Many-user multiple access with random user activity

Xiaoqi Liu
University of Cambridge, UK
xl394@cam.ac.uk

Pablo Pascual Cobo
University of Cambridge, UK
pp423@cam.ac.uk

Ramji Venkataramanan
University of Cambridge, UK
rv285@cam.ac.uk

***Abstract*—We study the Gaussian multiple access channel with random user activity, in the regime where the number of users is proportional to the code length. The receiver may know some statistics about the number of active users, but does not know the exact number nor the identities of the active users. We first derive achievability bounds on the error probabilities by analyzing Gaussian random codebooks with maximum-likelihood decoding. We then propose an efficient CDMA-type scheme based on a spatially coupled signature matrix and approximate message passing (AMP) decoding. Rigorous asymptotic guarantees on the error performance of the AMP decoder are derived. A numerical comparison indicates that the asymptotic error guarantees of the spatially coupled scheme are significantly better than those obtained via the finite-length achievability bounds.**

I. Introduction

We study the Gaussian multiple access channel (GMAC) with output of the form

$$\boldsymbol{y} = \sum_{\ell=1}^{L} \boldsymbol{c}_\ell + \boldsymbol{\varepsilon}, \tag{1}$$

where L is the number of users, $\boldsymbol{c}_\ell \in \mathbb{R}^n$ is the codeword of user ℓ, $\boldsymbol{\varepsilon} \sim \mathcal{N}_n(\boldsymbol{0}, \sigma^2 \boldsymbol{I})$ is random channel noise with spectral density $N_0 = 2\sigma^2$, and n is the number of channel uses. Motivated by applications such as the Internet of Things, there has been much interest in the *many-user* regime, where the number of users L grows with the code length n [1]–[5].

In this paper, we consider the many-user regime where the number of users L grows proportionally with n; the ratio $\mu := L/n$ is called the user density [3]. Each user transmits a fixed number of bits k (payload) under a constant energy-per-bit constraint $\|\boldsymbol{c}_\ell\|_2^2/k \leq E_b$. A key question in this regime is to characterize the tradeoff between user density μ, the signal-to-noise ratio E_b/N_0 and the error probability in decoding the codewords $\{\boldsymbol{c}_\ell\}$ from $\boldsymbol{y}$. The standard error metric is the per-user probability of error $\texttt{PUPE} := \frac{1}{L}\sum_{\ell=1}^{L} \mathbb{P}(\boldsymbol{c}_\ell \neq \widehat{\boldsymbol{c}_\ell})$ where $\widehat{\boldsymbol{c}_\ell}$ denotes the decoder estimate of $\boldsymbol{c}_\ell$. Achievability and converse bounds were derived in [3], [4], [6] in terms of the minimum E_b/N_0 needed to achieve a given target PUPE for a given user density μ. Efficient coding schemes were proposed in [7] in an attempt to approach the converse bound.

In practical GMAC settings, the users are seldom consistently active. Instead, they are active in a sporadic and uncoordinated manner. Motivated by recent work in this direction [1], [8]–[11], we study the GMAC with *random user activity*. Letting $K_\mathrm{a} \leq L$ denote the (random) number of active users, we consider the proportional regime with $\mathbb{E}[K_\mathrm{a}]/L \to \alpha$ and $\mathbb{E}[K_\mathrm{a}]/n \to \mu_\mathrm{a}$, where α and μ_a are both fixed constants. The receiver may know the distribution of K_a or some statistics, but it does not know the identities of the active users nor the exact value of K_a. We study the tradeoff between the *active* user density μ_a, the signal-to-noise ratio E_b/N_0 and the decoding performance, measured in terms of the probabilities of misdetection (MD), false alarm (FA), and active user error (AUE) (see Section II).

We emphasize that our GMAC setting is different from unsourced random-access [3], [11]–[14], where all the users share the same codebook, and a subset of them are active. In our setting, each user has a separate codebook. While unsourced random-access is particularly relevant for grant-free communication systems, coding schemes based on sparse regression for the standard AWGN channel and the GMAC [15] form the basis of several recent schemes for unsourced random-access [13], [14].

Our main contributions are two-fold:

1) In Theorem 1, we derive finite-length achievability bounds on the probabilities of misdetection, false alarm, and active user error. These are obtained by analyzing random Gaussian codebooks with (computationally infeasible) maximum-likelihood decoding, and are similar to the bounds obtained by Ngo et al. [11] for unsourced multiple-access with random user activity. However, the results are not comparable because we place different assumptions on the codebooks.

2) We propose an efficient coded-division multiple access (CDMA)-type coding scheme for the GMAC with random user activity, and rigorously characterize its asymptotic error performance for given values of E_b/N_0, μ_a and α (Theorem 2). The scheme uses a spatially coupled Gaussian matrix (whose columns are the users' signature sequences), and an iterative approximate message passing (AMP) algorithm for decoding. The AMP algorithm, which aims to recover a matrix-valued signal, can be viewed as a generalization of the spatially coupled AMP in [16] for vector-valued signals. To our knowledge, this is the first application of spatial coupling to random linear models with matrix-valued signals, which may be of independent interest. A numerical comparison indicates that the asymptotic error guarantees of the spatially coupled scheme are significantly better than those obtained via

This work was supported in part by a Schlumberger Cambridge International Scholarship and an EPSRC Doctoral Training Partnership (DTP).

the finite-length achievability bounds.

Related work: AMP, a class of iterative algorithms first proposed for estimation in random linear models [17], [18], has been applied to a variety of high-dimensional estimation problems [19]. Two appealing features of AMP are: i) it can be tailored to take advantage of the signal prior, and ii) under suitable model assumptions, its estimation performance in the high-dimensional limit can be characterized by a simple deterministic recursion called state evolution. Coding schemes based on spatially coupled random linear models with AMP decoding have been studied for communication over both point-to-point channels [20]–[22] and the many-user GMAC [7]. These works show that spatially coupled designs significantly improve on the performance of i.i.d. designs in certain regimes. More generally, spatially coupled designs with AMP decoding have been shown to achieve the Bayes-optimal estimation error for both linear models [16] and generalized linear models [23].

Notation: The indicator function of an event $\mathcal{A}$ is denoted by $\mathbb{1}\{\mathcal{A}\}$. For positive integers a and b, $[a] := \{1, ..., a\}$ and $[a : b] = \{a, a+1, \ldots, b\}$. We use boldface letters for vectors and matrices and plain font for scalars. The ℓ-th row of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}_\ell$ and is treated as a row vector; A_{ij} denotes the (i, j)-th entry. Let $x^+ = \max\{x, 0\}$. We write $\mathcal{N}_d(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ for a d-dimensional Gaussian with mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$.

II. Performance Metrics

Let $\mathcal{C}^{(\ell)} = \{\boldsymbol{c}_1^{(\ell)}, \boldsymbol{c}_2^{(\ell)}, \ldots, \boldsymbol{c}_M^{(\ell)}\}$ denote the codebook of user ℓ. Each active user transmits k bits, so $M = 2^k$. Let $w_\ell \in \{\emptyset, 1, 2, \ldots, M\}$ denote the index of the codeword chosen by user ℓ, where $w_\ell = \emptyset$ indicates that the user is silent (not active). We use index pair (i, j) to refer to the jth codeword of the ith codebook. Then the set of transmitted codewords can be defined as $\mathcal{W} := \{(\ell, w_\ell) : w_\ell \neq \emptyset\}$, and the GMAC channel output can be written as

$$\boldsymbol{y} = \sum_{\ell:(\ell, w_\ell)\in\mathcal{W}} \boldsymbol{c}_{w_\ell}^{(\ell)} + \boldsymbol{\varepsilon}, \tag{2}$$

where the number of active users is $K_\text{a} = |\mathcal{W}| \leq L$.

Given the channel output $\boldsymbol{y}$ and the codebooks $\{\mathcal{C}^{(\ell)}\}$, the decoder aims to recover the set of transmitted codewords. Let $\widehat{w_\ell} \in \{\emptyset, 1, 2, \ldots, M\}$ denote the decoded codeword in the ℓth codebook, and let $\widehat{\mathcal{W}} := \{(\ell, \widehat{w_\ell}) : \widehat{w_\ell} \neq \emptyset\}$ denote the set of decoded codewords, or the decoded set in short, with size $|\widehat{\mathcal{W}}| =: \widehat{K_\text{a}}$.

Such a decoder can make three types of errors, which we call misdetection (`MD`), false alarm (`FA`) and active user error (`AUE`). `MD` refer to an active user declared silent, and `FA` to a silent user declared active. An `AUE` occurs when an active users is correctly declared active, but the decoded codeword is incorrect. The error probabilities corresponding to these events are defined as follows:

$$p_\text{MD} := \mathbb{E}\left[\mathbb{1}\{K_\text{a} \neq 0\} \cdot \frac{1}{K_\text{a}} \sum_{\ell:(\ell, w_\ell)\in\mathcal{W}} \mathbb{1}\{\widehat{w_\ell} = \emptyset\}\right], \tag{3}$$

$$p_\text{FA} := \mathbb{E}\left[\mathbb{1}\{\widehat{K_\text{a}} \neq 0\} \cdot \frac{1}{\widehat{K_\text{a}}} \sum_{\ell:(\ell, \widehat{w_\ell})\in\widehat{\mathcal{W}}} \mathbb{1}\{w_\ell = \emptyset\}\right], \tag{4}$$

$$p_\text{AUE} := \mathbb{E}\left[\mathbb{1}\{K_\text{a} \neq 0\} \cdot \frac{1}{K_\text{a}} \sum_{\ell:(\ell, w_\ell)\in\mathcal{W}} \mathbb{1}\{\widehat{w_\ell} \neq w_\ell\}\right]. \tag{5}$$

III. Random Coding Achievability Bound

In this section, we derive achievability bounds on p_MD, p_FA and p_AUE using a random coding scheme under maximum-likelihood decoding, assuming the number of active users K_a is drawn from a known distribution p_{K_a} and given the signal-to-noise ratio E_b/N_0. Without loss of generality, we take $\sigma^2 = N_0/2 = 1$, so the constraint on E_b/N_0 is enforced via E_b.

Recall that each active user transmits k bits using a length n codeword. For $\ell \in [L]$, we construct the codebook $\mathcal{C}^{(\ell)}$ with codewords $\boldsymbol{c}_j^{(\ell)} = \tilde{\boldsymbol{c}}_j^{(\ell)} \mathbb{1}\{\|\tilde{\boldsymbol{c}}_j^{(\ell)}\|_2^2 \leq E_b k\}$, where $\tilde{\boldsymbol{c}}_j^{(\ell)} \overset{\text{i.i.d.}}{\sim} \mathcal{N}_n(\boldsymbol{0}, E_b' k/n \boldsymbol{I})$ and $E_b' \leq E_b$. The indicator function ensures the codewords $\boldsymbol{c}_j^{(\ell)}$ satisfy the energy-per-bit constraint $\|\boldsymbol{c}_j^{(\ell)}\|_2^2/k \leq E_b$. Note that the codewords across codebooks are distinct almost surely.

Given $\mathbf{y}$ is the realisation of $\boldsymbol{y}$, the decoder first computes the maximum-likelihood estimate K_a' of the number of active users K_a, via

$$K_\text{a}' = \underset{K\in[\kappa_l:\kappa_\text{u}]}{\arg\max}\, p_{\boldsymbol{y}|K_\text{a}}(\mathbf{y}|K), \tag{6}$$

where $\kappa_l, \kappa_\text{u}$ are deterministic values between 0 and L, chosen such that $\mathbb{P}(K_\text{a} \notin [\kappa_l, \kappa_\text{u}])$ is small (e.g., 10^{-6}). This step is similar to the scheme in [11].

The decoder then estimates the codewords $\boldsymbol{c}_{\widehat{w_\ell}}^{(\ell)}$ for $\ell \in [L]$ via a combinatorial least squares problem. Define dummy variables $w_\ell' \in \{\emptyset, 1, 2, \ldots, M\}$ for $\ell \in [L]$, and the set $\mathcal{W}' := \{(\ell, w_\ell') : w_\ell' \neq \emptyset\}$. Let $c(\mathcal{W}') := \sum_{j:(j,w_j')\in\mathcal{W}'} \boldsymbol{c}_{w_j'}^{(j)}$. The decoder computes:

$$\{\widehat{w_1}, \widehat{w_2}, \ldots, \widehat{w_L}\} = \underset{\substack{\{w_1', w_2', \ldots, w_L'\}: \\ \underline{K_\text{a}'} \leq |\mathcal{W}'| \leq \overline{K_\text{a}'}}}{\arg\min} \|c(\mathcal{W}') - \mathbf{y}\|_2^2 \tag{7}$$

where $\underline{K_\text{a}'} = \max\{\kappa_l, K_\text{a}' - r\}, \overline{K_\text{a}'} = \min\{\kappa_\text{u}, K_\text{a}' + r\}$, with $r \geq 0$ being the decoding radius. In words, the decoder searches over all decoded sets $\mathcal{W}'$ with size between $\underline{K_\text{a}'}$ and $\overline{K_\text{a}'}$, which contain at most one codeword from each codebook, to minimize the distance between the sum of the codewords in the decoded set $c(\mathcal{W}')$ and the observed channel output $\mathbf{y}$. The complexity of this decoder grows exponentially with n, making it computationally infeasible.

Theorem 1. *Assume the distribution of the number of active users p_{K_a} is known. Given $P = E_b k/n$, $P' = E_b' k/n$, $E_b \geq E_b', r \geq 0$, $\kappa_l, \kappa_\text{u}$ such that $0 \leq \kappa_l \leq \kappa_\text{u} \leq L$, $\underline{\kappa_\text{a}'} = \max\{\kappa_l, \kappa_\text{a}' - r\}$ and $\overline{\kappa_\text{a}'} = \min\{\kappa_\text{u}, \kappa_\text{a}' + r\}$, the random coding scheme with the decoder described in* (6)–(7)

has error probabilities that satisfy: $p_{\rm MD} \le \varepsilon_{\rm MD}$, $p_{\rm FA} \le \varepsilon_{\rm FA}$ *and* $p_{\rm AUE} \le \varepsilon_{\rm AUE}$, *where*

$$\varepsilon_{\rm MD} = \sum_{\kappa_{\rm a}=\kappa_l}^{\kappa_{\rm u}} p_{K_{\rm a}}(\kappa_{\rm a}) \sum_{\kappa'_{\rm a}=\kappa_l}^{\kappa_{\rm u}} \sum_{t\in\mathcal{T}} \sum_{\hat{t}\in\widehat{\mathcal{T}}_t} \min\{p_{t,\hat{t}}, \xi(\kappa_{\rm a},\kappa'_{\rm a})\} \cdot \sum_{\psi=0}^{\psi_{\rm u}} \frac{\nu(t_{\min},\psi)}{\nu(t_{\min})} \frac{(\kappa_{\rm a}-\overline{\kappa'_{\rm a}})^+ + (t-\hat{t})^+ + \psi}{\kappa_{\rm a}} + \tilde{p} \tag{8}$$

$$\varepsilon_{\rm FA} = \sum_{\kappa_{\rm a}=\kappa_l}^{\kappa_{\rm u}} p_{K_{\rm a}}(\kappa_{\rm a}) \sum_{\kappa'_{\rm a}=\kappa_l}^{\kappa_{\rm u}} \sum_{t\in\mathcal{T}} \sum_{\hat{t}\in\widehat{\mathcal{T}}_t} \min\{p_{t,\hat{t}}, \xi(\kappa_{\rm a},\kappa'_{\rm a})\} \cdot \sum_{\psi=0}^{\psi_{\rm u}} \frac{\nu(t_{\min},\psi)}{\nu(t_{\min})} \Delta_{\rm FA} + \tilde{p} \tag{9}$$

$$\Delta_{\rm FA} = \frac{(\underline{\kappa'_{\rm a}}-\kappa_{\rm a})^+ + (\hat{t}-t)^+ + \psi}{\kappa_{\rm a} - t - (\kappa_{\rm a}-\overline{\kappa'_{\rm a}})^+ + \hat{t} + (\underline{\kappa'_{\rm a}}-\kappa_{\rm a})^+}$$

$$\varepsilon_{\rm AUE} = \sum_{\kappa_{\rm a}=\kappa_l}^{\kappa_{\rm u}} p_{K_{\rm a}}(\kappa_{\rm a}) \sum_{\kappa'_{\rm a}=\kappa_l}^{\kappa_{\rm u}} \sum_{t\in\mathcal{T}} \sum_{\hat{t}\in\widehat{\mathcal{T}}_t} \min\{p_{t,\hat{t}}, \xi(\kappa_{\rm a},\kappa'_{\rm a})\} \cdot \sum_{\psi=0}^{\psi_{\rm u}} \frac{\nu(t_{\min},\psi)}{\nu(t_{\min})} \frac{t_{\min}-\psi}{\kappa_{\rm a}} + \tilde{p}\ , \tag{10}$$

for which

$$\tilde{p} = \mathbb{P}\big(K_{\rm a} \notin [\kappa_l,\kappa_{\rm u}]\big) + \mathbb{E}[K_{\rm a}] \frac{\Gamma(\frac{n}{2}, \frac{nP}{2P'})}{\Gamma(\frac{n}{2})}$$

$$p_{t,\hat{t}} = \exp\left(-\frac{n}{2}E(t,\hat{t})\right)$$

$$E(t,\hat{t}) = \max_{\rho,\rho_1\in[0,1]} -\rho\rho_1\hat{t}R_1 - \rho_1 R_2 + E_0(\rho,\rho_1)$$

$$R_1 = \frac{2}{n}\ln M, \quad R_2 = \frac{2}{n}\ln\binom{\min\{\kappa_{\rm a},\overline{\kappa'_{\rm a}}\}}{t}$$

$$E_0(\rho,\rho_1) = \max_{\lambda} \rho_1 a + \ln(1-\rho_1 P_1 b)$$

$$P_1 = 1 + \left((\kappa_{\rm a}-\overline{\kappa'_{\rm a}})^+ + (\underline{\kappa'_{\rm a}}-\kappa_{\rm a})^+\right)P'$$

$$a = \rho\ln(1+P'\hat{t}\lambda) + \ln(1+P't\mu)$$

$$b = \rho\lambda - \frac{\mu}{1+P't\mu}, \quad \mu = \frac{\rho\lambda}{1+P'\hat{t}\lambda}$$

$$\mathcal{T} = \big[0 : \min\{\kappa_{\rm a},\overline{\kappa'_{\rm a}}\}\big]$$

$$\widehat{\mathcal{T}}_t = \left[\ \left\{t + (\kappa_{\rm a}-\overline{\kappa'_{\rm a}})^+ - (\kappa_{\rm a}-\underline{\kappa'_{\rm a}})^+\right\}^+ : t_{\rm u}\right]$$

$$t_{\rm u} = \min\Big\{\overline{\kappa'_{\rm a}} - (\underline{\kappa'_{\rm a}}-\kappa_{\rm a})^+,\ t + (\overline{\kappa'_{\rm a}}-\kappa_{\rm a})^+ - (\underline{\kappa'_{\rm a}}-\kappa_{\rm a})^+\Big\}$$

$$t_{\min} = \min\{t,\hat{t}\}$$

$$\mathcal{R} = L - \kappa_{\rm a} + t_{\min} - (\underline{\kappa'_{\rm a}}-\kappa_{\rm a})^+ - (\hat{t}-t)^+ > 0$$

$$\psi_{\rm u} = \min\{t_{\min}, \mathcal{R}-t_{\min}\}$$

$$\nu(t_{\min},\psi) = \binom{\mathcal{R}-t_{\min}}{\psi} M^{\psi} \cdot \binom{t_{\min}}{t_{\min}-\psi} (M-1)^{t_{\min}-\psi}$$

$$\nu(t_{\min}) = \sum_{\psi=0}^{\psi_{\rm u}} \nu(t_{\min},\psi)$$

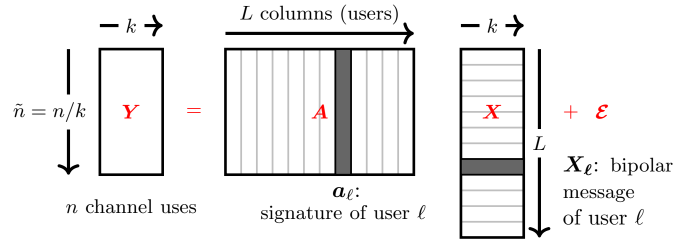

Figure 1. Linear coding scheme.

$$\xi(\kappa_{\rm a},\kappa'_{\rm a}) = \min_{\kappa\in[\kappa_l:\kappa_{\rm u}]\setminus\kappa'_{\rm a}} \left\{\mathbb{1}\{\kappa<\kappa'_{\rm a}\}\frac{\Gamma(\frac{n}{2},\zeta)}{\Gamma(\frac{n}{2})} + \mathbb{1}\{\kappa>\kappa'_{\rm a}\}\frac{\gamma(\frac{n}{2},\zeta)}{\Gamma(\frac{n}{2})}\right\}$$

$$\zeta = \frac{n}{2(1+\kappa_{\rm a}P')}\ln\left(\frac{1+\kappa P'}{1+\kappa'_{\rm a}P'}\right)\left[\frac{1}{1+\kappa'_{\rm a}P'} - \frac{1}{1+\kappa P'}\right]^{-1}.$$

The proof of Theorem 1 is provided in the longer version of this paper. It is analogous to the proof of the bounds in [11, Theorem 1], involving union bounds over error events via a change of measure, Chernoff bound and Gallager's ρ-trick.

Remark: Theorem 1 reduces to the achievability bound in [3, Sec. IV] when $K_{\rm a} = K'_{\rm a} = L$ with probability 1 and the decoding radius is $r = 0$, i.e., when all users are active and the number of users is known and fixed.

IV. Efficient CDMA-type Coding Scheme

In this section, we describe an efficient coding scheme and obtain performance guarantees that can be compared with the achievability bound in Theorem 1.

Coding Scheme: Let the row vector $\boldsymbol{X}_\ell \in \mathbb{R}^k$ denote the length-k message of user $\ell \in [L]$, recalling that k (the number of bits transmitted by each active user) is fixed, and does not grow with n and L. To construct the codeword of user ℓ, we first left-multiply $\boldsymbol{X}_\ell$ with a column vector (signature sequence) $\boldsymbol{a}_\ell \in \mathbb{R}^{\frac{n}{k}}$ to obtain the matrix $\boldsymbol{C}_\ell = \boldsymbol{a}_\ell\boldsymbol{X}_\ell \in \mathbb{R}^{\frac{n}{k}\times k}$. The codeword of user ℓ is then obtained as $\boldsymbol{c}_\ell = \text{vectorize}(\boldsymbol{C}_\ell) \in \mathbb{R}^n$.

For simplicity, we let $\tilde{n} = \frac{n}{k}$, and define the matrices $\boldsymbol{A} = [\boldsymbol{a}_1,\dots,\boldsymbol{a}_L] \in \mathbb{R}^{\tilde{n}\times L}$ and $\boldsymbol{X} = [\boldsymbol{X}_1^\top,\dots,\boldsymbol{X}_\ell^\top]^\top \in \mathbb{R}^{L\times k}$. Then the GMAC channel output in (1) can be rewritten in matrix form as

$$\boldsymbol{Y} = \sum_{\ell=1}^{L} \boldsymbol{C}_\ell + \boldsymbol{\mathcal{E}} = \boldsymbol{A}\boldsymbol{X} + \boldsymbol{\mathcal{E}}, \tag{11}$$

where $\boldsymbol{\mathcal{E}} \in \mathbb{R}^{\tilde{n}\times k}$ is the channel noise with $\mathcal{E}_{ij} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,\sigma^2)$. See Fig. 1 for an illustration. In this paper, we consider the simple case where each user ℓ is independently active with probability α: $\boldsymbol{X}_\ell \in \{\pm\sqrt{E}\}^k$ when the user is active, and $\boldsymbol{X}_\ell = \boldsymbol{0}$ when the user is not active. This is equivalent to considering $\boldsymbol{X}_\ell \overset{\text{i.i.d.}}{\sim} p_{\bar{\boldsymbol{X}}}$, for $\ell \in [L]$, where

$$p_{\bar{\boldsymbol{X}}} = (1-\alpha)\delta_{\boldsymbol{0}} + \alpha\, p_{\bar{\boldsymbol{X}}_{\rm a}}\,. \tag{12}$$

Here $\delta_{\boldsymbol{0}}$ is the unit mass at $\boldsymbol{0}$, and $p_{\bar{\boldsymbol{X}}_{\rm a}}$ is a uniform distribution on the set of messages $\mathcal{X}_{\rm a} = \{\pm\sqrt{E}\}^k$. Our framework can be adapted flexibly to more complicated priors $p_{\bar{\boldsymbol{X}}}$, which will be discussed in the extended version of this paper.

Spatially Coupled Signature Design: We use a spatially coupled design for $\boldsymbol{A}$, the matrix of signature sequences. The entries of $\boldsymbol{A} \in \mathbb{R}^{\tilde{n}\times L}$ are independent zero-mean Gaussian random variables whose variances are specified by a *base matrix* $\boldsymbol{W} \in \mathbb{R}^{\mathsf{R}\times\mathsf{C}}$. The matrix $\boldsymbol{A}$ is obtained by replacing each entry $W_{\mathsf{r},\mathsf{c}}$ of the base matrix by an $\frac{\tilde{n}}{\mathsf{R}} \times \frac{L}{\mathsf{C}}$ block with entries drawn $\stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, \frac{W_{\mathsf{r},\mathsf{c}}}{\tilde{n}/\mathsf{R}})$, for $\mathsf{r} \in [\mathsf{R}]$, $\mathsf{c} \in [\mathsf{C}]$. More precisely,

$$A_{i\ell} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}\left(0, \frac{1}{\tilde{n}/\mathsf{R}} W_{\mathsf{r}(i),\mathsf{c}(\ell)}\right), \quad \text{for } i \in [\tilde{n}],\ \ell \in [L]. \quad (13)$$

Here the operators $\mathsf{r}(\cdot) : [\tilde{n}] \to [\mathsf{R}]$ and $\mathsf{c}(\cdot) : [L] \to [\mathsf{C}]$ map a particular row or column index of $\boldsymbol{A}$ to its corresponding *row block* or *column block* index of $\boldsymbol{W}$. See Fig. 2 for an example. Similar designs have been used for sparse superposition codes for communication over AWGN channels [20], [22].

As in [16], we assume that entries of the base matrix $\boldsymbol{W}$ are scaled to satisfy:

$$\sum_{\mathsf{r}=1}^{\mathsf{R}} W_{\mathsf{r},\mathsf{c}} = 1 \ \text{ for } \mathsf{c} \in [\mathsf{C}], \qquad \kappa_1 \leq \sum_{\mathsf{c}=1}^{\mathsf{C}} W_{\mathsf{r},\mathsf{c}} \leq \kappa_2, \quad (14)$$

for some $\kappa_1, \kappa_2 > 0$. This scaling ensures that the columns of $\boldsymbol{A}$ (the signature sequences $\boldsymbol{a}_\ell$) have unit norm in expectation. The standard i.i.d. Gaussian design where $A_{i\ell} \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1/\tilde{n})$ is a special case of the spatially coupled design, obtained by using a base matrix with a single entry ($\mathsf{R} = \mathsf{C} = 1$).

A. AMP Decoder and State Evolution

In this subsection, we describe an AMP decoder to recover $\boldsymbol{X}$ from $\boldsymbol{Y}$ given the spatially coupled matrix $\boldsymbol{A}$, and provide asymptotic guarantees on its error performance.

Spatially Coupled AMP (SC-AMP): Starting with an initializer $\boldsymbol{X}^0 = \boldsymbol{0}_{L\times k}$, in iteration $t \geq 0$ the decoder computes:

$$\boldsymbol{Z}^t = \boldsymbol{Y} - \boldsymbol{A}\boldsymbol{X}^t + \tilde{\boldsymbol{Z}}^t, \qquad \boldsymbol{X}^{t+1} = \eta_t\left(\boldsymbol{X}^t + \boldsymbol{V}^t\right). \quad (15)$$

Here $\boldsymbol{X}^{t+1}$ is an updated estimate of $\boldsymbol{X}$ produced using a denoising function $\eta_t : \mathbb{R}^{L\times k} \to \mathbb{R}^{L\times k}$, and $\boldsymbol{Z}^t$ can be viewed as a modified residual. The denoiser is assumed to be Lipschitz and separable, acting row-wise on matrix inputs:

$$\eta_t(\boldsymbol{S}) = \begin{bmatrix} \eta_{t,\mathsf{c}(1)}(\boldsymbol{S}_1) \\ \vdots \\ \eta_{t,\mathsf{c}(L)}(\boldsymbol{S}_L) \end{bmatrix}, \quad (16)$$

where $\eta_{t,\mathsf{c}} : \mathbb{R}^{1\times k} \to \mathbb{R}^{1\times k}$ corresponds to the denoiser for users in block $\mathsf{c} \in [\mathsf{C}]$. We will specify our choice of $\eta_{t,\mathsf{c}}$ shortly.

For $t \geq 0$, $\tilde{\boldsymbol{Z}}^t$ and $\boldsymbol{V}^t$ in (15) are defined through a matrix $\boldsymbol{Q}^t \in \mathbb{R}^{k\mathsf{R}\times k\mathsf{C}}$, which consists of $\mathsf{R} \times \mathsf{C}$ submatrices, each of size $k \times k$. For $\mathsf{r} \in [\mathsf{R}], \mathsf{c} \in [\mathsf{C}]$, the submatrix $\boldsymbol{Q}^t_{\mathsf{r},\mathsf{c}} \in \mathbb{R}^{d\times d}$ is defined as:

$$\boldsymbol{Q}^t_{\mathsf{r},\mathsf{c}} = \left[\boldsymbol{\Phi}^t_{\mathsf{r}}\right]^{-1} \boldsymbol{T}^t_{\mathsf{c}},$$

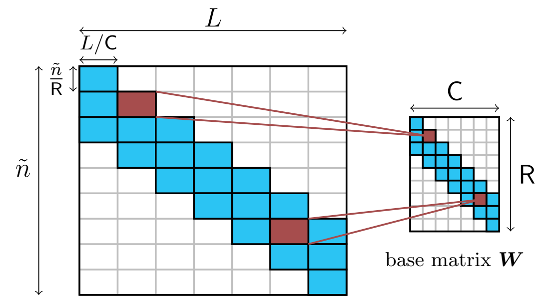

Figure 2. The entries of $\boldsymbol{A}$ are independent and distributed as $A_{i\ell} \sim \mathcal{N}(0, \frac{1}{\tilde{n}/R} W_{\mathsf{r}(i),\mathsf{c}(\ell)})$, where $\boldsymbol{W}$ is the base matrix. The white parts of $\boldsymbol{A}$ and $\boldsymbol{W}$ correspond to zeros.

where $\boldsymbol{\Phi}^t_{\mathsf{r}}, \boldsymbol{T}^t_{\mathsf{c}} \in \mathbb{R}^{k\times k}$ are deterministic matrices defined later in (17)–(18). The ith row of matrix $\tilde{\boldsymbol{Z}}^t \in \mathbb{R}^{\tilde{n}\times k}$ then takes the form:

$$\tilde{\boldsymbol{Z}}^t_i = k\mu_{\text{in}} \boldsymbol{Z}^{t-1}_i \sum_{\mathsf{c}=1}^{\mathsf{C}} W_{\mathsf{r}(i),\mathsf{c}} \boldsymbol{Q}^{t-1}_{\mathsf{r}(i),\mathsf{c}} \left\langle \eta'_{t-1,\mathsf{c}}\left(\boldsymbol{X}^{t-1}_\ell + \boldsymbol{V}^{t-1}_\ell\right)\right\rangle^{\mathsf{T}}_{\mathsf{c}},$$

where $\mu_{\text{in}} = \mu\frac{\mathsf{R}}{\mathsf{C}}$ and $\langle\cdot\rangle_{\mathsf{c}}$ is an average operator over the rows in the cth block, i.e.,

$$\left\langle \eta'_{t,\mathsf{c}}\left(\boldsymbol{X}^t_\ell + \boldsymbol{V}^t_\ell\right)\right\rangle_{\mathsf{c}} = \frac{1}{L/\mathsf{C}} \sum_{\ell\in\mathcal{L}_{\mathsf{c}}} \eta'_{t,\mathsf{c}(\ell)}(\boldsymbol{X}^t_\ell + \boldsymbol{V}^t_\ell)$$

with $\mathcal{L}_{\mathsf{c}} = \{(\mathsf{c}-1)L/\mathsf{C} + 1, \ldots, \mathsf{c}L/\mathsf{C}\}$. Here $\eta'_{t,\mathsf{c}}(\boldsymbol{s}) = \frac{\mathrm{d}\eta_{t,\mathsf{c}}(\boldsymbol{s})}{\mathrm{d}\boldsymbol{s}} \in \mathbb{R}^{k\times k}$ is the derivative (Jacobian) of $\eta_{t,\mathsf{c}}$, and quantities with negative iteration index are set to all-zero matrices. The ℓth row of $\boldsymbol{V}^t \in \mathbb{R}^{L\times k}$ takes the form:

$$\boldsymbol{V}^t_\ell = \sum_{i=1}^{\tilde{n}} A_{i,\ell} \boldsymbol{Z}^t_i \boldsymbol{Q}^t_{\mathsf{r}(i),\mathsf{c}(\ell)}, \ \text{ for } \ell \in [L].$$

State Evolution: As we shall prove in the longer version of this paper, for $t \geq 1$, in the limit as $L, n \to \infty$ with $L/n \to \mu$, the empirical distribution of the rows of $\boldsymbol{Z}^t$ in the block $\mathsf{r} \in [\mathsf{R}]$ converges to a Gaussian $\mathcal{N}_k(\boldsymbol{0}, \boldsymbol{\Phi}^t_{\mathsf{r}})$. Furthermore, the empirical distribution of the rows of $(\boldsymbol{X}^t + \boldsymbol{V}^t - \boldsymbol{X})$ in block $\mathsf{c} \in [\mathsf{C}]$ converges to another Gaussian $\mathcal{N}_k(\boldsymbol{0}, \boldsymbol{T}^t_{\mathsf{c}})$. The covariance matrices $\boldsymbol{\Phi}^t_{\mathsf{r}}, \boldsymbol{T}^t_{\mathsf{c}} \in \mathbb{R}^{k\times k}$ are iteratively defined via a deterministic recursion called the spatially coupled *state evolution* (SC-SE). Initialized with $\boldsymbol{\Psi}^0_{\mathsf{c}} = \alpha E \boldsymbol{I}_{k\times k}$, SC-SE is defined for $t \geq 0$ as follows, for $\mathsf{r} \in [\mathsf{R}]$ and $\mathsf{c} \in [\mathsf{C}]$:

$$\boldsymbol{\Phi}^t_{\mathsf{r}} = \sigma^2 \boldsymbol{I} + k\mu_{\text{in}} \sum_{c=1}^{\mathsf{C}} W_{\mathsf{r},\mathsf{c}} \boldsymbol{\Psi}^t_{\mathsf{c}}, \quad (17)$$

$$\boldsymbol{\Psi}^{t+1}_{\mathsf{c}} = \mathbb{E}\left\{\left[\eta_{t,\mathsf{c}}\left(\bar{\boldsymbol{X}} + \boldsymbol{G}^t_{\mathsf{c}}\right) - \bar{\boldsymbol{X}}\right]^{\mathsf{T}} \left[\eta_{t,\mathsf{c}}\left(\bar{\boldsymbol{X}} + \boldsymbol{G}^t_{\mathsf{c}}\right) - \bar{\boldsymbol{X}}\right]\right\},$$

$$\text{where } \boldsymbol{G}^t_{\mathsf{c}} \sim \mathcal{N}_k(\boldsymbol{0}, \boldsymbol{T}^t_{\mathsf{c}}), \quad \boldsymbol{T}^t_{\mathsf{c}} = \left[\sum_{r=1}^{\mathsf{R}} W_{\mathsf{r},\mathsf{c}} [\boldsymbol{\Phi}^t_{\mathsf{r}}]^{-1}\right]^{-1}. \quad (18)$$

Here $\boldsymbol{G}^t_{\mathsf{c}}$ is independent of $\bar{\boldsymbol{X}} \sim p_{\bar{\boldsymbol{X}}}$. We can interpret $\boldsymbol{\Psi}^t_{\mathsf{c}}$ as the asymptotic error covariance between the estimate $\boldsymbol{X}^t$ and the true message matrix $\boldsymbol{X}$, for the c-th block of users.

Bayes-optimal denoiser η_t: Since the empirical distribution of the rows of $(\boldsymbol{X}^t + \boldsymbol{V}^t)$ in block c converges to the law of $\bar{\boldsymbol{X}} + \boldsymbol{G}_{\mathsf{c}}^t$, the Bayes-optimal denoiser $\eta_{t,\mathsf{c}}^{\text{Bayes}}$ is the minimum mean squared error (MMSE) estimator for recovering a vector signal $\bar{\boldsymbol{X}}$ embedded in independent Gaussian noise with covariance $\boldsymbol{T}_{\mathsf{c}}^t$. This denoiser takes the following form. For $\mathsf{c} \in [\mathsf{C}]$ and $\boldsymbol{s} \in \mathbb{R}^{1\times k}$,

$$\eta_{t,\mathsf{c}}^{\text{Bayes}}(\boldsymbol{s}) = \mathbb{E}\{\bar{\boldsymbol{X}} | \bar{\boldsymbol{X}} + \boldsymbol{G}_{\mathsf{c}}^t = \boldsymbol{s}\} \tag{19}$$
$$= \frac{\frac{\alpha}{2^k}\sum_{\boldsymbol{x}'\in\mathcal{X}_{\text{a}}} \boldsymbol{x}' \exp\left(-\frac{1}{2}(\boldsymbol{x}' - 2\boldsymbol{s})\left(\boldsymbol{T}_{\mathsf{c}}^t\right)^{-1}\boldsymbol{x}'^{\mathsf{T}}\right)}{(1-\alpha) + \frac{\alpha}{2^k}\sum_{\tilde{\boldsymbol{x}}'\in\mathcal{X}_{\text{a}}} \exp\left(-\frac{1}{2}(\tilde{\boldsymbol{x}}' - 2\boldsymbol{s})\left(\boldsymbol{T}_{\mathsf{c}}^t\right)^{-1}\tilde{\boldsymbol{x}}'^{\mathsf{T}}\right)},$$

where we recall $\mathcal{X}_{\text{a}} = \{\pm\sqrt{E}\}^k$. While we use the denoiser in (19) for numerical simulations, Theorem 2 holds for any Lipschitz denoiser. The Bayes-optimal denoiser in (19) can be verified to be Lipschitz by direct differentiation. A limitation of this denoiser is that its computational cost is exponential in k, making it impractical for large k. We explore efficient alternatives in the extended version of this paper.

Hard-decision estimate: In each iteration $t \geq 1$, the decoder can produce a hard-decision maximum a posteriori (MAP) estimate of the message matrix, denoted by $\hat{\boldsymbol{X}}^{t+1} = h_t(\boldsymbol{X}^t + \boldsymbol{V}^t) \in \mathbb{R}^{L\times k}$. Analogous to η_t, h_t acts row-wise on matrix inputs, with $h_{t,\mathsf{c}}$ denoting the function acting on the c-th block of the matrix input, where $h_{t,\mathsf{c}}(\boldsymbol{s}) = \arg\max_{\boldsymbol{x}'\in\mathcal{X}_{\text{a}}\cup\boldsymbol{0}} \mathbb{P}\left(\bar{\boldsymbol{X}} = \boldsymbol{x}' \mid \bar{\boldsymbol{X}} + \boldsymbol{G}_{\mathsf{c}}^t = \boldsymbol{s}\right)$.

We now present Theorem 2, which characterizes the asymptotic performance of SC-AMP (in terms of $p_{\text{MD}}, p_{\text{FA}}$, p_{AUE}) via its corresponding state evolution recursion in (17)-(18).

Theorem 2. *Consider the spatially coupled coding scheme, with the rows of the message matrix $\boldsymbol{X}$ i.i.d. according to the prior $p_{\bar{\boldsymbol{X}}}$ in* (12)*. Assume that for $t \geq 1$, the denoisers η_t used in the AMP decoder* (15) *are of the form in* (16)*, with $\eta_{t,\mathsf{c}}$ Lipschitz continuous for $\mathsf{c} \in [\mathsf{C}]$. Let $\hat{\boldsymbol{X}}^{t+1} = h_t(\boldsymbol{X}^t + \boldsymbol{V}^t)$ be the hard-decision estimate in iteration t. In the limit as $L, n \to \infty$ with $L/n \to \mu$, the p_{MD}, p_{FA} and p_{AUE} in iteration $t \geq 1$ satisfy the following almost surely:*

$$\lim_{L\to\infty} p_{\text{MD}} = \lim_{L\to\infty} \mathbb{E}\Big[\frac{\mathbb{1}\{K_{\text{a}} \neq 0\}}{K_{\text{a}}} \sum_{\ell:\boldsymbol{X}_\ell\neq\boldsymbol{0}} \mathbb{1}\{\hat{\boldsymbol{X}}_\ell^{t+1} = \boldsymbol{0}\}\Big]$$
$$= \frac{1}{C}\sum_{c=1}^{C} \mathbb{P}\left(h_{t,c}\left(\bar{\boldsymbol{X}}_{\text{a}} + \boldsymbol{G}_c^t\right) = \boldsymbol{0}\right), \tag{20}$$

$$\lim_{L\to\infty} p_{\text{FA}} = \lim_{L\to\infty} \mathbb{E}\Big[\frac{\mathbb{1}\{\widehat{K_{\text{a}}} \neq 0\}}{\widehat{K_{\text{a}}}} \sum_{\ell:\hat{\boldsymbol{X}}_\ell^{t+1}\neq\boldsymbol{0}} \mathbb{1}\{\boldsymbol{X}_\ell = \boldsymbol{0}\}\Big]$$
$$= \left[\frac{\alpha\sum_{c=1}^{C}\mathbb{P}(h_{t,c}(\bar{\boldsymbol{X}}_{\text{a}} + \boldsymbol{G}_c^t) \neq \boldsymbol{0})}{(1-\alpha)\sum_{c=1}^{C}\mathbb{P}(h_{t,c}(\boldsymbol{G}_c^t) \neq \boldsymbol{0})} + 1\right]^{-1}, \tag{21}$$

$$\lim_{L\to\infty} p_{\text{AUE}} = \lim_{L\to\infty} \mathbb{E}\Big[\frac{\mathbb{1}\{K_{\text{a}} \neq 0\}}{K_{\text{a}}} \sum_{\ell:\boldsymbol{X}_\ell\neq\boldsymbol{0}} \mathbb{1}\{\hat{\boldsymbol{X}}_\ell^{t+1} \neq \boldsymbol{X}_\ell\}\Big]$$
$$= \frac{1}{C}\sum_{c=1}^{C} \mathbb{P}\left(h_{t,c}\left(\bar{\boldsymbol{X}}_{\text{a}} + \boldsymbol{G}_c^t\right) \neq \bar{\boldsymbol{X}}_{\text{a}}\right). \tag{22}$$

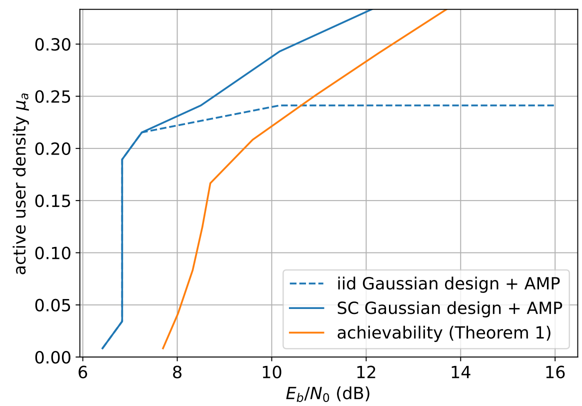

Figure 3. Minimum E_b/N_0 needed to achieve target error probability $\max\{p_{\text{MD}}, p_{\text{FA}}\} + p_{\text{AUE}} < 0.01$, for a range of active user densities μ_{a}. User payload $k = 6$, and $\alpha = 0.7$. Orange curve is the achievability bound in Theorem 1 computed with $L = 200$. Blue curves correspond to the guarantees computed via Theorem 2 with i.i.d. or SC design matrix.

Here $\bar{\boldsymbol{X}}_{\text{a}} \sim p_{\bar{\boldsymbol{X}}_{\text{a}}}$ and $\boldsymbol{G}_c^t \sim \mathcal{N}_k(\boldsymbol{0}, \boldsymbol{T}_c^t)$ are independent, with their laws given in (12) *and* (18) *respectively.*

The proof is given in the longer version of this paper.

B. Numerical Results

Fig. 3 plots the achievability bound of Theorem 1 (orange curve), along with the asymptotic guarantees of Theorem 2 for AMP decoding with i.i.d. Gaussian (dashed blue) or spatially coupled Gaussian design (solid blue). The user payload is $k = 6$ bits, and each user is active with probability $\alpha = 0.7$. For a range of active user densities $\mu_{\text{a}} = \alpha L/n$, we plot the minimum E_b/N_0 required for each coding scheme to achieve $\max\{p_{\text{MD}}, p_{\text{FA}}\} + p_{\text{AUE}} < 0.01$. We set the messages of active users to be $\boldsymbol{X}_\ell \in \{\pm 1\}^k$ $(E = 1)$ and vary E_b/N_0 via the noise variance σ^2.

The achievability bound in Theorem 1 was computed with $L = 200$ and decoding radius $r = L$. This means the maximum-likelihood decoder in (7) searches through all possible combinations of active users out of L, which gave the tightest achievability bounds in this setting. The curves for AMP decoding performance are obtained using the limiting error probabilities in Theorem 2, which can be computed using the state evolution recursion in (17)-(18). The state evolution depends on L, n only through the ratio $\mu = L/n$. The spatially coupled state evolution also depends on the parameters of the base matrix $\boldsymbol{W}$.

We observe that the asymptotic error probabilities of SC-AMP are consistently lower than those obtained from the finite-length achievability bounds. The achievability bound depends on n and L, and is challenging to compute for large n, L. A future direction is to derive asymptotic achievability bounds which depend only on $\mu = L/n$. The extended version of this paper will provide a more in-depth simulation study, including i) the empirical performance of the AMP decoder at finite n, L, and ii) the tradeoff between different types of errors `FA`, `MD`, `AUE` for different values of α.

REFERENCES

[1] X. Chen and D. Guo, "Many-access channels: The Gaussian case with random user activities," in *Proc. IEEE Int. Symp. Inf. Theory*, 2014, pp. 3127–3131.

[2] X. Chen, T.-Y. Chen, and D. Guo, "Capacity of Gaussian many-access channels," *IEEE Trans. Inf. Theory*, vol. 63, no. 6, pp. 3516–3539, 2017.

[3] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. IEEE Int. Symp. Inf. Theory*, 2017, pp. 2523–2527.

[4] I. Zadik, Y. Polyanskiy, and C. Thrampoulidis, "Improved bounds on Gaussian MAC and sparse regression via Gaussian inequalities," in *Proc. IEEE Int. Symp. Inf. Theory*, 2019, pp. 430–434.

[5] J. Ravi and T. Koch, "Scaling laws for Gaussian random many-access channels," *IEEE Trans. Inf. Theory*, vol. 68, no. 4, pp. 2429–2459, 2022.

[6] S. S. Kowshik, "Improved bounds for the many-user MAC," in *Proc. IEEE Int. Symp. Inf. Theory*, 2022, pp. 2874–2879.

[7] K. Hsieh, C. Rush, and R. Venkataramanan, "Near-optimal coding for many-user multiple access channels," *IEEE Journal on Selected Areas in Information Theory*, vol. 3, no. 1, pp. 21–36, 2022.

[8] O. Ordentlich and Y. Polyanskiy, "Low complexity schemes for the random access Gaussian channel," in *Proc. IEEE Int. Symp. Inf. Theory*, 2017, pp. 2528–2532.

[9] R. C. Yavas, V. Kostina, and M. Effros, "Gaussian multiple and random access channels: Finite-blocklength analysis," *IEEE Trans. Inf. Theory*, vol. 67, no. 11, pp. 6983–7009, 2021.

[10] ——, "Random access channel coding in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 67, no. 4, p. 2115–2140, 2021.

[11] K.-H. Ngo, A. Lancho, G. Durisi, and A. Graell i Amat, "Unsourced multiple access with random user activity," *IEEE Trans. Inf. Theory*, vol. 69, no. 7, pp. 4537–4558, 2023.

[12] V. K. Amalladinne, J.-F. Chamberland, and K. R. Narayanan, "A coded compressed sensing scheme for unsourced multiple access," *IEEE Trans. Inf. Theory*, vol. 66, no. 10, pp. 6509–6533, 2020.

[13] A. Fengler, P. Jung, and G. Caire, "SPARCs for unsourced random access," *IEEE Trans. Inf. Theory*, vol. 67, no. 10, pp. 6894–6915, 2021.

[14] V. K. Amalladinne, A. K. Pradhan, C. Rush, J.-F. Chamberland, and K. R. Narayanan, "Unsourced random access with coded compressed sensing: Integrating AMP and belief propagation," *IEEE Trans. Inf. Theory*, vol. 68, no. 4, pp. 2384–2409, 2022.

[15] R. Venkataramanan, S. Tatikonda, and A. Barron, "Sparse regression codes," *Foundations and Trends® in Communications and Information Theory*, vol. 15, no. 1-2, pp. 1–195, 2019.

[16] D. L. Donoho, A. Javanmard, and A. Montanari, "Information-theoretically optimal compressed sensing via spatial coupling and approximate message passing," *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7434–7464, 2013.

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

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

[19] O. Y. Feng, R. Venkataramanan, C. Rush, and R. J. Samworth, "A unifying tutorial on approximate message passing," *Foundations and Trends in Machine Learning*, vol. 15, no. 4, pp. 335–536, 2022.

[20] J. Barbier and F. Krzakala, "Approximate message-passing decoder and capacity achieving sparse superposition codes," *IEEE Trans. Inf. Theory*, vol. 63, no. 8, pp. 4894–4927, 2017.

[21] J. Barbier, M. Dia, and N. Macris, "Universal sparse superposition codes with spatial coupling and GAMP decoding," *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5618–5642, 2019.

[22] C. Rush, K. Hsieh, and R. Venkataramanan, "Capacity-achieving spatially coupled sparse superposition codes with amp decoding," *IEEE Trans. Inf. Theory*, vol. 67, no. 7, pp. 4446–4484, 2021.

[23] P. Pascual Cobo, K. Hsieh, and R. Venkataramanan, "Bayes-optimal estimation in generalized linear models via spatial coupling," in *Proc. IEEE Int. Symp. Inf. Theory*, 2023, pp. 773–778, Extended version at: https://arxiv.org/abs/2309.08404.