Abstract—We propose efficient and low complexity multiuser detection (MUD) algorithms for Gaussian multiple access channel (G-MAC) for short-packet transmission in massive machine type communications. We formulate the problem of MUD for the G-MAC as a compressive sensing problem with virtual sparsity induced by one-hot vector transmission of devices in code domain. Our proposed MUD algorithms rely on block and non-block approximate message passing (AMP) for soft decoding of the transmitted packets. Taking into account the induced sparsity leads to a known prior distribution on the transmit vector; hence, the optimal minimum mean squared error (MMSE) denoiser can be employed. We consider both separable and non-separable MMSE denoisers for AMP soft decoding. The effectiveness of the proposed MUD algorithms for a large number of devices is supported by simulation results. For packets of 8 information bits, while the state of the art MMSE denoiser achieves 5/100 of the converse bound at $E_b/N_0 = 4$ dB, the proposed algorithms reach 4/9 and 1/2 of the converse bound.

I. INTRODUCTION

MASSIVE uplink connectivity is one of the key factors in the realization of the Internet-of-Things (IoT). In many IoT applications, massive machine-type communications (mMTC) services are required, where a large number of devices transmit very short packets to a base station (BS). Typically, these short packets are transmitted whenever a measured value changes [1], [2].

Recently, bounds on the exact block error probability of Gaussian random codes in additive white Gaussian noise (AWGN) for massive uplink multiple access (MA) at finite message length have been derived [3]. The author in [3] has shown that a promising solution to support massive uplink connectivity for short packet transmission is the Gaussian multiple access channel (G-MAC). The challenging task in the realization of G-MAC is the design of efficient and low complexity multiuser detection (MUD) algorithms at the base station (BS) to achieve the desirable spectral efficiency and reliability for short-packet transmission [4].

In this paper, we propose two efficient and low-complexity MUD algorithms for G-MAC for short-packet mMTC. We first formulate the MUD detection problem as an under-determined compressive sensing problem. Then, we employ approximate message passing (AMP) for soft decoding of the short packets transmitted by the IoT devices in the network. The flexibility in the AMP algorithm design is attributed to the denoiser $\eta_k(\cdot)$. The denoisers in the AMP variants are typically designed to reduce the estimation error at each iteration [5].

For most problems, the prior distribution of the vector to be recovered is generally assumed to be unknown. In this case, the denoiser is designed under the minimax framework to optimize the AMP algorithm performance for the worst-case or least-favorable distribution, leading to a soft-thresholding denoiser. However, in the problem at hand, we have the exact and asymptotic empirical prior distribution of the sparse vector to be recovered. Hence, the optimal MMSE denoiser can be employed in AMP. By taking into account the exact and asymptotic prior distribution of the transmit vector, we employ separable and non-separable MMSE denoisers for Bayesian AMP. The decoupling effect results in low-complexity MMSE denoising in each AMP iteration [6], [7].

The main contributions of the paper are:

- We introduce two MUD algorithms called (i) the relaxed block sparsity AMP (RBS-AMP), and (ii) block sparsity AMP (BS-AMP) MUD algorithms for G-MAC suitable for short packet transmission in mMTC. We show that our BAMP MUD algorithm results in a new achievability bound that is closer to the converse bound in [3].
- The multiuser efficiency (MUE) of the proposed MUD algorithms is derived, and shown to be close to one.

Notation: The symbols $(\cdot)^T$, $\text{Tr}(\cdot)$, $\| \cdot \|_0$, and $\| \cdot \|$ show the transpose operator, trace of a matrix, 0-norm, and Euclidean norm of a vector, respectively. $\mathbb{E}\{ \cdot \}$ is the statistical expectation, and $\hat{x}$ is an estimate of vector $x$. $\delta(\cdot)$ is the Dirac delta function.

II. SYSTEM MODEL

We consider a single-hop Internet of Things (IoT) network, where $D$ IoT devices simultaneously communicate with a single BS. Each IoT device transmits $K$ bits of information by using Gaussian random coding with codewords of length $M = nD$. The aggregate rate of the multiuser coded system is $DK/M$. For short packet transmission, the received vector at the BS in vector form is

$$\mathbf{y} = \mathbf{C} \mathbf{x} + \mathbf{w},$$

where $\mathbf{C} \in \mathbb{R}^{M \times N}$ has independent and identically distributed (i.i.d.) entries, $C_{ij} \sim \mathcal{N}(0, 1/M)$, $N \triangleq 2^K D$, and $\mathbf{x}$ is a block sparse vector of length $N \times 1$ as

$$\mathbf{x} = [\mathbf{x}_1^T \mathbf{x}_2^T \ldots \mathbf{x}_D^T]^T,$$

where $\mathbf{x}_q$ is a $2^K \times 1$ vector as $||\mathbf{x}_q||_0 = 1, q = 1, 2, \ldots, D$. The $M \times 1$ vector $\mathbf{w}$ in (1) is the additive white Gaussian noise (AWGN) vector with covariance matrix $\sigma_w^2 \mathbf{I}$. Here, we assume that $\beta \triangleq \lim_{D \to \infty} M/N = n/(2^K) \in \mathbb{R}_+^+$. We consider that the $D$ IoT devices in the network transmit with equal power $P_1$. In this case, the prior distribution for the transmit vector $\mathbf{x}_q \triangleq [x_{q,1}, x_{q,2}, \ldots, x_{q,2^K}]^T, q = 1, 2, \ldots, D$, in (2) is given by

$$f_{\mathbf{x}}(\mathbf{x}_q; P_1, K) = \sum_{k=1}^{2^K} \frac{1}{2^K} \delta(\mathbf{x}_q - \sqrt{P_1} \mathbf{e}_k),$$

where $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{2^K}$ are orthonormal basis in right-handed Cartesian coordinate system, i.e., $\mathbf{e}_k \triangleq [e_1 e_2 \ldots e_{2^K}]^T$.
\[ \begin{align*}
\{0 0 \cdots 0 1 0 \cdots 0\}^T, \quad \text{where the } k\text{th element is one and the remaining elements are zero}, \quad \text{and } \delta(x_q) = \prod_{k=1}^{K} \delta(x_{q,k}).
\end{align*} \]

Our goal is to design a MUD algorithm at the BS to detect the \( K \) information bits transmitted by each IoT device. As \( x \) is a block-sparse vector, we have a sparse signal reconstruction problem. In this paper, we adopt the low-complexity Bayesian AMP algorithm for signal recovery.

### III. Review of AMP with Different Denoisers

In this section, we first briefly review AMP algorithm with separable and non-separable denoisers. Then, in Theorem 2, we show that the per-iteration mean-squared error of the reconstructed vector through AMP with non-separable denoiser, can be predicted using a state evolution (SE) in (5). Theorem 2 will later be used for the derivation of the MUE of the proposed BS-MUD algorithm.

#### A. AMP Algorithm with Separable Denoiser

For noisy linear measurements of the form

\[ s = A\mathbf{u} + \mathbf{w}, \quad (4) \]

where \( A \in \mathbb{R}^{M \times N}, \mathbf{s} \in \mathbb{R}^{M \times 1}, \mathbf{u} \in \mathbb{R}^{N \times 1}, \) and \( \mathbf{w} \in \mathbb{R}^{M \times 1} \). AMP algorithm with denoiser \( \eta_\beta : \mathbb{R}^N \rightarrow \mathbb{R}^N \) at the \((t+1)\)th iteration estimates vector \( \mathbf{u} \) as follows [8]

\[ \hat{\mathbf{u}}^{t+1} = \eta_\beta(\mathbf{v}^t) = \eta_\beta(\hat{\mathbf{u}}^t + \mathbf{A}^T \mathbf{r}^t), \quad (5) \]

where

\[ \mathbf{r}^t = \mathbf{s} - \mathbf{A} \hat{\mathbf{u}}^t + \frac{N}{M} \mathbf{r}^{t-1} \text{div} \left( \eta_\beta(1) \mathbf{v}^{t-1} \right), \quad (6) \]

and

\[ \mathbf{v}^t = \mathbf{u}^0 + \mathbf{A}^T \mathbf{r}^t, \quad (7) \]

where \( \text{div}(\eta_\beta(\mathbf{g})) = \frac{1}{N} \text{Tr} \left( \frac{\partial \eta_m(\mathbf{g})}{\partial \mathbf{g}} \right) \) with \( \mathbf{g} = [g_1 g_2 \cdots g_N]^T, \mathbf{u}^0 = \mathbf{0}, \) and \( \mathbf{r}^0 = \mathbf{s}. \) For identical separable Lipschitz continuous (LC) denoiser, \( \eta_\beta(\mathbf{g}) = [\eta_1(\mathbf{g}_1) \eta_2(\mathbf{g}_2) \cdots \eta_N(\mathbf{g}_N)]^T \), and div(\( \eta_\beta(\mathbf{g}) \)) = \( \sum_{j=1}^{N} \frac{1}{\beta} \partial \eta_j(\mathbf{g}_j)/\partial \mathbf{g}_j \). Similarly, for non-separable LC denoisers, we have [8]

\[ \eta_\beta(\mathbf{g}) = [\eta_1(\mathbf{g}_1) \eta_2(\mathbf{g}_2) \cdots \eta_L(\mathbf{g}_L)]^T, \quad (8) \]

where \( \mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_L \) denote \( L \) non-overlapping partitions of \( \mathbf{g}. \) The initial work on AMP focused on the case where \( \eta_\beta(\cdot) \) is a separable elementwise function with identical components [9], while the later work allowed non-separable blockwise denoisers [8].

For large, i.i.d., sub-Gaussian random matrices \( \mathbf{A} \) in (4), \( \lim_{M \rightarrow \infty} M/N = \beta > 0, \) and identical separable/non-separable LC denoiser, AMP displays decoupling behavior, meaning that \( \mathbf{v}^t \) in (7) behaves like an AWGN corrupted version of the true signal \( \mathbf{u} \) in (4) as

\[ \mathbf{v}^t = \mathbf{u} + \mathcal{N}(0, \sigma^2 I), \quad (10) \]

where an estimate of \( \sigma^2 \) can be obtained as \( \hat{\sigma}^2 = \frac{1}{M} \mathbf{v}^T \mathbf{v}. \) The larger \( M, \) the lower the estimation error.

### B. SE of AMP with Separable Denoiser

For large, i.i.d., sub-Gaussian random matrices \( \mathbf{A} \), the performance of AMP can be exactly predicted by a scalar SE, which also provides testable conditions for optimality. For identical separable LC denoiser \( \eta_\beta(\cdot), \) when the elements of \( \mathbf{u} = [u_1, u_2, \ldots, u_N] \) empirically converges to some random variable \( \mathbf{U} \) with probability density function (PDF) \( f_U, \) the SE i.e., the sequence of \( \{\sigma^2\} \geq 0 \) in (10) is expressed as

\[ \sigma^2_{t+1} = \sigma^2_w + \frac{1}{\beta} \mathbb{E} \{ (\eta_\beta(U + \sigma_Z Z) - U)^2 \}, \quad (11) \]

where \( \beta = M/N \)

\[ \sigma^2_0 = \sigma^2_w + \frac{1}{\beta} \mathbb{E} \{ U^2 \}, \quad (12) \]

where \( Z \sim \mathcal{N}(0, 1) \) and is independent from \( U. \) The term \( (\sigma^2_{t+1} - \sigma^2_w) \) in (11) can be estimated by the sample mean estimate of the denoiser error as it is expressed in Theorem 1.

#### Theorem 1

Let \( \mathbf{A} \in \mathbb{R}^{M \times N} \) be a matrix with i.i.d. entries \( A_{ij} \sim \mathcal{N}(0, 1/M), \) and assume \( M/N \rightarrow \beta > 0. \) Consider a random signal \( \mathbf{u} \in \mathbb{R}^N, \) whose empirical distributions converge weakly to a PDF \( f_u \) on \( \mathbb{R} \) with bounded moment of order \((2k-2)\) as \( N \rightarrow \infty \) for some \( k \geq 2. \) Then, for any pseudo-Lipschitz function \( \psi(\cdot, \cdot) : \mathbb{R}^2 \rightarrow \mathbb{R} \) of order \( k \) and all \( t \geq 0, \) we almost surely have (Proof in [13])

\[ \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^{N} \psi(\hat{u}_{i+1}^t, u_i) = \mathbb{E} \{ \psi(\eta_\beta(U + \sigma_1 Z, U) \}, \quad (13) \]

where \( \hat{u}_{i+1}^t \triangleq [\hat{u}_i^t \hat{u}_i^t + 1 \cdots \hat{u}_N^t + 1]^T \) is obtained using AMP in (5)-(7). \( \hat{u}_{i+1}^t \) is independent of \( U, \) and \( \sigma_1 \) follows the SE in (11) and (12). For \( \psi(x, y) = (x-y)^2, \) we have

\[ \lim_{N \rightarrow \infty} \frac{1}{N} \| \mathbf{u}^{t+1} - \mathbf{u}_t \|_2^2 = \mathbb{E} \{ (\eta_\beta(U + \sigma_1 Z) - U)^2 \} \]

\[ = (\sigma^2_{t+1} - \sigma^2_w) \beta. \]

#### C. AMP with Non-separable Denoiser

Let us consider a block structure for the vector \( \mathbf{u} \) in (4) as \( \mathbf{u} = [\mathbf{u}_1^T, \mathbf{u}_2^T, \ldots, \mathbf{u}_L^T]^T, \) where \( \mathbf{u}_i \) and \( \mathbf{u}_j, \) \( i \neq j \) are independent random vectors. We consider the following cases: 1) the joint PDF of \( \mathbf{u}_i \) \( \triangleq [u_{i,1}, u_{i,2}, \ldots, u_{i,L}]^T \) is \( U = [U_1 U_2 \ldots U_L]^T \sim f_U, \) and 2) the empirical joint PDF of \( \mathbf{u}_i \) converges weakly to \( f_U \), where \( \ell = N/L. \) AMP with identical non-separable LC denoiser in (8) with \( \eta_\beta(\cdot) : \mathbb{R}^{L \rightarrow \mathbb{R}^L} = \{ \eta_m(\cdot) = \eta(\cdot), m = 1, 2, \ldots, L \}, \) reconstructs the vector \( \mathbf{u} \) from the observation model in (4) through the relations in (5)-(7). Also, the decoupling behavior occurs for the \( mt \) block as

\[ \mathbf{v}^t = \mathbf{u} + \mathcal{N}(0, \sigma^2 I), \quad (10) \]

where an estimate of \( \sigma^2 \) can be obtained as \( \hat{\sigma}^2 = \frac{1}{M} \mathbf{v}^T \mathbf{v}. \) The larger \( M, \) the lower the estimation error.

\[ \begin{align*}
\{0 0 \cdots 0 1 0 \cdots 0\}^T, \quad \text{where the } k\text{th element is one and the remaining elements are zero}, \quad \text{and } \delta(x_q) = \prod_{k=1}^{K} \delta(x_{q,k}).
\end{align*} \]
\[ \rho_0^2 = \sigma_v^2 + \frac{1}{\ell \beta} \mathbb{E}\{\|U\|^2\}, \] (18)

where \( Z = [Z_1 Z_2 \ldots Z_\ell]^T \sim N(0, I) \) is independent of \( U = [U_1 U_2 \ldots U_\ell]^T \sim P_U \). A modified version of Theorem 1 for identical non-separable LC denoiser is given as follows:

**Theorem 2.** Let \( A \in \mathbb{R}^{M \times N} \) be a matrix with i.i.d. entries \( A_{ij} \sim \mathcal{N}(0, 1/M) \), and assume \( M/N \to \beta > 0 \). Consider a block sparse signal \( u = [u_1^T u_2^T \ldots u_\ell^T]^T \), where the joint distribution of each block of length \( \ell \) is \( f_U \) (or empirical joint PDF converges weakly to \( f_U \) on \( \mathbb{R}^{\ell} \). Then, we have

\[
\lim_{L \to \infty} \frac{1}{L} \sum_{n=1}^{L} \|\hat{u}_n^{t+1} - u_n\|^2 = \mathbb{E}\{\|\tilde{\eta}(U + \rho_1 Z) - U\|^2\},
\] (19)

where \( \hat{u} = [\hat{u}_1^T \hat{u}_2^T \ldots \hat{u}_\ell^T]^T \) is obtained by using (5)-(7), \( U \sim f_U \), \( Z \sim N(0, I) \) is independent of \( U \), and \( \rho_1 \) follows the SE in (17) and (18).

**Proof.** Proof in Appendix A \( \square \)

**IV. AMP-BASED MUD FOR SHORT-PACKET COMMUNICATIONS**

In this section, we propose RBS-AMP and BS-AMP MUD algorithms for G-MAC. The former AMP-based MUD algorithm employs separable MMSE denoiser and offers lower complexity, and the latter algorithm uses block non-separable MMSE denoiser and achieve higher decoding performance.

**A. RBS-AMP MUD**

The RBS-AMP MUD algorithm neglects the block-sparsity of \( x \) in (1) and consider the asymptotic semiempirical PDF. This block sparsity relaxation results in a low complexity MUD algorithm since the prior PDF in (3) is replaced with the simple semiempirical PDF. The empirical distribution of the block sparse vector \( x = [x_1, x_2, \ldots, x_N]^T \) in (2) as \( N \to \infty \), converges weakly to the scaled Bernoulli distribution with the PDF \( f_X(x; P_1; K) \) as

\[
f_X(x; P_1; K) = \frac{1}{2^K} \delta(x - \sqrt{P}) + \left(1 - \frac{1}{2^K}\right) \delta(x).
\] (20)

Let us consider \( \hat{x} = [\hat{x}_1^{t+1}, \hat{x}_2^{t+1}, \ldots, \hat{x}_N^{t+1}]^T \), the reconstructed vector by the AMP algorithm summarized in (5)-(7) at the \( (t+1) \)th iteration for the observation model in (1), i.e., \( s = y \), \( u = x \), and \( A = C \). For the identical separable MMSE denoiser \( \eta(\cdot) = \mathbb{E}\{\cdot | v_i^t\} \) with prior distribution \( U \sim X \sim f_X(x; P_1; K) \) in (20), by taking into account the decoding behavior in (10), the \( i \)th element of the soft decoded vector at the \( (t+1) \)th iteration is given by

\[
\hat{x}_i^{t+1} = \eta(v_i^t) = \mathbb{E}\{X_i | X_i + \sigma_i Z = v_i^t\}
\] (23)

\[= \frac{p T_i^t}{p T_i^t + 1 - p}, \]

where \( Z \sim N(0, 1) \),

\[
T_i^t \triangleq \exp\left(-\frac{-P_i + 2 v_i^t \sqrt{T_i^t}}{2 \sigma_i^2}\right), \] (24)

and \( p = 1/2^K \). Proof of the general case for the prior PDF in (3) is given in Appendix B. Moreover, the divergence term at the \( (t+1) \)th iteration of the RBS-AMP in (6) is obtained as

\[
\text{div}(\eta(v^t)) = \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial v_i^t} \eta(v_i^t) = \frac{1}{N} \sum_{i=1}^{N} \mathbb{E}\{X_i | X_i + \sigma_i Z = v_i^t\}
\] (25)

\[= \frac{p(1-p) v_i^t}{\sigma_i^2 N} \sum_{i=1}^{T_i^t} \frac{T_i^t + 1 - p}{\sigma_i^2 N}.
\]

In practice, the SE, i.e., \( \{\sigma_i^2\}_{i \geq 0} \) is tracked by the sample mean estimate \( \hat{\sigma}_i^2 \approx \frac{1}{M} \|r^t\|^2 \), where \( r^t \) is given in (6) for \( A = C \) and \( \tilde{u} = \tilde{x} \). The SE can also be theoretically obtained by using Theorem 1. In this case, from (14) and (20), we obtain (26) at the top of the next page.

**MUE of the RBS-AMP MUD Algorithm:** The degradation in bit error rate due to the presence of MA interference in AWGN channel can be measured by the multiuser asymptotic efficiency, which is defined as

\[
\xi = \frac{\sigma_v^2}{\sigma_u^2 + I},
\] (27)

where \( \sigma_u^2 \) is the variance of noise, and \( I \) is the interference after soft decoding.

**Theorem 3.** The MUE of the RBS-AMP MUD algorithm at the \( (t+1) \)th iteration is given by

\[
\xi_{RB}^{t+1} = \frac{\sigma_u^2}{\sigma_v^2 + (\sigma_{u_i}^{t+1} - \sigma_v^2)} = \frac{\sigma_v^2}{\sigma_{u_i}^{t+1}} \cdot (t+1), \] (28)

where \( \xi_{RB}^{t+1} \in [0, 1] \), and \( \sigma_{u_i}^{t+1} \) is given in (11) for \( U = X \).

**Proof.** Proof in Appendix C \( \square \)

In practice, the sample mean estimate of \( \{\sigma_i^2\}_{i \geq 0} \), i.e., \( \frac{1}{M} \|r^t\|^2 \) is employed to track MUE of the RBS-AMP MUD algorithm as \( \xi_{RB}^{t+1} = \frac{M \sigma_u^2}{M \sigma_u^2 + \sigma_v^2} \). In Fig. 1 we show \( \xi_{RB}^{t+1} \) for \( K = 8 \) and \( U = 1024 \) at 4 dB SNR. As seen, \( \xi_{RB}^{t+1} \) is achieved. Also, the lower spectral efficiency \( R \) the faster the decoding.

**B. BS-AMP MUD Algorithm with Non-separable Denoiser**

By considering the exact prior distribution of \( x \) in (3), AMP with block non-separable denoiser can be employed to reconstruct the vector \( x = [x_1, x_2, \ldots, x_N]^T \) from the underdetermined system of linear equations \( y = Cx + w \) in (1) by using the relations (5)-(7). For identical block non-separable MMSE denoiser \( \tilde{\eta}(\cdot) = \mathbb{E}\{\cdot | v_u^t\} : \mathbb{R}^{2^K} \to \mathbb{R}^{2^K} \), the reconstructed vector for the \( u \)th IoT device at the \( (t+1) \)th iteration is given as follows

\[
\tilde{x}_u^{t+1} = \mathbb{E}\{X | X + \rho_t Z = v_u^t\},
\] (29)

where \( X = [X_1, X_2, \cdots, X_{2^K}]^T \) is a random vector with the joint PDF in (3) and \( v_u^t = \tilde{x}_u^{t} + C^T r^t = x_u + N(0, \rho_t^2 I) \) due to the decoding affect as it was explained in (15) in (16).
\[\lim_{N \to \infty} \frac{1}{N} \|x_{t+1} - x\|^2_2 = \mathbb{E} \{(\eta_t(X + \sigma_t Z) - X)^2\} = \frac{1}{2^K \sqrt{2 \pi \sigma_t^2}} \int_{-\infty}^{\infty} \left( \frac{p T_t^i}{p T_t^i + 1 - p} - \sqrt{\rho} \right)^2 \exp \left\{ - \frac{(v_t^i - \sqrt{\rho} t)^2}{2 \sigma_t^2} \right\} dv_t^i + (1 - \frac{1}{2^K}) \frac{1}{\sqrt{2 \pi \sigma_t^2}} \int_{-\infty}^{\infty} \left( \frac{p T_t^i}{p T_t^i + 1 - p} \right)^2 \exp \left\{ - \frac{(v_t^i)^2}{2 \sigma_t^2} \right\} dv_t^i = (\sigma_{t+1}^2 - \sigma_w^2) \beta \] (26)

Proof. Proof in Appendix B.

By employing (5) and the fact that \(Z \sim \mathcal{N}(0, I)\), \(\hat{x}_t^{t+1}\) in (29) is obtained as

\[\hat{x}_t^{t+1} = \tilde{\eta}(v_t^i) = \mathbb{E} \{X|X + \rho_t Z = v_t^i\} = \sqrt{\rho_t} \sum_{k=1}^{2^K} \exp \left\{ -\frac{||v_t^i - \sqrt{\rho_t} e_k||_2^2}{2 \rho_t} \right\} e_k
\]

where \(e_1, e_2, \ldots, e_{2^K}\) are orthonormal basis.

Proof. Proof in Appendix B.

By using (9) and (30) and after some simplification, the divergence for the MMSE block non-separable denoiser is derived as in (31) at the top of the next page.

**MUE of the BS-AMP MUD algorithm:** At the \((t+1)\)th iteration of the BS-AMP algorithm, we have \(\hat{x}_t^{t+1} \triangleq [\hat{x}_t^{t+1} \hat{x}_2^{t+1} \ldots \hat{x}_{2^K}^{t+1}]^T\), where \(\hat{x}_0^{t+1}\) is given (30). Similar to the RBS-AMP algorithm, we can obtain the MUE for the BS-AMP algorithm. For \(D \to \infty\), by taking into account Theorem 2 for \(L = D\) and \(\ell = 2^K\), using the fact that \(C^T C = I (N \to \infty, M \to \infty)\), and employing (17) and (19), we can write the interference term as \(I_t^{t+1} = (\rho_t^{t+1} - \sigma_w^2)\) where \(\rho_t^{t+1} = \sigma_w^2 + \frac{1}{\sqrt{2 \pi \sigma_t^2}} \mathbb{E} \{||\tilde{\eta}(X + \rho_t Z) - X||_2^2\}\) and \(\rho_0^2 = \sigma_w^2 + \frac{1}{\sqrt{2 \pi \sigma_t^2}} \mathbb{E} \{||X||_2^2\}\). By substituting \(I_t^{t+1}\) into (27), we obtain the MUE of the BS-AMP MUD algorithm at the \((t+1)\)th iteration as follows

\[\xi_{t+1}^{t+1} = \frac{\sigma_w^2}{\sigma_w^2 + (\rho_t^{t+1} - \sigma_w^2)} = \frac{\sigma_w^2}{\rho_t^{t+1}}. \] (31)

**V. SIMULATION RESULTS**

In Fig. 2, we show the spectral efficiency \(R [\text{bits/sec/Hz}]\) versus \(10 \log_{10}(E_b/N_0)\) for average block error rate \(P_e = 10^{-3}\), \(K = 8\), and \(D = 512\). We also illustrate the AMP with soft-thresholding (9) and the converse bound in (3). As expected, the BS-AMP outperforms the RBS-AMP since the non-separable MMSE denoiser in more efficiently suppresses noise. Moreover, as seen, both algorithms outperform the AMP with soft-thresholding. For packets of 8 information bits, the RBS-AMP and BS-AMP MUD algorithms reach 4/9 and 1/2 of the converse bound at \(E_b/N_0 = 4 \text{ dB}\), respectively.

**VI. CONCLUSION**

Two new MUD algorithms for massive G-MAC with short-packet transmission were developed in this letter. Our proposed MUD algorithms were developed based on AMP with non-separable and separable MMSE denoisers. Our results show that the proposed MUD algorithms offer superior soft decoding performance compared to the state of the art AMP with soft-threshold denoising. This higher spectral efficiency is achieved with lower computational complexity for massive G-MAC with short-packet transmission due to decoupling effect.

**APPENDIX A**

For the proof, we need to show that the sample mean estimator in (19) is consistent [12]. Let us define random variable \(Q_t \triangleq ||\tilde{\eta}(U + \rho_t Z) - U||_2^2\) with finite mean \(\theta_t \triangleq \mathbb{E} \{Q_t\}\). The sample mean estimate of \(\theta_t\) is given by \(\hat{\theta}_{t,L} = \frac{1}{L} \sum_{l=1}^{L} q_{t,l}\) where \(q_{t,l}, l = 1, 2, \ldots, L\) are i.i.d. drawn from random variable \(Q_t\), and the second inequality comes from the decoupling behavior in (15) and (16) and the deionising operation in (5). By using Markov’s inequality, we can write

\[P \left( |\hat{\theta}_{t,L} - \theta_t| > \epsilon_t \right) \leq \frac{\mathbb{E} \{ |\hat{\theta}_{t,L} - \theta_t| \}}{\epsilon_t}. \] (32)

For the sample mean estimator \(\hat{\theta}_{t,L}\), we have

\[|\hat{\theta}_{t,L} - \theta_t|^2 = \frac{1}{L} \sum_{l=1}^{L} q_{t,l} - \theta_t|^2 \leq \frac{1}{L^2} \sum_{l=1}^{L} |q_{t,l} - \theta_t|^2. \] (33)
\[ \text{div}(\eta_i(v^t)) = \frac{1}{N} \sum_{u=1}^{D} T_i \frac{\partial \tilde{\eta}_i(v^t_u)}{\partial v^t_u} = \frac{1}{N} \sum_{u=1}^{D} \sum_{n=1}^{2^K} \exp \left( -\frac{|v^t_u - \sqrt{\rho_t} e_n|^2}{2\rho_t^2} \right) \sum_{k=1}^{2^K} \left( \frac{\sqrt{\rho_t} - v_{u,k}}{\rho_t^2} \right) \exp \left( -\frac{|v^t_u - \sqrt{\rho_t} e_k|^2}{2\rho_t^2} \right) \]

\[ - \frac{\sqrt{\rho_t}}{N} \sum_{u=1}^{D} \left[ \sum_{n=1}^{2^K} \exp \left( -\frac{|v^t_u - \sqrt{\rho_t} e_n|^2}{2\rho_t^2} \right) \right]^2 \sum_{k=1}^{2^K} \left( \frac{\sqrt{\rho_t} - v_{u,k}}{\rho_t^2} \right) \exp \left( -\frac{|v^t_u - \sqrt{\rho_t} e_k|^2}{2\rho_t^2} \right) \]

By applying the statistical expectation to both sides of (33), and then using (32), we obtain

\[ P(|\hat{\theta}_t, L - \theta|^2 > \epsilon^t_2) \leq \frac{\sigma^2_{Q_k}}{L \epsilon^t_2}, \] (34)

where \( \mathbb{E}[Q_k - \theta|^2] = \sigma^2_{Q_k} = \mathbb{E}[|\eta_t, L - \theta|^2] \). The right-hand side of (34) goes to zero as \( L \to \infty \). Hence, for any small positive values \( \epsilon_t \), we can write

\[ \lim_{L \to \infty} P \left( \left| \frac{1}{L} \sum_{u=1}^{L} \| u^t_1 - u_{-t} \|^2 - \theta_t \right|^2 > \epsilon^t_2 \right) = 0. \] (35)

APPENDIX B

For the proposed BS-AMP MUD with block non-separable MMSE denoiser with the prior PDF in (4), we have

\[ \hat{x}^t_1 = \mathbb{E}\{X | X + \rho_t Z = v^t_1\} = \int_{\mathbb{R}^2} f_{X|V}(x|v^t_1; K) \, dx, \]

\[ = \int_{\mathbb{R}^2} f_{V_1}(v^t_1|x; K) f_{X|V}(x|v^t_1; K) \, dx \]

\[ = \sqrt{\rho_t} \sum_{k=1}^{2^K} \sum_{n=1}^{2^K} f_{V_1}(v^t_1|e_n; K) f_{X|V}(x|e_n; K) e_k, \] (36)

where \( f_{V_1}|V(v^t_1|x; K) (f_{X|V}(v^t_1|x; K)) \) denotes the conditional PDF of the random vector \( v^t_1(x) \) given random vector \( X \) (\( v^t_1 \)), and \( e_k \triangleq \sqrt{\rho_t} e_k \). The last equality in (36) is obtained by replacing (3) in the third equality and then using \( \int_{\mathbb{R}^2} f(x-x_0) \, dx = 1 \). Moreover, due to the decoupling effect, we can write \( v^t_1 = x_u + \mathcal{N}(0, \rho_t^2 I) \); thus, we have

\[ f_{X|V}(v^t_1|x; K) = \frac{1}{\eta} \exp \left( -\frac{|v^t_1 - x_u|^2}{2\rho_t^2} \right), \] (37)

where \( \eta = \frac{1}{\sqrt{2\pi}} \frac{\sigma^2}{2} \). Finally, by substituting (37) into (36), we obtain (39). Similarly, for the RBS-AMP MUD algorithm, we can write

\[ \hat{x}^{t+1}_1 = \eta(v^{t+1}_1) = \mathbb{E}\{X | x_1 + \sigma_t Z = v^{t+1}_1\} \]

\[ = \int_{-\infty}^{\infty} \hat{f}_{X|V}(x^1_1|v^{t+1}_1; K) \, dx_1 \]

\[ = \frac{p f_{V_1}(v^{t+1}_1|x_1 = \sqrt{\rho_t} K) (1-p) f_{V_1}(v^{t+1}_1|x_1 = 0; K)}, \]

where \( p = 1/2^K \). By substituting \( f_{V_1}(v^{t+1}_1|x_1 = \sqrt{\rho_t} K) = \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{(v^{t+1}_1 - x_1)^2}{2\sigma^2_t} \right) \) into (38), and after some simplification, we obtain (23).

APPENDIX C

At the \((t+1)\)th iteration of the RBS-AMP algorithm, we have \( x^{t+1} \triangleright [x^{t+1}_1 x^{t+1}_2 \ldots x^{t+1}_N]^T \), where \( x^{t+1}_1 \) is given in (23). By employing (1) and the fact that \( w \) is independent of \( x \) and \( v^t_1 \) (due to decoupling), the average reconstruction error \( E \) at the \((t+1)\)th iteration of the RBS-AMP can be written as

\[ E^{t+1} = \frac{1}{M} \mathbb{E}\{y - CX^{t+1}\} \]

\[ = \frac{1}{M} \mathbb{E}\{w^T w\} + \frac{1}{M} \mathbb{E}\{(x - \hat{x}^{t+1})^T C^T C (x - \hat{x}^{t+1})\}. \] (39)

The first term in (39) denotes the noise power \( \sigma^2_w \), and the second term is the interference at the \((t+1)\)th iteration, i.e., \( I^{t+1} \). By using Theorem 1 for \( N \to \infty \), we can write

\[ I^{t+1} = \frac{1}{M} \mathbb{E}\{(x - \hat{x}^{t+1})^T C^T C (x - \hat{x}^{t+1})\}. \] (41)

By substituting (41) into (27), we obtain the MUE of the RBS-AMP MUD algorithm at the \((t+1)\)th iteration as in (28).

REFERENCES

[1] M. Mohammadkarimi, A. M. Raza, and O. A. Dobre, “Signature-based nonorthogonal massive multiple access for future wireless networks: Uplink massive connectivity for machine-type communications,” IEEE Veh. Technol. Mag., vol. 13, no. 4, pp. 40–50, Oct. 2018.

[2] M. Mohammadkarimi, O. A. Dobre, and M. Z. Win, “Massive uncoordinated multiple access for beyond 5G,” IEEE Trans. Wireless Commun., 2021.

[3] I. Zadik, Y. Polyanskiy, and C. Thrampoulidis, “Improved bounds on Gaussian MAC and sparse regression via Gaussian inequalities,” in proc. ISIT, 2019, pp. 430–434.

[4] J. Choi, “NOMA-based compressive random access using Gaussian spreading,” IEEE Trans. Commun., vol. 67, no. 7, pp. 5167–5177, Mar. 2019.

[5] S. Rangan, “Generalized approximate message passing for estimation with random linear mixing,” in Proc. ISIT, 2011, pp. 2168–2172.

[6] M. Mayer and N. Goertz, “Bayesian optimal approximate message passing to recover structured sparse signals,” arXiv preprint arXiv:1508.01104, 2015.

[7] M. A. Zhou, C. Rush, D. Baron, and D. Needell, “An approximate message passing framework for side information,” vol. 67, no. 7, pp. 1875–1888, Feb. 2019.

[8] R. Berthier, A. Montanari, and P.-M. Nguyen, “State evolution for approximate message passing with non-separable functions,” Information and Inference: A Journal of the IMA, vol. 9, no. 1, pp. 33–79, Jan. 2019.

[9] 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.

[10] C. Rush and R. Venkataramanan, “Finite sample analysis of approximate message passing algorithms,” IEEE Trans. Inf. Theory, vol. 64, no. 11, pp. 7264–7286, Mar. 2018.

[11] 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, Jan. 2011.

[12] A. Stuart, M. G. Kendall et al., The advanced theory of statistics. Griffin, 1963.