Spatially Coupled Sparse Regression Codes: Design and State Evolution Analysis

Kuan Hsieh
University of Cambridge, UK
kh525@cam.ac.uk

Cynthia Rush
Columbia University, USA
cynthia.rush@columbia.edu

Ramji Venkataramanan
University of Cambridge, UK
ramji.v@eng.cam.ac.uk

***Abstract*—We consider the design and analysis of spatially coupled sparse regression codes (SC-SPARCs), which were recently introduced by Barbier et al. for efficient communication over the additive white Gaussian noise channel. SC-SPARCs can be efficiently decoded using an Approximate Message Passing (AMP) decoder, whose performance in each iteration can be predicted via a set of equations called state evolution. In this paper, we give an asymptotic characterization of the state evolution equations for SC-SPARCs. For any given base matrix (that defines the coupling structure of the SC-SPARC) and rate, this characterization can be used to predict whether AMP decoding will succeed in the large system limit. We then consider a simple base matrix defined by two parameters (ω, Λ), and show that AMP decoding succeeds in the large system limit for all rates $R < \mathcal{C}$. The asymptotic result also indicates how the parameters of the base matrix affect the decoding progression. Simulation results are presented to evaluate the performance of SC-SPARCs defined with the proposed base matrix.**

I. Introduction

We consider communication over the memoryless additive white Gaussian noise (AWGN) channel, in which the output y is generated from input x according to $y = x + w$. The noise w is Gaussian with zero mean and variance σ^2, and the input x has an average power constraint P. If $x_1, x_2, \ldots, x_n$ are transmitted over n uses of the channel then

$$\frac{1}{n}\sum_{i=1}^{n} x_i^2 \leq P. \tag{1}$$

The Shannon capacity of this channel is given by $\mathcal{C} = \frac{1}{2}\ln\left(1+\frac{P}{\sigma^2}\right)$ nats/transmission.

Sparse superposition codes, or sparse regression codes (SPARCs), were introduced by Joseph and Barron [1], [2] for efficient communication over the AWGN channel. These codes have been proven to be reliable at rates approaching $\mathcal{C}$ with various low complexity iterative decoders [2]–[4]. As shown in Fig. 1, a SPARC is defined by a design matrix A of dimensions $n \times ML$, where n is the code length and M, L are integers such that A has L sections with M columns each. Codewords are generated as linear combinations of L columns of A, with one column from each section. Thus a codeword can be represented as $A\beta$, with β being an $ML \times 1$ *message vector* with exactly one non-zero entry in each of its L sections. The message is indexed by the locations of the non-zero entries in β. The values of the non-zero entries are fixed a priori.

This work was supported in part by a Marie Curie Career Integration Grant under grant agreement 631489, and by an EPSRC Doctoral Training Award.

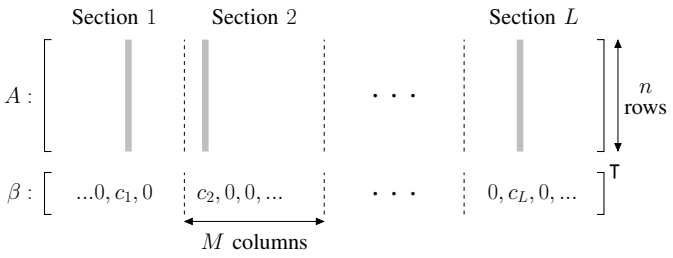

Fig. 1: A is an $n \times ML$ design matrix and β is an $ML \times 1$ message vector with one non-zero entry in each of its L sections. Codewords are of the form $A\beta$. The non-zero values $c_1, \ldots, c_L$ are fixed a priori.

Since there are M choices for the location of the non-zero entry in each of the L sections, there are M^L codewords. To achieve a communication rate of R nats/transmission, we therefore require

$$M^L = e^{nR} \quad \text{or} \quad nR = L\ln M. \tag{2}$$

In the standard SPARC construction introduced in [1], [2], the design matrix A is constructed with i.i.d. standard Gaussian entries. The values of the non-zero coefficients in the message vector β then define a *power allocation* across sections. With an appropriately chosen power allocation (e.g., one that is exponentially decaying across sections), the feasible decoders proposed in [2]–[4] have been shown to be asymptotically capacity-achieving. The choice of power allocation has also been shown to be crucial for obtaining good finite length performance with the standard SPARC construction [5].

Spatially coupled (SC) SPARCs, where the design matrix is composed of blocks with different variances, were recently proposed in [6]–[9]. In these works, an approximate message passing (AMP) algorithm was used for decoding, whose performance can be predicted via a recursion known as state evolution. The state evolution recursion was analyzed for a certain class of SC-SPARCs by Barbier et al. [6], using the potential function method introduced in [10], [11]. It was shown in [6] that for any $R < \mathcal{C}$, state evolution predicts vanishing probability of decoding error in the limit of large code length.

As in [6], we analyze the AMP decoder for SC-SPARCs via the associated state evolution recursion. However, the analysis in this paper does not use the potential function method;

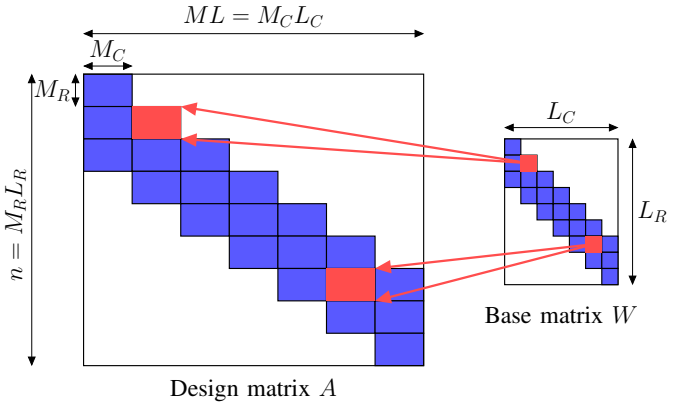

Fig. 2: A spatially coupled design matrix A is divided into blocks of size $M_R \times M_C$. There are L_R and L_C blocks in each column and row respectively. The independent matrix entries are normally distributed, $A_{ij} \sim \mathcal{N}(0, \frac{1}{L} W_{\mathsf{r}(i)\mathsf{c}(j)})$, where W is the base matrix. The base matrix shown here is an (ω, Λ) base matrix with parameters $\omega = 3$ and $\Lambda = 7$. The white parts of A and W correspond to zeros.

rather, it is based on a simple asymptotic characterization of the state evolution equations (Lemmas 1 and 2). This characterization gives insight into how the parameters defining the spatial coupling influence the decoding progression. For a given coupling matrix, the result can be used to determine whether reliable AMP decoding is possible in the large system limit. (In the rest of the paper, the terminology 'large system limit' or 'asymptotic limit' refers to (L, M, n) all tending to infinity such that $L \ln M = nR$.)

Using a simple base matrix inspired by protograph-based spatially coupled LDPC constructions, we show that state evolution predicts reliable AMP decoding for all rates $R < \mathcal{C}$, and bound the number of iterations required in the large system limit (Propostion 1). We also present numerical simulation to evaluate the finite length performances of SC-SPARCs constructed using the proposed base matrix.

We note that the results in this paper do not constitute a complete proof that SC-SPARCs are capacity-achieving. For this, one has to further show that the mean-squared error of the AMP estimates converges almost surely to the corresponding state evolution prediction. Obtaining such a result by extending the AMP analysis techniques for standard SPARCs [4], [12] is part of ongoing work.

II. Spatially coupled SPARC construction

In a spatially coupled (SC) SPARC, the design matrix A (see Fig. 1) consists of independent zero-mean normally distributed entries whose variances are specified by a *base matrix* W of dimension $L_R \times L_C$. The matrix A is obtained from the base matrix W by replacing each entry W_{rc}, for $r \in [L_R]$, $c \in [L_C]$, by an $M_R \times M_C$ block with i.i.d. entries $\sim \mathcal{N}(0, W_{rc}/L)$. (The set $\{1, 2, \ldots, k\}$ is denoted by $[k]$.) See Fig. 2 for an example, and note that $n = L_R M_R$ and $ML = L_C M_C$.

From the construction, the design matrix has independent normal entries

$$A_{ij} \sim \mathcal{N}\left(0, \frac{1}{L} W_{\mathsf{r}(i)\mathsf{c}(j)}\right) \;\; \forall\, i \in [n],\;\; j \in [ML]. \tag{3}$$

The operators $\mathsf{r}(\cdot) : [n] \to [L_R]$ and $\mathsf{c}(\cdot) : [ML] \to [L_C]$ in (3) map a particular row or column index to its corresponding *row block* or *column block* index. Conversely, we define operators $\mathsf{R}(\cdot)$ and $\mathsf{C}(\cdot)$ which map row and column block indices to the *set* of row and column indices they correspond to, i.e.,

$$\begin{aligned} \mathsf{R}(r) &= \{(r-1)M_R + 1, \ldots, rM_R\} \text{ for } r \in [L_R], \\ \mathsf{C}(c) &= \{(c-1)M_C + 1, \ldots, cM_C\} \text{ for } c \in [L_C]. \end{aligned} \tag{4}$$

Therefore, $|\mathsf{R}(r)| = M_R$ and $|\mathsf{C}(c)| = M_C$, for $r \in [L_R]$ and $c \in [L_C]$. We also require L_C to divide L, resulting in $\frac{L}{L_C}$ sections per column block.

The non-zero coefficients of β (see Fig. 1) are all set to 1, i.e., $c_1 = c_2 = \ldots = c_L = 1$. Then, for any base matrix W, it can be shown that the entries must satisfy

$$\frac{1}{L_R L_C} \sum_{r=1}^{L_R} \sum_{c=1}^{L_C} W_{rc} = P \tag{5}$$

in order to satisfy the average power constraint in (1).

The trivial base matrix with $L_R = L_C = 1$ corresponds to a standard (non-SC) SPARC without power allocation [1], while a single-row base matrix $L_R = 1$, $L_C = L$ is equivalent to standard SPARCs with power allocation [2], [4]. In this paper, we use the following base matrix inspired by the coupling structure of SC-LDPC codes constructed from protographs [13].

Definition 1: An (ω, Λ) base matrix W for SC-SPARCs is described by two parameters: coupling width $\omega \geq 1$ and coupling length $\Lambda \geq 2\omega - 1$. The matrix has $L_R = \Lambda + \omega - 1$ rows, $L_C = \Lambda$ columns, with each column having ω identical non-zero entries. For an average power constraint P, the (r, c)th entry of the base matrix, for $r \in [L_R], c \in [L_C]$, is given by

$$W_{rc} = \begin{cases} P \cdot \frac{\Lambda + \omega - 1}{\omega} & \text{if } c \leq r \leq c + \omega - 1, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

The base matrix in Fig. 2 has parameters $\omega = 3$ and $\Lambda = 7$.

Each non-zero entry in an (ω, Λ) base matrix W corresponds to an $M_R \times (ML/L_C)$ block in the design matrix A. Each block can be viewed as a standard (non-SC) SPARC with $\frac{L}{L_C}$ sections (with M columns in each section), code length M_R, and rate $R_{\text{inner}} = \frac{(L/L_C) \ln M}{M_R}$ nats. Using (2), the overall rate of the SC-SPARC is related to R_{inner} according to

$$R = \frac{\Lambda}{\Lambda + \omega - 1} R_{\text{inner}}. \tag{7}$$

With spatial coupling, ω is an integer greater than 1, so $R < R_{\text{inner}}$, which is often referred to as a rate loss. The rate loss depends on the ratio $(\omega - 1)/\Lambda$, which becomes negligible when Λ is large w.r.t. ω.

Remark 1: SC-SPARC constructions generally have a 'seed' to jumpstart decoding. In [6], a small fraction of sections of β are fixed a priori — this pinning condition is used to analyze the state evolution equations via the potential function method. Analogously, in the construction in [9], additional rows are introduced in the design matrix for the blocks corresponding

to the first row of the base matrix. In an (ω, Λ) base matrix, the fact that the number of rows in the base matrix exceeds the number of columns by $(\omega - 1)$ helps decoding start from both ends. The asymptotic state evolution equations derived in Sec. IV-A show how AMP decoding progresses in an (ω, Λ) base matrix.

III. AMP DECODER

The decoder aims to recover the message vector $\beta \in \mathbb{R}^{ML}$ from the channel output sequence $y = A\beta + w \in \mathbb{R}^n$. Approximate Message Passing (AMP) [14], [15] refers to a class of iterative algorithms that are Gaussian/quadratic approximations of loopy belief propagation for certain high-dimensional estimation problems (e.g., compressed sensing and low-rank matrix estimation). For decoding SC-SPARCs, we use the following AMP decoder, which is similar to the one used in [9], and can be derived from the Generalized Approximate Message Passing algorithm [15] by using the variances specified by the base matrix for the blocks of A.

Given the channel output sequence y, the AMP decoder generates successive estimates of the message vector, denoted by $\beta^t \in \mathbb{R}^{ML}$, for $t = 0, 1, \ldots$. It initialises β^0 to the all-zero vector, and for $t \geq 0$, iteratively computes

$$\begin{aligned} z^t &= y - A\beta^t + \widetilde{\mathsf{b}}^t \odot z^{t-1} \\ \beta^{t+1} &= \eta\left(\beta^t + \widetilde{\varsigma}^t \odot \left[A^{\mathsf{T}}\left(z^t \odot (\widetilde{\varphi}^t)^{-1}\right)\right], \widetilde{\varsigma}^t\right), \end{aligned} \tag{8}$$

where $\odot$ is the Hadamard (element-wise) product, and z^{-1} is set to the all zero vector. The vector $(\widetilde{\varphi}^t)^{-1}$ denotes the element-wise inverse of $\widetilde{\varphi}^t \in \mathbb{R}^n$. The vectors $\widetilde{\mathsf{b}}^t \in \mathbb{R}^n$ and $\widetilde{\varphi}^t \in \mathbb{R}^n$ are obtained by repeating M_R times each entry of $\mathsf{b}^t \in \mathbb{R}^{L_R}$ and $\varphi^t \in \mathbb{R}^{L_R}$. Similarly, $\widetilde{\varsigma}^t \in \mathbb{R}^{ML}$ is obtained by repeating M_C times each entry of $\varsigma^t \in \mathbb{R}^{L_C}$. The vectors $\varsigma^t, \varphi^t, \mathsf{b}^t$ in (8) are computed as follows:

$$\varsigma^t = \frac{L}{M_R}\left[W^{\mathsf{T}}(\varphi^t)^{-1}\right]^{-1}, \tag{9}$$

and for $r \in [L_R]$,

$$\varphi_r^t = \|z^t_{\mathsf{R}(r)}\|_2^2 / M_R, \tag{10}$$

$$\mathsf{b}_r^t = \frac{1}{L_C}\left[\sum_{c=1}^{L_C} W_{rc}\left(1 - \frac{\|\beta^t_{\mathsf{C}(c)}\|_2^2}{L/L_C}\right)\right](\varphi_r^{t-1})^{-1}. \tag{11}$$

Finally, let $\mathrm{sec}(\ell)$ denote the set of column indices in the ℓ^{th} section, i.e., $\mathrm{sec}(\ell) \coloneqq \{(\ell-1)M + 1, \ldots, \ell M\}$ for $\ell \in [L]$. The denoising function $\eta : \mathbb{R}^{ML} \times \mathbb{R}^{ML} \to \mathbb{R}^{ML}$ is written as $\eta = (\eta_1, \ldots, \eta_{ML})$, where for $j \in [ML]$ such that $j \in \mathrm{sec}(\ell)$,

$$\eta_j(s, \widetilde{\varsigma}) = \frac{e^{s_j/\widetilde{\varsigma}_j}}{\sum_{j' \in \mathrm{sec}(\ell)} e^{s_{j'}/\widetilde{\varsigma}_{j'}}}. \tag{12}$$

Notice that $\eta_j(s, \widetilde{\varsigma})$ depends on all the components of s and $\widetilde{\varsigma}$ in the section containing j.

When the change in β^t (or one of the other parameters) across successive iterations falls below a pre-specified tolerance, or the decoder reaches a maximum allowed iteration number, we take the latest AMP estimate, and set the largest entry in each section to 1 and the remaining entries to zero. This gives the decoded message vector, denoted by $\widehat{\beta}$.

Interpretation of the AMP decoder: The first argument of the $\eta(\cdot, \cdot)$ in (8), denoted by s^t, can be viewed as a noisy version of β. In particular, the entry s^t is approximately distributed as $\beta + \sqrt{\widetilde{\varsigma}^t} Z$, where Z is a standard normal random vector independent of β. Recall that $\beta_{\mathsf{C}(c)} \in \mathbb{R}^{M_C}$ is the part of the message vector corresponding to column block c of the design matrix. Then, for $c \in [L_C]$, the scalar ς_c^t is an estimate of the noise variance in block c of the effective observation s^t, i.e. $\varsigma_c^t \approx \frac{1}{M_C}\|s^t_{\mathsf{C}(c)} - \beta_{\mathsf{C}(c)}\|_2^2$. Under the above distributional assumption, the denoising function η_j in (12) is the minimum mean squared error (MMSE) estimator for β_j, i.e.,

$$\eta_j(s, \widetilde{\varsigma}) = \mathbb{E}\left[\beta_j | s = \beta + \widetilde{\varsigma} Z\right], \qquad \text{for } j \in [ML],$$

where the expectation is calculated over β and Z, with the location of the non-zero entry in each section of β being uniformly distributed within the section.

The vector $z^t \in \mathbb{R}^n$ in (8) is a residual vector, modified with the 'Onsager' term $\widetilde{\mathsf{b}}^t \odot z^{t-1}$. This term arises naturally in the derivation of the AMP algorithm, and is crucial for good decoding performance. For intuition about the role of the Onsager term, see [16, Sec. I-C] and [11, Sec. VI]. Finally, for $r \in [L_R]$, the scalar φ_r^t is an estimate of the variance of the rth block of the residual $z^t_{\mathsf{R}(r)}$. The residual has approximately zero mean, hence (10) is used to estimate its variance.

The key difference between the AMP decoder in (8) and the one for standard (non-SC) SPARCs in [4] is that in the latter case, the variance φ^t is a scalar that does not depend on the row index of the base matrix.

IV. PERFORMANCE OF THE AMP DECODER

The performance of a SPARC decoder is measured by the *section error rate*, defined as $\mathcal{E}_{\text{sec}} \coloneqq \frac{1}{L}\sum_{\ell=1}^{L} \mathbb{1}\{\widehat{\beta}_{\text{sec}(\ell)} \neq \beta_{\text{sec}(\ell)}\}$. Here $\mathbb{1}$ is the indicator function and $\beta_{\text{sec}(\ell)}$ is the length M vector corresponding to the ℓ^{th} section of the message vector. If the AMP decoder is run for T steps, then the section error rate can be bounded in terms of the *normalized mean square error* (NMSE) $\frac{1}{L}\|\beta^T - \beta\|^2$ of the AMP decoder [4]. The NMSE can be predicted using a recursion called state evolution. For an SC-SPARC defined by base matrix W, state evolution (SE) iteratively defines vectors $\phi^t \in \mathbb{R}^{L_R}$ and $\psi^t \in \mathbb{R}^{L_C}$ as follows. Initialize $\psi_c^0 = 1$ for $c \in [L_C]$, and for $t = 0, 1, \ldots$, compute

$$\phi_r^t = \sigma^2 + \frac{1}{L_C}\sum_{c=1}^{L_C} W_{rc}\psi_c^t, \qquad r \in [L_R], \tag{13}$$

$$\psi_c^{t+1} = 1 - \mathcal{E}(\tau_c^t), \qquad c \in [L_C], \tag{14}$$

where $\tau_c^t = \frac{R}{\ln M}\left[\frac{1}{L_R}\sum_r \frac{W_{rc}}{\phi_r^t}\right]^{-1}$, and

$$\mathcal{E}(\tau_c^t) = \mathbb{E}\left[\frac{e^{U_1/\sqrt{\tau_c^t}}}{e^{U_1/\sqrt{\tau_c^t}} + e^{-\frac{1}{\tau_c^t}}\sum_{j=2}^{M} e^{U_j/\sqrt{\tau_c^t}}}\right], \tag{15}$$

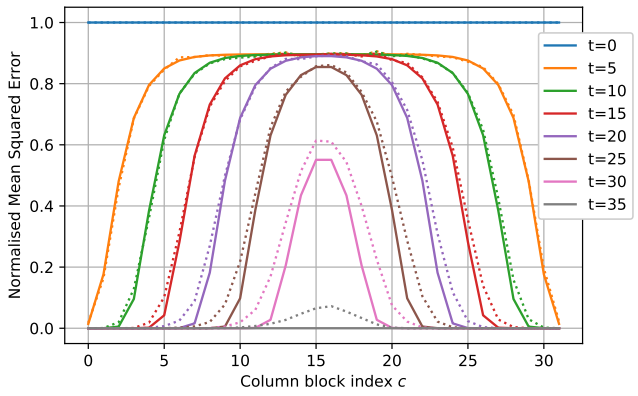

Fig. 3: NMSE $\frac{\|\beta^t_{\mathsf{C}(c)}-\beta_{\mathsf{C}(c)}\|_2^2}{L/L_C}$ vs. column block index $c \in [L_C]$ for several iteration numbers. The SC-SPARC with an (ω, Λ) base matrix uses parameters: $R = 1.5$ bits, $\mathcal{C} = 2$ bits, $\omega = 6$, $\Lambda = 32$, $M = 512$, $L = 2048$ and $n = 12284$. The solid lines are the SE predictions from (14), and the dotted lines are the average NMSE over 100 instances of AMP decoding.

with $U_1, \ldots, U_M \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. The SE equations in (13)-(14) are analogous to those for compressed sensing with spatially coupled measurement matrices [11, Eq. (32)-(33)], but modified to account for the section-wise structure of β.

As illustrated in Fig. 3, ψ^t closely tracks the NMSE of each block of the message vector, i.e., $\psi^t_c \approx \frac{\|\beta^t_{\mathsf{C}(c)}-\beta_{\mathsf{C}(c)}\|_2^2}{L/L_C}$ for $c \in [L_C]$. We additionally observe from the figure that as AMP iterates, the NMSE reduction propagates from the ends towards the center blocks. This decoding propagation phenomenon can be explained using the asymptotic state evolution analysis in the next subsection.

A. Asymptotic state evolution

Note that $\mathcal{E}(\tau^t_c)$ in (15) takes a value in $[0,1]$. If $\mathcal{E}(\tau^t_c) = 1$, then $\psi^{t+1}_c = 0$, which means that the sections with indices in $\mathsf{C}(c)$ will decode correctly. If we terminate the AMP decoder at iteration T, we want $\psi^T_c = 0$, for $c \in [L_C]$, so that the entire message vector is decoded correctly. The condition under which $\mathcal{E}(\tau^t_c)$ equals 1 in the large system limit is specified by the following lemma. The proof is omitted due to space constraints, and can be found in [17, Sec. IV].

Lemma 1: In the limit as the section size $M \to \infty$, the expectation $\mathcal{E}(\tau^t_c)$ in (15) converges to either 1 or 0 as follows.

$$\lim_{M\to\infty} \mathcal{E}(\tau^t_c) = \begin{cases} 1 & \text{if } \frac{1}{L_R}\sum_{r=1}^{L_R} \frac{W_{rc}}{\phi^t_r} > 2R \\ 0 & \text{if } \frac{1}{L_R}\sum_{r=1}^{L_R} \frac{W_{rc}}{\phi^t_r} < 2R. \end{cases} \tag{16}$$

This results in the following asymptotic state evolution recursion. Initialise $\bar{\psi}^0_c = 1$, for $c \in [L_C]$, and for $t = 0, 1, 2, \ldots$,

$$\bar{\phi}^t_r = \sigma^2 + \frac{1}{L_C}\sum_{c=1}^{L_C} W_{rc}\bar{\psi}^t_c, \qquad r \in [L_R], \tag{17}$$

$$\bar{\psi}^{t+1}_c = 1 - \mathbb{1}\left\{\frac{1}{L_R}\sum_{r=1}^{L_R} \frac{W_{rc}}{\bar{\phi}^t_r} > 2R\right\}, \qquad c \in [L_C], \tag{18}$$

where $\bar{\phi}, \bar{\psi}$ indicate asymptotic values.

Remark 2: The term $\frac{1}{L_R}\sum_r \frac{W_{rc}}{\phi^t_r}$ in (16) represents the average signal to effective noise ratio after iteration t for the column index c. If this quantity exceeds the prescribed threshold of $2R$, then the c^{th} block of the message vector, $\beta_{\mathsf{C}(c)}$, will be decoded at the next iteration in the large system limit, i.e., $\psi^{t+1}_c = 0$.

The asymptotic SE recursion (17)-(18) is given for a general base matrix W. We now apply it to the (ω, Λ) base matrix introduced in Definition 1. Recall that an (ω, Λ) base matrix has $L_R = \Lambda + \omega - 1$ rows and $L_C = \Lambda$ columns, with each column having ω non-zero entries, all equal to $P \cdot \frac{\Lambda+\omega-1}{\omega}$.

Lemma 2: The asymptotic SE recursion (17)-(18) for an (ω, Λ) base matrix W is as follows. Initialise $\bar{\psi}^0_c = 1 \ \forall \ c \in [\Lambda]$, and for $t = 0, 1, 2, \ldots$,

$$\bar{\phi}^t_r = \sigma^2\left(1 + \frac{\kappa \cdot \text{snr}}{\omega}\sum_{c=\underline{c}_r}^{\bar{c}_r} \bar{\psi}^t_c\right), \quad r \in [\Lambda + \omega - 1], \tag{19}$$

$$\bar{\psi}^{t+1}_c = 1 - \mathbb{1}\left\{\frac{P}{\omega}\sum_{r=c}^{c+\omega-1} \frac{1}{\bar{\phi}^t_r} > 2R\right\}, \quad c \in [\Lambda], \tag{20}$$

where $\kappa = \frac{\Lambda+\omega-1}{\Lambda}$, $\text{snr} = \frac{P}{\sigma^2}$, and

$$(\underline{c}_r, \bar{c}_r) = \begin{cases} (1, r) & \text{if } 1 \le r \le \omega \\ (r-\omega+1, r) & \text{if } \omega \le r \le \Lambda \\ (r-\omega+1, \Lambda) & \text{if } \Lambda \le r \le \Lambda+\omega-1. \end{cases} \tag{21}$$

Proof: Substitute the value of W_{rc} from (6), and $L_C = \Lambda$, $L_R = \Lambda + \omega - 1$ in (17)-(18). ∎

Observe that the $\bar{\phi}^t_r$'s and $\bar{\psi}^t_c$'s are symmetric about the middle indices, i.e. $\bar{\phi}^t_r = \bar{\phi}^t_{L_R-r+1}$ for $r \le \lfloor \frac{L_R}{2} \rfloor$ and $\bar{\psi}^t_c = \bar{\psi}^t_{L_C-c+1}$ for $c \le \lfloor \frac{L_C}{2} \rfloor$.

Lemma 2 gives insight into the decoding progression for a large SC-SPARC defined using an (ω, Λ) base matrix. On initialization ($t = 0$), the value of $\bar{\phi}^0_r$ for each r depends on the number of non-zero entries in row r of W, which is equal to $\bar{c}_r - \underline{c}_r + 1$, with $\bar{c}_r, \underline{c}_r$ given by (21). Therefore, $\bar{\phi}^0_r$ increases from $r = 1$ until $r = \omega$, is constant for $\omega \le r \le \Lambda$, and then starts decreasing again after $r = \Lambda$. As a result, $\bar{\psi}^1_c$ is smallest for c at either end of the base matrix ($c \in \{1, \Lambda\}$) and increases as c moves towards the middle, since the $\sum_{r=c}^{c+\omega-1}(\bar{\phi}^0_r)^{-1}$ term in (20) is largest for $c \in \{1, \Lambda\}$, followed by $c \in \{2, \Lambda - 1\}$, and so on. Therefore, we expect the blocks of the message vector corresponding to column index $c \in \{1, \Lambda\}$ to be decoded most easily, followed by $c \in \{2, \Lambda - 1\}$, and so on. Fig. 3 shows that this is indeed the case.

B. Asymptotic State Evolution analysis

With a slight abuse of terminology, we will use the phrase "column c is decoded in iteration t" to mean $\bar{\psi}^t_c = 0$.

Proposition 1: Consider a SC-SPARC constructed using an (ω, Λ) base matrix with rate $R < \frac{1}{2\kappa}\ln(1 + \kappa \cdot \text{snr})$, where $\kappa = \frac{\Lambda+\omega-1}{\Lambda}$. (Note that $\frac{1}{2\kappa}\ln(1 + \kappa \cdot \text{snr}) \in [\mathcal{C}/\kappa, \mathcal{C}]$.) Then, according to the asymptotic state evolution equations in

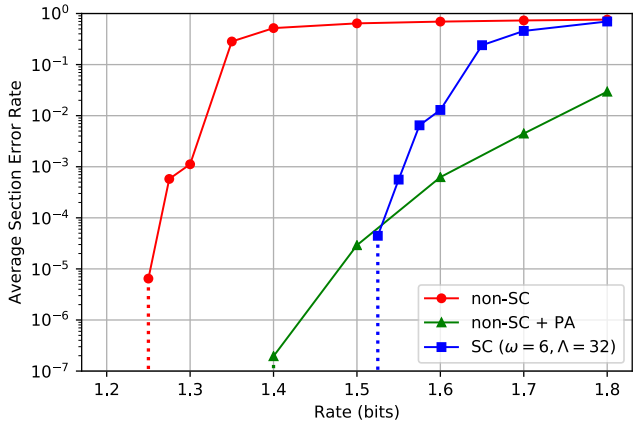

Fig. 4: Average section error rate (SER) vs. rate at snr $= 15$, $\mathcal{C} = 2$ bits, $M = 512$, $L = 1024$, $n \in [5100, 7700]$. The SERs are averaged over 10^4 trials. Plots are shown for non-SC SPARCs with and without power allocation, and SC-SPARCs with an (ω, Λ) base matrix with $\omega = 6, \Lambda = 32$. The code length is the same for the three cases. The dotted vertical lines indicate that no section errors were observed over 10^4 trials at smaller rates.

Lemma 2, the following statements hold in the large system limit:

1) The AMP decoder will be able to start decoding if
$$\omega > \left(\frac{1}{e^{2R\kappa} - 1} - \frac{1}{\kappa \cdot \text{snr}} \right)^{-1}. \tag{22}$$
2) If (22) is satisfied, then the sections in the first and last c^* blocks of the message vector will be decoded in the first iteration (i.e. $\bar{\psi}_c^1 = 0$ for $c \in \{1, 2, \ldots, c^*\} \cup \{\Lambda - c^* + 1, \Lambda - c^* + 2, \ldots, \Lambda\}$), where
$$c^* \geq \min \Bigg\{ (\omega - 1), \left\lfloor \omega \cdot \frac{1 + \kappa \cdot \text{snr}}{(\kappa \cdot \text{snr})^2} \cdot [\ln(1 + \kappa \cdot \text{snr}) - 2R\kappa] \right\rfloor \Bigg\}. \tag{23}$$
3) At least $2c^*$ additional columns will decode in each subsequent iteration until the message is fully decoded. Therefore, the AMP decoder will fully decode in at most $\left\lceil \frac{\Lambda}{2c^*} \right\rceil$ iterations.

The proof of the proposition is given in [17, Sec. IV].

Remark 3: The proposition implies that for any rate $R < \mathcal{C}$, AMP decoding is successful in the large system limit, i.e., $\bar{\psi}_c^T = 0$ for all $c \in [\Lambda]$. Indeed, consider a rate $R = \mathcal{C}/\kappa_0$, for any constant $\kappa_0 > 1$. Then choose ω to satisfy (22) (with κ replaced by κ_0), and Λ large enough that $\kappa = \frac{\Lambda + \omega - 1}{\Lambda} \leq \kappa_0$. With this choice of (ω, Λ) and rate R, the conditions of the proposition are satisfied, and hence, all the columns decode in the large system limit.

Numerical Simulations: We evaluate the empirical performance of SC-SPARCs constructed from (ω, Λ) base matrices. For the simulations, we used a Hadamard based design matrix instead of a Gaussian one as it gives significant reduction in running time and required memory, with very similar error performance [4], [9]. Fig. 4 compares the average section error rate (SER) of spatially coupled SPARCs with standard (non-SC) SPARCs, both with and without power allocation (PA). The code length is the same for all three codes, and the power allocation was designed using the algorithm proposed in [5]. Comparing standard SPARCs without PA and SC-SPARCs, we see that spatial coupling significantly improves the error performance: the rate threshold below which the SER drops steeply to a negligible value is higher for SC-SPARCs. We also observe that at rates close to $\mathcal{C}$, standard SPARCs with PA have lower SER than SC-SPARCs. However, as the rate decreases, the drop in SER for standard SPARCs with PA is not as steep as that for SC-SPARCs. Additional simulation results, showing the effect of changing ω while keeping Λ fixed, are given in [17, Sec. V].

REFERENCES

[1] A. Joseph and A. R. Barron, "Least squares superposition codes of moderate dictionary size are reliable at rates up to capacity," *IEEE Trans. Inf. Theory*, vol. 58, pp. 2541–2557, May 2012.

[2] A. Joseph and A. R. Barron, "Fast sparse superposition codes have near exponential error probability for $R < \mathcal{C}$," *IEEE Trans. Inf. Theory*, vol. 60, no. 2, pp. 919–942, 2014.

[3] S. Cho and A. R. Barron, "Approximate iterative Bayes optimal estimates for high-rate sparse superposition codes," in *The Sixth Workshop on Inf. Theoretic Methods in Sci. and Eng.*, pp. 35–42, 2013.

[4] C. Rush, A. Greig, and R. Venkataramanan, "Capacity-achieving sparse superposition codes via approximate message passing decoding," *IEEE Trans. Inf. Theory*, vol. 63, pp. 1476–1500, March 2017.

[5] A. Greig and R. Venkataramanan, "Techniques for improving the finite length performance of sparse superposition codes," *IEEE Trans. Commun.*, vol. 66, pp. 905–917, March 2018.

[6] J. Barbier, M. Dia, and N. Macris, "Proof of threshold saturation for spatially coupled sparse superposition codes," in *Proc. IEEE Int. Symp. Inf. Theory*, 2016.

[7] J. Barbier, M. Dia, and N. Macris, "Threshold saturation of spatially coupled sparse superposition codes for all memoryless channels," *Proc. IEEE Inf. Theory Workshop*, 2016.

[8] J. Barbier, M. Dia, and N. Macris, "Universal Sparse Superposition Codes with Spatial Coupling and GAMP Decoding," ArXiv e-prints arXiv:1707.04203, July 2017.

[9] J. Barbier and F. Krzakala, "Approximate message-passing decoder and capacity achieving sparse superposition codes," *IEEE Trans. Inf. Theory*, vol. 63, pp. 4894–4927, Aug 2017.

[10] A. Yedla, Y.-Y. Jian, P. S. Nguyen, and H. D. Pfister, "A simple proof of Maxwell saturation for coupled scalar recursions," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 6943–6965, 2014.

[11] 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, pp. 7434–7464, Nov. 2013.

[12] C. Rush and R. Venkataramanan, "The error exponent of sparse regression codes with AMP decoding," in *Proc. IEEE Int. Symp. Inf. Theory*, 2017.

[13] D. G. M. Mitchell, M. Lentmaier, and D. J. Costello, "Spatially coupled LDPC codes constructed from protographs," *IEEE Trans. Inf. Theory*, vol. 61, pp. 4866–4889, Sept 2015.

[14] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proceedings of the National Academy of Sciences*, vol. 106, no. 45, pp. 18914–18919, 2009.

[15] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, 2011.

[16] M. Bayati and A. Montanari, "The dynamics of message passing on dense graphs, with applications to compressed sensing," *IEEE Trans. Inf. Theory*, vol. 57, pp. 764–785, Feb 2011.

[17] K. Hsieh, C. Rush, and R. Venkataramanan, "Spatially coupled sparse regression codes: Design and state evolution analysis," Online: https://arxiv.org/abs/1801.01796, 2018.