Minimum Energy per Bit of Unsourced Multiple Access with Location-Based Codebook Partitioning

Deekshith Pathayappilly Krishnan*, Kaan Okumus*, Khac-Hoang Ngo†, and Giuseppe Durisi*
*Department of Electrical Engineering, Chalmers University of Technology, 41296 Gothenburg, Sweden
†Department of Electrical Engineering, Linköping University, 58183 Linköping, Sweden
Email: {deepat, okumus}@chalmers.se, khac-hoang.ngo@liu.se, durisi@chalmers.se

***Abstract*—We derive finite-blocklength bounds on the minimum achievable energy per bit over a Gaussian unsourced multiple access (UMA) channel in the presence of heterogeneous path-loss conditions. We consider a setting in which the path loss is known to the users, which enables the use of location-based codebook partitioning [Çakmak et al., 2025]. Through numerical simulations and a large-system analysis based on the replica method, we quantify the performance gain of this strategy relative to the conventional UMA approach in which all users employ a common codebook.**

I. Introduction

Massive machine-type communication (mMTC) is a rapidly growing use-case in wireless networks due to the fast expansion of Internet of Things (IoT) systems. mMTC is characterized by sporadic, uplink-centered transmissions of short packets from a high density of devices often operating under stringent energy-efficiency requirements [1]. As first formalized in [2], the mMTC problem has unique features, which makes it different from the traditional multiple access channel (MAC). One key difference is that, to capture the massiveness of the user population, it is fundamental to assume that it is impossible for the system to assign to each user a different codebook. The resulting scenario, in which all users are equipped with the same codebook, is commonly referred to in the literature as unsourced multiple-access (UMA).

Finite-blocklength bounds on the minimum energy per bit for the Gaussian UMA scenario were first obtained in [2], under the assumption that all users are received at the same power. These bounds have been recently improved and generalized to different scenarios, including the case in which the number of active users is random and unknown to the receiver [3] and the case in which transmission occurs over a quasi-static multiple-input multiple-output (MIMO) fading channel [4]. See [5] for a recent review of the area. Furthermore, coding schemes approaching these bounds have been designed. A particularly promising approach [6]–[9] leverages the similarity of UMA decoding and sparse signal recovery and relies on approximate message passing (AMP) for message recovery. This strategy, however, incurs a challenge when applied to realistic channel models in which signals from different users are received at different average-power levels because of path-loss effects. Indeed, since in UMA it is, in general, not possible to establish any association between the position of a user (and, hence, its path loss) and the codeword the user will transmit, one is forced to use at the AMP denoiser a *diffuse* prior, which involves an averaging over all possible user positions.

This problem has been recently sidestepped in [10] via the insightful observation that, in most cellular wireless networks, an estimate of the path loss between users and base-stations can be obtained at the users via the downlink control information transmitted by the cellular base stations. This information can be used in a UMA system as follows: in the system-design phase, one quantizes the possible path loss values using a pre-determined number Q of levels; then, one partitions the codebook in Q subcodebooks. In the operational phase, each user estimates its path loss, quantizes it, and uses the corresponding subcodebook to transmit its message. This strategy, referred in [10] as *location-based codebook partitioning*, allows one to establish a link between transmitted codewords and path loss, and results in a much more concentrated prior for the AMP denoiser. As shown in [10] both theoretically and experimentally, this yields better AMP decoder performance.

Contributions: The purpose of this paper is to assess the effectiveness of location-based codebook partitioning via finite-blocklength information-theoretic bounds similar to the ones developed in [2]. For simplicity, we restrict our attention to a toy-model version of the scenario considered in [10]. Specifically, we assume that the channel between each user and a single-antenna receiver is modeled as a nonfading Gaussian channel, with (deterministic) path loss that can take only two different values, which we denote by g_1 and g_2, respectively.

For this setup, we derive finite-blocklength bounds on the minimum energy per bit achievable when the codebook is partitioned into two subcodebooks and each user selects the subcodebook corresponding to its path loss. Furthermore, we compare this achievability bound with those obtained when all users select codewords from the same codebook, and when the decoder employs successive interference cancellation. Through numerical results, we show that the bound on the minimum energy per bit achieved via location-based codebook partitioning lies below the two alternative bounds. We also present simulation results obtained by combining the coded compressive sensing (CCS) coding scheme proposed in [8] with the multisource-AMP decoder introduced in [10]. These results exhibit a similar performance ordering.

To provide insights into the benefits of location-based codebook partitioning, we finally present a large-system characterization of the per-user probability of error (PUPE) via replica analysis [11], in the regime where both the number of users and the blocklength grow to infinity with a fixed ratio. This (non-rigorous) large-system analysis reveals that, with location-based codebook partitioning, the UMA large-system performance can be characterized by analyzing two equivalent scalar Gaussian channels with SNR proportional to g_1^2 and g_2^2, respectively. On the contrary, when a single codebook is used, the relevant scalar channel is a fading channel, with instantaneous fading value (not known to the receiver), taking value in $\{g_1, g_2\}$ with probability depending on the asymptotic fraction of users experiencing path loss g_1 and g_2, respectively.

Notation: We denote system parameters by uppercase non-italic letters, e.g., K. Uppercase italic letters, e.g., X, denote scalar random variables and their realizations are in lowercase, e.g., x. We also use Greek letters to denote random variables when appropriate; such choices are stated explicitly. Vectors are denoted likewise in boldface, e.g., a random vector $\boldsymbol{X}$ and its realization $\boldsymbol{x}$. We denote their ith entries as $[\boldsymbol{X}]_i$ and $[\boldsymbol{x}]_i$. We use a script font for random matrices, e.g., $\mathscr{C}$, and a sans-serif font for deterministic matrices, e.g., C. We denote the $n \times n$ identity matrix by I_n, and the all-zero vector by $\mathbf{0}$. The superscript $^{\mathrm{T}}$ stands for transposition. We denote sets with calligraphic letters, e.g., $\mathcal{S}$, the set of integers $\{m, \dots, n\}$, $m \le n$, as $[m:n]$, the set $[1:n]$ as $[n]$, and the set of all size-k subsets of $\mathcal{A}$ by $\binom{\mathcal{A}}{k}$. We denote the set of natural numbers by $\mathbb{N}$ and the set of reals by $\mathbb{R}$. We denote the real-valued Gaussian vector distribution with mean $\boldsymbol{\mu}$ and covariance matrix A by $\mathcal{N}(\boldsymbol{\mu}, \mathsf{A})$, and the Bernoulli distribution with parameter p by $\mathrm{Ber}(p)$. We denote convergence in distribution by $\xrightarrow{D}$. Finally, $x^+ = \max\{0, x\}$; $\otimes$ denotes the Kronecker product; $\mathbb{1}\{\cdot\}$ is the indicator function; $\mathrm{diag}(x_1, \dots, x_n)$ denotes the diagonal matrix with $(x_1, \dots, x_n)$ as the diagonal.

II. System Model

We consider a stationary memoryless Gaussian UMA scenario, in which $\mathrm{K_a}$ users transmit their messages to a receiver over N channel uses. The users are assumed to be clustered according to their path loss. Specifically, users within the same cluster ℓ experience the same path loss g_ℓ. For simplicity, we consider the case of 2 clusters. Our results can however be readily generalized to an arbitrary finite number of clusters. We focus on the impact of heterogeneous path loss, and do not model small-scale fading. Let $\boldsymbol{X}_{\ell,k} \in \mathbb{R}^{\mathrm{N}}$ be the signal transmitted by the kth user in cluster ℓ. The received signal is

$$\boldsymbol{Y} = g_1 \sum_{k=1}^{\mathrm{K_{a1}}} \boldsymbol{X}_{1,k} + g_2 \sum_{k=1}^{\mathrm{K_{a2}}} \boldsymbol{X}_{2,k} + \boldsymbol{Z} \tag{1}$$

where $\mathrm{K_{a1}}$ and $\mathrm{K_{a2}}$, with $\mathrm{K_{a1}} + \mathrm{K_{a2}} = \mathrm{K_a}$, are the number of active users that belong to clusters 1 and 2, respectively, and $\boldsymbol{Z} \sim \mathcal{N}(\mathbf{0}, \mathsf{I}_n)$ is the Gaussian noise, which is independent of the transmitted signals. We consider the power constraint $\|\boldsymbol{X}_{\ell,k}\|^2 \le \mathrm{NP}$ for all ℓ and k. We also assume that $\mathrm{K_a}, \mathrm{K_{a1}}$, and $\mathrm{K_{a2}}$ are fixed and known to the receiver. Finally, we assume that each user has perfect knowledge of its path loss. As already pointed out, this information can be obtained in practical systems via the downlink control information transmitted by cellular base stations.

For this channel, an $(\mathrm{M}, \mathrm{N}, \mathrm{P}, \epsilon)$ UMA code with codebook size M and codeword length N consists of

- two encoding functions $f_\ell : [\mathrm{M}] \to \mathbb{R}^{\mathrm{N}}$, $\ell \in \{1, 2\}$, that produce the transmitted codeword $\boldsymbol{X}_{\ell,k} = f_\ell(W_k)$, satisfying the power constraint, of user k in cluster ℓ, for a user message W_k uniformly distributed over $[\mathrm{M}]$;
- a decoding function $g : \mathbb{R}^{\mathrm{N}} \to \binom{[\mathrm{M}]}{\mathrm{K_a}}$ that provides an estimate $\widehat{\mathcal{W}} = \{\widehat{W}_1, \dots, \widehat{W}_{|\widehat{\mathcal{W}}|}\} = g(\boldsymbol{Y})$ of the list of transmitted messages.

The decoding function satisfies the following constraint on the PUPE:

$$P_e = \frac{1}{|\widehat{\mathcal{W}}|} \sum_{i=1}^{|\widehat{\mathcal{W}}|} \mathbb{P}[\widetilde{W}_i \notin \widehat{\mathcal{W}}] \le \epsilon, \tag{2}$$

where $\widetilde{\mathcal{W}} = \{\widetilde{W}_1, \dots, \widetilde{W}_{|\widetilde{\mathcal{W}}|}\}$ denotes the set of distinct elements of $\mathcal{W} = \{W_1, \dots, W_{\mathrm{K_a}}\}$. In (2), we use the convention $0/0 = 0$ to circumvent the case $|\widehat{\mathcal{W}}| = 0$.

The main difference between the definition of UMA coding scheme just provided and the one originally provided in [2] is that we allow the encoder to depend on the cluster. This allows us to model location-based codebook partitioning.

III. Random Coding Bound

In this section, we first derive a random-coding achievability bound on the PUPE achievable over the channel (1) that exploits location-based codebook partitioning. Then, we provide in Section III-B two bounds for the case in which $f_1 = f_2$, which we refer to as common codebook case. The first bound relies on joint decoding, whereas the second bound relies on interference cancellation.

A. Location-Based Codebook Partitioning

We assume that the codewords of each of the two codebooks $\mathcal{C}_\ell = \{\boldsymbol{C}_{\ell,m}\}_{m=1}^{\mathrm{M}}$, $\ell \in \{1, 2\}$ are drawn independently (across both ℓ and m) from a $\mathcal{N}(\mathbf{0}, \mathrm{P}'\mathsf{I}_{\mathrm{N}})$ distribution, for a fixed $\mathrm{P}' < \mathrm{P}$. To convey message W_k, an active user k in cluster ℓ transmits $\boldsymbol{C}_{\ell,W_k}$, provided that $\|\boldsymbol{C}_{\ell,W_k}\|^2 \le \mathrm{NP}$. Otherwise, the user transmits the all-zero codeword. That is,

$$f_\ell(W_k) = \boldsymbol{C}_{\ell,W_k} \mathbb{1}\left\{\|\boldsymbol{C}_{\ell,W_k}\|^2 \le \mathrm{NP}\right\}. \tag{3}$$

We consider a joint decoder, whose output is the set of estimated messages $\widehat{\mathcal{W}} = \widehat{\mathcal{W}}_1 \cup \widehat{\mathcal{W}}_2$, where

$$(\widehat{\mathcal{W}}_1, \widehat{\mathcal{W}}_2) = \operatorname*{arg\,min}_{\substack{\mathcal{W}_1', \mathcal{W}_2' \subset [\mathrm{M}] :\, \mathcal{W}_1' \cap \mathcal{W}_2' = \emptyset, \\ |\mathcal{W}_1'| = \mathrm{K_{a1}}, |\mathcal{W}_2'| = \mathrm{K_{a2}}}} \|\boldsymbol{Y} - g_1 c_1(\mathcal{W}_1') - g_2 c_2(\mathcal{W}_2')\| \tag{4}$$

with $c_\ell(\mathcal{W}') = \sum_{w \in \mathcal{W}'} \boldsymbol{C}_{\ell,w}$, $\ell \in \{1, 2\}$.

An error analysis of this random-coding scheme yields the following achievability bound.

Theorem 1 (Random-coding bound, location-based codebook partitioning, joint decoding)**.** *Fix* $\mathrm{P}' < \mathrm{P}$. *For the Gaussian*

UMA channel (1) *there exists an* $(\mathrm{M}, \mathrm{N}, \mathrm{P}, \epsilon)$ *random-access code for which*

$$\epsilon \leq \sum_{t=0}^{\mathrm{K_a}} \frac{t}{\mathrm{K_a}} p_t + p_0 \tag{5}$$

where

$$p_0 = 1 - \frac{\mathrm{M}!}{M^{\mathrm{K_a}}(M - \mathrm{K_a})!} + \mathrm{K_a} \frac{\Gamma(\frac{\mathrm{N}}{2}, \frac{\mathrm{N}P}{2\mathrm{P}'})}{\Gamma(\mathrm{N}/2)}, \tag{6}$$

$$p_t = \sum_{t_{\mathrm{MD1}} \in \mathcal{T}_t} \sum_{t_{\mathrm{FP1}} \in \mathcal{T}_t} \sum_{t_{\mathrm{AC1}} \in \mathcal{T}'_{t, t_{\mathrm{MD1}}, t_{\mathrm{FP1}}}} \exp(-\mathrm{N} E(t, t_{\mathrm{MD1}}, t_{\mathrm{FP1}}, t_{\mathrm{AC1}})), \tag{7}$$

$$\begin{aligned} E(t, t_{\mathrm{MD1}}, t_{\mathrm{FP1}}, t_{\mathrm{AC1}}) &= \max_{\rho_1, \rho_2, \rho_3 \in [0,1]} -\rho_1\rho_2\rho_3(R_1 + R_2) \\ &\quad - \rho_2\rho_3(R_3 + R_4) - \rho_3(R_5 + R_6) \\ &\quad + E_0(\rho_1, \rho_2, \rho_3), \end{aligned} \tag{8}$$

$$R_1 = \frac{1}{\mathrm{N}} \ln \binom{\mathrm{K_{a1}} - t_{\mathrm{MD1}}}{t_{\mathrm{AC1}}}, \tag{9}$$

$$R_2 = \frac{1}{\mathrm{N}} \ln \binom{\mathrm{K_{a2}} - t + t_{\mathrm{MD1}}}{t_{\mathrm{AC1}} + t_{\mathrm{MD1}} - t_{\mathrm{FP1}}}, \tag{10}$$

$$R_3 = \frac{1}{\mathrm{N}} \ln \binom{M - \mathrm{K_a}}{t_{\mathrm{FP1}}}, \tag{11}$$

$$R_4 = \frac{1}{\mathrm{N}} \ln \binom{M - \mathrm{K_a} - t_{\mathrm{FP1}}}{t - t_{\mathrm{FP1}}}, \tag{12}$$

$$R_5 = \frac{1}{\mathrm{N}} \ln \binom{\mathrm{K_{a1}}}{t_{\mathrm{MD1}}}, \tag{13}$$

$$R_6 = \frac{1}{\mathrm{N}} \ln \binom{\mathrm{K_{a2}}}{t - t_{\mathrm{MD1}}}, \tag{14}$$

$$E_0(\rho_1, \rho_2, \rho_3) = \max_{\lambda} \rho_3\rho_2 a_1 + \rho_3 a_2 + \frac{1}{2}\ln(1 - 2\rho_3 b), \tag{15}$$

$$a_1 = \frac{\rho_1}{2}\ln(1 + 2\lambda\kappa_{\mathrm{AC}}\mathrm{P}') + \frac{1}{2}\ln(1 + 2\mu_1\kappa_{\mathrm{FP}}\mathrm{P}'), \tag{16}$$

$$a_2 = \frac{1}{2}\ln(1 + 2\mu_2\kappa_{\mathrm{MD}}\mathrm{P}'), \tag{17}$$

$$b = \lambda\rho_1\rho_2 - \frac{\mu_2}{1 + 2\mu_2\kappa_{\mathrm{MD}}\mathrm{P}'}, \tag{18}$$

$$\mu_1 = \frac{\lambda\rho_1}{1 + 2\lambda\kappa_{\mathrm{AC}}\mathrm{P}'}, \tag{19}$$

$$\mu_2 = \frac{\mu_1\rho_2}{1 + 2\mu_1\kappa_{\mathrm{FP}}\mathrm{P}'}, \tag{20}$$

$$\kappa_{\mathrm{AC}} = (g_1^2 + g_2^2)(2t_{\mathrm{AC1}} + t_{\mathrm{MD1}} - t_{\mathrm{FP1}}), \tag{21}$$

$$\kappa_{\mathrm{MD}} = g_1^2 t_{\mathrm{MD1}} + g_2^2(t - t_{\mathrm{MD1}}), \tag{22}$$

$$\kappa_{\mathrm{FP}} = g_1^2 t_{\mathrm{FP1}} + g_2^2(t - t_{\mathrm{FP1}}), \tag{23}$$

$$\mathcal{T}_t = [(t - \mathrm{K_{a2}})^+ : \min\{t, \mathrm{K_{a1}}\}], \tag{24}$$

$$\mathcal{T}'_{t, t_{\mathrm{MD1}}, t_{\mathrm{FP1}}} = [\max\{0, t_{\mathrm{FP1}} - t_{\mathrm{MD1}}\} : \min\{\mathrm{K_{a1}} - t_{\mathrm{MD1}}, \mathrm{K_{a2}} - t + t_{\mathrm{FP1}}\}]. \tag{25}$$

Proof. The proof follows similar steps as the proof of [2, Th. 1], namely, a change of measure and a Gallager-type error exponent analysis that makes use of Chernoff bound, Gallager's ρ-trick, and Gaussian statistics. One fundamental difference compared to [2] is that the decoder, when analyzing codewords received at power $\mathrm{P}g_2^2$, may put out a set of messages, which we denote by $\mathcal{W}_{\mathrm{AC1}}$, that are false positives from the perspective of cluster 2, but happen to coincide with the messages from cluster 1, and vice versa. We refer to these messages as "accidentally correct" (AC) messages. Compared to [2], this results in an additional union bound over $t_{\mathrm{AC1}} = |\mathcal{W}_{\mathrm{AC1}}|$ (see (7)). See Appendix B for details. □

B. Common Codebook

We now consider the common codebook case $\mathcal{C}_1 = \mathcal{C}_2 = \{\boldsymbol{C}_1, \dots, \boldsymbol{C}_{\mathrm{M}}\}$.

1) Joint Decoding: We first obtain a bound for the case of joint decoding.

Corollary 1 (Random-coding bound, common codebook, joint decoding)**.** *Fix* $\mathrm{P}' < \mathrm{P}$. *For the Gaussian UMA channel* (1)*, there exists an* $(\mathrm{M}, \mathrm{N}, \mathrm{P}, \epsilon)$ *random-access code for which* ϵ *is bounded as in* (5) *with* (21) *replaced by*

$$\kappa_{\mathrm{AC}} = (g_1 - g_2)^2(2t_{\mathrm{AC1}} + t_{\mathrm{MD1}} - t_{\mathrm{FP1}}). \tag{26}$$

Proof. The proof follows along the same lines as the proof of Theorem 1, with the fundamental difference that the assumption of common codebook causes an increase in κ_{AC}, which represents the variance of a term related to the accidentally corrected messages. See Appendix C. □

2) Interference-Cancellation Decoding: Assume without loss of generality that $g_1 \geq g_2$. We now consider an interference-cancellation decoder that operates by first decoding messages from cluster 1 as

$$\widehat{\mathcal{W}}_1 = \operatorname*{arg\,min}_{\mathcal{W}'_1 \subset [M] :\, |\mathcal{W}'_1| = \mathrm{K_{a1}}} \|\boldsymbol{Y} - g_1 c(\mathcal{W}'_1)\| \tag{27}$$

with $c(\mathcal{W}) = \sum_{w \in \mathcal{W}} \boldsymbol{C}_w$, and then canceling the interference from cluster 1 to decode messages coming from cluster 2 as

$$\widehat{\mathcal{W}}_2 = \operatorname*{arg\,min}_{\mathcal{W}'_2 \subset [M] \setminus \widehat{\mathcal{W}}_1 :\, |\mathcal{W}'_2| = \mathrm{K_{a2}}} \|\boldsymbol{Y} - g_1 c(\widehat{\mathcal{W}}_1) - g_2 c(\mathcal{W}'_2)\|. \tag{28}$$

We state a random-coding bound for this strategy in the following theorem.

Theorem 2 (Random-coding bound, common codebooks, interference-cancellation decoding)**.** *Fix* $\mathrm{P}' < \mathrm{P}$. *For the Gaussian UMA channel* (1)*, there exists an* $(\mathrm{M}, \mathrm{N}, \mathrm{P}, \epsilon)$ *random-access code for which* ϵ *is bounded as in* (5) *with* $E_0(\rho_1, \rho_2, \rho_3)$ *replaced by*

$$\begin{aligned} E_0(\rho_1, \rho_2, \rho_3) = \max_{\lambda_1, \lambda_2} \Big(&\frac{1}{2}\rho_1\rho_2\rho_3 \ln\det(\mathbf{I}_2 - 2\Sigma_{\mathrm{AC}}\mathbf{U}_{\mathrm{AC}}) \\ &+ \frac{1}{2}\rho_2\rho_3 \ln\det(\mathbf{I}_2 - 2\Sigma_{\mathrm{FP}}\mathbf{U}_{\mathrm{FP}}) \\ &+ \frac{1}{2}\rho_3 \ln\det(\mathbf{I}_2 - 2\Sigma_{\mathrm{MD}}\mathbf{U}_{\mathrm{MD}}) \\ &+ \frac{1}{2}\ln\det(\mathbf{I}_2 - 2\Sigma_{\mathrm{ZC}}\mathbf{U}_{\mathrm{ZC}}) \Big), \end{aligned} \tag{29}$$

$$\Sigma_{\mathrm{AC}} = \operatorname{diag}(t_{\mathrm{AC1}}\mathrm{P}', (t_{\mathrm{AC1}} + t_{\mathrm{MD1}} - t_{\mathrm{FP1}})\mathrm{P}'), \tag{30}$$

$$\Sigma_{\rm FP} = {\rm diag}(t_{\rm FP1}{\rm P}', (t - t_{\rm FP1}){\rm P}'), \tag{31}$$

$$\Sigma_{\rm MD} = {\rm diag}(t_{\rm MD1}{\rm P}', (t - t_{\rm MD1}){\rm P}'), \tag{32}$$

$$\Sigma_{\rm ZC} = {\rm diag}(1, ({\rm K_{a2}} - t + t_{\rm FP1} - t_{\rm AC1}){\rm P}'), \tag{33}$$

$$\mathbf{U}_{\rm AC} = \begin{bmatrix} u_{\rm AC}^{(1)} & u_{\rm AC}^{(2)} \\ u_{\rm AC}^{(2)} & u_{\rm AC}^{(3)} \end{bmatrix}, \tag{34}$$

$$u_{\rm AC}^{(1)} = -\lambda_1 g_1^2 - \lambda_2 g_2(g_2 - 2g_1), \tag{35}$$

$$u_{\rm AC}^{(2)} = \lambda_1 g_1(g_1 - g_2) + \lambda_2 g_2(g_2 - 2g_1), \tag{36}$$

$$u_{\rm AC}^{(3)} = -\lambda_1 g_1(g_1 - 2g_2) - \lambda_2 g_2(g_2 - 2g_1), \tag{37}$$

$$\mathbf{U}_{\rm FP} = 2\rho_1 \Lambda_{\rm FP} \mathbf{Q}_{\rm FP} - \rho_1 \begin{bmatrix} \lambda_1 g_1^2 & \lambda_2 g_1 g_2 \\ \lambda_2 g_1 g_2 & \lambda_2 g_2^2 \end{bmatrix}, \tag{38}$$

$$\mathbf{Q}_{\rm FP} = \begin{bmatrix} \lambda_1 g_1^2 - \lambda_2 g_1 g_2 & \lambda_2 g_2(g_1 - g_2) \\ -\lambda_1 g_1(g_1 - g_2) + \lambda_2 g_1 g_2 & -\lambda_2 g_2(g_1 - g_2) \end{bmatrix}, \tag{39}$$

$$\Lambda_{\rm FP} = \mathbf{Q}_{\rm FP}^{\rm T}(\Sigma_{\rm AC}^{-1} - 2\mathbf{U}_{\rm AC})^{-1}, \tag{40}$$

$$\begin{aligned} \mathbf{U}_{\rm MD} &= 2\rho_2 \Lambda_{\rm MD} \mathbf{Q}_{\rm MD} + 2\rho_2\rho_1 \Lambda_{\mathbf{M}} \mathbf{M} \\ &\quad - \rho_2\rho_1 \begin{bmatrix} \lambda_1 g_1^2 & (\lambda_1 + \lambda_2) g_1 g_2 \\ (\lambda_1 + \lambda_2) g_1 g_2 & \lambda_2 g_2^2 \end{bmatrix}, \end{aligned} \tag{41}$$

$$\mathbf{M} = \lambda_1 g_1 g_2 \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} - \mathbf{Q}_{\rm FP}, \tag{42}$$

$$\mathbf{Q}_{\rm MD} = 2\rho_1 \Lambda_{\rm FP} \mathbf{M} + \rho_1 \begin{bmatrix} \lambda_1 g_1^2 & (\lambda_1 + \lambda_2) g_1 g_2 \\ \lambda_2 g_1 g_2 & \lambda_2 g_2^2 \end{bmatrix}, \tag{43}$$

$$\Lambda_{\rm MD} = \mathbf{Q}_{\rm MD}^{\rm T}(\Sigma_{\rm FP}^{-1} - 2\mathbf{U}_{\rm FP})^{-1}, \tag{44}$$

$$\Lambda_{\mathbf{M}} = \mathbf{M}^{\rm T}(\Sigma_{\rm AC}^{-1} - 2\mathbf{U}_{\rm AC})^{-1}, \tag{45}$$

$$\begin{aligned} \mathbf{U}_{\rm ZC} &= 2\rho_3 \mathbf{Q}_{\rm ZC}^{\rm T}(\Sigma_{\rm MD}^{-1} - 2\mathbf{U}_{\rm MD})^{-1}\mathbf{Q}_{\rm ZC} \\ &\quad + 2\rho_3\rho_2 \Lambda_{\rm ZCa}^{\rm T}(\Sigma_{\rm FP}^{-1} - 2\mathbf{U}_{\rm FP})^{-1}\Lambda_{\rm ZCa} \\ &\quad + 4\rho_3\rho_2\rho_1 \Lambda_{\rm ZCb}^{\rm T}(\Sigma_{\rm AC}^{-1} - 2\mathbf{U}_{\rm AC})^{-1}\Lambda_{\rm ZCb}, \end{aligned} \tag{46}$$

$$\begin{aligned} \mathbf{Q}_{\rm ZC} &= \rho_2\rho_1 \left(-\mathbf{I}_2 + (4\Lambda_{\rm MD}\Lambda_{\rm FP} + 2\Lambda_{\mathbf{M}}) \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} + 2\Lambda_{\rm MD} \right) \\ &\quad \cdot \begin{bmatrix} \lambda_1 g_1 & \lambda_1 g_1 g_2 \\ \lambda_2 g_2 & 0 \end{bmatrix}, \end{aligned} \tag{47}$$

$$\Lambda_{\rm ZCa} = \rho_1 \left(\mathbf{I}_2 + 2\Lambda_{\rm FP} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \right) \begin{bmatrix} \lambda_1 g_1 & \lambda_1 g_1 g_2 \\ \lambda_2 g_2 & 0 \end{bmatrix}, \tag{48}$$

$$\Lambda_{\rm ZCb} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \lambda_1 g_1 & \lambda_1 g_1 g_2 \\ \lambda_2 g_2 & 0 \end{bmatrix}. \tag{49}$$

Proof. The proof follows similar steps as the proof of Theorem 1, with some extensions to capture the sequential nature of the decoding operations (27) and (28). See Appendix D. □

IV. Replica Method Prediction

To obtain insights on system performance, we provide next a large-system limit characterization of the PUPE achievable over the channel (1). To derive the result, we rely on the replica method, in line with its application to multiuser detection systems [12]–[14], and the original UMA problem [15].

To state the main results of this section, we first introduce the following definition of multiuser spectral efficiency.

Definition 1. *Let*

$$B = \sqrt{a}A + N, \tag{50}$$

with $a > 0$, $A \sim {\rm Ber}(p)$, $p \in (0,1)$, and $N \sim \mathcal{N}(0,1)$. Let $I(a)$ denote the mutual information $I(A;B)$ and let G be a binary random variable taking values in $\{g_1, g_2\}$ with probability $\mathbb{P}[G = g_\ell] = \gamma_\ell$. Then, for every $\beta > 0$, the multiuser efficiency η is defined as

$$\eta = \arg\min_x \big(\beta\,\mathbb{E}_G[I(x{\rm P}G^2)] + \tfrac{1}{2}(x - 1 - \ln x)\big). \tag{51}$$

A. Location-Based Codebook Partitioning

We describe the random codebook corresponding to cluster ℓ as a matrix $\mathscr{C}_\ell$ with M columns drawn independently from a $\mathcal{N}(\mathbf{0}, \mathbf{I}_{\rm N}/{\rm N})$ distribution. For a fixed $\mu \in (0,1)$, we consider the regime in which the total number of active users satisfies ${\rm K_a} = \mu{\rm N}$. In contrast to Section III, where each user independently selects a message uniformly at random yielding a message selection probability $1 - (1 - 1/{\rm M})^{\rm K_a} \approx {\rm K_a}/{\rm M}$, we consider a setting in which, within each cluster ℓ with ${\rm K}_{{\rm a}\ell}$ active users, each message is selected independently according to a ${\rm Ber}({\rm K}_{{\rm a}\ell}/{\rm M})$ distribution. Let $\alpha_\ell = {\rm K}_{{\rm a}\ell}/{\rm K_a}$ and $\beta = 2{\rm M}/{\rm N}$. Furthermore, let $\boldsymbol{U}'_\ell$ denote an ${\rm M} \times 1$ vector with entries drawn independently from a ${\rm Ber}(\alpha_\ell\mu/\beta)$ distribution, representing message selection in cluster ℓ. Let finally $\boldsymbol{Z}' \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{\rm N})$. We analyze the large-system performance of location-based codebook partitioning achievable over the channel

$$\boldsymbol{Y}' = \sqrt{\rm P} g_1 \mathscr{C}_1 \boldsymbol{U}'_1 + \sqrt{\rm P} g_2 \mathscr{C}_2 \boldsymbol{U}'_2 + \boldsymbol{Z}'. \tag{52}$$

We have the following result.

Claim 1 (Replica decoupling, location-based codebook partitioning)**.** *Let $\ell \in \{1,2\}$. Fix $p_\ell = \alpha_\ell\mu/\beta$, and $\gamma_\ell = 1/2$. Denote by*

$$V'_{\ell,{\rm N}} = \mathbb{P}[[\boldsymbol{U}'_\ell]_1 = 1 \mid \boldsymbol{Y}', \mathscr{C}_1, \mathscr{C}_2] \tag{53}$$

the marginal posterior probability associated with (52)*. Consider also the following scalar channels for $\ell \in \{1,2\}$:*

$$B_\ell = \sqrt{{\rm P}\eta} g_\ell A_\ell + N_\ell. \tag{54}$$

Here, $A_\ell \sim {\rm Ber}(p_\ell\gamma_\ell)$, and $N_\ell \sim \mathcal{N}(0,1)$, independent of A_ℓ. Let ${\rm M}, {\rm N}, {\rm K_a} \to \infty$ with β, μ, and α_ℓ held constant. Then,

$$V'_{\ell,{\rm N}} \xrightarrow{D} \mathbb{P}[A_\ell = 1 \mid B_\ell]. \tag{55}$$

Proof. The proof follows from [14, Prop. 1]. See for details. □

Remark 1. *Claim 1 implies that, in the large-system limit, the channel* (52) *decouples into the two scalar channels in* (54)*. Note in particular that the two scalar channels have deterministic path losses known to the receiver. Following an approach similar to the one detailed in [15], we can use these two scalar channels to obtain a large-system characterization of the PUPE. Specifically, let ϵ_ℓ be the solution of*

$$\sqrt{{\rm P}\eta} g_\ell = Q^{-1}(\epsilon_\ell) + Q^{-1}(p_\ell\gamma_\ell\epsilon_\ell/(1 - p_\ell\gamma_\ell)). \tag{56}$$

In the large-system limit, the PUPE is given by $(\epsilon_1 + \epsilon_2)/2$.

B. Common Codebook

We denote the common codebook as a matrix $\mathscr{C}$ with M columns drawn independently from a $\mathcal{N}(\mathbf{0}, \mathbf{I}_{\mathrm{N}}/\mathrm{N})$ distribution. We let μ and α_ℓ to be defined as before, but set now $\beta = \mathrm{M}/\mathrm{N}$, and use a $\mathrm{Ber}(\mathrm{K_a}/\mathrm{M})$ distribution for message selection. Note that $\mathrm{K_a}/\mathrm{M} = \mu/\beta$. Let $\boldsymbol{U}''$ be the binary vector describing the message selection. Let also G be a random variable taking values in $\{g_1,\ g_2\}$ with $\mathbb{P}[G = g_\ell] = \alpha_\ell$. Finally, let $\boldsymbol{Z}'' \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{\mathrm{N}})$. We analyze the large-system performance achievable over the channel

$$\boldsymbol{Y}'' = \sqrt{P} G \mathscr{C} \boldsymbol{U}'' + \boldsymbol{Z}''. \tag{57}$$

We have the following result.

Claim 2 (Replica decoupling, common codebook)**.** *Let $\ell \in \{1, 2\}$. Fix $p = \mu/\beta$, and $\gamma_\ell = \alpha_\ell$. Denote by*

$$V''_{\mathrm{N}} = \mathbb{P}[[\boldsymbol{U}'']_1 = 1 \mid \boldsymbol{Y}'', \mathscr{C}''], \tag{58}$$

the marginal posterior probability associated with (57)*. Consider the scalar channel*

$$B'' = \sqrt{P\eta} G A'' + N''. \tag{59}$$

Here, $A'' \sim \mathrm{Ber}(p)$, G takes values in $\{g_1, g_2\}$ with $\mathbb{P}[G = g_\ell] = \alpha_\ell$, and $N'' \sim \mathcal{N}(0, 1)$. Furthermore, these three random variables are mutually independent. Let $\mathrm{M}, \mathrm{N}, \mathrm{K_a} \to \infty$ with β, μ, and α_ℓ held constant. Then,

$$V''_{\mathrm{N}} \xrightarrow{D} \mathbb{P}[A'' = 1 \mid B'']. \tag{60}$$

Proof. The proof follows from [13, Claim 1]. See for details. □

Remark 2. *Note that, in contrast to the previous case, the equivalent scalar channel in the common codebook case is a single fading channel, with fading coefficient G not known to the receiver. The large-system characterization of the PUPE for this scenario can again be carried out along the lines of [15]. However, the presence of G precludes a closed-form expression. Specifically, in the large-system limit, the PUPE is given by*

$$\epsilon = \frac{1-p}{p} \mathbb{P}_1 \left[\ln \frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1} \geq \theta \right] \tag{61}$$

where θ is determined by imposing that

$$\mathbb{P}_0 \left[\ln \frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1} \geq \theta \right] + \frac{1-p}{p} \mathbb{P}_1 \left[\ln \frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1} \geq \theta \right] = 1. \tag{62}$$

Here, $\mathbb{P}_0 = \mathcal{N}(0, 1)$ and $\mathbb{P}_1 = \gamma_1 \mathcal{N}(g_1, 1) + \gamma_2 \mathcal{N}(g_2, 1)$.

V. Numerical Results

We evaluate the minimum energy per bit $\mathrm{N}P/(2\mathrm{B})$ required to achieve a PUPE of 0.01 as a function of the number of active users $\mathrm{K_a}$ for the case in which each user map messages of $\mathrm{B} = 128$ bits to codewords of length $\mathrm{N} = 30\,000$. Throughout, we set $g_1 = 1$, $g_2 = 0.8$, and $\mathrm{K_{a2}} = 2\mathrm{K_{a1}}$. In Fig. 1, we depict

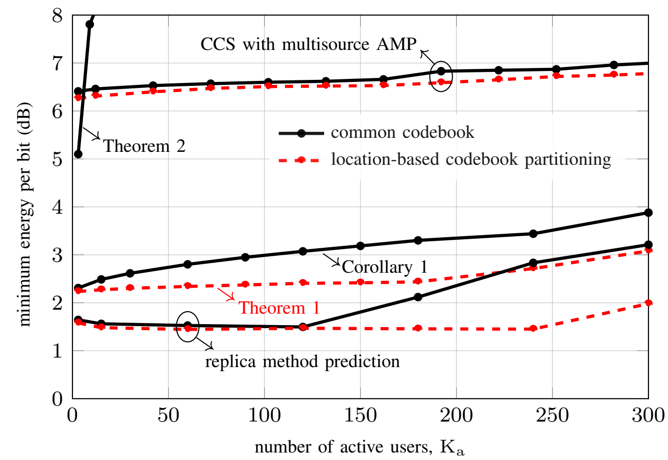

Fig. 1. Minimum energy per bit to achieve target PUPE of 0.01 vs. number of active users $\mathrm{K_a}$.

the random-coding bound with location-based codebook partitioning in Theorem 1 as well as those corresponding to common codebook with joint decoding (Corollary 1) and common codebook with interference-cancellation decoding (Theorem 2). In the figure, we also depict the replica method predictions, obtained via the expressions given in Remark 1 and Remark 2. We also use the normalization with respect to the effective number of bits suggested in [15]. Finally, we provide simulation results for a CCS scheme [8] in which the decoder operates according to the multisource-AMP framework proposed in [10]. Specifically, in the resulting scheme, each message is divided into 16 blocks of 16 bits each. The CCS inner decoder performs multisource-AMP signal reconstruction, while the outer decoder stitches together the reconstructed signals, to ensure that they form valid codewords. To do so, we assume that the inner decoder produces a list of candidate messages of size $\mathrm{K_a} + 10$.

As shown in the figure, for the channel model considered in the paper, location-based codeword partitioning results in a consistent reduction of the minimum energy per bit (although this reduction is marginal when $\mathrm{K_a}$ is small) across all types of curves depicted in the figure (random-coding bounds, replica-method predictions, performance of CCS schemes). Note also that, for the scenario considered in this section, i.e., $\mathrm{K_{a2}} = 2\mathrm{K_{a1}}$ with $g_2 < g_1$, interference cancellation exhibits poor performance (6 dB gap from the location-based codeword partitioning bound for $\mathrm{K_a} = 12$).

VI. Conclusion

For a Gaussian UMA channel characterized by heterogeneous path loss, we showed that location-based codebook partitioning outperforms, in terms of the minimum energy per bit required to meet a specific PUPE, the conventional UMA framework, which utilizes a common codebook for all users. These gains were validated through finite-blocklength random coding bounds, replica method large-system limit predictions, and empirical performance of a coding scheme based on CCS and multisource-AMP.

We anticipate that the energy efficiency gains from location-based codebook partitioning will be even more pronounced in wireless network architectures featuring distributed access points (distributed MIMO). Indeed, in such systems, the access points are located so as to ensure uniform quality of service, which should amplify the benefit of the location-based codebook partitioning, as illustrated in [10] for the case of multisource-AMP decoders. Such extension, as well as the inclusion of small-scale fading will be considered in future works.

References

[1] A. E. Kalor, G. Durisi, S. Coleri, S. Parkvall, W. Yu, A. Mueller, and P. Popovski, “Wireless 6G connectivity for massive number of devices and critical services,” *Proc. IEEE*, Nov. 2024.

[2] Y. Polyanskiy, “A perspective on massive random-access,” in *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, Aachen, Germany, Jun. 2017, pp. 2523–2527.

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

[4] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, “Energy efficient coded random access for the wireless uplink,” *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 4694–4708, Jun. 2020.

[5] K. Andreev, P. Rybin, and A. Frolov, “Unsourced random access,” *Foundations and Trends® in Commun. and Inf. Theory*, vol. 22, no. 5-6, pp. 605–842, Aug. 2025.

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

[7] V. K. Amalladinne, A. Vem, D. K. Soma, K. R. Narayanan, and J. Chamberland, “A coupled compressive sensing scheme for unsourced multiple access,” in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Proc. (ICASSP)*, Calgary, AB, Canada, Apr. 2018, pp. 6628–6632.

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

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

[10] B. Çakmak, E. Gkiouzepi, M. Opper, and G. Caire, “Joint message detection and channel estimation for unsourced random access in cell-free user-centric wireless networks,” *IEEE Trans. Inf. Theory*, vol. 71, no. 5, pp. 3614–3643, May 2025.

[11] S. F. Edwards and P. W. Anderson, “Theory of spin glasses,” *Journal of Physics F: Metal Physics*, vol. 5, no. 5, p. 965, May 1975.

[12] T. Tanaka, “A statistical-mechanics approach to large-system analysis of CDMA multiuser detectors,” *IEEE Trans. Inf. Theory*, vol. 48, no. 11, pp. 2888–2910, Nov. 2002.

[13] D. Guo and S. Verdú, “Randomly spread CDMA: Asymptotics via statistical physics,” *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 1983–2010, Jun. 2005.

[14] D. Guo, “Performance of synchronous multirate CDMA via statistical physics,” in *Proc. Int. Symp. Inf. Theory (ISIT)*, Adelaide, SA, Australia, Sept. 2005, pp. 199–203.

[15] Y. Polyanskiy, “Remarks on massive random access,” 2020, DLR–MIT–TUM Workshop, Munich, Germany, Feb. 2020.

[16] Y. Ohnishi and J. Honorio, “Novel change of measure inequalities with applications to PAC-Bayesian bounds and Monte Carlo estimation,” in *Proc. Int. Conf. Artificial Intelligence and Statistics (AISTATS)*. Virtual Conference: PMLR, Apr. 2021, pp. 1711–1719.

[17] M. H. DeGroot and M. J. Schervish, *Probability and Statistics*, 4th ed. Boston, MA, USA: Pearson, 2012.

[18] R. G. Gallager, *Information theory and reliable communication*. New York, USA: Wiley, 1968.

[19] Y. Kabashima, “Introduction to the replica method.” [Online]. Available: https://ithems.riken.jp/download/file/18/Intro_RM_rev.pdf

[20] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. Springer Berlin, Heidelberg, 2009.

[21] P. Billingsley, *Probability and Measure*, 3rd ed. New York: Wiley, 1995.

Appendix A
Mathematical Preliminaries for the Random Coding Bounds

The following results will be used in the proofs of Theorems 1, and 2.

Lemma 1 (Change of measure [16, Lemma 4])**.** *Let p and q be two probability measures. Consider a random variable X supported on $\mathcal{H}$ and a function $f\colon \mathcal{H} \to [0,1]$. It holds that*

$$\mathbb{E}_p[f(X)] \le \mathbb{E}_q[f(X)] + d_{\mathrm{TV}}(p,q) \tag{63}$$

where $d_{\mathrm{TV}}(p,q)$ denotes the total variation distance between p and q.

Lemma 2 (Chernoff bound [17, Th. 6.2.7])**.** *For a random variable X with moment-generating function $\mathbb{E}[e^{tX}]$ defined for all $|t| \le b$, it holds for all $\lambda \in [0,b]$ that*

$$\mathbb{P}[X \le x] \le e^{\lambda x}\,\mathbb{E}[e^{-\lambda X}]. \tag{64}$$

Corollary 2 (Chernoff joint bound)**.** *For two random variables X_1 and X_2 with joint moment-generating function $\mathbb{E}[e^{t_1 X_1 + t_2 X_2}]$ defined for all $|t_1| \le b_1$ and $|t_2| \le b_2$, it holds for all $\lambda_1 \in [0,b_1]$, $\lambda_2 \in [0,b_2]$ that*

$$\mathbb{P}[X_1 \le x_1, X_2 \le x_2] \le e^{\lambda_1 x_1 + \lambda_2 x_2}\,\mathbb{E}[e^{-\lambda_1 X_1 - \lambda_2 X_2}]. \tag{65}$$

Lemma 3 (Gallager's ρ-trick [18, p. 136])**.** *It holds that $\mathbb{P}[\cup_i A_i] \le (\sum_i \mathbb{P}[A_i])^\rho$ for every $\rho \in [0,1]$.*

Lemma 4 (Moment-generating function of quadratic forms of a Gaussian vector)**.** *Let $\boldsymbol{X} \in \mathbb{R}^{\mathrm{N}}$ and $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$. Let $\mathsf{A} \in \mathbb{R}^{\mathrm{N}\times\mathrm{N}}$ be a symmetric matrix and $\boldsymbol{b} \in \mathbb{R}^{\mathrm{N}}$. For every $\gamma \in \mathbb{R}$ such that $\Sigma^{-1} + 2\gamma\mathsf{A} \succ \mathbf{0}$, it holds that*

$$\begin{aligned}
&\mathbb{E}[\exp(-\gamma(\boldsymbol{X}^{\mathrm{T}}\mathsf{A}\boldsymbol{X} + 2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{X}))] \\
&= \det(\mathsf{I}_N + 2\gamma\Sigma\mathsf{A})^{-1/2} \\
&\quad \cdot \exp\big(2\gamma^2(\mathsf{A}\boldsymbol{\mu} + \boldsymbol{b})^{\mathrm{T}}(\Sigma^{-1} + 2\gamma\mathsf{A})^{-1}(\mathsf{A}\boldsymbol{\mu} + \boldsymbol{b}) \\
&\qquad - \gamma(\boldsymbol{\mu}^{\mathrm{T}}\mathsf{A}\boldsymbol{\mu} + 2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{\mu})\big).
\end{aligned} \tag{66}$$

In particular, for $\gamma = -1$ and $\boldsymbol{\mu} = \mathbf{0}$, it holds that

$$\begin{aligned}
&\mathbb{E}[\exp(\boldsymbol{X}^{\mathrm{T}}\mathsf{A}\boldsymbol{X} + 2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{X})] \\
&= \det(\mathsf{I}_N - 2\Sigma\mathsf{A})^{-1/2}\exp\big(2\boldsymbol{b}^{\mathrm{T}}(\Sigma^{-1} - 2\mathsf{A})^{-1}\boldsymbol{b}\big),
\end{aligned} \tag{67}$$

given that $\Sigma^{-1} \succ 2\mathsf{A}$; for $\Sigma = \sigma^2\mathsf{I}_{\mathrm{N}}$, $\mathsf{A} = \mathsf{I}_{\mathrm{N}}$, and $\boldsymbol{b} = \mathbf{0}$, it holds that

$$\mathbb{E}[e^{-\gamma\|\boldsymbol{X}\|^2}] = (1+2\gamma\sigma^2)^{-\mathrm{N}/2}\exp\bigg(-\frac{\gamma\|\boldsymbol{\mu}\|^2}{1+2\gamma\sigma^2}\bigg), \tag{68}$$

for every $\gamma > -\frac{1}{2\sigma^2}$.

Proof. Denote the quadratic form as $Q = \boldsymbol{X}^{\mathrm{T}}\mathsf{A}\boldsymbol{X} + 2\boldsymbol{b}^{\mathrm{T}}\boldsymbol{X}$. Using the density of $\boldsymbol{X} \sim \mathcal{N}(\boldsymbol{\mu}, \Sigma)$, we compute $\mathbb{E}[e^{-\gamma Q}]$ as

$$\begin{aligned}
\mathbb{E}[e^{-\gamma Q}] &= \frac{1}{(2\pi)^{\mathrm{N}/2}\det(\Sigma)^{1/2}} \\
&\cdot \int_{\mathbb{R}^n} \exp\big(-\tfrac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^{\mathrm{T}}\Sigma^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) - \gamma\boldsymbol{x}^{\mathrm{T}}\mathsf{A}\boldsymbol{x} - 2\gamma\boldsymbol{b}^{\mathrm{T}}\boldsymbol{x}\big)\,\mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{69}$$

We define $\mathsf{K} = \Sigma^{-1} + 2\gamma\mathsf{A}$ and $\boldsymbol{h} = \Sigma^{-1}\boldsymbol{\mu} - 2\gamma\boldsymbol{b}$. Since Σ^{-1} and A are symmetric, K is symmetric. The exponent on the right-hand side of (69) becomes $-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathsf{K}\boldsymbol{x} + \boldsymbol{h}^{\mathrm{T}}\boldsymbol{x} - \frac{1}{2}\boldsymbol{\mu}^{\mathrm{T}}\Sigma^{-1}\boldsymbol{\mu}$. Hence,

$$\begin{aligned}
\mathbb{E}\big[e^{-\gamma Q}\big] &= \frac{\exp\big(-\frac{1}{2}\boldsymbol{\mu}^{\mathrm{T}}\Sigma^{-1}\boldsymbol{\mu}\big)}{(2\pi)^{n/2}\det(\Sigma)^{1/2}} \\
&\quad \cdot \int_{\mathbb{R}^n} \exp\Big(-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathsf{K}\boldsymbol{x} + \boldsymbol{h}^{\mathrm{T}}\boldsymbol{x}\Big)\,\mathrm{d}\boldsymbol{x}.
\end{aligned} \tag{70}$$

Under the condition that K is positive definite, the standard Gaussian integral identity yields

$$\begin{aligned}
&\int_{\mathbb{R}^n} \exp\Big(-\frac{1}{2}\boldsymbol{x}^{\mathrm{T}}\mathsf{K}\boldsymbol{x} + \boldsymbol{h}^{\mathrm{T}}\boldsymbol{x}\Big)\,\mathrm{d}\boldsymbol{x} \\
&\qquad = (2\pi)^{n/2}\det(\mathsf{K})^{-1/2}\exp\Big(\frac{1}{2}\boldsymbol{h}^{\mathrm{T}}\mathsf{K}^{-1}\boldsymbol{h}\Big).
\end{aligned} \tag{71}$$

Substituting this into (70) gives

$$\begin{aligned}
\mathbb{E}\left[e^{-\gamma Q}\right] &= \det(\Sigma)^{-1/2}\det(\mathsf{K})^{-1/2} \\
&\quad \cdot \exp\Big(-\frac{1}{2}\boldsymbol{\mu}^{\mathrm{T}}\Sigma^{-1}\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{h}^{\mathrm{T}}\mathsf{K}^{-1}\boldsymbol{h}\Big),
\end{aligned} \tag{72}$$

which leads to (66) after some simplifications of the determinant and exponent terms. The particular cases (67) and (68) follow straightforwardly from (66).

□

Appendix B
Proof of Theorem 1

We analyze the PUPE achieved with the random-coding scheme introduced in Section III-A, averaged over the Gaussian code ensemble. We denote by $\mathcal{W}_\ell$ the set of messages transmitted by users in cluster $\ell \in \{1,2\}$, and define $(\widehat{\mathcal{W}}_1, \widehat{\mathcal{W}}_2)$ as in (4). Furthermore, we denote by $\mathcal{W}_{\mathrm{MD}}$ the set of misdetected messages, i.e., $\mathcal{W}_{\mathrm{MD}} = \widetilde{\mathcal{W}} \setminus \widehat{\mathcal{W}}$, and by $\mathcal{W}_{\mathrm{FP}}$ the set of false-positive messages, i.e., $\mathcal{W}_{\mathrm{FP}} = \widehat{\mathcal{W}} \setminus \widetilde{\mathcal{W}}$. The PUPE can be expressed as

$$P_{\mathrm{e}} = \mathbb{E}\left[\frac{|\mathcal{W}_{\mathrm{MD}}|}{|\widetilde{\mathcal{W}}|}\right]. \tag{73}$$

A. Change of Measure

We apply Lemma 1 to the random variable $\frac{|\mathcal{W}_{\mathrm{MD}}|}{|\widetilde{\mathcal{W}}|}$ to replace the measure under which the expectation is taken by the one under which: i) the active users transmit distinct messages, i.e., $|\widetilde{\mathcal{W}}| = \mathrm{K_a}$ and $\widetilde{W}_1, \dots, \widetilde{W}_{\mathrm{K_a}}$ are sampled uniformly without replacement from $[\mathrm{M}]$; ii) $\boldsymbol{X}_{\ell,k} = \boldsymbol{C}_{\ell,W_k}$, $\forall \ell, k$, instead of $\boldsymbol{X}_{\ell,k} = \boldsymbol{C}_{\ell,W_k}\mathbb{1}\{\|\boldsymbol{C}_{\ell,W_k}\| \le \mathrm{N}P\}$. The total variation between the original measure and the new one is upper-bounded by

$$\begin{aligned}
&\mathbb{P}\Big[|\widetilde{\mathcal{W}}| < \mathrm{K_a}\Big] + \mathbb{P}[\exists k \in [\mathrm{K_a}]\colon \|\boldsymbol{C}_{\ell,W_k}\| \ge \mathrm{N}P] \\
&\le 1 - \frac{\mathrm{M}!}{M^{\mathrm{K_a}}(M-\mathrm{K_a})!} + \mathrm{K_a}\frac{\Gamma(\mathrm{N}/2, \mathrm{N}P/(2\mathrm{P}'))}{\Gamma(\mathrm{N}/2)} \qquad (74) \\
&= p_0. \qquad (75)
\end{aligned}$$

The inequality (74) follows from the same analysis as in [3, App. A-A]. We consider implicitly the new measure hereafter

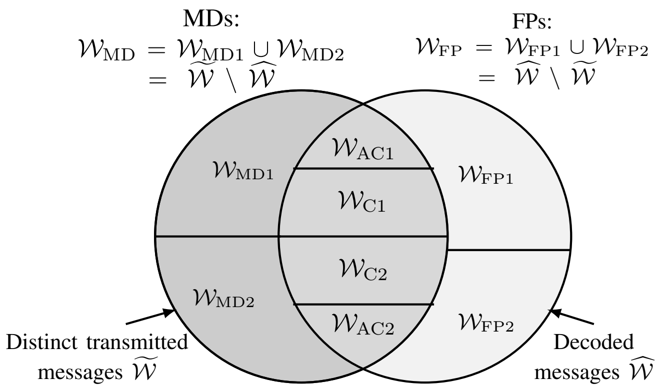

Fig. 2. A diagram depicting the relation between the sets of messages.

at a cost of adding p_0 to the original expectation in (73). Specifically, we expand this expectation as

$$P_{\rm e} \leq \sum_{t=0}^{\rm K_a} \frac{t}{\rm K_a} \mathbb{P}[|\mathcal{W}_{\rm MD}| = |\mathcal{W}_{\rm FP}| = t] + p_0, \tag{76}$$

and next focus on upper-bounding $\mathbb{P}[|\mathcal{W}_{\rm MD}| = |\mathcal{W}_{\rm FP}| = t]$ under the new measure.

B. Message Sets

We further split the sets of misdetected messages, false-positive messages, and correctly decoded messages, as depicted in Fig. 2. We first split $\mathcal{W}_{\rm MD}$ into two sets $\mathcal{W}_{\rm MD1}$ and $\mathcal{W}_{\rm MD2}$ that contain the misdetected messages of users in clusters 1 and 2, respectively. Similarly, we split $\mathcal{W}_{\rm FP}$ into $\mathcal{W}_{\rm FP1}$ and $\mathcal{W}_{\rm FP2}$. Recall that the decoded sets of messages for clusters 1 and 2 are $\widehat{\mathcal{W}}_1$ and $\widehat{\mathcal{W}}_2$, respectively. We denote $\mathcal{W}_{\rm AC1} = \mathcal{W}_1 \cap \widehat{\mathcal{W}}_2$, which is the set of messages that are false positives from the perspective of cluster 2 but happen to coincide with messages transmitted by cluster 1. We refer to these messages as "accidentally correct" messages (hence the subscript AC). Similarly, we denote the set of accidentally correct messages for cluster 2 as $\mathcal{W}_{\rm AC2} = \mathcal{W}_2 \cap \widehat{\mathcal{W}}_1$. Finally, we denote the sets of other correctly decoded messages as $\mathcal{W}_{\rm C1} = \widehat{\mathcal{W}}_1 \setminus \mathcal{W}_{\rm MD1} \setminus \mathcal{W}_{\rm AC1}$ and $\mathcal{W}_{\rm C2} = \widehat{\mathcal{W}}_2 \setminus \mathcal{W}_{\rm MD2} \setminus \mathcal{W}_{\rm AC2}$.

Exploiting symmetry, we assume without loss of generality that $\widetilde{\mathcal{W}} = [\rm K_a]$, $\widetilde{\mathcal{W}}_1 = [\rm K_{a1}]$, and thus $\widetilde{\mathcal{W}}_2 = [\rm K_{a1}+1 : K_a]$. Denote the cardinality of the message sets as $t_{\rm MD1} = |\mathcal{W}_{\rm MD1}|$, $t_{\rm FP1} = |\mathcal{W}_{\rm FP1}|$, $t_{\rm AC1} = |\mathcal{W}_{\rm AC1}|$, and $t_{\rm AC2} = |\mathcal{W}_{\rm AC2}|$. Under the event $|\mathcal{W}_{\rm MD}| = |\mathcal{W}_{\rm FP}| = t$, we have that $|\mathcal{W}_{\rm MD2}| = t - t_{\rm MD1}$ and $|\mathcal{W}_{\rm FP2}| = t - t_{\rm FP1}$. As $\mathcal{W}_1 = \mathcal{W}_{\rm C1} \cup \mathcal{W}_{\rm AC1} \cup \mathcal{W}_{\rm MD1}$, $\widehat{\mathcal{W}}_1 = \mathcal{W}_{\rm C1} \cup \mathcal{W}_{\rm AC2} \cup \mathcal{W}_{\rm FP1}$, and $|\mathcal{W}_1| = |\widehat{\mathcal{W}}_1| = \rm K_{a1}$, we obtain that

$$t_{\rm FP1} + t_{\rm AC2} = t_{\rm MD1} + t_{\rm AC1} = \rm K_{a1} - t_{C1} \leq K_{a1}. \tag{77}$$

Similarly, we have that

$$t_{\rm FP2} + t_{\rm AC1} = t_{\rm MD2} + t_{\rm AC2} = \rm K_{a2} - t_{C2} \leq K_{a2}. \tag{78}$$

Therefore, given t, $t_{\rm MD1}$ is upper-bounded by both t and $\rm K_{a1}$, and lower-bounded as $t_{\rm MD1} = t - t_{\rm MD2} \geq t - \rm K_{a2}$. That is, $t_{\rm MD1}$ is bounded in $\mathcal{T}_t$ defined in (24). Similarly, so is $t_{\rm FP1}$. Furthermore, given $t_{\rm MD1}$ and $t_{\rm FP1}$, $t_{\rm AC1}$ belongs to the set $\mathcal{T}'_{t, t_{\rm MD1}, t_{\rm FP1}}$ defined in (25).

C. Pairwise Error Event

We express the received signal as

$$\begin{aligned} \boldsymbol{Y} = g_1[c_1(\mathcal{W}_{\rm MD1}) + c_1(\mathcal{W}_{\rm AC1}) + c_1(\mathcal{W}_{\rm C1})] \\ + g_2[c_2(\mathcal{W}_{\rm MD2}) + c_2(\mathcal{W}_{\rm AC2}) + c_2(\mathcal{W}_{\rm C2})] + \boldsymbol{Z}. \end{aligned} \tag{79}$$

Furthermore, the sums of the codewords corresponding to the decoded message sets $(\widehat{\mathcal{W}}_1, \widehat{\mathcal{W}}_2)$ are expressed as

$$c_1(\widehat{\mathcal{W}}_1) = c_1(\mathcal{W}_{\rm FP1}) + c_1(\mathcal{W}_{\rm AC2}) + c_1(\mathcal{W}_{\rm C1}), \tag{80}$$

$$c_2(\widehat{\mathcal{W}}_2) = c_2(\mathcal{W}_{\rm FP2}) + c_2(\mathcal{W}_{\rm AC1}) + c_2(\mathcal{W}_{\rm C2}). \tag{81}$$

The pairwise error event $\widetilde{\mathcal{W}} \to \widehat{\mathcal{W}}$ occurs when

$$\|\boldsymbol{Y} - g_1 c_1(\widehat{\mathcal{W}}_1) - g_2 c_2(\widehat{\mathcal{W}}_2)\| < \|\boldsymbol{Z}\|. \tag{82}$$

Using (79), (80), and (81), we obtain that (82) is equivalent to

$$\begin{aligned} &\|\boldsymbol{Z} + g_1 c_1(\mathcal{W}_{\rm MD1}) + g_2 c_2(\mathcal{W}_{\rm MD2}) \\ &\quad - g_1 c_1(\mathcal{W}_{\rm FP1}) - g_2 c_2(\mathcal{W}_{\rm FP2}) \\ &\quad + g_1 c_1(\mathcal{W}_{\rm AC1}) - g_2 c_2(\mathcal{W}_{\rm AC1}) \\ &\quad - g_1 c_1(\mathcal{W}_{\rm AC2}) + g_2 c_2(\mathcal{W}_{\rm AC2})\| \leq \|\boldsymbol{Z}\|. \end{aligned} \tag{83}$$

We denote by $F(\mathcal{W}_{{\rm MD}\ell}, \mathcal{W}_{{\rm FP}\ell}, \mathcal{W}_{{\rm AC}\ell}, \ell \in \{1,2\})$ the set of $(\mathcal{W}_{{\rm MD}\ell}, \mathcal{W}_{{\rm FP}\ell}, \mathcal{W}_{{\rm AC}\ell}, \ell \in \{1,2\})$ such that (83) holds.

D. Error-Exponent Analysis

By writing the event $|\mathcal{W}_{\rm MD}| = |\mathcal{W}_{\rm FP}| = t$ as a union of the pairwise error events $F(\mathcal{W}_{{\rm MD}\ell}, \mathcal{W}_{{\rm FP}\ell}, \mathcal{W}_{{\rm AC}\ell}, \ell \in \{1,2\})$, we have that

$$\begin{aligned} &\mathbb{P}[|\mathcal{W}_{\rm MD}| = |\mathcal{W}_{\rm FP}| = t] \\ &= \mathbb{P}\Bigg[\bigcup_{t_{\rm MD1} \in \mathcal{T}_t} \bigcup_{t_{\rm FP1} \in \mathcal{T}_t} \bigcup_{t_{\rm AC1} \in \mathcal{T}'_{t, t_{\rm MD1}, t_{\rm FP1}}} \\ &\qquad \bigcup_{\mathcal{W}_{\rm MD1} \subset \binom{[\rm K_{a1}]}{t_{\rm MD1}}} \bigcup_{\mathcal{W}_{\rm MD2} \subset \binom{[\rm K_{a1}+1:K_a]}{t - t_{\rm MD1}}} \\ &\qquad \bigcup_{\mathcal{W}_{\rm FP1} \subset \binom{[\rm K_a+1:M]}{t_{\rm FP1}}} \bigcup_{\mathcal{W}_{\rm FP2} \subset \binom{[\rm K_a+1:M] \setminus \mathcal{W}_{\rm FP1}}{t - t_{\rm FP1}}} \\ &\qquad \bigcup_{\mathcal{W}_{\rm AC1} \subset \binom{[\rm K_{a1}] \setminus \mathcal{W}_{\rm MD1}}{t_{\rm AC1}}} \bigcup_{\mathcal{W}_{\rm AC2} \subset \binom{[\rm K_{a1}+1:K_a] \setminus \mathcal{W}_{\rm MD2}}{t_{\rm AC1} + t_{\rm MD1} - t_{\rm FP1}}} \\ &\qquad F(\mathcal{W}_{{\rm MD}\ell}, \mathcal{W}_{{\rm FP}\ell}, \mathcal{W}_{{\rm AC}\ell}, \ell \in \{1,2\}) \Bigg] \end{aligned} \tag{84}$$

$$\begin{aligned} &\leq \sum_{t_{\rm MD1} \in \mathcal{T}_t} \sum_{t_{\rm FP1} \in \mathcal{T}_t} \sum_{t_{\rm AC1} \in \mathcal{T}'_{t, t_{\rm MD1}, t_{\rm FP1}}} \\ &\quad \mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm MD1} \subset \binom{[\rm K_{a1}]}{t_{\rm MD1}}} \bigcup_{\mathcal{W}_{\rm MD2} \subset \binom{[\rm K_{a1}+1:K_a]}{t - t_{\rm MD1}}} \\ &\qquad \bigcup_{\mathcal{W}_{\rm FP1} \subset \binom{[\rm K_a+1:M]}{t_{\rm FP1}}} \bigcup_{\mathcal{W}_{\rm FP2} \subset \binom{[\rm K_a+1:M] \setminus \mathcal{W}_{\rm FP1}}{t - t_{\rm FP1}}} \end{aligned}$$

$$\bigcup_{\mathcal{W}_{\rm AC1}\subset\binom{[\rm K_{a1}]\setminus\mathcal{W}_{\rm MD1}}{t_{\rm AC1}}}\bigcup_{\mathcal{W}_{\rm AC2}\subset\binom{[\rm K_{a1}+1:K_a]\setminus\mathcal{W}_{\rm MD2}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}} F(\mathcal{W}_{\rm MD\ell},\mathcal{W}_{\rm FP\ell},\mathcal{W}_{\rm AC\ell},\ell\in\{1,2\})\Bigg]. \tag{85}$$

Given $c_\ell(\mathcal{W}_{\rm MD\ell})$, $c_\ell(\mathcal{W}_{\rm FP\ell})$, $\ell\in\{1,2\}$, and $\boldsymbol{Z}$, it holds for every $\lambda>0$ that

$$\begin{aligned}
&\mathbb{P}\left[F(\mathcal{W}_{\rm MD\ell},\mathcal{W}_{\rm FP\ell},\mathcal{W}_{\rm AC\ell},\ell\in\{1,2\})\right]\\
&\le \exp(\lambda\|\boldsymbol{Z}\|^2)\,\mathbb{E}_{\mathcal{W}_{\rm AC1},\mathcal{W}_{\rm AC2}}\big[\exp(-\lambda\|\boldsymbol{Z}\\
&\quad+g_1c_1(\mathcal{W}_{\rm MD1})+g_2c_2(\mathcal{W}_{\rm MD2})\\
&\quad-g_1c_1(\mathcal{W}_{\rm FP1})-g_2c_2(\mathcal{W}_{\rm FP2})+g_1c_1(\mathcal{W}_{\rm AC1})\\
&\quad-g_2c_2(\mathcal{W}_{\rm AC1})-g_1c_1(\mathcal{W}_{\rm AC2})+g_2c_2(\mathcal{W}_{\rm AC2})\|^2)\big] \qquad (86)\\
&=\exp(\lambda\|\boldsymbol{Z}\|^2)(1+2\lambda(g_1^2+g_2^2)(t_{\rm AC1}+t_{\rm AC2})\mathrm{P}')^{-\mathrm{N}/2}\\
&\quad\cdot\exp\big(-\lambda\|\boldsymbol{Z}+g_1c_1(\mathcal{W}_{\rm MD1})+g_2c_2(\mathcal{W}_{\rm MD2})\\
&\qquad-g_1c_1(\mathcal{W}_{\rm FP1})-g_2c_2(\mathcal{W}_{\rm FP2})\|^2)\\
&\qquad\cdot(1+2\lambda(g_1^2+g_2^2)(t_{\rm AC1}+t_{\rm AC2})\mathrm{P}')^{-1}\big), \qquad (87)
\end{aligned}$$

where we have applied the Chernoff bound in Lemma 2 in (86) and then used (66) in Lemma 4 to compute the expectation. Note that $(g_1^2+g_2^2)(t_{\rm AC1}+t_{\rm AC2})$ is equal to $\kappa_{\rm AC}$ defined in (21). Next, we apply Gallager's ρ-trick in Lemma 3 to get that, given $c_\ell(\mathcal{W}_{\rm MD\ell})$, $c_\ell(\mathcal{W}_{\rm FP\ell})$, $\ell\in\{1,2\}$, and $\boldsymbol{Z}$, it holds for every $\rho_1\in[0,1]$ that

$$\begin{aligned}
&\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm AC1}\subset\binom{[\rm K_{a1}]\setminus\mathcal{W}_{\rm MD1}}{t_{\rm AC1}}}\bigcup_{\mathcal{W}_{\rm AC2}\subset\binom{[\rm K_{a1}+1:K_a]\setminus\mathcal{W}_{\rm MD2}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}} F(\mathcal{W}_{\rm MD\ell},\mathcal{W}_{\rm FP\ell},\mathcal{W}_{\rm AC\ell},\ell\in\{1,2\})\Bigg]\\
&\le\binom{\mathrm{K_{a1}}-t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1}\binom{\mathrm{K_a}-\mathrm{K_{a1}}-t+t_{\rm MD1}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}^{\rho_1}\\
&\quad\cdot(1+2\lambda\kappa_{\rm AC}\mathrm{P}')^{-\mathrm{N}\rho_1/2}\\
&\quad\cdot\exp\Big(\mu_1\Big(\|\boldsymbol{Z}\|^2-\|\boldsymbol{Z}+g_1c_1(\mathcal{W}_{\rm MD1})+g_2c_2(\mathcal{W}_{\rm MD2})\\
&\qquad-g_1c_1(\mathcal{W}_{\rm FP1})-g_2c_2(\mathcal{W}_{\rm FP2})\|^2\Big)\Big)
\end{aligned}$$

where μ_1 is defined by (19). By taking the expectation over $\mathcal{W}_{\rm FP1}$ and $\mathcal{W}_{\rm FP2}$ using (66) in Lemma 4, we obtain that, for given $c_\ell(\mathcal{W}_{\rm MD\ell})$, $\ell\in\{1,2\}$, and given $\boldsymbol{Z}$,

$$\begin{aligned}
&\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm AC1}\subset\binom{[\rm K_{a1}]\setminus\mathcal{W}_{\rm MD1}}{t_{\rm AC1}}}\bigcup_{\mathcal{W}_{\rm AC2}\subset\binom{[\rm K_{a1}+1:K_a]\setminus\mathcal{W}_{\rm MD2}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}} F(\mathcal{W}_{\rm MD\ell},\mathcal{W}_{\rm FP\ell},\mathcal{W}_{\rm AC\ell},\ell\in\{1,2\})\Bigg]\\
&\le\binom{\mathrm{K_{a1}}-t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1}\binom{\mathrm{K_a}-\mathrm{K_{a1}}-t+t_{\rm MD1}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}^{\rho_1}\\
&\quad\cdot\exp\Big(\lambda\rho_1\|\boldsymbol{Z}\|^2-\mu_1\|\boldsymbol{Z}+g_1c_1(\mathcal{W}_{\rm MD1})+g_2c_2(\mathcal{W}_{\rm MD2})\|^2\\
&\qquad\cdot(1+2\mu_1(g_1^2t_{\rm FP1}+g_2^2t_{\rm FP2})\mathrm{P}')^{-1}-\mathrm{N}a_1\Big) \qquad (88)
\end{aligned}$$

where

$$\begin{aligned}
a_1&=\frac{\rho_1}{2}\ln(1+2\lambda\kappa_{\rm AC}\mathrm{P}')\\
&\quad+\frac{1}{2}\ln\big(1+2\mu_1(g_1^2t_{\rm FP1}+g_2^2t_{\rm FP2})\mathrm{P}'\big). \qquad (89)
\end{aligned}$$

Note that $g_1^2t_{\rm FP1}+g_2^2t_{\rm FP2}$ is equal to $\kappa_{\rm FP}$ defined in (23). Now we apply Gallager's ρ-trick again to obtain that, given $c_\ell(\mathcal{W}_{\rm MD\ell})$, $c_\ell(\mathcal{W}_{\rm FP\ell})$, $\ell\in\{1,2\}$, given $\boldsymbol{Z}$ and for every $\rho_2\in[0,1]$,

$$\begin{aligned}
&\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm FP1}\subset\binom{[\rm K_a+1:M]}{t_{\rm FP1}}}\bigcup_{\mathcal{W}_{\rm FP2}\subset\binom{[\rm K_a+1:M]\setminus\mathcal{W}_{\rm FP1}}{t-t_{\rm FP1}}}\\
&\quad\bigcup_{\mathcal{W}_{\rm AC1}\subset\binom{[\rm K_{a1}]\setminus\mathcal{W}_{\rm MD1}}{t_{\rm AC1}}}\bigcup_{\mathcal{W}_{\rm AC2}\subset\binom{[\rm K_{a1}+1:K_a]\setminus\mathcal{W}_{\rm MD2}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}}\\
&\quad F(\mathcal{W}_{\rm MD\ell},\mathcal{W}_{\rm FP\ell},\mathcal{W}_{\rm AC\ell},\ell\in\{1,2\})\Bigg] \qquad (90)\\
&\le\binom{M-\mathrm{K_a}}{t_{\rm FP1}}^{\rho_2}\binom{M-\mathrm{K_a}-t_{\rm FP1}}{t-t_{\rm FP1}}^{\rho_2}\\
&\quad\cdot\binom{\mathrm{K_{a1}}-t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1\rho_2}\binom{\mathrm{K_a}-\mathrm{K_{a1}}-t+t_{\rm MD1}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}^{\rho_1\rho_2}\\
&\quad\cdot\mathbb{E}_{c_1(\mathcal{W}_{\rm MD1}),c_2(\mathcal{W}_{\rm MD2})}\Big[\exp\Big(\lambda\rho_1\rho_2\|\boldsymbol{Z}\|^2\\
&\qquad-\mu_2\|\boldsymbol{Z}+g_1c_1(\mathcal{W}_{\rm MD1})+g_2c_2(\mathcal{W}_{\rm MD2})\|^2-\mathrm{N}\rho_2a_1\Big)\Big] \qquad (91)\\
&=\binom{M-\mathrm{K_a}}{t_{\rm FP1}}^{\rho_2}\binom{M-\mathrm{K_a}-t_{\rm FP1}}{t-t_{\rm FP1}}^{\rho_2}\\
&\quad\cdot\binom{\mathrm{K_{a1}}-t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1\rho_2}\binom{\mathrm{K_a}-\mathrm{K_{a1}}-t+t_{\rm MD1}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}^{\rho_1\rho_2}\\
&\quad\cdot(1+2\mu_2(g_1^2t_{\rm MD1}+g_2^2t_{\rm MD2})\mathrm{P}')^{-\mathrm{N}/2}\\
&\quad\cdot\exp\Big(\Big(\lambda\rho_1\rho_2-\frac{\mu_2}{1+2\mu_2(g_1^2t_{\rm MD1}+g_2^2t_{\rm MD2})\mathrm{P}'}\Big)\|\boldsymbol{Z}\|^2\\
&\qquad-N\rho_2a_1\Big) \qquad (92)\\
&=\binom{M-\mathrm{K_a}}{t_{\rm FP1}}^{\rho_2}\binom{M-\mathrm{K_a}-t_{\rm FP1}}{t-t_{\rm FP1}}^{\rho_2}\\
&\quad\cdot\binom{\mathrm{K_{a1}}-t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1\rho_2}\binom{\mathrm{K_a}-\mathrm{K_{a1}}-t+t_{\rm MD1}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}^{\rho_1\rho_2}\\
&\quad\cdot\exp(b\|\boldsymbol{Z}\|^2-\mathrm{N}\rho_2a_1-\mathrm{N}a_2), \qquad (93)
\end{aligned}$$

where we used (66) in Lemma 4 to obtain the first equality, and we define μ_2 by (20) and

$$b=\lambda\rho_1\rho_2-\frac{\mu_2}{1+2\mu_2(g_1^2t_{\rm MD1}+g_2^2t_{\rm MD2})\mathrm{P}'}, \tag{94}$$

$$a_2=\frac{1}{2}\ln\big(1+2\mu_2(g_1^2t_{\rm MD1}+g_2^2t_{\rm MD2})\mathrm{P}'\big). \tag{95}$$

Note that $g_1^2t_{\rm MD1}+g_2^2t_{\rm MD2}$ is equal to $\kappa_{\rm MD}$ defined in (22). We now apply Gallager's ρ-trick the third time to obtain that, given $\boldsymbol{Z}$ and for every $\rho_3\in[0,1]$,

$$=\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm MD1}\subset\binom{[\rm K_{a1}]}{t_{\rm MD1}}}\bigcup_{\mathcal{W}_{\rm MD2}\subset\binom{[\rm K_{a1}+1:K_a]}{t-t_{\rm MD1}}}$$

$$\bigcup_{\mathcal{W}_{\mathrm{FP}1}\subset\binom{[\mathrm{K_a}+1:\mathrm{M}]}{t_{\mathrm{FP}1}}}\ \bigcup_{\mathcal{W}_{\mathrm{FP}2}\subset\binom{[\mathrm{K_a}+1:\mathrm{M}]\setminus\mathcal{W}_{\mathrm{FP}1}}{t-t_{\mathrm{FP}1}}}\ \bigcup_{\mathcal{W}_{\mathrm{AC}1}\subset\binom{[\mathrm{K_{a1}}]\setminus\mathcal{W}_{\mathrm{MD}1}}{t_{\mathrm{AC}1}}}\ \bigcup_{\mathcal{W}_{\mathrm{AC}2}\subset\binom{[\mathrm{K_{a1}}+1:\mathrm{K_a}]\setminus\mathcal{W}_{\mathrm{MD}2}}{t_{\mathrm{AC}1}+t_{\mathrm{MD}1}-t_{\mathrm{FP}1}}} F(\mathcal{W}_{\mathrm{MD}\ell},\mathcal{W}_{\mathrm{FP}\ell},\mathcal{W}_{\mathrm{AC}\ell},\ell\in\{1,2\})\Bigg] \tag{96}$$

$$\begin{aligned}
&\le \binom{\mathrm{K_{a1}}}{t_{\mathrm{MD}1}}^{\rho_3}\binom{\mathrm{K_{a2}}}{t-t_{\mathrm{MD}1}}^{\rho_3}\\
&\quad\cdot\binom{M-\mathrm{K_a}}{t_{\mathrm{FP}1}}^{\rho_2\rho_3}\binom{M-\mathrm{K_a}-t_{\mathrm{FP}1}}{t-t_{\mathrm{FP}1}}^{\rho_2\rho_3}\\
&\quad\cdot\binom{\mathrm{K_{a1}}-t_{\mathrm{MD}1}}{t_{\mathrm{AC}1}}^{\rho_1\rho_2\rho_3}\binom{\mathrm{K_a}-\mathrm{K_{a1}}-t+t_{\mathrm{MD}1}}{t_{\mathrm{AC}1}+t_{\mathrm{MD}1}-t_{\mathrm{FP}1}}^{\rho_1\rho_2\rho_3}\\
&\quad\cdot\mathbb{E}_{\boldsymbol{Z}}[\exp(b\|\boldsymbol{Z}\|^2-\mathrm{N}\rho_2\rho_3 a)-\mathrm{N}\rho_3 a_1)] \qquad (97)\\
&=\binom{\mathrm{K_{a1}}}{t_{\mathrm{MD}1}}^{\rho_3}\binom{\mathrm{K_{a2}}}{t-t_{\mathrm{MD}1}}^{\rho_3}\\
&\quad\cdot\binom{M-\mathrm{K_a}}{t_{\mathrm{FP}1}}^{\rho_2\rho_3}\binom{M-\mathrm{K_a}-t_{\mathrm{FP}1}}{t-t_{\mathrm{FP}1}}^{\rho_2\rho_3}\\
&\quad\cdot\binom{\mathrm{K_{a1}}-t_{\mathrm{MD}1}}{t_{\mathrm{AC}1}}^{\rho_1\rho_2\rho_3}\binom{\mathrm{K_a}-\mathrm{K_{a1}}-t+t_{\mathrm{MD}1}}{t_{\mathrm{AC}1}+t_{\mathrm{MD}1}-t_{\mathrm{FP}1}}^{\rho_1\rho_2\rho_3}\\
&\quad\cdot(1-2b\rho_3)^{-\mathrm{N}/2}\exp(-\mathrm{N}\rho_2\rho_3 a_1-\mathrm{N}\rho_3 a_2), \qquad (98)
\end{aligned}$$

where in the last equality, we use (66) in Lemma 4 to compute the expectation over $\boldsymbol{Z}$.

We conclude that $\mathbb{P}\,|\mathcal{W}_{\mathrm{MD}}|=|\mathcal{W}_{\mathrm{FP}}|=t$ is upper-bounded by the right-hand side of (98). We complete the proof by substituting this bound into (76).

Appendix C
Proof of Corollary 1

When $\mathcal{C}_1=\mathcal{C}_2$, we have that $c_1(\mathcal{W})=c_2(\mathcal{W})$ for a given message set $\mathcal{W}$. The proof follows the same steps detailed in Appendix B, except that, in (87), the covariance matrix of $g_1c_1(\mathcal{W}_{\mathrm{AC}1})-g_2c_2(\mathcal{W}_{\mathrm{AC}1})-g_1c_1(\mathcal{W}_{\mathrm{AC}2})+g_2c_2(\mathcal{W}_{\mathrm{AC}2})$ is given by $\kappa_{\mathrm{AC}}\mathrm{P}'\mathbf{I}_{\mathrm{N}}$ with κ_{AC} defined in (26).

Appendix D
Proof of Theorem 2

By performing the same change of measure as in Appendix B, we obtain the bound (76). We next focus on upper-bounding $\mathbb{P}\,|\mathcal{W}_{\mathrm{MD}}|=|\mathcal{W}_{\mathrm{FP}}|=t$ under the new measure and for the interference-cancellation decoder.

A. Pairwise Error Event

For this decoder, the pairwise error event $\widetilde{\mathcal{W}}\to\widehat{\mathcal{W}}$ implies that

$$\|\boldsymbol{Y}-g_1c(\widehat{\mathcal{W}}_1)\|\le\|\boldsymbol{Y}-g_1c(\mathcal{W}_1)\|, \tag{99}$$

$$\|\boldsymbol{Y}-g_1c(\widehat{\mathcal{W}}_1)-g_2c(\widehat{\mathcal{W}}_2)\|\le\|\boldsymbol{Y}-g_1c(\widehat{\mathcal{W}}_1)-g_2c(\mathcal{W}_2)\| \tag{100}$$

where we recall that $c(\mathcal{W})=\sum_{w\in\mathcal{W}}\boldsymbol{C}_w$. Splitting message sets as in Appendix B, we express (99) and (100) as

$$\begin{aligned}
&\|\underbrace{\boldsymbol{Z}+g_1c(\mathcal{W}_{\mathrm{MD}1})+g_2c(\mathcal{W}_{\mathrm{MD}2})-g_1c(\mathcal{W}_{\mathrm{FP}1})+g_2c(\mathcal{W}_{\mathrm{C}2})}_{=\boldsymbol{A}}\\
&\quad+g_1c(\mathcal{W}_{\mathrm{AC}1})+(g_2-g_1)c(\mathcal{W}_{\mathrm{AC}2})\|\\
&\quad\le\|\underbrace{\boldsymbol{Z}+g_2c(\mathcal{W}_{\mathrm{C}2})+g_2c(\mathcal{W}_{\mathrm{MD}2})}_{=\boldsymbol{B}}+g_2c(\mathcal{W}_{\mathrm{AC}2})\|, \qquad (101)\\
&\|\underbrace{\boldsymbol{Z}+g_1c(\mathcal{W}_{\mathrm{MD}1})+g_2c(\mathcal{W}_{\mathrm{MD}2})-g_1c(\mathcal{W}_{\mathrm{FP}1})-g_2c(\mathcal{W}_{\mathrm{FP}2})}_{=\boldsymbol{C}}\\
&\quad+(g_1-g_2)(c(\mathcal{W}_{\mathrm{AC}1})-c(\mathcal{W}_{\mathrm{AC}2}))\|\\
&\quad\le\|\underbrace{\boldsymbol{Z}+g_1c(\mathcal{W}_{\mathrm{MD}1})-g_1c(\mathcal{W}_{\mathrm{FP}1})}_{=\boldsymbol{D}}\\
&\qquad+g_1c(\mathcal{W}_{\mathrm{AC}1})-g_1c(\mathcal{W}_{\mathrm{AC}2})\|, \qquad (102)
\end{aligned}$$

where we also defined the vectors $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ to make the notation more compact in the following. We denote by $F_{\mathrm{IC}}(\mathcal{W}_{\mathrm{MD}\ell},\mathcal{W}_{\mathrm{FP}\ell},\mathcal{W}_{\mathrm{AC}\ell},\ell\in\{1,2\})$ the set of $(\mathcal{W}_{\mathrm{MD}\ell},\mathcal{W}_{\mathrm{FP}\ell},\mathcal{W}_{\mathrm{AC}\ell},\ell\in\{1,2\})$ such that (101) and (102) hold. Note that $\mathcal{W}_{\mathrm{C}2}$ is fully determined given $\widetilde{\mathcal{W}}=[\mathrm{K_{a1}}+1:\mathrm{K_a}]$, $\mathcal{W}_{\mathrm{MD}2}$, and $\mathcal{W}_{\mathrm{AC}2}$. By writing the event $|\mathcal{W}_{\mathrm{MD}}|=|\mathcal{W}_{\mathrm{FP}}|=t$ as a union of the pairwise error events, we obtain a similar bound as in (85) with $F(\mathcal{W}_{\mathrm{MD}\ell},\mathcal{W}_{\mathrm{FP}\ell},\mathcal{W}_{\mathrm{AC}\ell},\ell\in\{1,2\})$ replaced by $F_{\mathrm{IC}}(\mathcal{W}_{\mathrm{MD}\ell},\mathcal{W}_{\mathrm{FP}\ell},\mathcal{W}_{\mathrm{AC}\ell},\ell\in\{1,2\})$.

B. Error Exponent Analysis

We apply the Chernoff joint bound in Corollary 2 to bound the pairwise error probability as

$$\begin{aligned}
&\mathbb{P}[F_{\mathrm{IC}}(\mathcal{W}_{\mathrm{MD}\ell},\mathcal{W}_{\mathrm{FP}\ell},\mathcal{W}_{\mathrm{AC}\ell},\ell\in\{1,2\})]\\
&=\mathbb{P}\Big[\|\boldsymbol{A}+g_1c(\mathcal{W}_{\mathrm{AC}1})+(g_2-g_1)c(\mathcal{W}_{\mathrm{AC}2})\|^2\\
&\qquad-\|\boldsymbol{B}+g_2c(\mathcal{W}_{\mathrm{AC}2})\|^2\le 0,\\
&\quad\ \|\boldsymbol{C}+(g_1-g_2)(c(\mathcal{W}_{\mathrm{AC}1})-c(\mathcal{W}_{\mathrm{AC}2}))\|^2\\
&\qquad-\|\boldsymbol{D}+g_1c(\mathcal{W}_{\mathrm{AC}1})-g_1c(\mathcal{W}_{\mathrm{AC}2})\|^2\le 0\Big] \qquad (103)\\
&\le\mathbb{E}\Big[\exp\Big(-\lambda_1(\|\boldsymbol{A}+g_1c(\mathcal{W}_{\mathrm{AC}1})+(g_2-g_1)c(\mathcal{W}_{\mathrm{AC}2})\|^2\\
&\qquad-\|\boldsymbol{B}+g_2c(\mathcal{W}_{\mathrm{AC}2})\|^2)\\
&\qquad-\lambda_2(\|\boldsymbol{C}+(g_1-g_2)(c(\mathcal{W}_{\mathrm{AC}1})-c(\mathcal{W}_{\mathrm{AC}2}))\|^2\\
&\qquad-\|\boldsymbol{D}+g_1c(\mathcal{W}_{\mathrm{AC}1})-g_1c(\mathcal{W}_{\mathrm{AC}2})\|^2)\Big)\Big] \qquad (104)
\end{aligned}$$

for all $\lambda_1>0$ and $\lambda_2>0$.

1) Expectation over $c(\mathcal{W}_{\mathrm{AC}1})$ and $c(\mathcal{W}_{\mathrm{AC}2})$ and the First Gallager-ρ Trick: The exponent in (104) can be written as

$$[c(\mathcal{W}_{\mathrm{AC}1})^{\mathrm{T}}\ c(\mathcal{W}_{\mathrm{AC}2})^{\mathrm{T}}](\mathbf{U}_{\mathrm{AC}}\otimes\mathbf{I}_{\mathrm{N}})\begin{bmatrix}c(\mathcal{W}_{\mathrm{AC}1})\\ c(\mathcal{W}_{\mathrm{AC}2})\end{bmatrix}+2\boldsymbol{V}_{\mathrm{AC}}^{\mathrm{T}}\begin{bmatrix}c(\mathcal{W}_{\mathrm{AC}1})\\ c(\mathcal{W}_{\mathrm{AC}2})\end{bmatrix}+R_{\mathrm{AC}} \tag{105}$$

where

$$\mathbf{U}_{\mathrm{AC}}=\begin{bmatrix}u_{\mathrm{AC}}^{(1)} & u_{\mathrm{AC}}^{(2)}\\ u_{\mathrm{AC}}^{(2)} & u_{\mathrm{AC}}^{(3)}\end{bmatrix}, \tag{106}$$

$$u_{\mathrm{AC}}^{(1)} = -\lambda_1 g_1^2 - \lambda_2 g_2(g_2 - 2g_1), \tag{107}$$
$$u_{\mathrm{AC}}^{(2)} = \lambda_1 g_1(g_1 - g_2) + \lambda_2 g_2(g_2 - 2g_1), \tag{108}$$
$$u_{\mathrm{AC}}^{(3)} = -\lambda_1 g_1(g_1 - 2g_2) - \lambda_2 g_2(g_2 - 2g_1), \tag{109}$$
$$\boldsymbol{V}_{\mathrm{AC}} = \begin{bmatrix} -\lambda_1 g_1 \boldsymbol{A} - \lambda_2(g_1 - g_2)\boldsymbol{C} + \lambda_2 g_1 \boldsymbol{D} \\ \lambda_1(g_1-g_2)\boldsymbol{A} + \lambda_1 g_2 \boldsymbol{B} + \lambda_2(g_1-g_2)\boldsymbol{C} - \lambda_2 g_1 \boldsymbol{D} \end{bmatrix}, \tag{110}$$
$$R_{\mathrm{AC}} = -\lambda_1(\|\boldsymbol{A}\|^2 - \|\boldsymbol{B}\|^2) - \lambda_2(\|\boldsymbol{C}\|^2 - \|\boldsymbol{D}\|^2). \tag{111}$$

Note that $\begin{bmatrix} c(\mathcal{W}_{\mathrm{AC1}}) \\ c(\mathcal{W}_{\mathrm{AC2}}) \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathrm{AC}} \otimes \mathbf{I}_{\mathrm{N}})$ with $\Sigma_{\mathrm{AC}} = \mathrm{diag}(t_{\mathrm{AC1}}\mathrm{P}', t_{\mathrm{AC2}}\mathrm{P}')$. We apply (67) in Lemma 4 to compute the expectation in (104) over $c(\mathcal{W}_{\mathrm{AC1}})$, $c(\mathcal{W}_{\mathrm{AC2}})$ as

$$\det(\mathbf{I}_2 - 2\Sigma_{\mathrm{AC}}\mathbf{U}_{\mathrm{AC}})^{-\mathrm{N}/2} \cdot \exp\left(2\boldsymbol{V}_{\mathrm{AC}}^{\mathrm{T}}[(\Sigma_{\mathrm{AC}}^{-1} - 2\mathbf{U}_{\mathrm{AC}})^{-1} \otimes \mathbf{I}_{\mathrm{N}}]\boldsymbol{V}_{\mathrm{AC}} + R_{\mathrm{AC}}\right) \tag{112}$$

where we require λ_1, λ_2 to satisfy that $\Sigma_{\mathrm{AC}}^{-1} \succ 2\mathbf{U}_{\mathrm{AC}}$.

Now, we apply Gallager's ρ-trick in Lemma 3 to get that, given $c(\mathcal{W}_{\mathrm{MD}\ell})$, $c(\mathcal{W}_{\mathrm{FP}\ell})$, $\ell \in \{1,2\}$, $c(\mathcal{W}_{\mathrm{C2}})$, and $\boldsymbol{Z}$, it holds for every $\rho_1 \in [0,1]$ that

$$\begin{aligned}
&\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\mathrm{AC1}} \subset \binom{[\mathrm{K_{a1}}]\setminus \mathcal{W}_{\mathrm{MD1}}}{t_{\mathrm{AC1}}}} \bigcup_{\mathcal{W}_{\mathrm{AC2}} \subset \binom{[\mathrm{K_{a1}}+1:\mathrm{K_a}]\setminus \mathcal{W}_{\mathrm{MD2}}}{t_{\mathrm{AC1}}+t_{\mathrm{MD1}}-t_{\mathrm{FP1}}}} F_{\mathrm{IC}}(\mathcal{W}_{\mathrm{MD}\ell}, \mathcal{W}_{\mathrm{FP}\ell}, \mathcal{W}_{\mathrm{AC}\ell}, \ell \in \{1,2\})\Bigg] \\
&\le \binom{\mathrm{K_{a1}} - t_{\mathrm{MD1}}}{t_{\mathrm{AC1}}}^{\rho_1} \binom{\mathrm{K_a} - \mathrm{K_{a1}} - t + t_{\mathrm{MD1}}}{t_{\mathrm{AC1}} + t_{\mathrm{MD1}} - t_{\mathrm{FP1}}}^{\rho_1} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\mathrm{AC}}\mathbf{U}_{\mathrm{AC}})^{-\mathrm{N}\rho_1/2} \\
&\quad \cdot \exp\left(2\rho_1\boldsymbol{V}_{\mathrm{AC}}^{\mathrm{T}}[(\Sigma_{\mathrm{AC}}^{-1} - 2\mathbf{U}_{\mathrm{AC}})^{-1} \otimes \mathbf{I}_{\mathrm{N}}]\boldsymbol{V}_{\mathrm{AC}} + \rho_1 R_{\mathrm{AC}}\right).
\end{aligned} \tag{113}$$

2) Expectation over $c(\mathcal{W}_{\mathrm{FP1}})$ and $c(\mathcal{W}_{\mathrm{FP2}})$ and the Second Gallager-ρ Trick: Next, we denote

$$\boldsymbol{E} = \boldsymbol{Z} + g_1 c(\mathcal{W}_{\mathrm{MD1}}) + g_2 c(\mathcal{W}_{\mathrm{MD2}}) + g_2 c(\mathcal{W}_{\mathrm{C2}}), \tag{114}$$
$$\boldsymbol{F} = \boldsymbol{Z} + g_1 c(\mathcal{W}_{\mathrm{MD1}}) + g_2 c(\mathcal{W}_{\mathrm{MD2}}), \tag{115}$$
$$\boldsymbol{G} = \boldsymbol{Z} + g_1 c(\mathcal{W}_{\mathrm{MD1}}). \tag{116}$$

It follows that $\boldsymbol{A} = \boldsymbol{E} - g_1 c(\mathcal{W}_{\mathrm{FP1}})$, $\boldsymbol{C} = \boldsymbol{F} - g_1 c(\mathcal{W}_{\mathrm{FP1}}) - g_2 c(\mathcal{W}_{\mathrm{FP2}})$, and $\boldsymbol{D} = \boldsymbol{G} - g_1 c(\mathcal{W}_{\mathrm{FP1}})$. We can express the exponent in (113) as

$$\begin{aligned}
&2\rho_1\boldsymbol{V}_{\mathrm{AC}}^{\mathrm{T}}[(\Sigma_{\mathrm{AC}}^{-1} - 2\mathbf{U}_{\mathrm{AC}})^{-1} \otimes \mathbf{I}_{\mathrm{N}}]\boldsymbol{V}_{\mathrm{AC}} + \rho_1 R_{\mathrm{AC}} \\
&= [c(\mathcal{W}_{\mathrm{FP1}})^{\mathrm{T}}\ c(\mathcal{W}_{\mathrm{FP2}})^{\mathrm{T}}](\mathbf{U}_{\mathrm{FP}} \otimes \mathbf{I}_{\mathrm{N}}) \begin{bmatrix} c(\mathcal{W}_{\mathrm{FP1}}) \\ c(\mathcal{W}_{\mathrm{FP2}}) \end{bmatrix} \\
&\quad + 2\boldsymbol{V}_{\mathrm{FP}}^{\mathrm{T}} \begin{bmatrix} c(\mathcal{W}_{\mathrm{FP1}}) \\ c(\mathcal{W}_{\mathrm{FP2}}) \end{bmatrix} + R_{\mathrm{FP}}
\end{aligned} \tag{117}$$

where

$$\mathbf{U}_{\mathrm{FP}} = 2\rho_1\Lambda_{\mathrm{FP}}\mathbf{Q}_{\mathrm{FP}} - \rho_1\begin{bmatrix} \lambda_1 g_1^2 & \lambda_2 g_1 g_2 \\ \lambda_2 g_1 g_2 & \lambda_2 g_2^2 \end{bmatrix}, \tag{118}$$
$$\begin{aligned}\boldsymbol{V}_{\mathrm{FP}} &= 2\rho_1[\Lambda_{\mathrm{FP}} \otimes \mathbf{I}_{\mathrm{N}}]\boldsymbol{Q}_{\mathrm{FP}} \\ &\quad + \rho_1\begin{bmatrix} \lambda_1 g_1\boldsymbol{E} + \lambda_2 g_1\boldsymbol{F} - \lambda_2 g_1\boldsymbol{G} \\ \lambda_2 g_2 \boldsymbol{F}, \end{bmatrix},\end{aligned} \tag{119}$$
$$\begin{aligned}R_{\mathrm{FP}} &= 2\rho_1\boldsymbol{Q}_{\mathrm{FP}}^{\mathrm{T}}\left[(\Sigma_{\mathrm{AC}}^{-1} - 2\mathbf{U}_{\mathrm{AC}})^{-1} \otimes \mathbf{I}_{\mathrm{N}}\right]\boldsymbol{Q}_{\mathrm{FP}} \\ &\quad - \rho_1\lambda_1(\|\boldsymbol{E}\|^2 - \|\boldsymbol{B}\|^2) - \rho_1\lambda_2(\|\boldsymbol{F}\|^2 - \|\boldsymbol{G}\|^2),\end{aligned} \tag{120}$$

with

$$\mathbf{Q}_{\mathrm{FP}} = \begin{bmatrix} \lambda_1 g_1^2 - \lambda_2 g_1 g_2 & \lambda_2 g_2(g_1 - g_2) \\ -\lambda_1 g_1(g_1 - g_2) + \lambda_2 g_1 g_2 & -\lambda_2 g_2(g_1 - g_2) \end{bmatrix}, \tag{121}$$
$$\boldsymbol{Q}_{\mathrm{FP}} = \begin{bmatrix} -\lambda_1 g_1\boldsymbol{E} - \lambda_2(g_1 - g_2)\boldsymbol{F} + \lambda_2 g_1\boldsymbol{G} \\ \lambda_1(g_1-g_2)\boldsymbol{E} + \lambda_1 g_2\boldsymbol{B} + \lambda_2(g_1-g_2)\boldsymbol{F} - \lambda_2 g_1\boldsymbol{G} \end{bmatrix}, \tag{122}$$
$$\Lambda_{\mathrm{FP}} = \mathbf{Q}_{\mathrm{FP}}^{\mathrm{T}}(\Sigma_{\mathrm{AC}}^{-1} - 2\mathbf{U}_{\mathrm{AC}})^{-1}. \tag{123}$$

Note that $\begin{bmatrix} c(\mathcal{W}_{\mathrm{FP1}}) \\ c(\mathcal{W}_{\mathrm{FP2}}) \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathrm{FP}} \otimes \mathbf{I}_{\mathrm{N}})$ with $\Sigma_{\mathrm{FP}} = \mathrm{diag}(t_{\mathrm{FP1}}\mathrm{P}', t_{\mathrm{FP2}}\mathrm{P}')$. By applying (67) in Lemma 4 to compute the expectation of (113) over $c(\mathcal{W}_{\mathrm{FP1}}), c(\mathcal{W}_{\mathrm{FP2}})$, we obtain that, given $c(\mathcal{W}_{\mathrm{MD}\ell})$, $\ell \in \{1,2\}$, $c(\mathcal{W}_{\mathrm{C2}})$, and $\boldsymbol{Z}$,

$$\begin{aligned}
&\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\mathrm{AC1}} \subset \binom{[\mathrm{K_{a1}}]\setminus \mathcal{W}_{\mathrm{MD1}}}{t_{\mathrm{AC1}}}} \bigcup_{\mathcal{W}_{\mathrm{AC2}} \subset \binom{[\mathrm{K_{a1}}+1:\mathrm{K_a}]\setminus \mathcal{W}_{\mathrm{MD2}}}{t_{\mathrm{AC1}}+t_{\mathrm{MD1}}-t_{\mathrm{FP1}}}} F_{\mathrm{IC}}(\mathcal{W}_{\mathrm{MD}\ell}, \mathcal{W}_{\mathrm{FP}\ell}, \mathcal{W}_{\mathrm{AC}\ell}, \ell \in \{1,2\})\Bigg] \\
&\le \binom{\mathrm{K_{a1}} - t_{\mathrm{MD1}}}{t_{\mathrm{AC1}}}^{\rho_1} \binom{\mathrm{K_a} - \mathrm{K_{a1}} - t + t_{\mathrm{MD1}}}{t_{\mathrm{AC1}} + t_{\mathrm{MD1}} - t_{\mathrm{FP1}}}^{\rho_1} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\mathrm{AC}}\mathbf{U}_{\mathrm{AC}})^{-\mathrm{N}\rho_1/2} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\mathrm{FP}}\mathbf{U}_{\mathrm{FP}})^{-\mathrm{N}/2} \\
&\quad \cdot \exp\left(2\boldsymbol{V}_{\mathrm{FP}}^{\mathrm{T}}[(\Sigma_{\mathrm{FP}}^{-1} - 2\mathbf{U}_{\mathrm{FP}})^{-1} \otimes \mathbf{I}_{\mathrm{N}}]\boldsymbol{V}_{\mathrm{FP}} + R_{\mathrm{FP}}\right)
\end{aligned} \tag{124}$$

under the condition that $\Sigma_{\mathrm{FP}}^{-1} \succ 2\mathbf{U}_{\mathrm{FP}}$.

Now, we apply Gallager's ρ-trick again to get that, given $c(\mathcal{W}_{\mathrm{MD}\ell})$, $\ell \in \{1,2\}$, $c(\mathcal{W}_{\mathrm{C2}})$, and $\boldsymbol{Z}$, and for every $\rho_2 \in [0,1]$, it holds that

$$\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\mathrm{FP1}} \subset \binom{[\mathrm{K_a}+1:\mathrm{M}]}{t_{\mathrm{FP1}}}} \bigcup_{\mathcal{W}_{\mathrm{FP2}} \subset \binom{[\mathrm{K_a}+1:\mathrm{M}]\setminus \mathcal{W}_{\mathrm{FP1}}}{t-t_{\mathrm{FP1}}}} \bigcup_{\mathcal{W}_{\mathrm{AC1}} \subset \binom{[\mathrm{K_{a1}}]\setminus \mathcal{W}_{\mathrm{MD1}}}{t_{\mathrm{AC1}}}} \bigcup_{\mathcal{W}_{\mathrm{AC2}} \subset \binom{[\mathrm{K_{a1}}+1:\mathrm{K_a}]\setminus \mathcal{W}_{\mathrm{MD2}}}{t_{\mathrm{AC1}}+t_{\mathrm{MD1}}-t_{\mathrm{FP1}}}} F_{\mathrm{IC}}(\mathcal{W}_{\mathrm{MD}\ell}, \mathcal{W}_{\mathrm{FP}\ell}, \mathcal{W}_{\mathrm{AC}\ell}, \ell \in \{1,2\})\Bigg] \tag{125}$$

$$\begin{aligned}
&\le \binom{M - \mathrm{K_a}}{t_{\mathrm{FP1}}}^{\rho_2} \binom{M - \mathrm{K_a} - t_{\mathrm{FP1}}}{t - t_{\mathrm{FP1}}}^{\rho_2} \\
&\quad \cdot \binom{\mathrm{K_{a1}} - t_{\mathrm{MD1}}}{t_{\mathrm{AC1}}}^{\rho_1\rho_2} \binom{\mathrm{K_a} - \mathrm{K_{a1}} - t + t_{\mathrm{MD1}}}{t_{\mathrm{AC1}} + t_{\mathrm{MD1}} - t_{\mathrm{FP1}}}^{\rho_1\rho_2} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\mathrm{AC}}\mathbf{U}_{\mathrm{AC}})^{-\mathrm{N}\rho_1\rho_2/2} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\mathrm{FP}}\mathbf{U}_{\mathrm{FP}})^{-\mathrm{N}\rho_2/2} \\
&\quad \cdot \exp\left(2\rho_2\boldsymbol{V}_{\mathrm{FP}}^{\mathrm{T}}[(\Sigma_{\mathrm{FP}}^{-1} - 2\mathbf{U}_{\mathrm{FP}})^{-1} \otimes \mathbf{I}_{\mathrm{N}}]\boldsymbol{V}_{\mathrm{FP}} + \rho_2 R_{\mathrm{FP}}\right).
\end{aligned} \tag{126}$$

3) Expectation over $c(\mathcal{W}_{\rm MD1})$ and $c(\mathcal{W}_{\rm MD2})$ and the Third Gallager-ρ Trick: We express the exponent in (126) as

$$\begin{aligned}
&2\rho_2 \boldsymbol{V}_{\rm FP}^{\rm T}[(\Sigma_{\rm FP}^{-1} - 2\mathbf{U}_{\rm FP})^{-1}\otimes \mathbf{I}_{\rm N}]\boldsymbol{V}_{\rm FP} + \rho_2 R_{\rm FP} \\
&= [c(\mathcal{W}_{\rm MD1})^{\rm T}\ c(\mathcal{W}_{\rm MD2})^{\rm T}](\mathbf{U}_{\rm MD}\otimes \mathbf{I}_{\rm N})\begin{bmatrix} c(\mathcal{W}_{\rm MD1}) \\ c(\mathcal{W}_{\rm MD2})\end{bmatrix} \quad (127)\\
&\quad + 2\boldsymbol{V}_{\rm MD}^{\rm T}\begin{bmatrix} c(\mathcal{W}_{\rm MD1}) \\ c(\mathcal{W}_{\rm MD2})\end{bmatrix} + R_{\rm MD} \quad (128)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{U}_{\rm MD} &= 2\rho_2\Lambda_{\rm MD}\mathbf{Q}_{\rm MD} + 2\rho_2\rho_1\Lambda_{\mathbf{M}}\mathbf{M} \\
&\quad - \rho_2\rho_1\begin{bmatrix} \lambda_1 g_1^2 & (\lambda_1+\lambda_2)g_1g_2 \\ (\lambda_1+\lambda_2)g_1g_2 & \lambda_2 g_2^2\end{bmatrix}, \quad (129)\\
\mathbf{M} &= \begin{bmatrix} -\lambda_1 g_1^2 + \lambda_2 g_1 g_2 & -\lambda_1 g_1 g_2 - \lambda_2 g_2(g_1-g_2) \\ \lambda_1 g_1(g_1-g_2) - \lambda_2 g_1 g_2 & \lambda_1 g_1 g_2 + \lambda_2 g_2 (g_1-g_2)\end{bmatrix}, \quad (130)\\
\mathbf{Q}_{\rm MD} &= 2\rho_1\Lambda_{\rm FP}\mathbf{M} + \rho_1\begin{bmatrix} \lambda_1 g_1^2 & (\lambda_1+\lambda_2)g_1g_2 \\ \lambda_2 g_1 g_2 & \lambda_2 g_2^2\end{bmatrix}, \quad (131)\\
\boldsymbol{V}_{\rm MD} &= 2\rho_2(\Lambda_{\rm MD}\otimes\mathbf{I}_{\rm N})\boldsymbol{Q}_{\rm MD} + 2\rho_2\rho_1(\Lambda_{\mathbf{M}}\otimes \mathbf{I}_{\rm N})\boldsymbol{M} \\
&\quad - \rho_2\rho_1\boldsymbol{R}, \quad (132)\\
\boldsymbol{Q}_{\rm MD} &= 2\rho_1(\Lambda_{\rm FP}\otimes\mathbf{I}_{\rm N})\boldsymbol{M} + \rho_1\boldsymbol{R}, \quad (133)\\
\boldsymbol{M} &= \left(\begin{bmatrix} 1 & -1 \\ -1 & 1\end{bmatrix}\otimes\mathbf{I}_{\rm N}\right)\boldsymbol{R}, \quad (134)\\
\boldsymbol{R} &= \left(\begin{bmatrix} \lambda_1 g_1 & \lambda_1 g_1 g_2 \\ \lambda_2 g_2 & 0\end{bmatrix}\otimes\mathbf{I}_{\rm N}\right)\begin{bmatrix} \boldsymbol{Z} \\ c(\mathcal{W}_{\rm C2})\end{bmatrix}, \quad (135)\\
R_{\rm MD} &= 2\rho_2\boldsymbol{Q}_{\rm MD}^{\rm T}[(\Sigma_{\rm FP}^{-1} - 2\mathbf{U}_{\rm FP})^{-1}]\boldsymbol{Q}_{\rm MD} \\
&\quad + 4\rho_1\rho_2\boldsymbol{M}^{\rm T}(\Sigma_{\rm AC}^{-1} - 2\mathbf{U}_{\rm AC})^{-1}\boldsymbol{M}, \quad (136)
\end{aligned}$$

with

$$\begin{aligned}
\Lambda_{\rm MD} &= \mathbf{Q}_{\rm MD}^{\rm T}(\Sigma_{\rm FP}^{-1} - 2\mathbf{U}_{\rm FP})^{-1}, \quad (137)\\
\Lambda_{\mathbf{M}} &= \mathbf{M}^{\rm T}(\Sigma_{\rm AC}^{-1} - 2\mathbf{U}_{\rm AC})^{-1}. \quad (138)
\end{aligned}$$

Note that $\begin{bmatrix} c(\mathcal{W}_{\rm MD1}) \\ c(\mathcal{W}_{\rm MD2})\end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\rm MD}\otimes\mathbf{I}_{\rm N})$ with $\Sigma_{\rm MD} = {\rm diag}(t_{\rm MD1}{\rm P}', t_{\rm MD2}{\rm P}')$. By applying (67) in Lemma 4 to compute the expectation of (126) over $c(\mathcal{W}_{\rm MD1}), c(\mathcal{W}_{\rm MD2})$, we obtain that, given $c(\mathcal{W}_{\rm C2})$ and $\boldsymbol{Z}$,

$$\begin{aligned}
&\mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm FP1}\subset\binom{[{\rm K_a}+1:{\rm M}]}{t_{\rm FP1}}}\ \bigcup_{\mathcal{W}_{\rm FP2}\subset\binom{[{\rm K_a}+1:{\rm M}]\setminus\mathcal{W}_{\rm FP1}}{t-t_{\rm FP1}}} \\
&\qquad \bigcup_{\mathcal{W}_{\rm AC1}\subset\binom{[{\rm K_{a1}}]\setminus\mathcal{W}_{\rm MD1}}{t_{\rm AC1}}}\ \bigcup_{\mathcal{W}_{\rm AC2}\subset\binom{[{\rm K_{a1}}+1:{\rm K_a}]\setminus\mathcal{W}_{\rm MD2}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}} \\
&\qquad F_{\rm IC}(\mathcal{W}_{{\rm MD}\ell}, \mathcal{W}_{{\rm FP}\ell}, \mathcal{W}_{{\rm AC}\ell}, \ell\in\{1,2\})\Bigg] \\
&\le \binom{{\rm K_{a1}} - t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1}\binom{{\rm K_a} - {\rm K_{a1}} - t + t_{\rm MD1}}{t_{\rm AC1} + t_{\rm MD1} - t_{\rm FP1}}^{\rho_1} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm AC}\mathbf{U}_{\rm AC})^{-{\rm N}\rho_1\rho_2/2} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm FP}\mathbf{U}_{\rm FP})^{-{\rm N}\rho_2/2} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm MD}\mathbf{U}_{\rm MD})^{-{\rm N}/2} \\
&\quad \cdot \exp\left(2\boldsymbol{V}_{\rm MD}^{\rm T}[(\Sigma_{\rm MD}^{-1} - 2\mathbf{U}_{\rm MD})^{-1}\otimes\mathbf{I}_{\rm N}]\boldsymbol{V}_{\rm MD} + R_{\rm MD}\right) \quad (139)
\end{aligned}$$

under the condition that $\Sigma_{\rm FP}^{-1} \succ 2\mathbf{U}_{\rm FP}$.

We now apply Gallager's ρ-trick the third time to obtain that, given $c(\mathcal{W}_{\rm C2})$ and $\boldsymbol{Z}$, and for every $\rho_3\in[0,1]$,

$$\begin{aligned}
&= \mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm MD1}\subset\binom{[{\rm K_{a1}}]}{t_{\rm MD1}}}\ \bigcup_{\mathcal{W}_{\rm MD2}\subset\binom{[{\rm K_{a1}}+1:{\rm K_a}]}{t-t_{\rm MD1}}} \\
&\qquad \bigcup_{\mathcal{W}_{\rm FP1}\subset\binom{[{\rm K_a}+1:{\rm M}]}{t_{\rm FP1}}}\ \bigcup_{\mathcal{W}_{\rm FP2}\subset\binom{[{\rm K_a}+1:{\rm M}]\setminus\mathcal{W}_{\rm FP1}}{t-t_{\rm FP1}}} \\
&\qquad \bigcup_{\mathcal{W}_{\rm AC1}\subset\binom{[{\rm K_{a1}}]\setminus\mathcal{W}_{\rm MD1}}{t_{\rm AC1}}}\ \bigcup_{\mathcal{W}_{\rm AC2}\subset\binom{[{\rm K_{a1}}+1:{\rm K_a}]\setminus\mathcal{W}_{\rm MD2}}{t_{\rm AC1}+t_{\rm MD1}-t_{\rm FP1}}} \\
&\qquad F_{\rm IC}(\mathcal{W}_{{\rm MD}\ell}, \mathcal{W}_{{\rm FP}\ell}, \mathcal{W}_{{\rm AC}\ell}, \ell\in\{1,2\})\Bigg] \quad (140)\\
&\le \binom{{\rm K_{a1}}}{t_{\rm MD1}}^{\rho_3}\binom{{\rm K_{a2}}}{t - t_{\rm MD1}}^{\rho_3} \\
&\quad \cdot \binom{M - {\rm K_a}}{t_{\rm FP1}}^{\rho_2\rho_3}\binom{M - {\rm K_a} - t_{\rm FP1}}{t - t_{\rm FP1}}^{\rho_2\rho_3} \\
&\quad \cdot \binom{{\rm K_{a1}} - t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1\rho_2\rho_3}\binom{{\rm K_a} - {\rm K_{a1}} - t + t_{\rm MD1}}{t_{\rm AC1} + t_{\rm MD1} - t_{\rm FP1}}^{\rho_1\rho_2\rho_3} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm AC}\mathbf{U}_{\rm AC})^{-{\rm N}\rho_1\rho_2\rho_3/2} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm FP}\mathbf{U}_{\rm FP})^{-{\rm N}\rho_2\rho_3/2} \\
&\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm MD}\mathbf{U}_{\rm MD})^{-{\rm N}\rho_3/2} \\
&\quad \cdot \exp\left(2\rho_3\boldsymbol{V}_{\rm MD}^{\rm T}[(\Sigma_{\rm MD}^{-1} - 2\mathbf{U}_{\rm MD})^{-1}\otimes\mathbf{I}_{\rm N}]\boldsymbol{V}_{\rm MD} + \rho_3 R_{\rm MD}\right). \quad (141)
\end{aligned}$$

4) Expectation over $\boldsymbol{Z}$ and $c(\mathcal{W}_{\rm C2})$: We express the exponent in (141) as

$$\begin{aligned}
&2\rho_3\boldsymbol{V}_{\rm MD}^{\rm T}[(\Sigma_{\rm MD}^{-1} - 2\mathbf{U}_{\rm MD})^{-1}\otimes\mathbf{I}_{\rm N}]\boldsymbol{V}_{\rm MD} + \rho_3 R_{\rm MD} \\
&= [\boldsymbol{Z}^{\rm T}\ c(\mathcal{W}_{\rm C2})^{\rm T}](\mathbf{U}_{\rm ZC}\otimes\mathbf{I}_{\rm N})\begin{bmatrix} \boldsymbol{Z} \\ c(\mathcal{W}_{\rm C2})\end{bmatrix} \quad (142)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{U}_{\rm ZC} &= 2\rho_3\mathbf{Q}_{\rm ZC}^{\rm T}(\Sigma_{\rm MD}^{-1} - 2\mathbf{U}_{\rm MD})^{-1}\mathbf{Q}_{\rm ZC} \\
&\quad + 2\rho_3\rho_2\Lambda_{\rm ZCa}^{\rm T}(\Sigma_{\rm FP}^{-1} - 2\mathbf{U}_{\rm FP})^{-1}\Lambda_{\rm ZCa} \\
&\quad + 4\rho_3\rho_2\rho_1\Lambda_{\rm ZCb}^{\rm T}(\Sigma_{\rm AC}^{-1} - 2\mathbf{U}_{\rm AC})^{-1}\Lambda_{\rm ZCb} \quad (143)
\end{aligned}$$

with

$$\begin{aligned}
\mathbf{Q}_{\rm ZC} &= \rho_2\rho_1\left(-\mathbf{I}_2 + (4\Lambda_{\rm MD}\Lambda_{\rm FP} + 2\Lambda_{\mathbf{M}})\begin{bmatrix} -1 & 1 \\ 1 & -1\end{bmatrix} + 2\Lambda_{\rm MD}\right) \\
&\quad \cdot \begin{bmatrix} \lambda_1 g_1 & \lambda_1 g_1 g_2 \\ \lambda_2 g_2 & 0\end{bmatrix}, \quad (144)\\
\Lambda_{\rm ZCa} &= \rho_1\left(\mathbf{I}_2 + 2\Lambda_{\rm FP}\begin{bmatrix} -1 & 1 \\ 1 & -1\end{bmatrix}\right)\begin{bmatrix} \lambda_1 g_1 & \lambda_1 g_1 g_2 \\ \lambda_2 g_2 & 0\end{bmatrix}, \quad (145)\\
\Lambda_{\rm ZCb} &= \begin{bmatrix} -1 & 1 \\ 1 & -1\end{bmatrix}\begin{bmatrix} \lambda_1 g_1 & \lambda_1 g_1 g_2 \\ \lambda_2 g_2 & 0\end{bmatrix}. \quad (146)
\end{aligned}$$

Note that $\begin{bmatrix} \boldsymbol{Z} \\ c(\mathcal{W}_{\rm C2}) \end{bmatrix} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\rm ZC} \otimes \mathbf{I}_{\rm N})$ with $\Sigma_{\rm ZC} = \operatorname{diag}(1, (\mathrm{K}_{\rm a2} - t_{\rm MD2} - t_{\rm AC2})\mathrm{P}')$. By applying (67) in Lemma 4 to compute the expectation of (126) over $\boldsymbol{Z}$ and $c(\mathcal{W}_{\rm C2})$, we obtain that

$$= \mathbb{P}\Bigg[\bigcup_{\mathcal{W}_{\rm MD1} \subset \binom{[\mathrm{K}_{\rm a1}]}{t_{\rm MD1}}} \bigcup_{\mathcal{W}_{\rm MD2} \subset \binom{[\mathrm{K}_{\rm a1}+1:\mathrm{K}_{\rm a}]}{t - t_{\rm MD1}}} \bigcup_{\mathcal{W}_{\rm FP1} \subset \binom{[\mathrm{K}_{\rm a}+1:\mathrm{M}]}{t_{\rm FP1}}} \bigcup_{\mathcal{W}_{\rm FP2} \subset \binom{[\mathrm{K}_{\rm a}+1:\mathrm{M}] \setminus \mathcal{W}_{\rm FP1}}{t - t_{\rm FP1}}} \bigcup_{\mathcal{W}_{\rm AC1} \subset \binom{[\mathrm{K}_{\rm a1}] \setminus \mathcal{W}_{\rm MD1}}{t_{\rm AC1}}} \bigcup_{\mathcal{W}_{\rm AC2} \subset \binom{[\mathrm{K}_{\rm a1}+1:\mathrm{K}_{\rm a}] \setminus \mathcal{W}_{\rm MD2}}{t_{\rm AC1} + t_{\rm MD1} - t_{\rm FP1}}} F_{\rm IC}(\mathcal{W}_{{\rm MD}\ell}, \mathcal{W}_{{\rm FP}\ell}, \mathcal{W}_{{\rm AC}\ell}, \ell \in \{1,2\}) \Bigg] \tag{147}$$

$$\begin{aligned} &\le \binom{\mathrm{K}_{\rm a1}}{t_{\rm MD1}}^{\rho_3} \binom{\mathrm{K}_{\rm a2}}{t - t_{\rm MD1}}^{\rho_3} \\ &\quad \cdot \binom{M - \mathrm{K}_{\rm a}}{t_{\rm FP1}}^{\rho_2\rho_3} \binom{M - \mathrm{K}_{\rm a} - t_{\rm FP1}}{t - t_{\rm FP1}}^{\rho_2\rho_3} \\ &\quad \cdot \binom{\mathrm{K}_{\rm a1} - t_{\rm MD1}}{t_{\rm AC1}}^{\rho_1\rho_2\rho_3} \binom{\mathrm{K}_{\rm a} - \mathrm{K}_{\rm a1} - t + t_{\rm MD1}}{t_{\rm AC1} + t_{\rm MD1} - t_{\rm FP1}}^{\rho_1\rho_2\rho_3} \\ &\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm AC}\mathbf{U}_{\rm AC})^{-\mathrm{N}\rho_1\rho_2\rho_3/2} \\ &\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm FP}\mathbf{U}_{\rm FP})^{-\mathrm{N}\rho_2\rho_3/2} \\ &\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm MD}\mathbf{U}_{\rm MD})^{-\mathrm{N}\rho_3/2} \\ &\quad \cdot \det(\mathbf{I}_2 - 2\Sigma_{\rm ZC}\mathbf{U}_{\rm ZC})^{-\mathrm{N}/2}, \end{aligned} \tag{148}$$

which is an upper-bound for $\mathbb{P}\,|\mathcal{W}_{\rm MD}| = |\mathcal{W}_{\rm FP}| = t$. Finally, by substituting this bound into (76) and rearranging the terms, we complete the proof.

Appendix E
Mathematical Preliminaries for Replica Analysis

We will make use of the definitions and results presented in this section for the proofs of Claim 1 and Claim 2. We will consider a generic channel model

$$\tilde{\boldsymbol{Y}} = \tilde{\mathscr{C}} \tilde{\boldsymbol{U}}^{(0)} + \tilde{\boldsymbol{Z}} \tag{149}$$

where $\tilde{\boldsymbol{Y}}, \tilde{\boldsymbol{Z}} \in \mathbb{R}^{\rm N}$, $\tilde{\boldsymbol{U}}^{(0)}$ is a random (column) vector of length m', $\tilde{\mathscr{C}} \in \mathbb{R}^{\mathrm{N} \times m'}$, $\tilde{\mathscr{C}}$ has i.i.d. $\mathcal{N}(0, 1/\mathrm{N})$ entries, $\tilde{\boldsymbol{Z}} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{\rm N})$, and will specialize the value of m' along with the definition of $\tilde{\boldsymbol{U}}^{(0)}$, to prove each claim. We fix $\beta = m'/\mathrm{N}$.

First, we introduce the definition of the partition function.

Definition 2 (Partition Function)**.** *Fix positive integers ℓ and r, and a real number t. Let $\{\tilde{\boldsymbol{U}}^{(i)}, i \ge 1\}$ be an i.i.d. process, where $\tilde{\boldsymbol{U}}^{(i)} \overset{D}{=} \tilde{\boldsymbol{U}}^{(0)}$. Then, the partition function is defined as*

$$\zeta^{(r)}(t; \tilde{\boldsymbol{Y}}, \tilde{\mathscr{C}}) = \mathbb{E}\left[e^{t \prod_{i=1}^{\ell} [\tilde{\boldsymbol{U}}^{(i)}]_1} \prod_{j=1}^{r} \mathbb{P}\Big[\tilde{\boldsymbol{Y}} \mid \tilde{\boldsymbol{U}}^{(j)}, \tilde{\mathscr{C}}\Big] \,\middle|\, \tilde{\boldsymbol{Y}}, \tilde{\mathscr{C}} \right]. \tag{150}$$

We have the following result.

Assumption 1 (Replica Trick for Moment Calculation)**.** *Let*

$$\kappa(t, r) = \log \mathbb{E}_{\tilde{\boldsymbol{Y}}, \tilde{\mathscr{C}}}\Big[\zeta^{(r)}(t; \tilde{\boldsymbol{Y}}, \tilde{\mathscr{C}})\Big] \tag{151}$$

denote the cumulant of the partition function in (150)*. Then, the ℓ^{th} moment of the random variable*

$$V_{\rm N} = \mathbb{P}\Big[[\tilde{\boldsymbol{U}}^{(0)}]_1 = 1 | \tilde{\boldsymbol{Y}}, \tilde{\mathscr{C}}\Big] \tag{152}$$

is given by

$$\mathbb{E}_{\tilde{\boldsymbol{Y}}, \tilde{\mathscr{C}}}\big[V_{\rm N}^{\ell}\big] = \lim_{r \to 0} \lim_{t \to 0} \frac{\mathrm{d}}{\mathrm{d}t} \kappa(t, r). \tag{153}$$

Remark 3. *In applying the replica trick in Assumption 1, we will compute in closed form $\mathbb{E}[\zeta^{(r)}]$ for $r \in \mathbb{N}$. We will then* assume *that the result holds for every $r \in \mathbb{R}$. This is a common practice in the application of the replica method (see [12] for such an application in the context of multiuser communication). A rigorous justification involves finding a unique analytic continuation of $\mathbb{E}[\zeta^{(r)}]$ from the set of positive integers to the entire real line; see, for instance, the discussion in [19].*

Lemma 5. *(Gaussian Integral Identity) Fix a positive integer r. For $i \in [0:r]$, define*

$$W_i = \sqrt{\frac{\mathrm{P}}{\mathrm{MN}}} \sum_{m=1}^{\mathrm{M}} [\tilde{\mathscr{C}}]_m [\tilde{\boldsymbol{U}}^{(i)}]_m. \tag{154}$$

Then,

$$\int \mathbb{E}_{\tilde{\mathscr{C}}}\left[\prod_{j=0}^{r} e^{-\frac{(y - \sqrt{\beta} W_j)^2}{2}} \right] \frac{\mathrm{d}y}{\sqrt{2\pi}} = e^{G^{(r)}(\mathscr{Q})} \tag{155}$$

where

$$\begin{aligned} G^{(r)}(\mathscr{Q}) = &-\frac{1}{2} \log\det(\mathbf{I}_{r+1} + \Sigma\mathscr{Q}) - \frac{1}{2}\log(1+r) \\ &- \frac{r}{2}\log(2\pi). \end{aligned} \tag{156}$$

Here, Σ and $\mathscr{Q}$ are $(r+1) \times (r+1)$ matrices with the (i,j)th entry given by $\beta\Big(\mathbb{1}\{i = j\} - \frac{1}{r+1}\Big)$ and $\mathbb{E}[W_i W_j]$, respectively.

The following lemma is used to evaluate the limiting value of the logarithm of the moment-generating function that arises in the analysis.

Lemma 6. *(Varadhan's Integral Lemma, Theorem 4.3.1 [20]) Suppose that measures $\{\mu_n, n \ge 1\}$ satisfy the large deviation principle with a good rate function $I : \mathcal{X} \mapsto [0, \infty]$ (see [20, Section 1.2]), and let $\phi : \mathcal{X} \mapsto \mathbb{R}$ be a continuous function, where $\mathcal{X}$ is a regular, topological space. Assume that, for some $\gamma > 1$,*

$$\limsup_{n \to \infty} \frac{\log \mathbb{E}\left[e^{\frac{\gamma \phi(Z_n)}{n}}\right]}{n} < \infty. \tag{157}$$

Then,

$$\limsup_{n\to\infty} \frac{\log \mathbb{E}\left[e^{\frac{\gamma\phi(Z_n)}{n}}\right]}{n} = \sup_{x\in\mathcal{X}}(\phi(x) - I(x)). \tag{158}$$

Lemma 7 (Moment convergence implies convergence in distribution [21, Theorem 30.2])**.** *Let $\{X_n\}_{n\geq 1}$ be a sequence of real-valued random variables and X a real-valued random variable. Suppose that:*

1) *For every integer $k \geq 1$, the kth moment $\mathbb{E}[X_n^k]$ exists for all sufficiently large n.*
2) *For every integer $k \geq 1$,*
$$\lim_{n\to\infty} \mathbb{E}[X_n^k] = \mathbb{E}[X^k].$$
3) *The distribution of X is uniquely determined by its moments (equivalently, Carleman's condition [21, Theorem 30.1] is satisfied).*

Then X_n converges in distribution to X, i.e.,

$$X_n \xrightarrow{D} X.$$

Appendix F
Proof of Claim 1

The proof follows the line of treatment outlined in Appendix G. The key difference involves modifying the definition of W_i in Lemma 5 to

$$\sqrt{\frac{\mathrm{P}g_1^2}{\mathrm{MN}}}\sum_{m=1}^{\mathrm{M}}[\tilde{\mathscr{C}}^{(1)}]_m[\tilde{\boldsymbol{U}}_1^{(i)}]_m + \sqrt{\frac{\mathrm{P}g_2^2}{\mathrm{MN}}}\sum_{m=\mathrm{M}+1}^{2\mathrm{M}}[\tilde{\mathscr{C}}^{(2)}]_m[\tilde{\boldsymbol{U}}_2^{(i)}]_m \tag{159}$$

in conducting the central limit theorem analysis. Here, $\tilde{\mathscr{C}}^{(i)} \overset{D}{=} \tilde{\mathscr{C}}$, $\tilde{\mathscr{C}}^{(1)}$ and $\tilde{\mathscr{C}}^{(2)}$ are independent, and $\tilde{\boldsymbol{U}}_\ell^{(i)}$ has i.i.d. $\mathrm{Ber}(\alpha_\ell\mu/\beta)$ entries. The rest of the proof is handled as in the proof of Proposition 1 in [14], which is built along the lines outlined in Appendix G.

Appendix G
Proof of Claim 2

To establish the convergence in distribution of the random variables V_{N}, by Lemma 7, it suffices to show that the moments converge, i.e., $\mathbb{E}[V_{\mathrm{N}}^\ell] \to \mathbb{E}[V^\ell]$ for all positive integers ℓ. To compute these moments, we invoke the replica method stated in Assumption 1.

The expectation of the partition function can be expressed as

$$\begin{aligned}&\mathbb{E}_{\tilde{\boldsymbol{Y}},\tilde{\mathcal{C}}}\left[\zeta^{(r)}(t;\tilde{\boldsymbol{Y}},\tilde{\mathcal{C}})\right]\\ &= \mathbb{E}\left[\int e^{t\prod_{i=1}^{\ell}[\tilde{\boldsymbol{U}}^{(i)}]_1}\Pr\left[\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{U}},\tilde{\mathcal{C}}\right]\prod_{j=1}^{r}\Pr\left[\tilde{\boldsymbol{y}}|\tilde{\boldsymbol{U}}^{(j)},\tilde{\mathcal{C}}\right]\mathrm{d}\tilde{\mathbf{y}}\right] \end{aligned}\tag{160}$$

$$= \mathbb{E}\left[\int e^{t\prod_{i=1}^{\ell}[\tilde{\boldsymbol{U}}^{(i)}]_1}(2\pi)^{-\frac{n(r+1)}{2}}\prod_{j=0}^{r}e^{-\frac{\|\tilde{\boldsymbol{y}}-\tilde{\mathcal{C}}\tilde{\boldsymbol{U}}_j\|^2}{2}}\mathrm{d}\tilde{\boldsymbol{y}}\right]. \tag{161}$$

We observe that, since the channel is memoryless, and the codebook $\tilde{\mathcal{C}}$ is i.i.d., the integrals over each of the coordinate of $\tilde{\mathbf{y}}$ are identical. Using this observation, we can express (161) as follows:

$$\begin{aligned}&\mathbb{E}\left[\zeta^{(r)}(t;\tilde{\boldsymbol{Y}},\tilde{\mathcal{C}})\right] =\\ &\mathbb{E}\left[e^{t\prod_{i=1}^{\ell}\tilde{\boldsymbol{U}}_1^{(i)}}\left\{(2\pi)^{-\frac{r}{2}}\int\mathbb{E}_{\tilde{\mathcal{C}}}\left[\prod_{j=0}^{r}e^{-\frac{\left(\tilde{y}-\sqrt{\mathrm{P}}\tilde{c}^T\tilde{\boldsymbol{U}}^{(i)}\right)^2}{2}}\right]\frac{\mathrm{d}\tilde{y}}{\sqrt{2\pi}}\right\}^{\mathrm{N}}\right].\end{aligned}\tag{162}$$

We define W_i as in Lemma 5. Using the central limit theorem, Edgeworth expansion as argued in [12, Appendix B], and invoking Lemma 5, we conclude that

$$\mathbb{E}\left[\zeta^{(r)}(t;\tilde{\boldsymbol{Y}},\tilde{\mathscr{C}})\right] = \mathbb{E}\left[e^{t\prod_{i=1}^{\ell}[\tilde{\boldsymbol{U}}]_1^{(i)}+G^{(r)}(\mathscr{Q})+O(M^{-1})}\right] \tag{163}$$

where $G^{(r)}(\mathscr{Q})$ is as defined in (156). We can further use Lemma 6 to obtain

$$\lim_{\mathrm{M}\to\infty}\log\mathbb{E}\left[\zeta^{(r)}(t;\tilde{\boldsymbol{Y}},\tilde{\mathscr{C}})\right] = \sup_{\mathscr{Q}}\left[\frac{1}{\beta}G^{(r)}(\mathscr{Q}) - I^{(r)}(t;\mathscr{Q})\right] \tag{164}$$

where

$$I^{(r)}(t;\mathscr{Q}) = \sup_{\tilde{\mathscr{Q}}}\left[\mathrm{Tr}(\tilde{\mathscr{Q}}\mathscr{Q}) - \log M^{(r)}(\tilde{\mathscr{Q}})\right] \tag{165}$$

is the rate function corresponding to the measure

$$\mathbb{E}\left[\prod_{i=0}^{r}\prod_{j\leq i}\delta\left(\sum_{m=1}^{\mathrm{M}}[\tilde{\boldsymbol{U}}^{(i)}]_m[\tilde{\boldsymbol{U}}^{(j)}]_m - \frac{\mathrm{M}\mathscr{Q}}{\mathrm{P}}\right)e^{t\prod_{k=1}^{\ell}[\tilde{\boldsymbol{U}}^{(k)}]_1}\right] \tag{166}$$

and $\mathscr{Q}$ is defined in Lemma 5. Also,

$$M^{(r)}(\tilde{\mathscr{Q}}) = \mathbb{E}\left[\exp\left(t\prod_{i=1}^{\ell}[\tilde{\boldsymbol{U}}^{(i)}]_1\right)e^{\mathscr{U}'^{\top}\tilde{\mathscr{Q}}\mathscr{U}'}\right] \tag{167}$$

where $\mathscr{U}' = [[\tilde{\boldsymbol{U}}^{(0)}]_1,\ldots,[\tilde{\boldsymbol{U}}^{(r)}]_1]^\top$.

We next substitute (165) to (164), which yield the following optimization problem.

$$\sup_{\mathscr{Q}}\inf_{\tilde{\mathscr{Q}}}\left[\frac{1}{\beta}G^{(r)}(\mathscr{Q}) - \mathrm{Tr}(\tilde{\mathscr{Q}}\mathscr{Q}) + \log M^{(r)}(\tilde{\mathscr{Q}})\right]. \tag{168}$$

Let

$$\begin{aligned}T^{(r)}(\mathscr{Q},\tilde{\mathscr{Q}}) = &-\frac{1}{2\beta}\log\det(\mathbf{I}_r + \Sigma\mathscr{Q}) - \mathrm{Tr}(\tilde{\mathscr{Q}}\mathscr{Q})\\ &+\log\mathbb{E}\left[e^{\mathscr{U}'^T\tilde{\mathscr{Q}}\mathscr{U}'}\right] - \frac{1}{2\beta}\log(1+r)\\ &-\frac{r}{2\beta}\log(2\pi).\end{aligned}\tag{169}$$

Fix a $\mathscr{Q}$. Then, the infimum with respect to $\tilde{\mathscr{Q}}$ satisfies

$$\mathscr{Q} = \frac{\mathbb{E}\left[\mathscr{U}'\mathscr{U}'^{\top}e^{\mathscr{U}'^{\top}\tilde{\mathscr{Q}}\mathscr{U}'}\right]}{\mathbb{E}\left[e^{\mathscr{U}'^{\top}\tilde{\mathscr{Q}}\mathscr{U}'}\right]}. \tag{170}$$

Let the solution of (170) be $\tilde{\mathscr{Q}}^*(\mathscr{Q})$. By defining the expectation with respect to the tilted measure

$$\mu(\tilde{\mathscr{Q}}) = \frac{e^{\mathscr{U}'^\top \tilde{\mathscr{Q}} \mathscr{U}'}}{\mathbb{E}\left[e^{\mathscr{U}'^\top \tilde{\mathscr{Q}} \mathscr{U}'}\right]} \tag{171}$$

we can rewrite (170) as

$$\mathscr{Q} = \mathbb{E}\left[\mathscr{U}'\mathscr{U}'^\top \middle| \tilde{\mathscr{Q}}\right]. \tag{172}$$

Now, with $\tilde{\mathscr{Q}} = \tilde{\mathscr{Q}}^*(\mathscr{Q})$, the supremum of the function $T^{(r)}(\mathscr{Q}, \tilde{\mathscr{Q}}^*(\mathscr{Q}))$ satisfies

$$\tilde{\mathscr{Q}} = -\frac{(\mathbf{I}_r + \Sigma\mathscr{Q})^{-1}\Sigma}{\beta}. \tag{173}$$

Equations (170) and (173) together form saddle point equations.

Now, we assume that the solutions of the saddle point equations satisfy some symmetry assumption. Such a symmetry assumption is referred to as the *replica symmetry assumption*. The work of [13], among others, make such a symmetry assumption to solve the saddle point equations. Specifically, in our case, we will assume that the joint solutions to (170) and (173) have the following structure:

$$\mathscr{Q}^* = \begin{bmatrix} p & q & q & \cdots & q \\ q & p & q & \cdots & q \\ q & q & p & \cdots & q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ q & q & q & \cdots & p \end{bmatrix}_{(r+1)\times(r+1)} \tag{174}$$

$$\tilde{\mathscr{Q}}^* = \begin{bmatrix} g & f & f & \cdots & f \\ f & g & f & \cdots & f \\ f & f & g & \cdots & f \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f & f & f & \cdots & g \end{bmatrix}_{(r+1)\times(r+1)}. \tag{175}$$

Under replica symmetry assumption,

$$G^{(r)}(\mathscr{Q}^*) = -\frac{r}{2}\log(2\pi) - \frac{r-1}{2}\log\left(1+\beta(p-q)\right) - \frac{1}{2}\log\left(1+\beta(p-q)+r(p-q)\right). \tag{176}$$

Next, we evaluate the moment generating function $M^{(r)}(\tilde{\mathscr{Q}}^*)$ as

$$M^{(r)}(\tilde{\mathscr{Q}}^*) = \mathbb{E}\left[e^{\left(\sum_{i=0}^r [\tilde{\boldsymbol{U}}]_i\right)^2 f + (g-f)\sum_{i=0}^r [\tilde{\boldsymbol{U}}]_i^2}\right] \tag{177}$$

which can be expressed as

$$M^{(r)}(\tilde{\mathscr{Q}}^*) = \mathbb{E}\left[\sqrt{\frac{f}{\pi}}\int e^{-fu^2 + (2f\sum_{i=0}^r [\tilde{\boldsymbol{U}}]_i)u + (g-f)\sum_{i=0}^r [\tilde{\boldsymbol{U}}]_i^2} \mathrm{d}u\right]. \tag{178}$$

Now, the rate function admits the form

$$I^{(r)}(\mathscr{Q}^*) = (r+1)(pg+rqf) - \log\mathbb{E}\left[\int\sqrt{\frac{f}{\pi}}\mathbb{E}\left[e^{-f(-(u-[\boldsymbol{U}]_0)^2 + g[\boldsymbol{U}]_0^2)}\right] \cdot \left(\mathbb{E}\left[e^{2f[\boldsymbol{U}]_1 + (g-f)[\boldsymbol{U}]_1^2}\right]\right)^r \mathrm{d}u\right] \tag{179}$$

Next, using the replica trick, we obtain

$$-\lim_{\mathrm{M}\to\infty}\mathbb{E}\left[\frac{1}{\mathrm{M}}\log\zeta^{(r)}(t;\tilde{\boldsymbol{Y}},\tilde{\mathcal{C}})\right] = -\lim_{\mathrm{M}\to\infty}\frac{1}{\mathrm{M}}\lim_{r\to 0}\frac{\mathrm{d}}{\mathrm{d}t}\left[\beta^{-1}G^{(r)}(\mathscr{Q}^*) - I^{(r)}(t;\tilde{\mathscr{Q}}^*)\right]\Bigg|_{t=0}. \tag{180}$$

Using (173), we conclude that

$$f = \frac{1}{2(1+\beta(p-q))} \tag{181}$$

and

$$g = 0. \tag{182}$$

To establish convergence to the limiting random variable, we next express the parameters above in terms of the corresponding conditional expectations. This analysis follows the approach outlined in [13], leading to relationships analogous to those in [13, eq. (132)–(135)]. Carleman's condition for moment determinacy [21, Chapter 30, Theorem 30.1], are satisfied as the random variables involved are bounded in $[0,1]$. These together prove the convergence in distribution to a limiting random variable. The expression for the limiting random variable involves calculation of the parameter by solving the final resultant saddle point equation

$$\frac{1}{\eta} = 1 + \mathrm{P}\beta\mathbb{E}\left[\left([\tilde{\boldsymbol{U}}]_1 - \mathbb{E}\left[[\tilde{\boldsymbol{U}}]_1 | \sqrt{\mathrm{P}}[\tilde{\boldsymbol{U}}]_1 + \frac{1}{\sqrt{\eta}}N\right]\right)^2\right] \tag{183}$$

with $N \sim \mathcal{N}(0,1)$, and $\eta = d^2/f$.