Ravi, J.; Koch, T. A. Scaling laws for many-access channels and unsourced random access. In: *2021 55th Asilomar Conference on Signals, Systems, and Computers, October 31-November 3, 2021 Pacific Grove, California, USA. IEEE, 2022, Pp. 1482-1487*

DOI: 10.1109/IEEECONF53345.2021.9723116

Scaling Laws for Many-Access Channels and Unsourced Random Access

Jithin Ravi[†‡] and Tobias Koch[†‡]

[†]Signal Theory and Communications Department, Universidad Carlos III de Madrid, 28911, Leganés, Spain
[‡]Gregorio Marañón Health Research Institute, 28007, Madrid, Spain.
Emails: {rjithin,koch}@tsc.uc3m.es

***Abstract*—In the emerging Internet of Things, a massive number of devices may connect to one common receiver. Consequently, models that study this setting are variants of the classical multiple-access channel where the number of users grows with the blocklength. Roughly, these models can be classified into three groups based on two criteria: the notion of probability of error and whether users use the same codebook. The first group follows the classical notion of probability of error and assumes that users use different codebooks. In the second group, users use different codebooks, but a new notion of probability of error called per-user probability of error is considered. The third group considers the per-user probability of error and that users are restricted to use the same codebook. This group is also known as *unsourced random access*. For the first and second groups of models, scaling laws that describe the capacity per unit-energy as a function of the order of growth of users were characterized by Ravi and Koch (arxiv.org/abs/2012.10350). In this paper, we first review these results. We then present scaling laws for the third group of models, i.e., unsourced random access.**

I. Introduction

In the emerging Internet of Things (IoT), a massive number of devices may connect to one common receiver. In such a scenario, the number of transmitters may be comparable to or even larger than the blocklength. Many transmitters communicating with a single receiver are studied in information theory under the name of multiple-access channel (MAC). In the classical MAC setting, the number of transmitters is fixed and will not vary with the blocklength. To model a massive number of transmitters, a variant of the MAC, referred to as the many-access channel (MnAC), has been introduced in [1]. The main feature of this channel is that the number of transmitters may increase with the blocklength.

In the information theory literature, different settings of the MnAC have been considered. These settings can be roughly classified into three groups, which differ from one another based on two criteria: the notion of probability of error and whether users use the same codebook or not. The first group, referred to as Group I, was introduced in [1]. This model assumes that:

J. Ravi and T. Koch have received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant No. 714161). T. Koch has further received funding from the Spanish Ministerio de Ciencia e Innovación under Grant PID2020-116683GB-C21 / AEI / 10.13039/501100011033.

1) Users use different codebooks.
2) The probability of a decoding error is defined as

$$P_{e,J}^{(n)} \triangleq \Pr\{(\hat{W}_1,\ldots,\hat{W}_{k_n}) \neq (W_1,\ldots,W_{k_n})\}. \quad (1)$$

Here, W_i denotes the message transmitted by user i and $\hat{W}_i$ denotes the decoder's estimate of W_i. We shall refer to the error probability $P_{e,J}^{(n)}$ as the *joint probability of error*.

The second group, referred to as Group II, was introduced in [2]. In this model:

1) Users use different codebooks.
2) The probability of a decoding error is defined as

$$P_{e,A}^{(n)} \triangleq \frac{1}{k_n}\sum_{i=1}^{k_n} \Pr\{\hat{W}_i \neq W_i\}. \quad (2)$$

The error probability $P_{e,A}^{(n)}$ is sometimes referred to as *per-user probability of error*. In this paper, we shall refer to it as the *average probability of error*.

The third group, referred to as Group III, was also introduced in [2]. The main difference between Group II and Group III is that, in the latter group, all users use the same codebook. Specifically, Group III assumes that:

1) All users use the same codebook.
2) The decoder outputs a list $\mathcal{L}(\mathbf{Y})$ of messages whose cardinality satisfies $|\mathcal{L}(\mathbf{Y})| \leq k_n$. The probability of error is defined as

$$P_{e,U}^{(n)} \triangleq \frac{1}{k_n}\sum_{i=1}^{k_n} \Pr(E_i) \quad (3)$$

where $E_i = F_i \cup G_i$, with $F_i = \{W_i \notin \mathcal{L}(\mathbf{Y})\}$ and $G_i = \{W_i = W_j \text{ for some } i \neq j\}$.

Group III is also known as *unsourced random access*. Recently, several works have addressed different aspects of unsourced random access [3]–[6].

This paper aims to understand the behavior of the capacity per unit-energy, defined as the largest number of bits per unit-energy that can be transmitted reliably, for the three groups described above. Specifically, we are interested in its scaling laws, i.e., in a characterization of the capacity per unit-energy as a function of the order of growth of k_n. The scaling laws for Group I and Group II were characterized in [7] and [8], respectively. Specifically, we showed in [7] that, for Group I, if the order of growth of k_n is strictly below $n/\log n$, then all users can achieve the single-user capacity per unit-energy. In contrast, if the order of growth of k_n is strictly above $n/\log n$, then the capacity per unit-energy is zero. Thus, there is a sharp transition between orders of growth where interference-free communication is feasible and orders of growth where no reliable communication is feasible. Similarly, we showed in [8] that the scaling laws for Group II exhibit the same behavior, but with the transition threshold located at n instead of at $n/\log n$. In this paper, we discuss the scaling laws for Group III. We demonstrate that, for this group, if the order of growth of k_n is sublinear, then all users can achieve the single-user capacity per unit-energy. In contrast, if the order of growth of k_n is strictly above $n \log\log n$, then the capacity per unit-energy is zero. This suggests that the scaling laws for Group III behave similarly as those for Groups I and II.

The rest of the paper is organized as follows. In Section II, we introduce the problem and define the three groups of models. In Section III, we review the scaling laws for Group I and Group II. In Section IV, we present the main result of this paper: the scaling laws for Group III. We conclude the paper in Section V.

II. Problem Formulation and Definitions

A. Three models of MnACs

Consider k users that wish to transmit their messages $W_i, i = 1, \ldots, k$, which are assumed to be independent and uniformly distributed on $\{1, \ldots, M_n^{(i)}\}$, to one common receiver. To achieve this, they send a codeword of n symbols over the channel, where n is referred to as the *blocklength*. We consider a many-access scenario where the number of users k grows with n, hence, we denote it as k_n. We further consider a Gaussian channel model where, for k_n users and blocklength n, the received vector $\mathbf{Y}$ is given by

$$\mathbf{Y} = \sum_{i=1}^{k_n} \mathbf{x}_i(W_i) + \mathbf{Z}.$$

Here, $\mathbf{x}_i(W_i)$ is the length-n codeword by user i for transmitting message W_i, and $\mathbf{Z}$ is a vector of n i.i.d. Gaussian components $Z_j \sim \mathcal{N}(0, N_0/2)$ (where $\mathcal{N}(\mu, \sigma^2)$ denotes the Gaussian distribution with mean μ and variance σ^2) independent of $\mathbf{X}_i \triangleq \mathbf{x}_i(W_i)$. We denote the vector of all transmitted codewords by $\mathbf{X} \triangleq (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_{k_n})$.

We next introduce the notion of an $\left(n, \{M_n^{(\cdot)}\}, \{E_n^{(\cdot)}\}, \epsilon\right)_\xi$ code for Groups I-III. We shall use subscripts "$\xi = J$" and "$\xi = A$" to indicate that a code is for Group I and Group II, respectively. We shall further use subscript "$\xi = U$" to indicate that a code is for Group III. (Here, "J" stands for joint probability of error, "A" stands for average probability of error, and "U" stands for unsourced MnAC.)

Definition 1: For $0 \leq \epsilon < 1$, an $\left(n, \{M_n^{(\cdot)}\}, \{E_n^{(\cdot)}\}, \epsilon\right)_\xi$ code ($\xi \in \{J, A\}$) for the Gaussian MnAC consists of:

1) k_n encoding functions $f_i : \{1, \ldots, M_n^{(i)}\} \to \mathcal{X}^n$, which map user i's message W_i to the codeword $\mathbf{x}_i(W_i)$, satisfying the energy constraint
$$\sum_{j=1}^{n} x_{ij}^2(W_i) \leq E_n^{(i)}, \quad \text{with probability one}$$
where x_{ij} is the j-th symbol of the transmitted codeword.
2) Decoding function $g : \mathcal{Y}^n \to \{M_n^{(\cdot)}\}$, which maps the received vector $\mathbf{Y}$ to the messages of all users.

The probability of error $P_{e,J}^{(n)}$ of an $\left(n, \{M_n^{(\cdot)}\}, \{E_n^{(\cdot)}\}, \epsilon\right)_J$ code satisfies

$$P_{e,J}^{(n)} = \Pr\{g(\mathbf{Y}) \neq (W_1, \ldots, W_{k_n})\} \leq \epsilon.$$

The probability of error $P_{e,A}^{(n)}$ of an $\left(n, \{M_n^{(\cdot)}\}, \{E_n^{(\cdot)}\}, \epsilon\right)_A$ code satisfies

$$P_{e,A}^{(n)} = \frac{1}{k_n} \sum_{i=1}^{k_n} \Pr\{\hat{W}_i \neq W_i\} \leq \epsilon.$$

Definition 2: For $0 \leq \epsilon < 1$, an $\left(n, \{M_n^{(\cdot)}\}, \{E_n^{(\cdot)}\}, \epsilon\right)_U$ code for the Gaussian MnAC consists of:

1) An encoding function $f : \{1, \ldots, M_n\} \to \mathcal{X}^n$, which maps user i's message to the codeword $\mathbf{x}(W_i)$, satisfying the energy constraint
$$\sum_{j=1}^{n} x_j^2(W_i) \leq E_n^{(i)}, \quad \text{with probability one}$$
where $x_j(W_i)$ is the j-th symbol of the transmitted codeword.
2) A decoder that outputs a list $\mathcal{L}(\mathbf{Y})$ of messages, whose cardinality is bounded by k_n, i.e., $|\mathcal{L}(\mathbf{Y})| \leq k_n$. The probability of error satisfies
$$P_{e,U}^{(n)} = \frac{1}{k_n} \sum_{i=1}^{k_n} \Pr(E_i) \leq \epsilon$$
where $E_i = F_i \cup G_i$, with $F_i = \{W_i \notin \mathcal{L}(\mathbf{Y})\}$ and $G_i = \{W_i = W_j \text{ for some } i \neq j\}$.

We shall say that the $\left(n, \{M_n^{(\cdot)}\}, \{E_n^{(\cdot)}\}, \epsilon\right)_\xi$ code, $\xi \in \{J, A, U\}$, is symmetric if $M_n^{(i)} = M_n$ and $E_n^{(i)} = E_n$ for all $i = 1, \ldots, k_n$. For compactness, we denote a symmetric code by $(n, M_n, E_n, \epsilon)_\xi, \xi \in \{J, A, U\}$. In this paper, we restrict ourselves to symmetric codes.

Definition 3: Let $\xi \in \{J, A, U\}$. For a symmetric $(n, M_n, E_n, \epsilon)_\xi$ code, the rate per unit-energy $\dot{R}^\xi$ is said to be ϵ-achievable if, for every $\alpha > 0$, there exists an n_0 such that if $n \geq n_0$, then an $(n, M_n, E_n, \epsilon)_\xi$ code can be found whose rate per unit-energy satisfies $\frac{\log M_n}{E_n} > \dot{R}^\xi - \alpha$. Furthermore, $\dot{R}^\xi$ is said to be achievable if it is ϵ-achievable for all $0 < \epsilon < 1$. The capacity per unit-energy $\dot{C}^\xi$ is the supremum of all achievable rates per unit-energy.

B. Order Notations

Let $\{a_n\}$ and $\{b_n\}$ be two sequences of nonnegative real numbers. We write $a_n = O(b_n)$ if $\limsup_{n\to\infty} \frac{a_n}{b_n} < \infty$. Similarly, we write $a_n = o(b_n)$ if $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$ and $a_n = \Omega(b_n)$ if $\liminf\limits_{n\to\infty} \frac{a_n}{b_n} > 0$. The notation $a_n = \Theta(b_n)$ indicates that $a_n = O(b_n)$ and $a_n = \Omega(b_n)$. Finally, we write $a_n = \omega(b_n)$ if $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \infty$.

III. Capacity Per Unit-Energy of Sourced Gaussian Many-Access Channels

In this section, we discuss the behavior of the capacity per unit-energy as a function of the order of growth of k_n for Groups I and II. In particular, in Theorem 1 we review the scaling laws for Group I, and in Theorem 2 we review the scaling laws for Group II.

Theorem 1: The capacity per unit-energy $\dot{C}^J$ for Group I has the following behavior:

1) If $k_n = o(n/\log n)$, then $\dot{C}^J = \frac{\log e}{N_0}$.
2) If $k_n = \omega(n/\log n)$, then $\dot{C}^J = 0$.
3) If $k_n = \Theta(n/\log n)$, then $0 < \dot{C}^J < \frac{\log e}{N_0}$.

Proof: Theorem 1 is [9, Theorem 1]. ■

Parts 1) and 2) of Theorem 1 were first presented in [7], and Part 3) was obtained in our recent work [9]. Theorem 1 shows that, if k_n is of an order strictly below $n/\log n$, then each user can achieve the single-user capacity per unit-energy $\frac{\log e}{N_0}$. Conversely, if k_n is of an order strictly above $n/\log n$, then users cannot achieve a positive rate per unit-energy. Furthermore, if the order of k_n is exactly $n/\log n$, then the capacity per unit-energy is strictly between zero and $\frac{\log e}{N_0}$. Thus, there is a sharp transition between orders of growth where interference-free communication is feasible and orders of growth where reliable communication at a positive rate per unit-energy is infeasible.

Theorem 2: The capacity per unit-energy $\dot{C}^A$ for Group II has the following behavior:

1) If $k_n = o(n)$, then $\dot{C}^A = \frac{\log e}{N_0}$.
2) If $k_n = \Omega(n)$, then $\dot{C}^A = 0$.

Proof: Theorem 2 is [9, Theorem 3]. ■

Theorem 2 was first presented in [8]. It shows that, if the order of growth of k_n is sublinear in n, then all users can achieve the single-user capacity per unit-energy $\frac{\log e}{N_0}$. Conversely, if the order of growth of k_n is linear or above, then users cannot achieve a positive rate per unit-energy. Thus, the behavior of the scaling law for Group II is the same as that of Group I: there is a sharp transition between orders of growth where interference-free communication is feasible and orders of growth where reliable communication at a positive rate per unit-energy is infeasible. The difference is that the transition threshold is at n instead of at $n/\log n$.

IV. Unsourced Many-Access Channels

It can be observed from the definition of the unsourced random access model that there are two important aspects. The first aspect is that all users are restricted to use the same codebook. Benefits of this assumption are two-fold: it is more practical in settings with a large number of users, and many well-known random access schemes, such as ALOHA and slotted ALOHA, are valid schemes under this model. If all users are restricted to use the same codebook, then decoding is possible only up to permutations of messages. Indeed, the decoder cannot distinguish between $\mathbf{x}(w_i)+\mathbf{x}(w_j)$ and $\mathbf{x}(w_j) + \mathbf{x}(w_i)$, so it is not able to decide on which user transmitted w_i and which user transmitted w_j. As a consequence, in unsourced many-access channels, the decoder outputs a list of transmitted messages, which is denoted by $\mathcal{L}(\mathbf{Y})$.

The second aspect is the definition of the probability of error given in (3). In this definition, there are two error events: F_i and G_i. The event F_i indicates whether user i's message is in the list or not. The second event G_i indicates whether two users transmit the same message. There are practical as well as theoretical considerations for including G_i in the definition of the probability of error. From a practical point of view, in random access schemes such as ALOHA, two users transmitting the same message at the same time will result in a collision. The event G_i accounts for such collisions. From a theoretical point of view, including G_i as an error event avoids certain trivial cases. To illustrate this, let us first consider the error event E_i only with F_i. In this case, the probability of error may vanish even if $\frac{\log M_n}{E_n}$ is infinite for every n. Indeed, assume that $M_n = 2$ and $E_n = 0$ for all n, and consider a decoding function that always outputs a list with both messages. Since, by assumption, $k_n \geq 2$, we have $|\mathcal{L}(\mathbf{Y})| \leq k_n$, and $W_i \in \mathcal{L}(\mathbf{Y})$ for all $i = 1, \ldots, k_n$. Thus, in this case, we have that $P_{e,U}^{(n)} = 0$, from which we obtain that $\dot{C}^U = \infty$ since $\frac{\log M_n}{E_n} = \infty$. We conclude that the definition of the error probability without G_i may lead to uninteresting cases. The following proposition shows that including G_i as an error event ensures that the probability of error vanishes only if $M_n \to \infty$, thereby excluding trivial cases such as the one above.

Proposition 1: $\Pr(G_i)$ vanishes as $n \to \infty$ if, and only if, $M_n = \omega(k_n)$.

Proof: We have

$$\begin{aligned}
1 - \Pr(G_i) &= \Pr(W_i \neq W_j \text{ for any } j \neq i) \\
&= \frac{M_n(M_n - 1)^{k_n - 1}}{M_n^{k_n}} \\
&= (1 - 1/M_n)^{k_n - 1}, \quad i = 1, \ldots, k_n
\end{aligned}$$

which tends to one as $n \to \infty$ if, and only if, $M_n = \omega(k_n)$. ■

The following theorem discusses the capacity per unit-energy for the unsourced random-access case.

Theorem 3: For unsourced random access, the capacity per unit-energy has the following behavior:

1) If $k_n = o(n)$, then $\dot{C}^U = \frac{\log e}{N_0}$.
2) If $k_n = \omega(n \log\log n)$, then $\dot{C}^U = 0$.

Proof: See Subsection IV-A. ■

Theorem 3 shows that, for any sublinear growth of k_n, the capacity per unit-energy is equal to the single-user capacity per unit-energy $\frac{\log e}{N_0}$. Conversely, if the order of k_n is above $n \log \log n$, then the capacity per unit-energy is zero. This suggests that the behavior of the capacity per unit-energy for Group III is similar to the ones for Groups I and II, in the sense that there is a sharp transition between orders of growth where interference-free communication is feasible and the orders of growth where no reliable communication is possible. Note, however, that the behavior of the capacity per unit-energy is still open for the regime where the order of growth of k_n is between n and $n \log \log n$.

To prove Part 1) of Theorem 3, we use an orthogonal codebook, i.e., a codebook for which the inner products between any two codewords are zero. For such a codebook, if the messages of any two users are different, then these users will be transmitting in different slots, resulting in an orthogonal-access scheme, i.e., in an access scheme where codewords transmitted by different users are orthogonal to each other. To prove Part 2), we show that, for every $\dot{R} > 0$, if $k_n = \Omega(n)$, then the energy E_n needs to satisfy two conditions: $E_n = O(\log k_n)$ and $E_n = \omega(2^{\frac{k_n}{n}})$. However, if $k_n = \omega(n \log \log n)$, then no sequence $\{E_n\}$ can simultaneously satisfy these two conditions, hence Part 2) follows.

A. Proof of Theorem 3

1) Proof of Part 1): The achievability part uses an orthogonal codebook with $M_n = n$ messages. To send message i, user $j, j = 1, \ldots, k_n$, sets the i-th position of the n-length codeword to $\sqrt{E_n}$ and the remaining positions to zero, i.e., for $i = 1, \ldots, k_n$,

$$x_j(W_i) = \begin{cases} \sqrt{E_n}, & \text{if } j = W_i \\ 0, & \text{otherwise.} \end{cases}$$

The decoder finds the k_n indices of the received vector of largest values. It outputs the messages corresponding to these indices as the list of messages, i.e.,

$$\mathcal{L}(\mathbf{Y}) = \{i_1, i_2, \ldots, i_{k_n}\}.$$

Thus, the received symbols $Y_{i_1}, Y_{i_2}, \ldots, Y_{k_n}$ denote the largest k_n components in $\mathbf{Y}$.

Next, we derive an upper bound on the probability $\Pr(F_i \cup G_i)$. To this end, we write $\Pr(F_i \cup G_i)$ as $\Pr(G_i) + \Pr(F_i \cap G_i^c)$ and upper-bound $\Pr(F_i \cap G_i^c)$. Together with the expression for $\Pr(G_i)$ from Proposition 1, the desired upper bound on $\Pr(F_i \cup G_i)$ follows. Let

$$\overline{\mathbf{X}} \triangleq \sum_{i=1}^{k_n} \mathbf{X}_i$$

where $\mathbf{X}_i$ denotes the transmitted codeword of user i. Further let $\overline{\mathbf{X}}(i)$ denote the i-th component of $\overline{\mathbf{X}}$, and let

$$\mathcal{Z}(\overline{\mathbf{X}}) \triangleq \{i : \overline{\mathbf{X}}(i) = 0\}.$$

Since the number of distinct messages of k_n users can vary from 1 to k_n, it follows that

$$n - k_n \leq |\mathcal{Z}(\overline{\mathbf{X}})| \leq n - 1.$$

The event $F_i \cap G_i^c$ happens only if, for some $j \in \mathcal{Z}(\overline{\mathbf{X}})$, we have $Y_j \geq Y_{W_i}$. So, by defining

$$K(y) \triangleq \Pr(Y_j > y, \text{ for some } j \in \mathcal{Z}(\overline{\mathbf{X}})),$$

we obtain

$$\Pr(F_i \cap G_i^c) \leq \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{\frac{-(y-\sqrt{E_n})^2}{N_0}} K(y) dy.$$

Applying the union bound on $K(y)$ gives that $K(y) \leq |\mathcal{Z}(\overline{\mathbf{X}})| Q(y)$. Consequently, we have for $i = 1, \ldots, k_n$ that

$$\begin{aligned}\Pr(F_i \cap G_i^c) &\leq |\mathcal{Z}(\overline{\mathbf{X}})| \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{\frac{-(y-\sqrt{E_n})^2}{N_0}} Q(y) dy \\ &\leq (M_n - 1) \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{\frac{-(y-\sqrt{E_n})^2}{N_0}} Q(y) dy\end{aligned} \tag{4}$$

where, in the last step, we used that $M_n = n$. Since $\Pr(E_i) = \Pr(G_i) + \Pr(F_i \cap G_i^c)$, it follows from Proposition 1 and (4) that

$$\begin{aligned}\frac{1}{k_n} \sum_{i=1}^{k_n} \Pr(E_i) &\leq 1 - (1 - 1/M_n)^{k_n - 1} \\ &\quad + (M_n - 1) \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi N_0}} e^{\frac{-(y-\sqrt{E_n})^2}{N_0}} Q(y) dy.\end{aligned} \tag{5}$$

The first term on the right-hand side (RHS) of (5) vanishes since, by assumption, $k_n = o(n)$ and $M_n = n$. The second term on the RHS of (5) is an upper bound on the probability of error for transmitting M_n messages using orthogonal codewords of energy E_n over a single-user Gaussian channel. It was shown [10, Section 8.2] that this upper bound vanishes as $n \to \infty$ if $\frac{\log M_n}{E_n} \leq \frac{\log e}{N_0}$. This proves that $\dot{C}^U \geq \frac{\log e}{N_0}$.

To prove that $\dot{C}^U \leq \frac{\log e}{N_0}$, we first note that

$$P_{e,U}^{(n)} \geq \min_i \Pr(E_i).$$

We further note that $\Pr(E_i) \geq \Pr(F_i \cap G_i^c)$. Next, consider a genie who informs the receiver about the codewords transmitted by users $j \neq i$. Then, $\Pr(F_i \cap G_i^c)$ is lower-bounded by the error probability P_G of the single-user Gaussian channel, since a single-user channel can be obtained from the MnAC by subtracting the codewords of all users $j \neq i$. By [11, eq. (30)], we have

$$P_G \geq Q\left(\sqrt{\frac{2E_n}{N_0}}\right), \quad M_n \geq 2,$$

which, in turn, yields that

$$P_{e,U}^{(n)} \geq Q\left(\sqrt{\frac{2E_n}{N_0}}\right), \quad M_n \geq 2. \tag{6}$$

Hence, $P_{e,U}^{(n)} \to 0$ only if $E_n \to \infty$. However, if $E_n \to \infty$ as $n \to \infty$, then the rate per unit-energy of the single-user Gaussian channel is upper-bounded by $\frac{\log e}{N_0}$ (see [8, Remark 1]). Thus, it follows that $\dot{C}^U \leq \frac{\log e}{N_0}$. This proves Part 1).

2) Proof of Part 2): Let $\mathbf{1}(\cdot)$ denote the indicator function. By following similar steps as in the proof of Fano's inequality, we obtain that

$$\begin{aligned} H(W_i) &= H(W_i|\mathcal{L}(\mathbf{Y})) + I(W_i;\mathcal{L}(\mathbf{Y})) \\ &\leq H(W_i, \mathbf{1}(E_i)|\mathcal{L}(\mathbf{Y})) + I(W_i;\mathcal{L}(\mathbf{Y})) \\ &= H(\mathbf{1}(E_i)|\mathcal{L}(\mathbf{Y})) + H(W_i|\mathbf{1}(E_i), \mathcal{L}(\mathbf{Y})) \\ &\quad + I(W_i;\mathcal{L}(\mathbf{Y})) \\ &\leq 1 + \Pr(\mathbf{1}(E_i)=0) H(W_i|\mathbf{1}(E_i)=0, \mathcal{L}(\mathbf{Y})) \\ &\quad + \Pr(\mathbf{1}(E_i)=1) H(W_i|\mathbf{1}(E_i)=1, \mathcal{L}(\mathbf{Y})) \\ &\quad + I(W_i;\mathcal{L}(\mathbf{Y})) \\ &\leq 1 + (1-\Pr(E_i))\log k_n + \Pr(E_i)\log M_n \\ &\quad + I(W_i;\mathcal{L}(\mathbf{Y})). \end{aligned}$$

Averaging over i, and using the definition of the error probability (3), we obtain that

$$\begin{aligned} &\log M_n \\ &\leq 1 + \left(1 - P_{e,U}^{(n)}\right)\log k_n + P_{e,U}^{(n)}\log M_n + \frac{I(\mathbf{W};\mathcal{L}(\mathbf{Y}))}{k_n} \\ &\leq 1 + \left(1 - P_{e,U}^{(n)}\right)\log k_n + P_{e,U}^{(n)}\log M_n + \frac{I(\mathbf{X};\mathbf{Y})}{k_n} \\ &\leq 1 + \left(1 - P_{e,U}^{(n)}\right)\log k_n + P_{e,U}^{(n)}\log M_n \\ &\quad + \frac{n}{2k_n}\log\left(1 + \frac{2k_nE_n}{nN_0}\right). \end{aligned}$$

Here, the first inequality follows from the independence of the messages W_i, the second inequality follows from the data processing inequality and the Markov chain $\mathbf{W}-\mathbf{X}-\mathbf{Y}-\mathcal{L}(\mathbf{Y})$, and the third inequality follows by upper-bounding $I(\mathbf{X};\mathbf{Y})$ by $\frac{n}{2}\log(1+\frac{2k_nE_n}{nN_0})$. Subtracting $P_{e,U}^{(n)}\log M_n$, and dividing by $\left(1-P_{e,U}^{(n)}\right)$, this yields

$$\log M_n \leq \frac{1}{1-P_{e,U}^{(n)}} + \log k_n + \frac{n\log\left(1+\frac{2k_nE_n}{nN_0}\right)}{2k_n(1-P_{e,U}^{(n)})}. \quad (7)$$

Dividing both sides of (7) by E_n gives

$$\dot{R} \leq \frac{1}{E_n\left(1-P_{e,U}^{(n)}\right)} + \frac{\log k_n}{E_n} + \frac{n\log\left(1+\frac{2k_nE_n}{nN_0}\right)}{2k_nE_n(1-P_{e,U}^{(n)})}. \quad (8)$$

As we noted before, $P_{e,U}^{(n)}$ vanishes as $n \to \infty$ only if $E_n \to \infty$ (cf. (6)). This implies that the first term on the RHS of (8) tends to zero if $P_{e,U}^{(n)} \to 0$ as $n \to \infty$. Next note that, by the theorem's assumption, $k_n = \Omega(n)$. Then, $E_n \to \infty$ implies that the last term on the RHS of (8) also tends to zero as $n \to \infty$. So, a positive rate per unit-energy $\dot{R}$ is possible only if

$$E_n = O(\log k_n), \quad (9)$$

since otherwise the second term on the RHS of (8) would vanish, too.

Subtracting $\log k_n$ on both sides of (7), we also obtain that

$$\log(M_n/k_n) \leq \frac{1}{1-P_{e,U}^{(n)}} + \frac{n\log\left(1+\frac{2k_nE_n}{nN_0}\right)}{2k_n(1-P_{e,U}^{(n)})}. \quad (10)$$

Recall that

$$\begin{aligned} P_{e,U}^{(n)} &\geq \min_i \Pr(E_i) \\ &\geq \min_i \Pr(G_i). \end{aligned} \quad (11)$$

Furthermore, Proposition 1 demonstrates that $\Pr(G_i)$ vanishes only if $M_n = \omega(k_n)$. Hence, $P_{e,U}^{(n)}$ vanishes only if $M_n = \omega(k_n)$. It follows that, if $P_{e,U}^{(n)} \to 0$, then the left-hand side of (10) tends to infinity as $n \to \infty$, so the last term on the RHS of (10) must tend to infinity as $n \to \infty$, too. We then have that

$$\begin{aligned} \frac{k_n}{n} &= o\left(\log\frac{k_nE_n}{n}\right) \\ &= o\left(\log\frac{k_n}{n} + \log E_n\right), \end{aligned}$$

which is possible only if $\frac{k_n}{n} = o(\log E_n)$, since for two sequences a_n and b_n, $\Theta(a_n+b_n) = \max\{\Theta(a_n), \Theta(b_n)\}$, and $\frac{k_n}{n} = o(\log\frac{k_n}{n})$ is not possible if $k_n = \Omega(n)$. The condition $\frac{k_n}{n} = o(\log E_n)$ is equivalent to

$$E_n = \omega(2^{\frac{k_n}{n}}). \quad (12)$$

Since we have assumed that $k_n = \Omega(n)$, there exists a sequence $a_n = \Omega(1)$ such that $k_n = a_n n$. Then, it follows from (9) and (12) that $\log k_n = \omega(2^{a_n})$, which is equivalent to

$$\begin{aligned} 2^{a_n} &= o(\log k_n) \\ &= o(\log a_n + \log n). \end{aligned}$$

This, in turn, implies that

$$2^{a_n} = o(\log n) \quad (13)$$

by the same argument that led to (12). It thus follows from (13) that a positive $\dot{R}$ is only achievable if $a_n = O(\log\log n)$. However, a_n cannot be $O(\log\log n)$ since $k_n = \omega(n\log\log n)$ by the theorem's assumption, so we conclude that $\dot{C} = 0$. This completes the proof of Theorem 3.

Remark 1: The main step in the proof of Part 2) of Theorem 3 is that, if $k_n = \Omega(n)$, then, to achieve a positive rate per unit-energy, E_n has to satisfy two constraints: $E_n = O(\log k_n)$ and $E_n = \omega(2^{\frac{k_n}{n}})$. The result then follows by noting that no sequence $\{E_n\}$ can satisfy these two constraints simultaneously if $k_n = \omega(n\log\log n)$. We used similar arguments to obtain the converse results for Group I and Group II. Indeed, for Group I, E_n must satisfy $E_n = \Omega(\log k_n)$ and $E_n = O(n/k_n)$ to achieve a positive rate per unit-energy. When $k_n = \omega(n/\log n)$, no sequence $\{E_n\}$ can simultaneously satisfy both constraints. Similarly, for Group II, E_n must satisfy $E_n = O(n/k_n)$ and $E_n = \omega(1)$ to achieve a positive rate per unit-energy. If $k_n = \Omega(n)$, no sequence $\{E_n\}$ can simultaneously satisfy both constraints. This demonstrates that the constraints on E_n vary from model to model, resulting in different transition thresholds.

V. Conclusion

We discussed the behavior of the capacity per unit-energy as a function of the order of growth of the number of users k_n for three groups of channel models:

- Group I considers the classical notion of probability of error and assumes that users employ different codebooks;
- Group II considers the per-user probability of error and assumes that users employ different codebooks;
- Group III considers the per-user probability of error and assumes that all users employ the same codebook.

We demonstrated that the capacities per unit-energy of these groups exhibit the same qualitative behavior in the sense that there is a sharp transition between orders of growth where interference-free communication is feasible and orders of growth where no reliable communication is possible. The choice of the channel model merely determines the order of growth where this transition occurs. One may argue that, from a scientific point of view, the most interesting regime is the one where the capacity per unit-energy depends on the exact growth of k_n, rather than solely on its order. Our results show that this regime is very narrow. For example, for Group II, we demonstrated in [8] that the capacity per unit-energy is different from 0 or $\frac{\log e}{N_0}$ only if k_n grows linearly in n, and only if we allow for a positive probability of error. This is precisely the regime that is most prevalent in the literature.

References

[1] X. Chen, T. Y. Chen, and D. Guo, "Capacity of Gaussian many-access channels," *IEEE Transactions on Information Theory*, vol. 63, no. 6, pp. 3516–3539, June 2017.

[2] Y. Polyanskiy, "A perspective on massive random-access," in *Proc. 2017 IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, June 2017, pp. 2523–2527.

[3] O. Ordentlich and Y. Polyanskiy, "Low complexity schemes for the random access Gaussian channel," in *Proc. 2017 IEEE International Symposium on Information Theory (ISIT)*, Aachen, Germany, June 2017, pp. 2528–2532.

[4] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient coded random access for the wireless uplink," *IEEE Transactions on Communications*, vol. 68, no. 8, pp. 4694–4708, August 2020.

[5] V. K. Amalladinne, J. F. Chamberland, and K. R. Narayanan, "A coded compressed sensing scheme for unsourced multiple access," *IEEE Transactions on Information Theory*, vol. 66, no. 10, pp. 6509–6533, October 2020.

[6] A. Fengler, P. Jung, and G. Caire, "SPARCs for unsourced random access," *IEEE Transactions on Information Theory*, vol. 67, no. 10, pp. 6894–6915, October 2021.

[7] J. Ravi and T. Koch, "Capacity per unit-energy of Gaussian many-access channels," in *Proc. 2019 IEEE International Symposium on Information Theory (ISIT)*, Paris, France, July 2019.

[8] J. Ravi and T. Koch, "On the per-user probability of error in Gaussian many-access channels," in *Proc. 2020 International Zurich Seminar on Information and Communication (IZS)*, February 2020, pp. 139–143.

[9] J. Ravi and T. Koch, "Scaling laws for Gaussian random many-access channels," `arxiv: 2012.10350[cs.IT]`, September 2021.

[10] R. G. Gallager, *Information Theory and Reliable Communication*, 1st ed. New York, NY, USA: John Wiley & Sons, 1968.

[11] Y. Polyanskiy, H. V. Poor, and S. Verdú, "Minimum energy to send k bits through the Gaussian channel with and without feedback," *IEEE Transactions on Information Theory*, vol. 57, no. 8, pp. 4880–4902, August 2011.