Clipping Can Improve the Performance of Spatially Coupled Sparse Superposition Codes

Shansuo Liang, Junjie Ma, and Li Ping, *Fellow, IEEE*

***Abstract*— Gaussian signaling of spatially coupled sparse superposition codes incurs a high peak-to-average-power ratio (PAPR) problem. Clipping can be used to reduce PAPR. At the receiver side, generalized approximate message passing algorithm is employed to handle the non-linear clipping distortion. Surprisingly, we observe empirically that clipping can improve the performance in low-to-medium rate region. We provide an explanation of this observation based on potential function analysis.**

***Index Terms*— Sparse superposition codes, spatial coupling, generalized approximate message passing, clipping, potential function analysis.**

I. Introduction

FOR a sparse superposition (SS) code [1]–[4], information symbols are first modulated in position and then linearly transformed by a compression matrix. As a result, codewords are Gaussian like and have a high peak-to-average-power ratio (PAPR). Clipping is a straightforward technique to reduce PAPR [5].

At the receiver side, approximate message passing (AMP) [6] has been applied to decode the message for SS codes [2], [4]. Clipping introduces non-linear distortion, which can be handled by a generalized approximate message passing (GAMP) decoder [7], [8]. Further, spatial coupling (SC) can be employed to circumvent the poor fixed-points of the above iterative processors [9], [10].

It is well-known that Gaussian signaling maximizes mutual information for an additive white Gaussian noise (AWGN) channel. Therefore, we might expect that clipping causes performance loss as it destroys Gaussianity. However, our simulations show the opposite: clipping may actually improve the performance at code rates that are of practical interests.

In this letter, we will provide an explanation based on potential function analysis [10] for the above observation. A brief intuitive explanation is as follows. The performance of a GAMP decoder can be characterized by a recursion between two transfer functions [7]. We show that clipping will reshape one of the transfer functions. It results in better matching between two transfer functions involved and hence improves the performance.

II. SCSS Codes With Clipping

A. Encoder Structure

For SCSS codes, a codeword $\mathbf{x}$ is given by [3] and [11]:

$$\underbrace{\begin{bmatrix} \mathbf{x}_1 \\ \vdots \\ \mathbf{x}_i \\ \vdots \\ \mathbf{x}_{K+W-1} \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} \mathbf{A}_{1,1} & \mathbf{0} & \cdots & & \mathbf{0} \\ \vdots & \mathbf{A}_{1,2} & \mathbf{0} & \cdots & \vdots \\ \mathbf{A}_{W,1} & \vdots & \ddots & & \\ \mathbf{0} & \mathbf{A}_{W,2} & \vdots & \ddots & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \vdots & \mathbf{A}_{1,K} \\ & & & \ddots & \vdots \\ \mathbf{0} & \cdots & & \mathbf{0} & \mathbf{A}_{W,K} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_k \\ \vdots \\ \mathbf{c}_K \end{bmatrix}}_{\mathbf{c}} \quad (1)$$

where $\mathbf{A}$ is a sensing matrix. Each $\mathbf{c}_k$ is an information-carrying sparse vector that contains L sections. Each section is of length-B and has a single nonzero entry with amplitude $\sqrt{B}$ [2], [4]. An example of $\mathbf{c}_k$ is as follows

$$\mathbf{c}_k^T = [\overbrace{0, \ldots, 0, \sqrt{B}}^{\text{section 1}}, |\overbrace{\ldots, 0, \sqrt{B}, 0}^{\text{section 2}}, |\cdots, |\overbrace{\sqrt{B}, 0, \ldots, 0}^{\text{section } L}] \quad (2)$$

For theoretical analysis in [2] and [4], each subblock $\mathbf{A}_{w,k} \in \mathbb{R}^{M\times N}(M < N)$ in (1) is assumed to have independent identically distributed (IID) Gaussian entries over $N(0, \beta N^{-1})$. In our simulations, each $\mathbf{A}_{w,k}$ is generated by randomly selecting M rows from the $N \times N$ discrete cosine transform (DCT) matrix multiplied by the scaling factor β. Such partial DCT matrices are also used in [4] to reduce complexity. Orthogonal AMP [12] can be applied to such non-IID sensing matrices. The performance difference between IID Gaussian matrices and partial DCT matrices is typically negligible unless B is very small [4]. The scaling factor is chosen as $\beta = 1/W$. For such defined $\mathbf{A}$, the elements of $\{\mathbf{x}_i, W \le i \le K\}$ have unit average power. When K is large, the overall average power can be approximated by $\mathbb{E}[x^2] = 1$ if we ignore the boundary effects for $i < W$ and $i > K$.

By the central limit theorem, the entries of $\mathbf{x}$ are approximately Gaussian distributed and thus have high PAPR.

Manuscript received June 7, 2017; revised July 30, 2017; accepted September 13, 2017. Date of publication September 19, 2017; date of current version December 8, 2017. This work was jointly supported by grants from the University Grants Committee of the Hong Kong Special Administrative Region, China (Projects AoE/E-02/08, CityU 11217515 and CityU 11280216). The associate editor coordinating the review of this letter and approving it for publication was M. Naeini. *(Corresponding author: Junjie Ma.)*

S. Liang and L. Ping are with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong (e-mail: ssliang3-c@my.cityu.edu.hk; eeliping@cityu.edu.hk).

J. Ma was with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong. He is now with the Department of Statistics, Columbia University, New York City, NY 10027-6902 USA (e-mail: junjiema2-c@my.cityu.edu.hk).

Digital Object Identifier 10.1109/LCOMM.2017.2754262

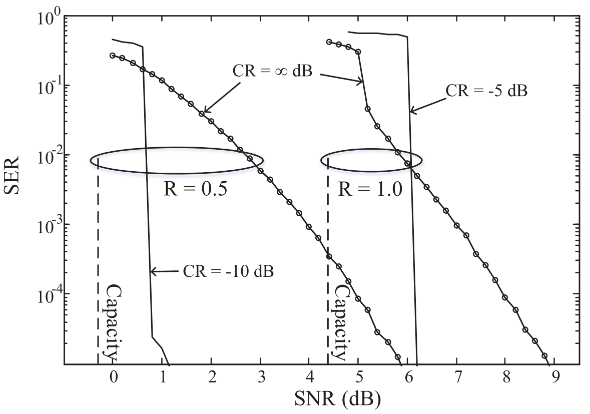

Fig. 1. Performance of SCSS codes with different clipping ratios. CR = ∞ dB represents the SCSS code without clipping. $B = 128$ and $L = 2048$. $K = 40$ and $W = 3$. The codeword length for $R = 0.5$ and $R = 1.0$ are 1204224 and 602112, respectively.

Clipping is a simple technique to reduce PAPR [5]. For real-valued input, the clipping function is defined as

$$f(x) = \begin{cases} Z, & x > Z \\ x, & -Z \leq x \leq Z \\ -Z, & x < -Z \end{cases} \tag{3}$$

where $Z > 0$ is a clipping threshold and we further define the clipping ratio (CR) as $\text{CR} \triangleq 10\log_{10}(Z^2/\mathbb{E}[x^2])$.

Consider an AWGN channel, the received signal for a clipped SCSS code can be written as

$$\mathbf{y} = \alpha f(\mathbf{x}) + \mathbf{n} = \alpha f(\mathbf{Ac}) + \mathbf{n}, \tag{4}$$

where $\mathbf{n} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ is the Gaussian noise and $\alpha \triangleq \sqrt{P/\mathbb{E}[f(x)^2]}$ is a normalizing factor that ensures the transmission power to be P. Signal-to-noise ratio (SNR) is defined as $\text{SNR} \triangleq P/\sigma^2$.

B. Impact of Clipping: A Numerical Example

The clipping operation introduces non-linear distortion, which can be handled by the GAMP algorithm [7], [8]. Fig. 1 compares the simulated section error rate (SER) performance. We consider two settings: $R = 0.5$ and $R = 1$. Here, the code rates of $\mathbf{x}$ are calculated as $R \cdot K/(K + W - 1)$, where $R = \log_2(B)/B/\delta$ ($\delta \triangleq M/N$) is the information rate of each $\mathbf{A}_{w,k}\mathbf{c}_k$ and the fraction $K/(K + W - 1)$ accounts for the boundary effect. The codeword length is given by $(K + W - 1) \cdot \delta N$. The effect of the number K and W on coupling construction is very limited when B is small.

It is shown in [4] that SCSS codes are capacity approaching as $B \to \infty$. However, the SER performance is poor for finite B in low-to-medium rate region as shown in Fig. 1. Similar results have also been reported in [13].

Interestingly, we see from Fig. 1 that clipping improves the performance in both rate settings, and the performance gain is more noticeable for $R = 0.5$. This is somewhat different from our expectation that clipping degrades the performance in general communication systems. For example, clipping incurs performance loss in an orthogonal frequency division multiplexing (OFDM) system [14]. A rough explanation is as follows. Clipping saves power but causes information loss, which affects performance in opposite ways. Therefore a more careful analysis is required for the overall effect, which will be studied later via the potential function analysis [10].

III. Potential Function Analysis

In this section, we first introduce state evolution (SE) [6] recursions for GAMP decoder and formulate the associated potential function [10]. Then, the impact of clipping on the performance of SCSS codes will be analyzed.

A. State Evolution for Coupled System

Under the assumption of IID Gaussian sensing matrices, the asymptotic performance of GAMP algorithm can be tracked by a scalar recursion called state evolution (SE) [6], [7], [15]. The vector valued SE operator has been derived and numerically verified for SCSS codes over generic memoryless channels in [8] and [16]. Note that the combined effect of clipping and an AWGN channel in (4) can be treated as an effective memoryless channel. Thus, the SE recursion derived in [16] can be readily applied here with a few modifications that account for the slight difference between the terminations in (1) and that in [8] and [16].

Let $\{v_i^t\}$ and $\{\rho_k^t\}$ be two state vectors at the t^{th} iteration. Initializing with $\{v_i^t = 1, \forall i\}$, the SE given by [16, Definition 3.9] is shown to be equivalent to the following recursion:

$$\rho_k^t = \frac{1}{W}\sum_{w=0}^{W-1}\phi\left(v_{k+w}^t\right), \quad 1 \leq k \leq K \tag{5a}$$

$$v_i^{t+1} = \frac{1}{W}\sum_{w=0}^{W-1}\psi\left(\rho_{i-w}^t\right), \quad 1 \leq i < K + W, \tag{5b}$$

where $\rho_k = 0$ for $k < 0$ and $k > K$ due to the assumed termination. Function $\phi(v)$ and $\psi(\rho)$ are defined as [7], [16]:

$$\phi(v) \triangleq \delta \cdot \mathbb{E}\left[\frac{1}{v}\left(1 - \frac{\text{Var}(x|\hat{x}, y, v)}{v}\right)\right], \tag{6}$$

$$\psi(\rho) \triangleq \frac{1}{B}\mathbb{E}\left[\left\|\mathbf{C} - \mathbb{E}\left[\mathbf{C}|\mathbf{C} + \rho^{-1/2}\mathbf{Z}\right]\right\|^2\right], \tag{7}$$

where $\delta = M/N$. In (6), $\text{Var}(x|\hat{x}, y, v)$ denotes the conditional variance of x with respect to $p(x|\hat{x}, y, v) \propto p(y|x)N(x; \hat{x}, v)$. The expectation is taken over the joint distribution $p(\hat{x}, x, y) = N(\hat{x}; 0, 1-v)N(x; \hat{x}, v)p(y|x)$. The function $\phi(v)$ depends on the equivalent memoryless channel $p(y|x)$ (see (4)). A special case of $\phi(v)$ for CR = ∞ dB is given by

$$\phi_\infty(v) \equiv \frac{\delta}{v + 1/\text{SNR}}, \tag{8}$$

which corresponds to an AWGN channel without clipping.

In (7), $\mathbf{C} \in \mathbb{R}^{B\times 1}$ is distributed as a component section of $\mathbf{c}_k$ and $\mathbf{Z} \in \mathbb{R}^{B\times 1}$ is the standard Gaussian noise independent of $\mathbf{C}$. The function $\psi(\rho)$ is the minimum mean square error (MMSE) for estimating $\mathbf{C}$ in an AWGN channel with SNR ρ. Both expectations in $\phi(v)$ and $\psi(\rho)$ can be evaluated numerically.

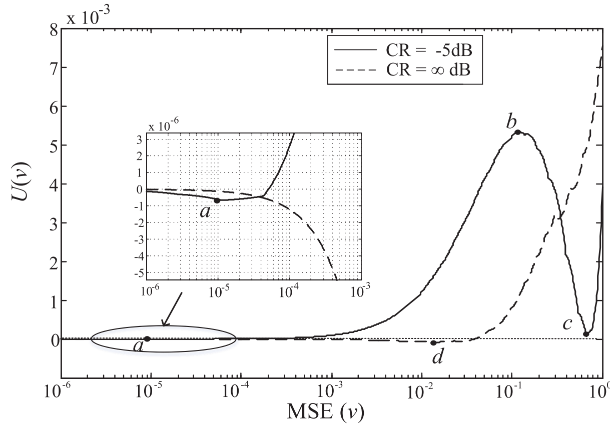

Fig. 2. The uncoupled potential function $U(\upsilon)$ with different clipping ratios. $B = 128$, $R = 1$ and SNR = 6.2 dB.

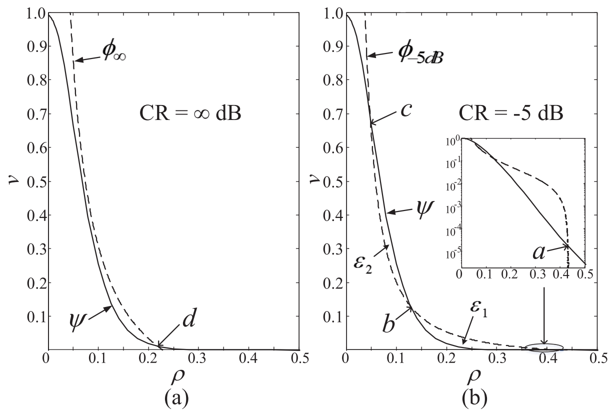

Fig. 3. Transfer functions $\psi(\rho)$ and $\phi(\upsilon)$ with different clipping ratios. Curves of ψ in two cases are identical. $B = 128$, $R = 1$ and SNR = 6.2 dB.

B. Potential Function Analysis

The behavior of the coupled SE in (5) can be determined by the following uncoupled potential function [10]:

$$U(\upsilon) = \int_{\phi(1)}^{\phi(\upsilon)} \left[\psi(\rho) - \phi^{-1}(\rho)\right] d\rho, \quad \text{for } \upsilon \in [0, 1], \quad (9)$$

where $\phi^{-1}(\cdot)$ is the inverse function of $\phi(\cdot)$. It is proven in [10, Th. 1] that the fixed point of the coupled SE recursion in (5) is determined by the minimizer of function $U(\upsilon)$[1] when $K, W \to \infty$.

Fig. 2 shows an example of $U(\upsilon)$ with two different clipping ratios: CR $= \infty$ dB and CR $= -5$ dB. The dashed curve (for CR $= \infty$ dB) has a unique minimum, marked with point d. The solid curve (for CR $= -5$ dB) has two local minima: the rightmost minimum is marked with c and the global minimum marked with a. Without spatial coupling, the GAMP decoder will converge to the rightmost minimum. From Fig. 2, we see that the MSE of c is larger than that of d, which means that clipping deteriorates the performance. In contrast, when spatial coupling is applied, the performance of SCSS codes is determined by the global minimum of the uncoupled potential function $U(\upsilon)$ [10]. From Fig. 2, we see that the MSE of a is much smaller than that of d, which explains why the SCSS code with CR $= -5$ dB significantly outperforms that with CR $= \infty$ dB in Fig. 1 at SNR = 6.2 dB. Note that MSE can be uniquely transformed into SER.

C. Area Interpretation

We now provide an intuitive explanation of the impact of clipping based on the area interpretation of the potential function [11], [17].

Fig. 3 plots the transfer functions $\phi(\upsilon)$ and $\psi(\rho)$ for the example in Fig. 2. We see that for CR $= \infty$ dB shown in the left sub-figure, the two curves have a single intersection marked as d. This point corresponds to the global minimum of $U(\upsilon)$ shown in Fig. 2, and it characterizes the performance of the SCSS code. For CR $= -5$ dB in Fig. 3(b), the two curves have three intersections: a, b and c. As shown in Fig. 2, a and c are minima of $U(\upsilon)$ while b is a local maximum. According to the potential function analysis, the performance of SCSS codes is determined by the global minimum of the uncoupled potential function $U(\upsilon)$. By the area interpretation [11], a is the global minimum if the area ε_1 is larger than that of ε_2, otherwise c is the global minimum. We verified numerically that the area ε_1 is larger than ε_2 and thus the global minimum is a. From Fig. 3, the MSE performance associated with a (shown by its vertical coordinate) is much smaller than that of d. So the SCSS code with CR $= -5$ dB significantly outperforms that with CR $= \infty$ dB for the setting considered here.

For more insights, recall that the performance of a turbo process can be improved by matching the EXIT functions of the two component processors involved [18]. We can treat ϕ and ψ as two EXIT functions. From Fig. 3, we observe ϕ_{-5} "matches" ψ better than ϕ_∞, in that both ϕ_{-5} and ψ are relatively flat at small υ while ϕ_∞ is steeper. This implies that CR = -5 dB can lead to an improved performance.

The above discussions are based on the assumption that ψ (determined by the structure of $\mathbf{c}_k$) is fixed. Now we consider the case that the structure $\mathbf{c}_k$ can be well optimized [2] so that ψ can be perfectly matched to ϕ. Then the achievable rate is determined by the area covered by ϕ [8]. It can be shown that the area covered by ϕ_{-5} is smaller than that by ϕ_∞. In this case, clipping will cause a loss of the achievable rate. Note that the optimization of $\mathbf{c}_k$ is beyond scope of this letter.

In summary, the statement that clipping can improve performance of SCCS codes is under the assumption that the structure of $\mathbf{c}_k$ is fixed as in (2). Without this restriction, clipping may deteriorate performance.

D. Impact of Clipping Ratio

We now examine the impact of the clipping ratio (CR) more carefully.

Fig. 4 shows the MSE performance versus CR for various code rates. The MSE is predicted via potential function analysis [10]. From Fig. 4, we see that the dashed line for $R = 0.5$ is a monotonically increasing function of CR, implying that

[1]Function $U(\upsilon)$ is defined based on the underlying uncoupled system. The definition here differs from [10, eq. (4)] by an additive constant. Such difference does not affect the minimizer of $U(\upsilon)$.

[2]Here, for a SCSS code, $\mathbf{c}_k$ is a sparse vector modulated by position. Generally, this restriction can be removed and $\mathbf{c}_k$ can have any structure, sparse or non-sparse. Then optimization can be performed over $\mathbf{c}_k$ [13].

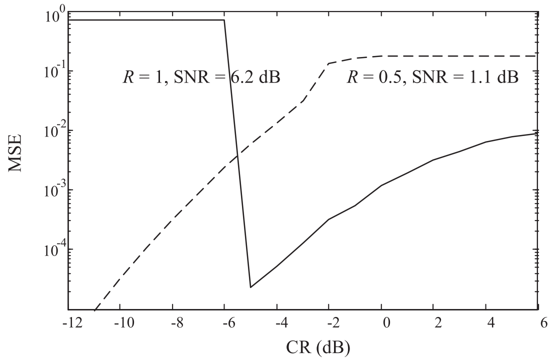

Fig. 4. Relationship between MSE performance and clipping ratio for different information rates at different SNRs.

TABLE I

THE RECOMMENDED CLIPPING RATIOS FOR TYPICAL RATES OF COMMUNICATIONS

CR (dB) R / B	0.5	0.6	0.7	0.8	0.9	1	1.1	1.2	1.3
128	$-\infty$	-15	-11	-9.0	-7.0	-5.0	-3.8	-3.0	-2.0
32	$-\infty$	$-\infty$	-16	-9.2	-7.2	-5.6	-4.4	-3.2	-2.2

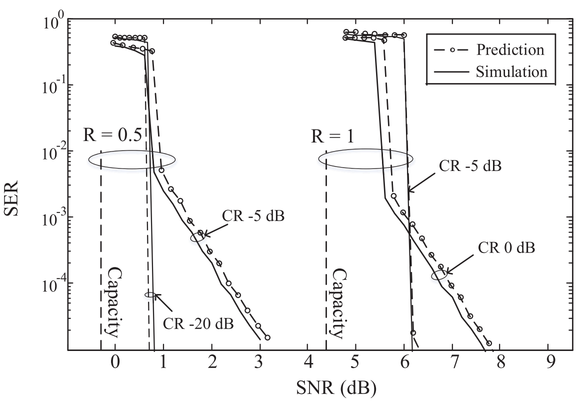

Fig. 5. Performance of SCSS codes with different clipping ratios. $B = 128$ and $L = 2048$. $K = 40$ and $W = 3$. The codeword length for $R = 0.5$ and $R = 1.0$ are 1204224 and 602112, respectively.

CR $= -\infty$ dB (which becomes biplolar signaling) may be optimal. Similar results have been observed in [8]. For solid line ($R = 1$), the MSE performance is not a monotonic function of CR, and the optimal CR is around -5 dB.

In general, given R and B, we could optimize the value of CR for the target MSE numerically via potential function analysis. Some examples are given in Table (I). We see that the recommended clipping ratio is decreasing as the rate decreases. For rate less than 0.5, CR $= -\infty$ dB seems to be optimal.

We now provide simulation results to examine the accuracy of potential function analysis. Fig. 5 shows the simulated SER performances of SCSS codes with two different rates: $R = 0.5$ and $R = 1$. We see that the simulated MSE performance are slightly better than those predicted by the potential function. The reason is that the actual code rate is $R \cdot K/(K + W - 1)$, which is smaller than R due to the boundary effect. Such rate loss will vanish when $W, K \to \infty$ in a proper order [10]. But overall, the prediction is quite accurate.

From Fig. 5, we see that clipping significantly improves the performances of SCSS codes. For $R = 1$, CR $= -5$ dB outperforms CR $= 0$ dB, which is consistent with the solid line in Fig. 4. For $R = 0.5$, the SCSS code with CR $= -20$ dB (approximately equivalent to bipolar signaling) can even approach channel capacity within 1 dB.

Empirically, when $R > 1.5$, SCSS codes with moderate B already have a quite satisfying performance under AMP decoding. For $R \leq 1.5$, we observed that clipping can be beneficial when $B \leq 512$.

IV. CONCLUSIONS

In this letter, we showed that clipping can improve the performance of SCSS codes in low-to-medium rate region. We provided an explanation based on potential function analysis and area interpretation.

REFERENCES

[1] A. R. Barron and A. Joseph, "Least squares superposition codes of moderate dictionary size are reliable at rates up to capacity," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 2541–2557, Feb. 2012.

[2] C. Rush, A. Greig, and R. Venkataramanan, "Capacity-achieving sparse superposition codes via approximate message passing decoding," *IEEE Trans. Inf. Theory*, vol. 63, no. 3, pp. 1476–1500, Mar. 2017.

[3] J. Barbier, C. Schülke, and F. Krzakala, "Approximate message-passing with spatially coupled structured operators, with applications to compressed sensing and sparse superposition codes," *J. Stat. Mech., Theory Experim.*, vol. 63, no. 5, p. P05013, May 2015.

[4] 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.

[5] S. H. Han and J. H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," *IEEE Wireless Commun.*, vol. 12, no. 2, pp. 56–65, Apr. 2005.

[6] D. L. Donoho, A. Maleki, and A. Montanari, "Message-passing algorithms for compressed sensing," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 45, pp. 18914–18919, Nov. 2009.

[7] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *Proc. IEEE Int. Symp. Inf. Theory*, Jul. 2011, pp. 2168–2172.

[8] E. Biyik, J. Barbier, and M. Dia, "Generalized approximate message-passing decoder for universal sparse superposition codes," in *Proc. IEEE Int. Symp. Inf. Theory*, Jun. 2017, pp. 1593–1597.

[9] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Spatially coupled ensembles universally achieve capacity under belief propagation," *IEEE Trans. Inf. Theory*, vol. 59, no. 12, pp. 7761–7813, Dec. 2013.

[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, Nov. 2014.

[11] C. Liang, J. Ma, and L. Ping, "Towards Gaussian capacity, universality and short block length," in *Proc. 9th Int. Symp. Turbo Codes*, Sep. 2016, pp. 412–416.

[12] J. Ma and L. Ping, "Orthogonal AMP," *IEEE Access*, vol. 5, pp. 2020–2033, 2017.

[13] C. Liang, J. Ma, and L. Ping, "Compressed FEC codes with spatial-coupling," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 987–990, May 2017.

[14] J. Tong, L. Ping, Z. Zhang, and V. K. Bhargava, "Iterative soft compensation for OFDM systems with clipping and superposition coded modulation," *IEEE Trans. Commun.*, vol. 58, no. 10, pp. 2861–2870, Oct. 2010.

[15] 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.

[16] J. Barbier, M. Dia, and N. Macris, "Threshold saturation of spatially coupled sparse superposition codes for all memoryless channels," in *Proc. IEEE Inf. Theory Workshop (ITW)*, Sep. 2016, pp. 76–80.

[17] S. Kudekar, T. J. Richardson, and R. L. Urbanke, "Wave-like solutions of general 1-D spatially coupled systems," *IEEE Trans. Inf. Theory*, vol. 61, no. 8, pp. 4117–4157, Aug. 2015.

[18] S. ten Brink, "Convergence behavior of iteratively decoded parallel concatenated codes," *IEEE Trans. Commun.*, vol. 49, no. 10, pp. 1727–1737, Oct 2001.