\tau^t = \frac{1}{\sum_\mu |\partial\mu|}\sum_\mu |\partial\mu|\sum_{j\in\partial\mu}(\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s}M_s q(1-q) \tag{20d}$$

where (20a) and (20b) calculate the conditional expectation and variance, respectively. The detailed derivation is relegated to Appendix A. At the Tth iteration, the approximated posterior mean of x_k can be expressed as

$$m_k = \mathsf{E}\left\{X \,|\, Y = \frac{\sum_{\nu\in\partial k} z_{\nu k}^{T-1}}{|\partial k| - 1} \,;\, \frac{\tau^{T-1}}{|\partial k| - 1}\right\}. \tag{21}$$

It is time-consuming to compute (20a) and (20b) for all (μ, k) pairs with $s_{\mu k} \neq 0$, especially in the case of large matrix $\boldsymbol{S}$. We can use the following two approximation techniques to further reduce the computational complexity. First, $|\partial k|$ in (20a), (20b), and (21) is replaced by its mean value $M_s q(1-q)$. Second, we use interpolation and extrapolation to further reduce the computation complexity of (20a), (20b), and (21). Specifically, in each iteration t, we only compute the conditional mean and variance for some chosen y's, i.e., we choose $y_1^t < y_2^t < \cdots < y_n^t$, which is a partition of an interval depending on τ^{t-1}, and compute

$$a_j^t = \mathsf{E}\left\{X \,\middle|\, Y = y_j^t; \frac{\tau^{t-1}}{M_s q(1-q)-1}\right\} \tag{22}$$

$$b_j^t = \mathsf{var}\left\{X \,\middle|\, Y = y_j^t; \frac{\tau^{t-1}}{M_s q(1-q)-1}\right\} \tag{23}$$

for $j = 1, 2, \ldots, n$, and then use those values to calculate (20a), (20b), and (21) by interpolation or extrapolation. To be more precise, for any pair μ, k with $s_{\mu k} \neq 0$, suppose y_j^t and y_{j+1}^t are chosen to be the closest to

$$y = \frac{\sum_{\nu\in\partial k\backslash\mu} z_{\nu k}^{t-1,i}}{|\partial k| - 1} \tag{24}$$

then $m_{k\mu}^t$ in (20a) and $(\sigma_{k\mu}^t)^2$ in (20b) can be approximated by

$$m_{k\mu}^t = a_j^t + \frac{y - y_j^t}{y_{j+1}^t - y_j^t}\left(a_{j+1}^t - a_j^t\right) \tag{25}$$

$$(\sigma_{k\mu}^t)^2 = b_j^t + \frac{y - y_j^t}{y_{j+1}^t - y_j^t}\left(b_{j+1}^t - b_j^t\right). \tag{26}$$

Similarly, m_k in (21) can be approximately calculated.

We now revisit the assumption that $\boldsymbol{X}$ has independent elements. In fact, $\boldsymbol{X}$ consists of K subvectors of length 2^l, where the entries of each subvector are all zero except for one position corresponding to the transmitted message. After obtaining the approximated posterior mean $\widetilde{m}_k$ by incorporating both real and imaginary parts calculated from (21), Algorithm 1 outputs the position of the element with the largest magnitude in each of the K subvectors of $[\widetilde{m}_1, \ldots, \widetilde{m}_N]$. In fact, the factor graph Fig. 2 can be modified to include K additional nodes, each of which

puts a constraint on one subvector. Slight improvement over Algorithm 1 may be obtained by carrying out message passing on the modified graph.

The performance of Algorithm 1 has been analyzed in [23, Sec. 4.4.3] in the so-called *large-system limit*, where the frame length and the number of messages from which a node can choose both tend to infinity with a fixed ratio. The evolution of the error rate achieved by Algorithm 1 with different numbers of iterations is asymptotically characterized by a fixed-point equation. The analysis uses techniques developed for statistical inference through noisy large linear systems. It is not the focus of this paper and thus is omitted.

The number of iterations, T, needed to achieve good performance in practice is typically not large. In the simulation results given in Section VI, we use $T = 16$ iterations.

V. RANDOM ACCESS SCHEMES

In this section, we describe two random access schemes, namely slotted ALOHA and CSMA, and provide lower bounds on the message error probability. The results will be used in Section VI to compare to the performance of RODD.

Suppose node transmissions are synchronized. The nominal SNR in each slot is the same as in (1), also denoted by γ. The channel model, network model, and propagation model (including Rayleigh fading) are as introduced in Section II.

Let L denote the total number of bits encoded into a frame, which includes an l-bit message and a few additional bits that identify the sender. This is in contrast to broadcasting via compressed sensing, where the signature itself identifies the sender (and carries the message). Each broadcasting period consists of a number of frames to allow for retransmissions. A message is assumed to be decoded correctly if the signal-to-interference-plus-noise ratio (SINR) in the corresponding frame transmission exceeds a threshold δ (multipacket reception is possible only if $\delta < 1$). Over the additive white noise channel with SINR δ, in order to send L bits reliably through the channel, the number of symbols in a frame must exceed $L/\log_2(1+\delta)$. Therefore, the number of frames in a period of M_r symbol intervals should satisfy

$$N_r \leq M_r \log_2(1+\delta)/L. \quad (27)$$

Without loss of generality, we still consider the typical node Z_0 at the origin. An error event is defined as that node Z_0 cannot correctly recover the message from one specific neighbor after a period of M_r symbol intervals. The corresponding error probabilities achieved by slotted ALOHA and CSMA are denoted by P_a^e and P_c^e, respectively. For ease of discussion, we allow the total number of symbol intervals, M_r, to be any positive integer, which may not be a multiple of the frame length. This results in underestimated number of intervals needed by the random access schemes to attain the desired performance.

A. Slotted ALOHA

In slotted ALOHA, suppose each node chooses independently with the same probability p to transmit in every frame interval. Fig. 1(c) illustrates signals transmitted by a typical node over 20 frame intervals, separated by dash-dotted lines.

In each of the six active frame intervals, capacity-achieving Gaussian signaling is used. The node listens to the channel over the remaining frame intervals to receive the signal shown in Fig. 1(d). The received signals during the node's own transmitting frames are erased. During some frame intervals, the received signals appear to be strong, which implies that one or more neighbors have transmitted. During some other frame intervals, the received signals appear to be weak, which consist of only noise and interference from non-neighbors.

Let Z denote one specific neighbor of node Z_0 and G denote the fading coefficient between them. Suppose the mark of Z is denoted by $\mathcal{G}$. Given that $(Z, \mathcal{G}) \in \hat{\Phi}$ where $\hat{\Phi}$ is given by (7), denote $\hat{\Phi}_1 = \hat{\Phi} \backslash \{(Z, \mathcal{G})\}$, which is also a marked p.p.p. with intensity λ. For a given realization of (Z, G) and $\hat{\Phi}_1$, define $\mathsf{P}_a^s(Z, G, \hat{\Phi}_1)$ as the probability that the received SINR from Z to Z_0 exceeds the threshold δ conditioning on that Z transmits in a given frame. In any given frame, the probability of the event that Z transmits, Z_0 listens, and the transmission is successful is thus $p(1-p)\mathsf{P}_a^s(Z, G, \hat{\Phi}_1)$. Therefore, the probability that the message from Z has not been successfully received by Z_0 after N_r consecutive frame intervals can be expressed as

$$\mathsf{P}_a^e = \mathsf{E}\left\{\left(1 - p(1-p)\mathsf{P}_a^s(Z, G, \hat{\Phi}_1)\right)^{N_r}\right\} \quad (28)$$

where the expectation is over the joint distribution $(Z, G, \hat{\Phi}_1)$. Due to the convexity of function $(\max\{0, 1-z\})^n, z \geq 0, n \in \{1, 2, \ldots\}$, P_a^e in (28) can be lower-bounded as

$$\mathsf{P}_a^e \geq \left(\max\left\{0, 1 - p(1-p)\mathsf{E}\left\{\mathsf{P}_a^s(Z, G, \hat{\Phi}_1)\right\}\right\}\right)^{N_r}. \quad (29)$$

In Appendix B, the expectation of $\mathsf{P}_a^s(Z, G, \hat{\Phi}_1)$ is upper-bounded using the known Laplace transform of the distribution of the interference [13]. For a period of M_r symbol intervals, the lower bound on P_a^e is presented in the following result.

Proposition 1: Consider an arbitrary neighbor Z of node Z_0. The probability that Z_0 cannot successfully receive the message from Z after a period of M_r symbol intervals is lower-bounded as follows:

$$\mathsf{P}_a^e \geq \left(\max\left\{0, 1 - \frac{1}{\pi}p(1-p)\left(\frac{\theta}{\delta}\right)^b \sin\left(\frac{b\pi}{2}\right)\Gamma(1-b)\right.\right.$$
$$\left.\left. \times \int_{-\infty}^{\infty} |\omega|^{b-1}\mathsf{E}\left\{-\lambda p\frac{b\pi^2}{\sin(b\pi)}(\iota\omega)^b - \iota\frac{\omega}{\gamma}\right\}d\omega\right\}\right)^{n_r} \quad (30)$$

where $\iota = \sqrt{-1}, b = 2/\alpha$ and $n_r = M_r \log_2(1+\delta)/L$.

Although (30) appears to be complicated, computing it only involves a straightforward single-variable integral (the outcome of the integral is in fact real-valued).

In the slotted ALOHA scheme, despite repeated transmissions, a given link may still fail to deliver the message due to the half-duplex constraint (the receiver happens to transmit during the same frame) and consistently weak received SINR due to random interference from other links.

B. CSMA

As an improvement over ALOHA, CSMA lets nodes use a brief contention period to negotiate a schedule in such a way

that nodes in a small neighborhood do not transmit data simultaneously. We analyze the performance of CSMA by using the Matérn hard core model [13]. To be specific, consider the following generic scheme: Each node senses the channel continuously; if the channel is busy, the node remains silent and disables its timer; as soon as the channel becomes available, the node starts its timer with a random offset and waits until the timer expires to transmit its frame. Clearly, the node whose timer expires first in its neighborhood captures the channel and transmits its frame.

Mathematically, let $\{T_i\}$ be i.i.d. random variables with uniform distribution on $[0,1]$, which represent the timer offsets for all nodes $\{Z_i\}$ in Φ, respectively. Node Z_i will transmit its frame if and only if $T_i < T_j$ for all $X_j \in \mathcal{N}(Z_i)$.

By viewing T_i as a mark of node Z_i, we redefine $\tilde{\Phi}$ and $\hat{\Phi}$ as $\tilde{\Phi} = \{(Z_i, \mathcal{G}_i, T_i)\}_i$ and $\hat{\Phi} = \tilde{\Phi} \backslash (Z_0, \mathcal{G}_0, T_0)$, respectively. Let Z be one specific neighbor of Z_0 and T denote its time offset. Define G and $\mathcal{G}$ as in Section V-A. Given that $(Z, \mathcal{G}, T) \in \hat{\Phi}$, denote $\hat{\Phi}_1 = \hat{\Phi} \backslash \{(Z, \mathcal{G}, T)\}$, which is still a marked p.p.p. with intensity λ. For a given realization of (Z, G, T) and $\hat{\Phi}_1$, define $\mathsf{P}_c^s(Z, G, T, \hat{\Phi}_1)$ as the probability that node Z transmits its frame and the received SINR from Z to Z_0 exceeds the threshold δ. Therefore, the probability that the message from Z has not been successfully received after N_r consecutive frame intervals can be expressed as

$$\mathsf{P}_c^e = \mathsf{E}\left\{\left(1 - \mathsf{P}_c^s(Z, G, T, \hat{\Phi}_1)\right)^{N_r}\right\}$$
$$\geq \left(\max\left\{0, 1 - \mathsf{E}\left\{\mathsf{P}_c^s(Z, G, T, \hat{\Phi}_1)\right\}\right\}\right)^{N_r} \quad (31)$$

where the expectation is over the joint distribution of $(Z, G, T, \hat{\Phi}_1)$, and (31) is due to the convexity of function $(\max\{0, 1-z\})^n, z \geq 0, n \in \{1, 2 \ldots\}$.

For a period of M_r symbol intervals, the lower bound on error probability P_c^e is given by the following result, which is proved in Appendix C.

Proposition 2: Consider an arbitrary neighbor Z of node Z_0. The probability that Z_0 cannot successfully receive the message from Z after a period of M_r symbol intervals is lower-bounded as follows:

$$\mathsf{P}_c^e \geq \left(\max\left\{0, 1 - \frac{1}{c^2}\left(\frac{\theta\gamma}{\delta}\right)^{\frac{2}{\alpha}}\left(e^{-c} + c - 1\right)\right\}\right)^{n_r} \quad (32)$$

where c is defined in (8) and $n_r = M_r \log_2(1+\delta)/L$.

In contrast to slotted ALOHA, frame loss due to the half-duplex constraint is eliminated through contention. However, a given link may still fail to deliver the message after repeated transmissions because the received SINRs were consistently weak due to random interference outside the neighborhood.

VI. NUMERICAL RESULTS

In order for a fair comparison, we assume the same power constraint for both the compressed sensing scheme and random access schemes, i.e., the average transmit power in each active slot (in which the node transmits energy) is the same. We choose the same transmission probability in each slot for the
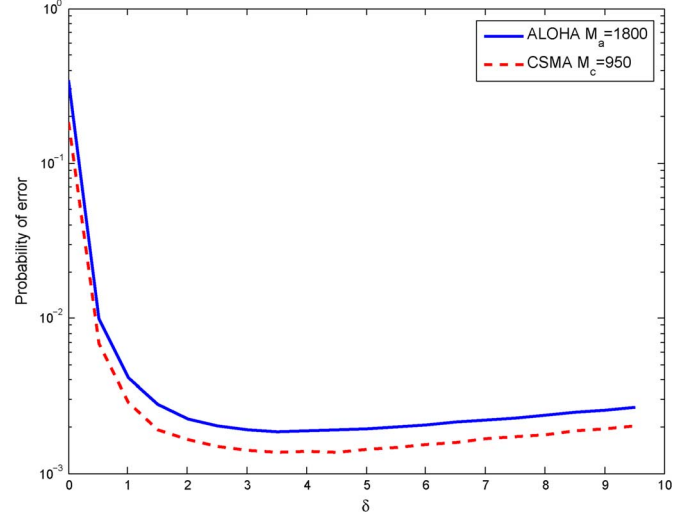

Fig. 3. Lower bounds for error probability in slotted ALOHA and CSMA for different threshold δ in the case of $l = 10$.

compressed sensing and slotted ALOHA schemes, i.e., $q = p = 1/(c+1)$. Also, the transmission probability in each slot for CSMA is $(1 - e^{-c})/c$ (see Appendix C), which is close to $1/(c+1)$ when c is large. The three schemes consume approximately the same amount of average power over any period of time.

Without loss of generality, let one unit of distance be 1 m. Consider a wireless network of 1000 nodes uniformly distributed in a square with side length of 500 m. The nodes form a Poisson point process in the square conditioned on the node population. Suppose the path-loss exponent $\alpha = 4$. The threshold of channel gain to define neighborhood is set to $\theta = 10^{-6}$. It means that if the SNR from a node 1 m away is 60 dB, then the SNR attenuates to 0 dB (the neighborhood boundary) at $10^{6/\alpha} \approx 31$ m due to path loss only. As both path loss and fading are considered, a node near the center of the square (without boundary effect) has on average $c \approx 11$ neighbors according to (8).

We consider two cases for the length of broadcasting, with $l = 5$ and 10 bits, respectively. In random access schemes, a packet of L bits consists of l-bit message and $\lceil \log_2 c \rceil$ additional bits to identify the sender. Fig. 3 shows that $\delta = 3.5$ minimizes the lower bounds for P_a^e in (30) and P_c^e in (32) in the case of $l = 10$.

The metric for performance comparison is the probability for one node to miss one specific neighbor, averaged over all pairs of neighboring nodes in the network. Suppose the transmit SNR of each node is $\gamma = 60$ dB. First, consider one realization of the network where each node has $c \approx 11$ neighbors on average and $l = 5$ bits to broadcast, so that on average $cl \approx 55$ bits are to be collected by each node. In slotted ALOHA, at least four additional bits are needed to identify a sender out of on average 11 neighbors, so we let $L = 9$. In Fig. 4, the error performance of slotted ALOHA and CSMA for $\delta = 0.5$ is compared to that of the compressed sensing scheme with the message-passing algorithm. The simulation result shows that the compressed sensing scheme significantly outperforms slotted ALOHA and CSMA, even compared to the minimum of the lower bounds computed
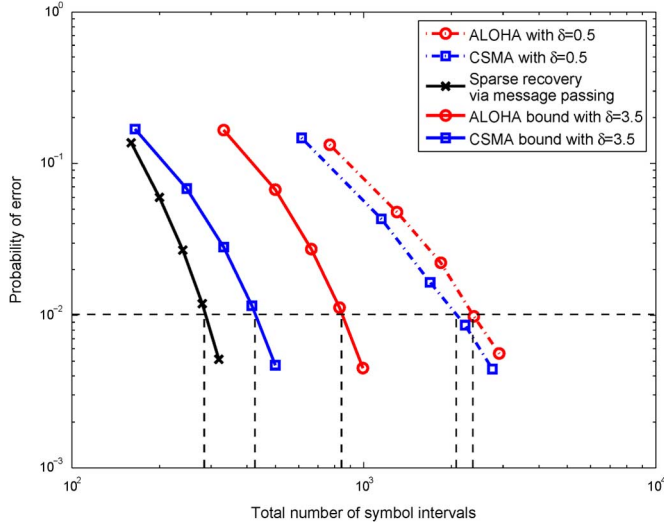
This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ZHANG AND GUO: VIRTUAL FULL DUPLEX WIRELESS BROADCASTING VIA COMPRESSED SENSING

9

Fig. 4. Performance comparison between the compressed sensing and random access schemes. Each node transmits a 5-bit message.

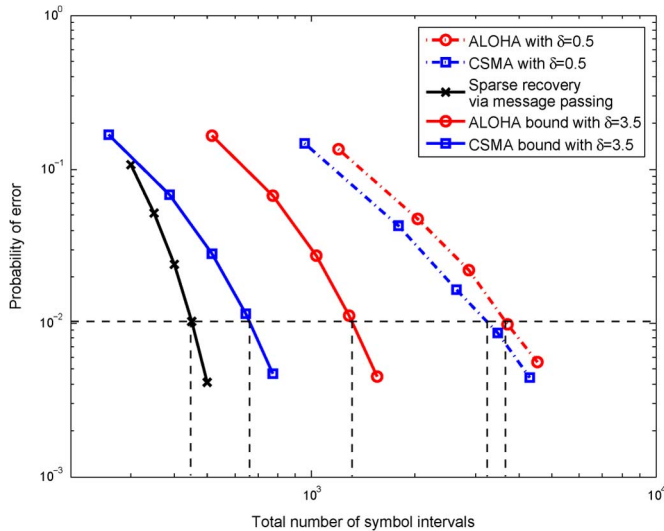Fig. 6. Performance of the compressed sensing scheme with different nominal SNRs (γ).

Fig. 5. Performance comparison between the compressed sensing and random access schemes. Each node transmits a 10-bit message.

Fig. 7. Performance of the compressed sensing scheme with different node densities.

from (30) and (32) for $\delta = 3.5$. For example, to achieve 1% error rate, the compressed sensing scheme takes fewer than 300 symbols. Slotted ALOHA and CSMA take no less than 800 and 400 symbols according to the bounds in (30) and (32) for $\delta = 3.5$, respectively. In fact, slotted ALOHA and CSMA with threshold $\delta = 0.5$ take more than 2000 symbols. Similar comparison is observed for several other SINR thresholds δ around 0.5, and the performance of ALOHA and CSMA are not good for $\delta \geq 1$ because the messages from weaker neighbors may never be successfully delivered. Some additional supporting numerical evidence is, however, omitted due to space limitations.

Fig. 5 repeats the experiment of Fig. 4 with 10-bit messages. The compressed sensing scheme has significant gain compared to slotted ALOHA and CSMA. For example, to achieve the error rate of 1%, the compressed sensing scheme takes about
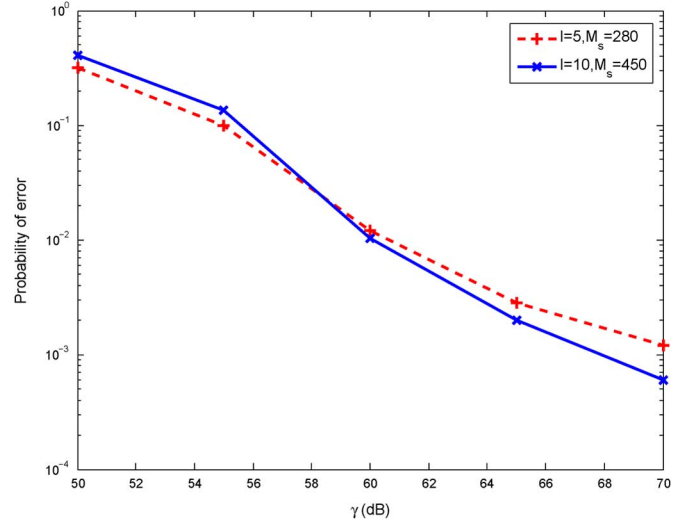
450 symbols, whereas slotted ALOHA and CSMA take at least 1000 and 650 symbols, respectively.

In Fig. 6, we simulate the same network with different nominal SNRs, i.e., γ varies from 50 to 70 dB. In the case that each node transmits a 5-bit message, the frame length is chosen to be 280 symbols. It can be seen from the figure that the probability of error decreases with the increase of SNR. The performance is similar when each node transmits a 10-bit message and the frame consists of 450 symbols.

Fig. 7 plots the probability of error achieved by our compressed sensing scheme as a function of the number of nodes on the 500×500-m^2 square where all other parameters are held constant. In particular, the nominal SNR $\gamma = 60$ dB. The frame is of 280 symbol intervals and each message consists of 5 bits. As the number of nodes increases from 500 to 1500, the average number of neighbors a node has increases from about 5 to about

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

10

IEEE/ACM TRANSACTIONS ON NETWORKING

16, where the frame error rate increases gracefully from about 0.1% to 7%.

VII. DISCUSSION AND CONCLUDING REMARKS

This study consists of two integrated components: One is the compressed sensing scheme for nodes to simultaneously broadcast to one-hop neighbors; the other is the on–off signaling that enables virtual full-duplex communication. Importantly, if alternative techniques are used to enable full duplex, the compressed sensing scheme applies equally well, where the codewords need not be on–off.

In the following, we provide further discussion of the proposed technology for wireless broadcasting, especially the advantages of RODD over related well-known schemes.

A. RODD Versus CDMA

In both RODD and direct sequence CDMA, nodes simultaneously transmit data bearing sequences. Their similarity ends there. In particular, their different timescales, access modes, and duplex schemes set them apart: 1) A CDMA spreading sequence spans over one *symbol interval* to carry one symbol, while an RODD sequence or signal spans over one *frame interval* to represent one frame of data. Typically, frames are the units of error control coding; that is, all symbols in a given frame form a codeword, whereas different frames are coded separately. Thus, the RODD sequence is one codeword, whereas in CDMA, one codeword spans over many spreading sequences (often repetitions of the same sequence); 2) RODD is designed for many-to-many transmission, whereas CDMA is a many-to-one multiaccess scheme; 3) RODD signaling enables full-duplex communication at frame level, whereas CDMA is half-duplex in the absence of self-interference cancellation.

B. RODD Versus TDMA

Time-division multiple access (TDMA), which is suitable for many-to-one communication, is difficult to apply in many-to-many communication in a large network, where different nodes see different neighborhoods. In the scenario considered in this paper, every node in a large network wishes to broadcast its message to all neighbors, and also wishes to receive messages from all neighbors. To generate a time-division schedule for all nodes to avoid collision in every neighborhood that is throughput optimal is an NP-hard problem. Using *ad hoc* solutions leads to a highly conservative schedule with low throughoput, whereas using a more aggressive schedule causes collisions and require complicated scheduling of retransmissions. Furthermore, a time-division schedule needs to be recomputed if a node moves in or out of a neighborhood, whereas the proposed RODD scheme is robust to network topology changes.

C. RODD Versus Random Access

The on–off signaling of RODD resembles that of ALOHA at a much faster timescale. There are, however, crucial differences. As frames are usually the units of error control coding, each frame (or a large on-slot) is coded separately in ALOHA or CSMA. In RODD, each frame consists of many (short) slots, so that we code over all the on-slots. From an individual node's viewpoint, other nodes' transmissions are seen as interference. If nodes use RODD signaling, the node experiences ergodic interference, whereas if nodes use random access schemes, the node experiences nonergodic interference. The former channel is ergodic because the interference fluctuates at slot level, but statistically the channel remains the same in every frame. The latter channel is nonergodic because the channel fluctuates at frame level and appears very different in different frames. The node may see no interference in some frames, but a lot of interference in other frames. When the channel is nonergodic, the node cannot predict the interference level, hence the node does not know the best data rate to transmit. If the node transmits at high rate, the frame may be lost when other nodes also transmit. If the node transmits at low rate, channel resources are wasted if no other nodes transmit at the same time. This problem is entirely overcome by RODD signaling. Since the channels of all frames look statistically the same for a given node, all frames can be coded at the same rate up to the capacity of the channel and be decoded reliably by the receiving node.

One might argue that it would be just as good to code over many frames when random access is used to also yield an ergodic channel. The problem with this is that many frames have to be received before decoding them, hence the decoding delay would be exceedingly large. For the same maximum decoding delay, random access achieves much lower rate compared to RODD.

D. RODD Versus Interference Cancellation

State-of-the-art MAC protocols are mostly designed based on the packet collision model for wireless networks, where if multiple nodes simultaneously transmit, their transmissions fail due to collision at the receiver. In contrast, RODD signaling takes full advantage of the superposition nature of the wireless medium. By coding over the entire frame of on- and off-slots and joint decoding of multiple users, signals in colliding slots are also fully utilized.

Other recent works such as [24] and [25] break away from the collision model. The basic idea is that when two senders transmit simultaneously, their packets superpose at the receiver, so that if the receiver already knows the content of one of the packets, it can cancel the interference and decode the other packet. Not only can RODD take advantage of such known-interference cancellation techniques, it is also suitable for many-to-many communication, whereas it is difficult to perform interference cancellation for more than two users by the techniques in [24] and [25].

E. Decoding Delay

The decoding delay of RODD is fixed to one frame interval, which is typically a few hundred to a few thousand symbols. As shown through simulations in Section VI, the delay of random access schemes for achieving the same error rate is many times larger than that of RODD. Admittedly, the delay of random access can be as short as a short frame (tens of symbols), if luckily no other nodes happen to transmit at the same time. However, the short frame is very likely to be lost due to collision, and many retransmissions are needed to achieve a desired performance. RODD has an advantage if a fixed small delay is more

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ZHANG AND GUO: VIRTUAL FULL DUPLEX WIRELESS BROADCASTING VIA COMPRESSED SENSING
11

desirable than a variable delay that has much larger expected value.

F. Computational Complexity

The computational complexity of the message-passing algorithm is linear in the frame length, the number of neighbors, and the number of messages from which each node can choose. RODD based on compressed sensing with random signatures is most suitable for the situation where the broadcasts consist of a small number of bits. If each node has many bits to send, a structured code with low decoding complexity is needed for the scheme to be practical. One such code is the Reed–Muller code considered in [18].

G. Other Applications

RODD can serve as a highly desirable sublayer of any network protocol stack to provide the important function of simultaneous message exchange among neighbors. This sublayer provides the missing link in many advanced resource allocation schemes, where it is often *assumed* that nodes are provided the state and/or demand of their neighbors.

Finally, the idea of using on–off signaling to achieve full-duplex communication using half-duplex radios applies to general peer-to-peer networks and is not limited to mutual broadcasting traffic focused on in this paper.

APPENDIX A
DERIVATION OF MESSAGE COMPUTATION (20)

We derive (20) from (19). Denote $\Delta_{\mu k} = \sum_{j \in \partial \mu \setminus k} s_{\mu j} x_j$. The key to the simplification is to recognize that $\Delta_{\mu k}$ is approximately Gaussian. To be precise, if $\{x_j\}_{j \in \partial \mu \setminus k}$ were independent (conditioned on the observations traversed so far on the graph), then by central limit theorem, $\Delta_{\mu k}$ converges weakly to a Gaussian random variable, whose mean is

$$v_{\mu k}^t = \sum_{j \in \partial \mu \setminus k} s_{\mu j} m_{j\mu}^t \tag{33}$$

and variance is

$$(\sigma_{\mu k}^t)^2 = \sum_{j \in \partial \mu \setminus k} s_{\mu j}^2 (\sigma_{j\mu}^t)^2 . \tag{34}$$

Using the preceding Gaussian approximation, (19b) can be calculated by a change of probability measure as

$$U_{\mu k}^t(x) \propto \int_{-\infty}^{\infty} \exp \left[-(y_\mu - \sqrt{\gamma_s} s_{\mu k} x - \sqrt{\gamma_s} \Delta)^2 \right]$$
$$\cdot \frac{1}{\sqrt{2\pi(\sigma_{\mu k}^t)^2}} \exp \left[-\frac{1}{2(\sigma_{\mu k}^t)^2} (\Delta - v_{\mu k})^2 \right] \mathrm{d}\Delta$$
$$\propto \exp \left[-\frac{1}{2\tau_{\mu k}^t} (x - z_{\mu k}^t)^2 \right] \tag{35}$$

where $z_{\mu k}^t$ is defined in (20c) and

$$\tau_{\mu k}^t = \sum_{j \in \partial \mu \setminus k} (\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s s_{\mu k}^2} . \tag{36}$$

Using law of large numbers, we further approximate $\tau_{\mu k}^t$ by its average over all (μ, k) pairs with $s_{\mu k} \neq 0$, i.e., $\tau_{\mu k}^t$ is replaced by

$$\tau^t = \frac{1}{\sum_\mu |\partial \mu|} \sum_k \sum_{\mu \in k} \sum_{j \in \partial \mu \setminus k} (\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s s_{\mu k}^2}$$
$$\approx \frac{1}{\sum_\mu |\partial \mu|} \sum_\mu |\partial \mu| \sum_{j \in \partial \mu} (\sigma_{j\mu}^t)^2 + \frac{1}{2\gamma_s} M_s q(1-q) \tag{37}$$

as shown in (20d).

Now we have $U_{\mu k}^t \sim \mathcal{N}(z_{\mu k}^t, \tau^t)$, so it is easy to see that

$$\prod_{\nu \in \partial k \setminus \mu} U_{\nu k}^t \sim \mathcal{N} \left(\frac{\sum_{\nu \in \partial k \setminus \mu} z_{\nu k}^t}{|\partial k| - 1}, \frac{\tau^t}{|\partial k| - 1} \right) . \tag{38}$$

According to (19a), $V_{k\mu}^{t+1}(x)$ can be viewed as the conditional distribution $p_{X|Y} \left(x \, \middle| \, \frac{\sum_{\nu \in \partial k \setminus \mu} z_{\nu k}^t}{|\partial k| - 1} \right)$, where Y is the output of a Gaussian channel with noise $W \sim \mathcal{N}(0, \frac{\tau^t}{|\partial k| - 1})$. Therefore, by definition, $m_{k\mu}^{t+1}$ and $(\sigma_{k\mu}^{t+1})^2$ can be expressed as the conditional mean and conditional variance as in (20a) and (20b).

APPENDIX B
PROOF OF PROPOSITION 1

Let $\hat{\Phi}_1^p$ be an independent thinning of $\hat{\Phi}_1$ with retention probability p to represent the transmitting nodes. It is easy to see that $\hat{\Phi}_1^p$ is an independent marked p.p.p. with intensity λp. Denote the interference by

$$I = \sum_{(Z_i, \mathcal{G}_i) \in \hat{\Phi}_1^p} G_{0i} |Z_i|^{-\alpha} \tag{39}$$

then we have

$$\mathsf{E} \left\{ \mathsf{P}_a^s(Z, \mathcal{G}, \hat{\Phi}_1) \right\} = \mathsf{E} \left\{ \mathsf{E} \left\{ 1 \left(\frac{\gamma G |Z|^{-\alpha}}{\gamma I + 1} \geq \delta \right) \, \middle| \, \hat{\Phi}_1^p \right\} \right\}$$
$$= \mathsf{E} \left\{ \mathsf{P} \left\{ G |Z|^{-\alpha} \geq \delta \left(I + \frac{1}{\gamma} \right) \, \middle| \, \hat{\Phi}_1^p \right\} \right\}$$
$$\leq \mathsf{E} \left\{ \left(\frac{\theta}{\delta} \right)^{\frac{2}{\alpha}} \left(I + \frac{1}{\gamma} \right)^{-\frac{2}{\alpha}} \right\} \tag{40}$$
$$= \left(\frac{\theta}{\delta} \right)^{\frac{2}{\alpha}} \int_{-\infty}^{\infty} \left| i + \frac{1}{\gamma} \right|^{-\frac{2}{\alpha}} p_I(i) \mathrm{d}i \tag{41}$$

where (40) is derived from (5) and p_I is the pdf of I.

Using the Laplace transform of p_I given in [13], the Fourier transform[5] of p_I is obtained as

$$\mathcal{F}_{p_I}(\omega) = \exp \left\{ -\lambda p(\iota \omega)^{2/\alpha} \frac{2\pi^2}{\alpha \sin(2\pi/\alpha)} \right\} . \tag{42}$$

[5]The reasons to work with Fourier transform in lieu of Laplace transform are: 1) The inverse Fourier transform here is easier to calculate; 2) the Fourier transform of $|i + \frac{1}{\gamma}|^{-\frac{2}{\alpha}}$ has a closed form.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

12   IEEE/ACM TRANSACTIONS ON NETWORKING

Since the Fourier transform of $|x|^a$ for $-1 < a < 0$ is

$$\mathcal{F}_{|x|^a}(\omega) = -\frac{2\sin(a\pi/2)\Gamma(a+1)}{|\omega|^{a+1}} \quad (43)$$

the Fourier transform of

$$q_I(i) = \left| i + \frac{1}{\gamma} \right|^{-\frac{2}{\alpha}} \quad (44)$$

for $\alpha > 2$ can be expressed as

$$\mathcal{F}_{q_I}(\omega) = e^{\iota\omega/\gamma} \frac{2\sin(\pi/\alpha)\Gamma(1-2/\alpha)}{|\omega|^{1-2/\alpha}}. \quad (45)$$

Since the integral in (41) can be viewed as the Fourier transform of $p_I(i)q_I(i)$ at $\omega = 0$, it can be calculated as the convolution of $\mathcal{F}_{p_I}(\omega)$ and $\mathcal{F}_{q_I}(\omega)$ at $\omega = 0$ [26]. Therefore, by (42) and (45), we have

$$\int_{-\infty}^{\infty} \left| i + \frac{1}{\gamma} \right|^{-\frac{2}{\alpha}} p_I(i)\mathrm{d}i = \frac{1}{2\pi} \mathcal{F}_{p_I}(\omega) * \mathcal{F}_{q_I}(\omega) \bigg|_{\omega=0} \quad (46)$$

where $*$ is the convolution operator. Therefore, according to (29), (41), and (46), the error probability $\mathsf{P}_\mathrm{a}^\mathrm{e}$ can be lower-bounded as

$$\mathsf{P}_\mathrm{a}^\mathrm{e} \geq \left(\max\left\{ 0, 1 - \frac{1}{2\pi}p(1-p)\left(\frac{\theta}{\delta}\right)^{\frac{2}{\alpha}} \right. \right.$$
$$\left. \left. \times \mathcal{F}_{p_I}(\omega) * \mathcal{F}_{q_I}(\omega) \bigg|_{\omega=0} \right\} \right)^{N_r}. \quad (47)$$

Therefore, (30) in Proposition 1 follows from (47) and (27).

APPENDIX C
PROOF OF PROPOSITION 2

For any $(Z_i, \mathcal{G}_i, T_i) \in \hat{\Phi}_1$, denote G_i as the fading coefficient from node Z_i to node Z. Define the following indicators for node X

$$F_1 = 1(T_0 > T) \quad (48)$$
$$F_2 = 1(T_i > T, \forall (Z_i, \mathcal{G}_i, T_i) \in \hat{\Phi}_1 \text{ with } G_i|Z_i - Z|^{-\alpha} \geq \theta) \quad (49)$$
$$F_3 = 1(\gamma G|Z|^{-\alpha} \geq \delta) \quad (50)$$

where $F_1 = 1$ if and only if the timer of Z expires before that of Z_0, $F_2 = 1$ if and only if Z's timer expires sooner than those of all its neighbors excluding Z_0, $F_3 = 1$ if and only if the received SNR from node Z to node Z_0 exceeds the threshold δ. In order for the transmission to be successful, we must have $F_1 = F_2 = F_3 = 1$. That is

$$\mathsf{E}\left\{ \mathsf{P}_\mathrm{c}^\mathrm{s}(Z, \mathcal{G}, \hat{\Phi}_1) \right\} = \mathsf{E}\left\{ F_1 F_2 F_3 \right\}. \quad (51)$$

Conditioned on $T = \varsigma$, we express the indicator F_2 as the value of some extremal shot-noise [13, Sec. 2.4]. For fixed ς, define the indicator of the event that Z_i is a neighbor of Z and it has a timer smaller than ς:

$$L(Z, Z_i, G_i, T_i) = 1(G_i|Z_i - Z|^{-\alpha} \geq \theta \text{ and } T_i < \varsigma) \quad (52)$$

for all $(Z_i, \mathcal{G}_i, T_i) \in \hat{\Phi}_1$. Define the extremal shot-noise at node Z as

$$\mathcal{Z}_{\hat{\Phi}_1}(Z) = \max_{(Z_i, \mathcal{G}_i, T_i) \in \hat{\Phi}_1} L(Z, Z_i, G_i, T_i). \quad (53)$$

Note that $\mathcal{Z}_{\hat{\Phi}_1}(Z)$ takes only two values 0 or 1 and consequently

$$\mathsf{E}\{F_2|T = \varsigma\} = \mathsf{P}\{\mathcal{Z}_{\hat{\Phi}_1}(Z) \leq 0|T = \varsigma\}. \quad (54)$$

By [13, Proposition 2.4.2], (54) can be further calculated as

$$\mathsf{E}\{F_2|T = \varsigma\}$$
$$= \exp\left\{ -\lambda \int_{\mathbb{R}^2} \int_0^\infty \int_0^1 1\left(L(Z, z, g, t) = 1\right) e^{-g}\mathrm{d}t\mathrm{d}g\mathrm{d}z \right\}$$
$$= \exp\left\{ -2\pi\lambda\varsigma \int_0^\infty \int_0^\infty 1\left(gr^{-\alpha} \geq \theta\right) re^{-g}\mathrm{d}g\mathrm{d}r \right\}$$
$$= e^{-c\varsigma} \quad (55)$$

where c is the average number of neighbors defined in (8).

Therefore, according to (51), we have

$$\mathsf{E}\left\{ \mathsf{P}_\mathrm{c}^\mathrm{s}(Z, \mathcal{G}, \hat{\Phi}_1) \right\} \leq \left(\frac{\theta\gamma}{\delta}\right)^{\frac{2}{\alpha}} \int_0^1 (1-\varsigma)e^{-c\varsigma}\mathrm{d}\varsigma \quad (56)$$
$$= \frac{1}{c^2}\left(\frac{\theta\gamma}{\delta}\right)^{\frac{2}{\alpha}}\left(e^{-c} + c - 1\right) \quad (57)$$

where (56) is derived from the the uniform distribution of T_0, (5), and (54). Proposition 2 then follows by combining (57) and (27).

As a by-product, by averaging over ς in (55), which is uniformly distributed on $[0, 1]$, the probability that a given node captures the channel to transmit in each slot can be calculated as $(1 - e^{-c})/c$.

REFERENCES

[1] D. Guo and L. Zhang, "Virtual full-duplex wireless communication via rapid on-off-division duplex," in *Proc. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, 2010, pp. 412–419.
[2] D. Needell and J. A. Tropp, "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Appl. Comput. Harmonic Anal.*, vol. 26, pp. 301–321, 2009.
[3] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci.*, vol. 106, no. 45, pp. 18914–18919, 2009.
[4] A. Maleki, "Approximate message passing algorithms for compressed sensing," Ph.D. dissertatioin, Stanford Univ., Stanford, CA, USA, 2011.
[5] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Trans. Inf. Theory*, vol. 55, no. 5, pp. 2230–2249, May 2009.
[6] D. Baron, S. Sarvotham, and R. G. Baraniuk, "Bayesian compressive sensing via belief propagation," *IEEE Trans. Signal Process.*, vol. 58, no. 1, pp. 269–280, Jan. 2010.

[7] I. D. Schizas, A. Ribeiro, and G. B. Giannakis, "Consensus in ad hoc WSNs with noisy links-part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 350–364, Jan. 2008.

[8] I. D. Schizas, G. B. Giannakis, S. I. Roumeliotis, and A. Ribeiro, "Consensus in ad hoc WSNs with noisy links—Part II: Distributed estimation and smoothing of random signals," *IEEE Trans. Signal Process.*, vol. 56, no. 4, pp. 1650–1666, Apr. 2008.

[9] O. Simeone, U. Spagnolini, Y. Bar-Ness, and S. Strogatz, "Distributed synchronization in wireless networks," *IEEE Signal Process. Mag.*, vol. 25, no. 5, pp. 81–97, Sep. 2008.

[10] M. Jain, J. I. Choi, T. M. Kim, D. Bharadia, S. Set, K. Srinivasan, P. Levis, S. Katti, and P. Sinha, "Practical, real-time, full duplex wireless," in *Proc. MobiCom*, 2011, pp. 301–312.

[11] A. Sahai, G. Patel, and A. Sabharwal, "Pushing the limits of full-duplex: Design and real-time implementation," arXiv:1107.0607v1, 2011.

[12] L. Applebaum, W. U. Bajwa, M. F. Duarte, and R. Calderbank, "Asynchronous code-division random access using convex optimization," *Phys. Commun.*, vol. 5, no. 2, pp. 129–147, 2012.

[13] F. Baccelli and B. Błaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume I Theory and Volume II Applications*. Boston, MA, USA: Now, 2009, vol. 4, Foundations and Trends in Networking.

[14] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge, U.K.: Cambridge Univ. Press, 2012.

[15] J. F. C. Kingman, *Poisson Processes*. London, U.K.: Oxford Univ. Press, 1993.

[16] S. A. Borbash, A. Ephremides, and M. J. McGlynn, "An asynchronous neighbor discovery algorithm for wireless sensor networks," *Ad Hoc Netw.*, vol. 5, pp. 998–1016, Sep. 2007.

[17] J. Luo and D. Guo, "Neighbor discovery in wireless ad hoc networks based on group testing," in *Proc. Allerton Conf. Commun., Control, Comput.*, Monticello, IL, USA, 2008, pp. 791–797.

[18] L. Zhang, J. Luo, and D. Guo, "Neighbor discovery for wireless networks via compressed sensing," *Perform. Eval.*, vol. 70, pp. 457–471, 2013.

[19] P. H. Tan and L. K. Rasmussen, "Belief propagation for coded multiuser detection," in *Proc. IEEE Int. Symp. In. Theory*, 2006, pp. 1919–1923.

[20] D. Guo and C.-C. Wang, "Multiuser detection of sparsely spread CDMA," *IEEE J. Sel. Areas Commun.*, vol. 26, pp. 421–431, Apr. 2008.

[21] Y. Kabashima, "A CDMA multiuser detection algorithm on the basis of belief propagation," *J. Phys. A, Math. Gen.*, vol. 26, pp. 11111–11121, Oct. 2003.

[22] T. Tanaka and M. Okada, "Approximate belief propagation, density evolution, and statistical neurodynamics for CDMA multiuser detection," *IEEE Trans. Inf. Theory*, vol. 51, no. 2, pp. 700–706, Feb. 2005.

[23] L. Zhang, "Virtual full duplex wireless networks," Ph.D. dissertation, Northwestern Univ., Evanston, IL, USA, 2012.

[24] D. Halperin, T. Anderson, and D. Wetherall, "Taking the sting out of carrier sense: Interference cancellation for wireless LANs," in *Proc. ACM MobiCom*, Sep. 2008, pp. 339–350.

[25] S. Gollakota and D. Katabi, "Zigzag decoding: Combating hidden terminals in wireless networks," in *Proc. ACM SIGCOMM*, Aug. 2008, pp. 159–170.

[26] A. V. Oppenheim, A. S. Willsky, and S. H. Nawab, *Signals & Systems*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.

Lei Zhang received the B.E. and M.S. degrees in electronics engineering from Tsinghua University, Beijing, China, in 2005 and 2007, respectively, and the Ph.D. degree in electrical engineering and computer science from Northwestern University, Evanston, IL, USA, in 2012.

He is currently with Qualcomm Research, Bridgewater, NJ, USA. His research interests include wireless ad hoc networks, compressed sensing, and information theory.

Dr. Zhang is a recipient of the Student Paper Award of the IEEE International Symposium on Information Theory in 2011. He also received the Walter P. Murphy Fellowship and Terminal Year Fellowship from Northwestern University in 2007 and 2011, respectively.
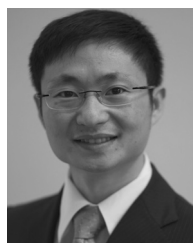
Dongning Guo (S'97–M'04–SM'11) received the B.Eng. degree from the University of Science and Technology of China, Hefei, China, in 1995, the M.Eng. degree from the National University of Singapore, Singapore, in 1999, and the M.A. and Ph.D. degrees from Princeton University, Princeton, NJ, USA in 2001 and 2004, respectively, all in electrical engineering.

He joined the faculty of Northwestern University, Evanston, IL, in 2004, where he is currently an Associate Professor with the Department of Electrical Engineering and Computer Science. He was an R&D Engineer with the Center for Wireless Communications (now the Institute for Infocom Research), Singapore, from 1998 to 1999. He has held visiting positions with Norwegian University of Science and Technology, Trondheim, Norway, in 2006 and the Institute of Network Coding, Chinese University of Hong Kong, Hong Kong, from 2010 to 2011. His research interests are in information theory, communications, and networking.

Dr. Guo has served as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION THEORY and is currently an Editor of *Foundations and Trends in Communications and Information Theory*, and a Guest Editor of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS Special Issue on Full-duplex Wireless Communications and Networks. He received the Huber and Suhner Best Student Paper Award in the International Zurich Seminar on Broadband Communications in 2000 and is a co-recipient of the IEEE Marconi Prize Paper Award in Wireless Communications in 2010 (with Y. Zhu and M. L. Honig). He is also a recipient of the National Science Foundation Faculty Early Career Development (CAREER) Award in 2007.