LONG-MEMORY MESSAGE-PASSING FOR SPATIALLY COUPLED SYSTEMS

Keigo Takeuchi

Toyohashi University of Technology, Dept. Electrical and Electronic Inf. Eng., Toyohashi, Japan

ABSTRACT

This paper addresses the reconstruction of sparse signals from spatially coupled, linear, and noisy measurements. A unified framework of rigorous state evolution is established for developing long-memory message-passing (LM-MP) in spatially coupled systems. LM-MP utilizes all previous messages to compute the current message while conventional MP only uses the latest messages. The unified framework is utilized to propose orthogonal approximate message-passing (OAMP) for spatially coupled systems. The framework for LM-MP is used as a technical tool to prove the convergence of state evolution for OAMP: Numerical results show that OAMP for spatially coupled systems is superior to that for systems without spatial coupling in the so-called waterfall region.

Index Terms—Compressed sensing, spatial coupling, long-memory message-passing, state evolution

1. INTRODUCTION

This paper addresses the reconstruction of $L$ unknown $N$-dimensional sparse signal vectors $\{x[\ell] \in \mathbb{R}^N : \ell \in \mathcal{L}_0 = \{0, \ldots, L-1\}\}$ from spatially coupled $M$-dimensional measurements $\{y[\ell] \in \mathbb{R}^M : \ell \in \mathcal{L}_W = \{0, \ldots, L+W-1\}\}$ with coupling width $W$ [1, 2, 3, 4], given by

$$y[\ell] = \sum_{w=0}^{W} \gamma[\ell][\ell-w]A[\ell][\ell-w]x[\ell-w] + w[\ell]. \quad (1)$$

In (1), $A[\ell][\ell] \in \mathbb{R}^{M \times N}$ denotes a known sensing matrix in row section $\ell \in \mathcal{L}_W$ and column section $l \in \mathcal{L}$. The additive white Gaussian noise (AWGN) vector $w[\ell] \sim \mathcal{N}(0,\sigma^2 I_M)$ has independent zero-mean Gaussian elements with variance $\sigma^2$. The coupling coefficient $\gamma[\ell][\ell] \in \mathbb{R}$ satisfies the normalization condition $L^{-1} \sum_{\ell=0}^{L-1} \sum_{w=0}^{W} \gamma^2[\ell+w][\ell] = 1$. The spatially coupled system (1) may be regarded as a mathematical model of cell-free massive multiple-input multiple-output (MIMO) [5] or sparse superposition codes [6, 7, 8].

Approximate message-passing (AMP) [2, 3, 4] is low-complexity message-passing (MP) for signal recovery in spatially coupled systems. When the sensing matrices $\{A[\ell][\ell]\}$ have independent and identically distributed (i.i.d.) elements with zero mean, Bayes-optimal AMP was proved to achieve the information-theoretic compression limit in the noiseless case [3, 9, 10]. However, AMP fails to converge for the other sensing matrices, such as the ill-conditioned [11] or non-zero mean [12] case.

Orthogonal AMP [13] or equivalently vector AMP [14] solves this convergence issue in AMP: Bayes-optimal orthogonal AMP was proved to converge [15, 16] and achieve the Bayes-optimal performance [14, 17] for orthogonally invariant sensing matrices—a generalization of zero-mean i.i.d. Gaussian matrices. However, Bayes-optimal orthogonal AMP requires the high-complexity linear minimum mean-square error (LMMSE) filter.

Long-memory (LM) MP [18, 19, 20, 21, 22, 23, 24] is an attractive approach to realize the advantages of both AMP and orthogonal/vector AMP: low complexity and Bayes-optimality. LM-MP in [21, 22] utilizes the same low-complexity filter as AMP and all previous messages to approximate the LMMSE filter while orthogonal/vector AMP uses the LMMSE filter to construct a sufficient statistic only from the latest messages [15, 16]. As a result, LM-MP in [21] was proved to achieve the Bayes-optimal performance for orthogonally invariant sensing matrices. Since the complexity issue is outside the main scope of this paper, however, it is left as future research for spatially coupled systems.

Another advantage of LM-MP is the convergence guarantee: Use of LM damping [21] guarantees the convergence of state evolution for LM-MP [15, 16]. This property of LM-MP was utilized as a technical tool to prove the convergence of state evolution for Bayes-optimal orthogonal/vector AMP without memory [15, 16]. This paper uses LM-MP as a tool for the convergence guarantee.

The purpose of this paper is to establish a unified framework of state evolution for LM-MP in spatially coupled systems. The framework is a generalization of state evolution for spatially coupled i.i.d. Gaussian matrices [10] to the orthogonally invariant and LM-MP cases. It is also a generalization of state evolution for the orthogonal invariance and LM-MP cases [20] to the spatial coupling case.

As a memoryless instance of the unified framework, this paper proposes orthogonal AMP (OAMP) for spatially coupled systems. The framework for LM-MP is utilized to prove the convergence of state evolution for the proposed OAMP.
Numerical results are presented to show the superiority of OAMP for spatially coupled systems to that for conventional systems without spatial coupling.

2. UNIFIED FRAMEWORK OF STATE EVOLUTION

To present a unified framework of state evolution for the spatially coupled system (1), we transform (1) into a vector system. Let $\mathcal{W}[\ell] = \{\max(\ell - (L - 1), 0), \ldots, \min(\ell, W)\}$ denote the set of indices for row section $\ell$ such that $\{A[\ell]\}[\ell - w] : w \in \mathcal{W}[\ell]$ are non-zero matrices in row section $\ell$ of the spatially coupled system (1). For $N_c[\ell] = |\mathcal{W}[\ell]|N$, the matrix $A[\ell] \in \mathbb{R}^{M \times N_c[\ell]}$ consists of the normalized matrices $|\mathcal{W}[\ell]|^{-1/2}A[\ell][\ell - w]$ for all $w \in \mathcal{W}[\ell]$. Then, the spatially coupled system (1) is transformed into

$$y[\ell] = A[\ell]x[\ell] + w[\ell],$$

with

$$x[\ell] = \sqrt{|\mathcal{W}[\ell]|} \text{vec}\left\{\gamma[\ell][\ell - w]x[\ell - w] : w \in \mathcal{W}[\ell]\right\},$$

where the notation $\text{vec}\{x_i : i = 1, \ldots, n\}$ denotes the column vector of the $x_i$'s. The system (2) looks like parallel systems without spatial coupling when we ignore the dependencies between $\{x[\ell]\}$ through the signal vectors $\{x[\ell]\}$ in (3).

We present the evolution used in the proposed framework of state evolution. Consider the singular-value decomposition (SVD) $A[\ell] = U[\ell]\Sigma[\ell]V[\ell]$. We define $\hat{X}[\ell] \in \mathbb{R}^{N_c[\ell]}$ as the vector that consists of all diagonal elements of $\Sigma[\ell]^T \Sigma[\ell]$. For $N_c = (W + 1)N$, we define the set of extended vectors $\Lambda = \{A[\ell]^T[0] \in \mathbb{R}^{N_c[\ell]} : \ell \in \mathcal{L}_W\}$. The extended vectors are used for a unified treatment of $\{X[\ell]\}$ in both bulk and boundary sections.

Similarly, we define the set of extended signal vectors as $\Omega = \{\{U[\ell]^T]\{w[\ell]\}^T \in \mathbb{R}^{N_c[\ell]} : \ell \in \mathcal{L}_W\}$. The set of extended signal vectors is written as $\mathcal{X} = \{\{S[\ell][\ell]\}^{\text{vec}}[x[\ell]] : \ell \in \mathcal{L}_W\}$. For deterministic selection matrices $\{S[\ell][\ell]\} \in \{0, 1\}^{N_c[\ell] \times LN}$, which select $N_c[\ell]$ different elements from an $LN$-dimensional vector multiplied from the right side. The selection matrices are used for flexibility of the unified framework.

Suppose that the dynamics of estimation errors for LM-MP can be described with a dynamical system with respect to four vectors $\{\vec{b}_t[\ell], \vec{m}^\text{ext}[\ell], \vec{h}_t[\ell], \vec{q}^\text{ext}[\ell]\} \in \mathbb{R}^{N_c[\ell]}$ for iteration $t$ and row section $\ell \in \mathcal{L}_W$. In particular, $\vec{h}_t[\ell]$ represents estimation errors just before denoising in LM-MP, which are proved to be asymptotically Gaussian-distributed. We define the matrix $\vec{B}[\ell] = \{\vec{b}_0[\ell], \ldots, \vec{b}_{t-1}[\ell]\} \in \mathbb{R}^{N_c[\ell] \times t}$ for all previous iterations and the set of extended matrices $\vec{B}_t = \{\{\vec{B}[\ell]\}^T \in \mathbb{R}^{N_c[\ell] \times t} : \ell \in \mathcal{L}_W\}$. Similarly, we define $\vec{H}_t[\ell] \in \mathbb{R}^{N_c[\ell] \times t}$ and $\vec{H}_t$. For two vector-valued functions $\phi_1[\ell] : \mathbb{R}^{N_c[\ell] \times (t+3)} \mathcal{L}_W \rightarrow \mathbb{R}^{N_c}$ and $\psi_1[\ell] : \mathbb{R}^{N_c[\ell] \times (t+2)} \mathcal{L}_W \rightarrow \mathbb{R}^{N_c}$, the proposed dynamical system with a general initial condition $\vec{q}^\text{ext}_0[\ell] = (I_{N_c[\ell]}, C)\psi_{-1}[\ell](\mathcal{X})$ is given by

$$\vec{b}_t[\ell] = V^T[\ell] \vec{q}^\text{ext}[\ell],$$

$$\vec{m}^\text{post}[\ell] = (I_{N_c[\ell]}, C)\phi_1[\ell](\vec{B}_{t+1}, \Lambda),$$

$$\vec{m}^\text{ext}[\ell] = \vec{m}^\text{post}[\ell] - \sum_{\tau=0}^t \xi_{A,\tau}[\ell] \vec{b}_\tau[\ell],$$

$$\vec{H}_t[\ell] = V[\ell] \vec{m}^\text{ext}[\ell],$$

$$\vec{q}^\text{post}_t[\ell] = (I_{N_c[\ell]}, C)\psi_1[\ell](\mathcal{H}_{t+1}[\ell], \mathcal{X}),$$

$$\vec{q}^\text{ext}_{t+1}[\ell] = \vec{q}^\text{post}_t[\ell] - \sum_{\tau=0}^t \xi_{B,\tau}[\ell] \vec{h}_\tau[\ell],$$

The coefficients $\xi_{A,\tau}[\ell]$ and $\xi_{B,\tau}[\ell]$ in the ONSAGER correction (6) and (9) are defined as

$$\xi_{A,\tau}[\ell] = \frac{\partial}{\partial \vec{b}_\tau[\ell]} \phi_1[\ell](\vec{B}_{t+1}, \Lambda),$$

$$\xi_{B,\tau}[\ell] = \frac{\partial}{\partial \vec{h}_\tau[\ell]} \psi_1[\ell](\mathcal{H}_{t+1}[\ell], \mathcal{X}),$$

where for $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ we have used $[\partial f(x)/\partial x] = \partial f(x)/\partial x$, and $[\partial f(x)/\partial x] = n^{-1} \sum_{i=1}^n \partial f_i(x)/\partial x$.

The general error model (4)–(9) for the spatial coupling case is a generalization of a conventional error model without spatial coupling [20]. The conventional error model was utilized to propose existing LM-MP algorithms [19, 20, 21, 22]. By designing the two functions $\phi_1[\ell]$ and $\psi_1[\ell]$ appropriately, the general error model (4)–(9) can represent the dynamics of estimation errors for LM-MP.

To present state evolution for the general error model (4)–(9), we postulate the following assumptions:

**Assumption 1.** For some $\epsilon > 0$, the signal vector $x[\ell]$ has i.i.d. elements with zero mean, unit variance, and a bounded $(2 + \epsilon)$th moment.

**Assumption 2.** The sensing matrices $A[\ell]$ in (2) are independent for all $\ell$. Each $A[\ell]$ is right-orthogonally invariant: $V[\ell]$ in the SVD is independent of $U[\ell]\Sigma[\ell]$ and Haar-distributed. Furthermore, the empirical eigenvalue distribution of $|\mathcal{W}[\ell]|A[\ell]^T A[\ell]$ converges almost surely to a compactly supported deterministic distribution with unit mean in the large system limit, where both $M$ and $N$ tends to infinity with the compression rate $\delta = M/N \in [0, 1]$ kept constant.

**Assumption 3.** The function $\phi_1[\ell]$ is separable with respect to all variables and proper2 Lipschitz-continuous with respect to $B_{t+1}$ and $\Lambda$ while $\psi_1[\ell]$ is separable and proper Lipschitz-continuous with respect to all variables. Furthermore, $||m^\text{ext}[\ell]|| \neq 0$ and $||q^\text{ext}[\ell]|| \neq 0$ hold for all $t$.

Let $L_t > 0$ denote a Lipschitz-constant for a $k$th-order pseudo-Lipschitz function $f_k : \mathbb{R}^n \rightarrow \mathbb{R}$: For all $x, y \in \mathbb{R}^n$, $f_k(x) - f_k(y) \leq L_t(1 + ||x||^{k-1} + ||y||^{k-1})||x - y||$ holds. The function $f = (f_1, \ldots, f_n)$ is said to be proper if $\lim_{t \to \infty} \sup_{n \to \infty} t^{-1} \sum_{i=1}^n L_t < \infty$ holds for all $j \in \mathbb{N}$.
Theorem 1. Postulate Assumptions 1, 2, and 3, and suppose that \( \phi_t(B_{t+1}, \Omega, \Lambda) : \mathbb{R}^{N_e \times (t+2)} \rightarrow \mathbb{R}^{N_e} \) is separable, second-order pseudo-Lipschitz with respect to \( B_{t+1} \) and \( \Omega, \Lambda \), and proper. Similarly, suppose that \( \psi_t(H_{t+1}, \mathcal{X}) : \mathbb{R}^{N_e \times (t+2)} \rightarrow \mathbb{R}^{N_e} \) is separable, second-order pseudo-Lipschitz, and proper function. Then,

\[
\begin{align*}
\langle \phi_t(B_{t+1}, \Omega, \Lambda) \rangle & \rightarrow 0, \\
\langle \psi_t(H_{t+1}, \mathcal{X}) \rangle & \rightarrow 0.
\end{align*}
\]

In (12), the set \( \mathcal{Z}_{A,t+1} = \{ [\mathbf{Z}_{A,t+1}^T, \mathbf{O}]^T : \ell \in \mathcal{L}_W \} \) is composed of independent random matrices for all \( \ell \). Each matrix \( \mathbf{Z}_{A,t+1}^T = (\mathbf{z}_{A,0}^T, \ldots, \mathbf{z}_{A,t}^T) \in \mathbb{R}^{N_e \times (t+1)} \) has zero-mean Gaussian random vectors with covariance \( \mathbb{E}[\mathbf{z}_{A,\tau}^T][\mathbf{z}_{A,\tau'}] = \mathbf{c}_{\tau,\tau'}[\mathbf{I}_{N_e}] \) for all \( \tau, \tau' \in \{0, \ldots, t\} \), with \( N_c^{-1}[\mathbf{q}_{\tau}^\mathsf{T}][\mathbf{q}_{\tau'}] \rightarrow \mathbf{c}_{\tau,\tau'} \). In (13), \( \mathbf{Z}_{B,t+1} \) is defined in the same manner as \( \mathbf{Z}_{A,t+1} \) with the exception of \( N_c^{-1}[\mathbf{m}_{\tau}^\mathsf{T}][\mathbf{m}_{\tau'}] \rightarrow \mathbf{c}_{\tau,\tau'} \).

Proof. See [25, Theorem 7].

Theorem 1 implies asymptotic Gaussianity for \( \tilde{B}_t \) and \( \tilde{H}_t \), which is an important property in deriving state evolution recursions for LMP. By defining the two functions \( \phi_t \) and \( \psi_t \) in (5) and (8) appropriately, Theorem 1 allows us to derive state evolution recursions for LMP.

3. ORTHOGONAL AMP

We propose Bayes-optimal OAMP for the spatially coupled system (2) on the basis of Theorem 1. See Algorithm 1 for the details of the proposed algorithm. Lines 9–12 of the algorithm correspond to the LMMSE estimation—module A—and element-wise nonlinear estimation—called module B—respectively.

Module A computes the LMMSE estimator of \( \hat{x}[t] \) in line 5. Then, the ONSager correction in lines 7 and 8 is performed to realize the asymptotic Gaussianity for the error vector \( \tilde{h}_t = \hat{x}_{A,t} - [\hat{W}[t]]^{-1/2} \hat{x} \).

Module B transforms the message \( \hat{x}_{A,t+1} \) in the extended signal space \( \mathbb{R}^{N_e} \) into the sufficient statistic \( x_{A+}^\text{opt} \) for the original signal vector \( x[t] \in \mathbb{R}^N \) in line 12. After denoising in line 13 and transforming \( x_{B,t+1}^\text{est} \in \mathbb{R}^N \) into \( x_{B,t+1} \) for the extended signal space in line 17, the ONSager correction is computed in lines 19 and 20. Lines 21 and 22 represent the damping steps with \( \zeta \in (0,1) \) to improve the convergence property of OAMP for finite \( M \) and \( N \).

The functions with the indexes \( f_{\text{opt}}, \text{Var} \) in lines 13 and 14 denote the Bayes-optimal denoiser \( f_{\text{opt}}(u,v) = \mathbb{E}[x_1[0]|x_1[0] = u] \) and the corresponding posterior variance \( \text{Var}(u) = \mathbb{E}[x_1[0]|x_1[0] = u] \mathbb{E}[x_1[0]|x_1[0] = u] \), with \( z \sim N(0, 1) \). The validity of these definitions is justified in the large system limit via Theorem 1. We have used a popular notation in the MP community for lines 13 and 14: For a scalar function \( f : \mathbb{R} \rightarrow \mathbb{R} \) the notation \( f(x) \) denotes the element-wise application of \( f \) to the vector \( x \), i.e., \( f([x]) = f([x])_i \). Assume no damping \( \zeta = 1 \) in lines 21 and 22 for theoretical analysis. As proved in [25, Lemma 8], the dynamics of the error vectors \( \tilde{h}_t = \tilde{x}_{A,t} - [\hat{W}[t]]^{-1/2} \hat{x} \) and \( \tilde{d}_t^{\mathsf{est}} = \tilde{x}_{B,t} - \hat{x} \) can be described with the general error model (4)–(9). State evolution recursions for OAMP are obtained using the unified framework of state evolution in Theorem 1.

We first define state evolution recursions for OAMP. To distinguish variables in the state evolution recursions from the variance parameters in OAMP, we use the notation \( \tilde{v}_{B,t+1} = \tilde{v}_{B,t+1}^\text{est} \).
recursively defined in the same manner as in lines 1, 4, 6, 8, 11, 14, 18, and 20. The exceptional steps are in lines 6 and 14, which are replaced with
\[
\tilde{\eta}_{A,t}[\ell] = \lim_{M=\delta N \to \infty} \frac{1}{N} \text{Tr} \left( I_{N,\ell} \right) - W_t^T[\ell] A[\ell] \tag{14}
\]
in the large system limit and
\[
\tilde{\eta}_{B,t+1}[\ell] = E \left[ (x_1[0] - f_{\text{opt}}(x_1[0] + \sqrt{v}; v))^2 \right] \tag{15}
\]
with \( v = \tilde{v}^{\text{opt}}_A - \tilde{v}^{\text{opt}}_B \), respectively. The limit in \( \tilde{\eta}_{A,t}[\ell] \) can be evaluated in closed form when the asymptotic eigenvalue distribution of \( A^T[\ell] A[\ell] \) is available. Note that the obtained state evolution recursions are deterministic.

**Theorem 2.** Consider no damping \( \zeta = 1 \) and postulate Assumptions 1 and 2. Furthermore, suppose that the Bayes-optimal denoiser \( f_{\text{opt}} \) is Lipschitz-continuous and nonlinear:

- The empirical error covariance \( N^{-1}(x_{B,t+1}[\ell] - x[\ell])^T (x_{B,t+1}[\ell] - x[\ell]) \) for OAMP converges almost surely to \( \tilde{v}_{B,t+1}[\ell] \) given in (15) for all \( t' \in \{0, \ldots, t\} \) in the large system limit.
- The state evolution recursions for Bayes-optimal OAMP converge to a fixed point as \( t \to \infty \).

**Proof.** The assumptions on the Bayes-optimal denoiser, as well as Assumption 2, are used to prove Assumption 3. The former statement is proved by confirming the inclusion of the error model for OAMP into the general error model (4)–(9). The proof of the latter statement is based on the LM-MP strategy in [15]; LM-OAMP is constructed and evaluated via Theorem 1. Furthermore, the obtained state evolution recursions for LM-OAMP are proved to converge and to be equivalent to those for OAMP without memory. See [25, Theorem 4] for the details.

Theorem 2 justifies the definitions of the variance parameters in OAMP. The asymptotic performance of OAMP can be evaluated by solving the state evolution recursions for OAMP.

### 4. NUMERICAL RESULTS

OAMP for the spatially coupled system (1) is numerically compared to that for the conventional system without spatial coupling. We used the uniform coupling coefficient \( \gamma[\ell]\big|\ell - w\big| = (W + 1)^{-1/2} \) in (1). In particular, \( W = 0 \) implies no spatial coupling.

The elements of the signal vector \( x[\ell] \) were sampled from \( \mathcal{N}(0, 1/\rho) \) with probability \( \rho \in [0, 1] \) uniformly and randomly. Otherwise, they took zero with probability \( 1 - \rho \).

The sensing matrices \( \{ A[l][\ell] \} \) or equivalently \( A[\ell] \) in (2) was postulated to have the SVD \( A[\ell] = \Sigma[\ell] V^T[\ell] \). The singular values in \( \Sigma[\ell] \) are uniquely determined from condition number \( \kappa \geq 1 \) and power normalization. See [25, Corollary 6] for the details. The \( |W[\ell]| N \times |W[\ell]| N \) orthogonal matrix \( V[\ell] \) is the Hadamard matrix with random column permutation, which is a low-complexity alternative of Haar-distributed orthogonal matrices in Assumption 2.

Figure 1 shows the largest mean-square error (MSE) of OAMP over all sections. \( 10^4 \) independent trials were simulated for the spatially coupled system with \( (L, W) = (16, 1) \) while \( 10^5 \) independent trials were simulated for \( (L, W) = (1, 0) \). The damping factor \( \zeta \) in lines 21 and 22 of Algorithm 1 was optimized for each \( \delta = M/N \) via exhaustive search.

OAMP for the spatial coupling case \( W = 1 \) is superior to for \( W = 0 \) in the so-called waterfall region, where the MSE decreases rapidly as the compression rate increases slightly. This result is consistent with existing results on spatial coupling [1, 2, 3, 4]. On the other hand, the MSE in the spatial coupling case degrades slightly for large \( \delta \). The latter result is a peculiar phenomenon for the spatially coupled system with orthogonally invariant sensing matrices.

This phenomenon results from two reasons: A minor reason is due to the loss in the overall compression rate \( (1 + W/L)\delta \). This influence vanishes in the limit \( L \to \infty \), as observed in [1, 2, 3, 4]. The other major reason is in the \( |W[\ell]| \)-dependencies of the empirical eigenvalue distribution of \( A^T[\ell] A[\ell] \in \mathbb{R}^{|W[\ell]| N \times |W[\ell]| N} \) in (2). To reduce the latter influence—remains even in the limit \( L \to \infty \)—small coupling width \( W \) should be used.

### 5. REFERENCES

[1] S. Kudekar, T. Richardson, and R. Urbanke, “Threshold saturation via spatial coupling: Why convolutional LDPC ensembles perform so well over the BEC,” IEEE
[2] F. Krzakala, M. Mézard, F. Sausset, Y. F. Sun, and L. Zdeborová, “Statistical-physics-based reconstruction in compressed sensing,” *Phys. Rev. X*, vol. 2, pp. 021005–1–18, May 2012.

[3] 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, Nov. 2013.

[4] K. Takeuchi, T. Tanaka, and T. Kawabata, “Performance improvement of iterative multiuser detection for large sparsely-spread CDMA systems by spatial coupling,” *IEEE Trans. Inf. Theory*, vol. 61, no. 4, pp. 1768–1794, Apr. 2015.

[5] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, “Cell-free massive MIMO versus small cells,” *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[6] 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, no. 5, pp. 2541–2557, May 2012.

[7] 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, Aug. 2017.

[8] 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, Jul. 2021.

[9] 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, Feb. 2011.

[10] A. Javanmard and A. Montanari, “State evolution for general approximate message passing algorithms, with applications to spatial coupling,” *Inf. Inference: A Journal of the IMA*, vol. 2, no. 2, pp. 115–144, Dec. 2013.

[11] S. Rangan, P. Schniter, A. Fletcher, and S. Sarkar, “On the convergence of approximate message passing with arbitrary matrices,” *IEEE Trans. Inf. Theory*, vol. 65, no. 9, pp. 5339–5351, Sep. 2019.

[12] J. Vila, P. Schniter, S. Rangan, F. Krzakala, and L. Zdeborová, “Adaptive damping and mean removal for the generalized approximate message passing algorithm,” in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, South Brisbane, Australia, Apr. 2015, pp. 2021–2025.