Random Coding Error Exponent for the Bee-Identification Problem

Anshoo Tandon
National University of Singapore

Vincent Y. F. Tan
National University of Singapore

Lav R. Varshney
University of Illinois at Urbana-Champaign

***Abstract*—Consider the problem of identifying a massive number of bees, uniquely labeled with barcodes, using noisy measurements. We introduce this "bee-identification problem", characterize the random coding exponent, and derive efficiently computable bounds for this exponent. We demonstrate that joint decoding of barcodes has much better exponent than separate decoding followed by permutation inference.**

I. Introduction: The Bee-Identification Problem

Consider a group of m different bees, where each bee is tagged with a unique barcode for identification purposes in order to understand interaction patterns in honeybee social networks [1]. Assume that a camera is employed to picture the beehive to study the interactions among bees. The image output (see Fig. 1) can be considered as a noisy and unordered set of m barcodes. We pose the problem of bee-identification from beehive image as an information-theoretic problem (Sec. I-A).

In a related work motivated by Internet of Things (IoT) setting, the identification of users in strongly asynchronous massive access channels was studied [2]. The identification of the underlying distributions of a set of observed sequences (where each sequence is generated i.i.d. by a distinct distribution) was analyzed in [3]. The effective channel for the bee-identification problem (see Fig. 2) is also related to the DNA storage channel [4].

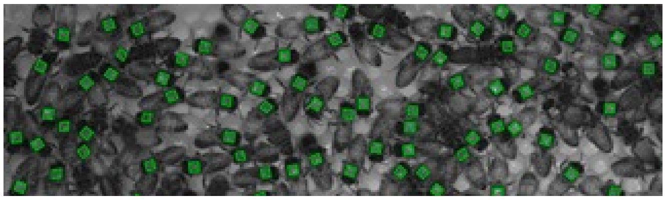

Fig. 1: Bees tagged with barcodes (adapted from [1]).

A. Problem Formulation

The barcode for each bee may be represented by a binary vector of length n. The collection of all the barcodes on bees may be viewed as a codebook C comprising m rows and n columns, with each row corresponding to a unique bee barcode. The channel output is a row-permuted and noisy version of the codebook. If π denotes a given permutation of m-letters, then the channel first permutes the m rows of codebook C, based on π, to produce C_π (see Fig. 2). Therefore, if $j = \pi(i)$ and $\boldsymbol{c}_i$ denotes the i-th row of C, then the jth row of C_π is equal to $\boldsymbol{c}_i$. The channel then applies noise on the permuted codebook C_π to produce $\tilde{C}_\pi$, where noise is modeled by a binary symmetric channel (BSC) having crossover probability p, denoted BSC(p), with $0 < p < 0.5$. If $j = \pi(i)$, and the j-th row of $\tilde{C}_\pi$ is denoted $\tilde{\boldsymbol{c}}_{\pi(i)}$, then

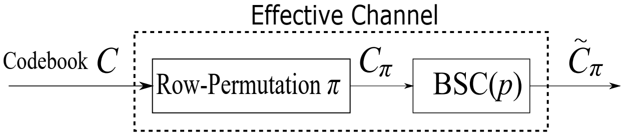

Fig. 2: Effective channel for the bee-identification problem.

$$\Pr\{\tilde{\boldsymbol{c}}_{\pi(i)}|\boldsymbol{c}_i, \pi\} = p^{d_i}(1-p)^{n-d_i}, \quad 1 \le i \le m,$$
$$\Pr\left\{\tilde{C}_\pi | C, \pi\right\} = \prod_{i=1}^{m} \Pr\{\tilde{\boldsymbol{c}}_{\pi(i)}|\boldsymbol{c}_i, \pi\} = \prod_{i=1}^{m} p^{d_i}(1-p)^{n-d_i}, \tag{1}$$

where $d_i \triangleq \mathrm{d_H}(\tilde{\boldsymbol{c}}_{\pi(i)}, \boldsymbol{c}_i)$ denotes the Hamming distance between vectors $\tilde{\boldsymbol{c}}_{\pi(i)}$ and $\boldsymbol{c}_i$. Let $\mathcal{M} \triangleq \{1, 2, \ldots, m\}$, and let the decoder correspond to a function ϕ which takes $\tilde{C}_\pi$ as an input and produces a map $\nu : \mathcal{M} \to \mathcal{M}$ where $\nu(k)$ corresponds to the index of the transmitted codeword which produced the received word $\tilde{\boldsymbol{c}}_k$, for $1 \le k \le m$. We assume that the decoder has knowledge of codebook C, and its task is to *recover the row-permutation* π introduced by the channel.

In this paper, we derive efficiently computable lower bounds for the random coding bee-identification exponent when (i) each received barcode is decoded *independently* (Sec. III), and (ii) when all barcodes are decoded *jointly* (Sec. IV). We also provide an explicit upper bound on the bee-identification exponent that holds for all possible codebook designs (Sec. V).

II. Bee-Identification Error

The task of the decoder in the bee-identification problem is to recover the row-permutation π introduced by the channel. The bee-identification error indicator is defined as

$$\mathcal{D}\left(\phi(\tilde{C}_\pi), \pi^{-1}\right) = \mathcal{D}\left(\nu, \pi^{-1}\right) \triangleq \begin{cases} 1, & \text{if } \nu \ne \pi^{-1}, \\ 0, & \text{if } \nu = \pi^{-1}. \end{cases}$$

For a given codebook C and decoding function ϕ, the expected bee-identification error probability over the BSC(p) is

$$D(C, p, \phi) \triangleq \mathbb{E}_\pi\left[\mathbb{E}\left[\mathcal{D}\left(\phi(\tilde{C}_\pi), \pi^{-1}\right)\right]\right], \tag{2}$$

This work was supported in part by a Singapore Ministry of Education Tier 2 grant (R-263-000-C83-112).

978-1-5386-6900-6/19/$31.00 ©2019 IEEE

where the inner expectation is over the distribution of $\tilde{C}_\pi$ given C and π (see (1)), and the outer expectation is over a uniform distribution of π over all m-letter permutations. Note that (2) can be equivalently expressed as

$$D(C,p,\phi) = \Pr\left\{\phi(\tilde{C}_\pi) \neq \pi^{-1}\right\} = \Pr\left\{\nu \neq \pi^{-1}\right\}. \tag{3}$$

For a given $R > 0$, let the number of barcodes m scale exponentially with blocklength n as $m = 2^{nR}$. Let $\mathscr{C}(n,R)$ be the set of all binary matrices with $m = 2^{nR}$ rows and n columns. When the codebook C is uniformly distributed over $\mathscr{C}(n,R)$, for given values of n and R, we define the random coding bee-identification error probability as

$$D_{\mathrm{RC}}(n,R,p,\phi) \triangleq \frac{1}{|\mathscr{C}(n,R)|} \sum_{C \in \mathscr{C}(n,R)} D(C,p,\phi). \tag{4}$$

For a given decoding function ϕ, the *random coding bee-identification exponent* is defined as

$$E_{D_{\mathrm{RC}}}(R,p,\phi) = \limsup_{n\to\infty} \frac{-\log D_{\mathrm{RC}}(n,R,p,\phi)}{n}. \tag{5}$$

III. Naïve Decoding Strategy

The naïve decoding strategy is one where each barcode is decoded independently. In this case, for $1 \leq i \leq m$, the decoder picks $\tilde{\boldsymbol{c}}_i$, the ith row of $\tilde{C}_\pi$, and then assigns $\nu(i) = \arg\min_k \mathrm{d_H}(\tilde{\boldsymbol{c}}_i, \boldsymbol{c}_k)$. If there is more than one codeword at the same minimum Hamming distance from $\tilde{\boldsymbol{c}}_i$, then any one of the corresponding codeword indices is chosen at random. From (3), (4), and the union bound, we have

$$D_{\mathrm{RC}}(n,R,p,\phi) \leq \frac{1}{|\mathscr{C}(n,R)|} \sum_{C \in \mathscr{C}(n,R)} \sum_{i=1}^{m} \Pr\left\{\nu(i) \neq \pi^{-1}(i)\right\}.$$

Let $P(n,R,p) \triangleq \frac{1}{|\mathscr{C}(n,R)|} \sum_{C \in \mathscr{C}(n,R)} \Pr\left\{\nu(i) \neq \pi^{-1}(i)\right\}$, which, for $j = \pi^{-1}(i)$, corresponds to the probability of error, averaged over the ensemble of codebooks uniformly distributed over $\mathscr{C}(n,R)$, when jth codeword is transmitted over BSC(p), with $0 < p < 0.5$. Further, $P(n,R,p)$ is independent of the index of the transmitted codeword due to the averaging over the ensemble, and we have $D_{\mathrm{RC}}(n,R,p,\phi) \leq mP(n,R,p)$. Further, the random coding bee-identification error probability is upper bounded by 1, and so

$$D_{\mathrm{RC}}(n,R,p,\phi) \leq \min\left\{1,\, mP(n,R,p)\right\}. \tag{6}$$

The random coding exponent over BSC(p), denoted $E_r(R,p)$, is defined as [5] $E_r(R,p) \triangleq \limsup_{n\to\infty}(-1/n)\log P(n,R,p)$. The channel capacity over BSC(p) is $1 - H(p)$, where $H(\cdot)$ is the binary entropy function, and $E_r(R,p) > 0$ for $R < 1 - H(p)$ while $E_r(R,p) = 0$ for $R \geq 1 - H(p)$ [5]. The exact value of $E_r(R,p)$ is given by [5], [6]

$$E_r(R,p) = \begin{cases} R_0(p) - R, & 0 \leq R \leq R_{\mathrm{cr}}(p) \\ D(\delta_{\mathrm{GV}}(R)||p), & R_{\mathrm{cr}}(p) \leq R \leq 1 - H(p) \\ 0, & R \geq 1 - H(p), \end{cases} \tag{7}$$

where

$$R_0(p) \triangleq 1 - \log\left(1 + \sqrt{4p(1-p)}\right), \tag{8}$$

$$R_{\mathrm{cr}}(p) = 1 - H\left(\frac{\sqrt{p}}{\sqrt{p} + \sqrt{1-p}}\right), \tag{9}$$

$D(x||y) \triangleq x \log \frac{x}{y} + (1-x)\log\frac{1-x}{1-y}$ and $\delta_{\mathrm{GV}}(R)$ denotes the value of δ in the interval $[0, 0.5]$ where $H(\delta) = 1 - R$. The next theorem is an explicit lower bound on $E_{D_{\mathrm{RC}}}(R,p,\phi)$.

Theorem 1. *We have*

$$E_{D_{\mathrm{RC}}}(R,p,\phi) \geq |R_0(p) - 2R|^{+}, \tag{10}$$

where $|x|^{+} \triangleq \max(0,x)$.

Proof: Combining (5), (6), and the fact that $m = 2^{nR}$,

$$E_{D_{\mathrm{RC}}}(R,p,\phi) \geq |E_r(R,p) - R|^{+}. \tag{11}$$

Using explicit numerical computation, it can be shown that $R_0(p) \leq 2R_{\mathrm{cr}}(p)$, and we obtain (10) using (11) and (7). Note that $R_0(p) \leq 2R_{\mathrm{cr}}(p)$ implies $|E_r(R,p) - R|^{+} = 0$ when $R \geq R_{\mathrm{cr}}(p)$, as $E_r(R,p)$ is non-increasing in R. ■

This bound on $E_{D_{\mathrm{RC}}}(R,p,\phi)$ is obtained by applying a naïve decoding strategy where each barcode is decode independently. In the next section, we characterize $E_{D_{\mathrm{RC}}}(R,p,\phi)$ when the decoding function ϕ *jointly decodes* all the barcodes.

IV. Joint Decoding of Barcodes

Let S_m denote the set of all m-letter permutations. For joint maximum likelihood (ML) decoding of barcodes, the decoding function ϕ takes the noisy row-permuted codebook $\tilde{C}_\pi$ as input, and produces permutation $\nu = \rho^{-1}$ as output, where $\rho = \arg\min_{\sigma \in S_m} \mathrm{d_H}(\tilde{C}_\pi, C_\sigma)$, and $\mathrm{d_H}(\tilde{C}_\pi, C_\sigma) \triangleq |\{(i,j) : \tilde{C}_\pi(i,j) \neq C_\sigma(i,j), 1 \leq i \leq m, 1 \leq j \leq n\}|$. We aim to provide bounds on $\Pr\{\nu \neq \pi^{-1}\} = \Pr\{\rho \neq \pi\}$.

For any two permutations $\pi_1, \pi_2 \in S_m$, the set of distances $\{\mathrm{d_H}(\tilde{C}_{\pi_1}, C_\sigma)\}_{\sigma \in S_m}$ and $\{\mathrm{d_H}(\tilde{C}_{\pi_2}, C_\sigma)\}_{\sigma \in S_m}$ are equal. Therefore, the performance of the joint ML decoder is independent of the channel permutation π, and we assume, without loss of generality, that the *permutation induced by the channel is the identity permutation, denoted* π_0.

For a given codebook C at the transmitter, let $\tilde{C}_{\pi_0}$ denote the received noisy codebook at the output of the effective channel, and for $\sigma \in S_m$ with $\sigma \neq \pi_0$, we define

$$\Pr\{\pi_0 \to \sigma\} \triangleq \Pr\left\{\mathrm{d_H}(\tilde{C}_{\pi_0}, C_\sigma) \leq \mathrm{d_H}(\tilde{C}_{\pi_0}, C_{\pi_0})\right\}. \tag{12}$$

For any two functions $f(n)$ and $g(n)$, we use the notation $f(n) \doteq g(n)$ if $\lim_{n\to\infty} n^{-1}\log\left(f(n)/g(n)\right) = 0$. Similarly, we write $f(n) \mathrel{\dot{\leq}} g(n)$ (resp. $\mathrel{\dot{\geq}} g(n)$) if $\limsup_{n\to\infty} n^{-1}\log\left(f(n)/g(n)\right) \leq 0$ (resp. ≥ 0).

Now consider two codewords $\boldsymbol{c}_{\hat{\imath}}$, $\boldsymbol{c}_{\hat{\jmath}}$ at distance d from each other. Given that $\boldsymbol{c}_{\hat{\imath}}$ is transmitted over BSC(p), the probability that the Hamming distance of the received word from $\boldsymbol{c}_{\hat{\jmath}}$ is not more than its distance from $\boldsymbol{c}_{\hat{\imath}}$ is upper bounded as $\Pr\{\boldsymbol{c}_{\hat{\imath}} \to \boldsymbol{c}_{\hat{\jmath}}\} \leq 2^{-d\,\alpha_p}$, where $\alpha_p \triangleq -\log\sqrt{4p(1-p)}$ [6]. Therefore, for a given codebook $C = C_{\pi_0}$ and permutation $\sigma \in S_m$ with $\sigma \neq \pi_0$, if $d_\sigma \triangleq \mathrm{d_H}(C_{\pi_0}, C_\sigma)$, then it follows that $\Pr\{\pi_0 \to \sigma\} \leq 2^{-d_\sigma \alpha_p}$. From (3) and (12), we get $D(C,p,\phi) = \Pr\left\{\bigcup_{\sigma \in S_m, \sigma \neq \pi_0} \{\pi_0 \to \sigma\}\right\} \leq \sum_{\sigma \in S_m, \sigma \neq \pi_0} \Pr\{\pi_0 \to \sigma\}$, where the last inequality follows from the union bound. Let

$$P_{\mathrm{RC},\sigma} \triangleq \mathbb{E}\left[\Pr\{\pi_0 \to \sigma\}\right], \tag{13}$$

where the expectation is over an ensemble of random binary codebooks. From (4) and (13), we get

$$D_{\mathrm{RC}}(n,R,p,\phi) \leq \sum_{\sigma \in S_m, \sigma \neq \pi_0} P_{\mathrm{RC},\sigma}. \tag{14}$$

Next, we quantify $P_{\mathrm{RC},\sigma}$ for different $\sigma \in S_m$.

A. σ is a transposition

We first consider the case where σ is a *transposition*, i.e. a permutation that interchanges only two indices. For indices $\hat{\imath}, \hat{\jmath}$, with $1 \leq \hat{\imath} < \hat{\jmath} \leq m$, the Hamming distance between codewords $\boldsymbol{c}_{\hat{\imath}}$ and $\boldsymbol{c}_{\hat{\jmath}}$ in a random codebook satisfies [6]

$$\Pr\left\{\mathrm{d_H}(\boldsymbol{c}_{\hat{\imath}}, \boldsymbol{c}_{\hat{\jmath}}) = d\right\} \leq 2^{-n(1-H(d/n))}. \tag{15}$$

Note that if $\mathrm{d_H}(\boldsymbol{c}_{\hat{\imath}}, \boldsymbol{c}_{\hat{\jmath}}) = d$ and $\sigma = (\hat{\imath}\ \hat{\jmath})$ denotes the permutation that only interchanges indices $\hat{\imath}$ and $\hat{\jmath}$, then $\mathrm{d_H}\left(C_{\pi_0}, C_{(\hat{\imath}\ \hat{\jmath})}\right) = 2d$, and it follows from (15) that

$$\Pr\left\{\mathrm{d_H}\left(C_{\pi_0}, C_{(\hat{\imath}\ \hat{\jmath})}\right) = 2d\right\} \leq 2^{-n(1-H(d/n))}. \tag{16}$$

Further, when $\mathrm{d_H}\left(C_{\pi_0}, C_{(\hat{\imath}\ \hat{\jmath})}\right) = 2d$, we have $\Pr\{\pi_0 \to (\hat{\imath}\ \hat{\jmath})\} \leq 2^{-2d\alpha_p}$. Now the probability $P_{\mathrm{RC},(\hat{\imath}\ \hat{\jmath})}$ can be characterized using (13) and (16) as

$$P_{\mathrm{RC},(\hat{\imath}\ \hat{\jmath})} \leq \sum_{d=0}^{n} 2^{-n(1-H(d/n)+2(d/n)\alpha_p)}. \tag{17}$$

If $\delta = d/n$ is treated as a continuous variable, then the exponent $E_2(\delta) \triangleq 1 - H(\delta) + 2\delta\alpha_p$ is a convex function with a unique minimum at $\delta = \hat{\delta}_p \triangleq \frac{4p(1-p)}{1+4p(1-p)}$. If we define $R_1(p)$ as

$$R_1(p) \triangleq 1 - \log(1 + 4p(1-p)), \tag{18}$$

then it can be verified that $E_2(\hat{\delta}_p) = R_1(p)$, and it follows from (17) that when σ is a transposition, we have

$$P_{\mathrm{RC},\sigma} \leq 2^{-n(R_1(p)-c_n)}, \tag{19}$$

where $c_n \triangleq (\log(n+1))/n$.

B. σ is a product (composition) of disjoint transpositions

We now consider the case where $\sigma = \sigma_1\sigma_2$, where σ_1 and σ_2 are disjoint transpositions with $\sigma_1 = (i\ j)$ and $\sigma_2 = (\hat{\imath}\ \hat{\jmath})$. As the codewords in a random codebook are independent, then using (15), we have $\Pr\left\{\{\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_j) = d_1\} \cap \{\mathrm{d_H}(\boldsymbol{c}_{\hat{\imath}}, \boldsymbol{c}_{\hat{\jmath}}) = d_2\}\right\} \leq \prod_{i=1}^{2} 2^{-n(1-H(d_i/n))}$. Further, if $\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_j) = d_1$ and $\mathrm{d_H}(\boldsymbol{c}_{\hat{\imath}}, \boldsymbol{c}_{\hat{\jmath}}) = d_2$, then $\mathrm{d_H}\left(C_{\pi_0}, C_\sigma\right) = 2(d_1 + d_2)$, and $\Pr\{\pi_0 \to \sigma\} \leq 2^{-2(d_1+d_2)\alpha_p}$. Therefore, if σ is a product of two disjoint transpositions, then

$$\begin{aligned} P_{\mathrm{RC},\sigma} &\leq \sum_{d_1,d_2} 2^{-n\left(\sum_{i=1}^{2}(1-H(d_i/n)+2(d_i/n)\alpha_p)\right)}, \\ &\leq 2^{-2n(R_1(p)-c_n)}. \end{aligned} \tag{20}$$

In general, when σ is a product of s disjoint transpositions, the above argument can be readily extended to show that

$$P_{\mathrm{RC},\sigma} \leq 2^{-sn(R_1(p)-c_n)}. \tag{21}$$

Now define $\lambda_p \triangleq \min\left\{\frac{R_1(p)}{2}, \frac{2R_0(p)}{3}\right\}$, where $R_0(p)$ and $R_1(p)$ are defined in (8) and (18), respectively. As $2\lambda_p \leq R_1(p)$, it follows from (21) that

$$P_{\mathrm{RC},\sigma} \leq 2^{-sn2(\lambda_p-c_n)}. \tag{22}$$

We remark that when σ is just a transposition, then from (19) we have $P_{\mathrm{RC},\sigma} \leq 2^{-n(R_1(p)-c_n)} \leq 2^{-n2(\lambda_p-c_n)}$, which is only a special case of (22) with $s = 1$.

C. σ is a k-cycle with $k > 2$

We will apply the following proposition towards characterizing $P_{\mathrm{RC},\sigma}$ when σ is a k-cycle with $k > 2$.

Proposition 1. *Let $\mathbb{F}_{2^n}$ denote the space of all n-length binary vectors. Let $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_k$ be $k > 2$ i.i.d. random vectors, uniformly distributed over $\mathbb{F}_{2^n}$, and let $d_1, d_2, \ldots, d_{k-1}$ be given non-negative integers. Then the following holds*

$$\Pr\left\{\bigcap_{i=1}^{k-1}\{\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_{i+1}) = d_i\}\right\} \leq \prod_{i=1}^{k-1} 2^{-n(1-H(d_i/n))}. \tag{23}$$

Proof: See Appendix A. ■

Let $\sigma \in S_m$ be a k-cycle $(i_1\ i_2\ \cdots\ i_k)$ where $i_{l+1} = \sigma(i_l)$ for $1 \leq l \leq k-1$, and $i_1 = \sigma(i_k)$. Let $d_1, \ldots, d_k$ be non-negative integers. For a given codebook C, if $\mathrm{d_H}(\boldsymbol{c}_{i_l}, \boldsymbol{c}_{i_{l+1}}) = d_l$ for $1 \leq l \leq k-1$, and $\mathrm{d_H}(\boldsymbol{c}_{i_k}, \boldsymbol{c}_{i_1}) = d_k$, then $\mathrm{d_H}(C_{\pi_0}, C_\sigma) = \sum_{l=1}^{k} d_l$, and therefore

$$\Pr\{\pi_0 \to \sigma\} \leq 2^{-\left(\sum_{l=1}^{k} d_l\right)\alpha_p}. \tag{24}$$

Further, if codebook C is uniformly distributed over $\mathscr{C}(n,R)$,

$$\begin{aligned} &\Pr\left\{\left(\bigcap_{l=1}^{k-1}\left\{\mathrm{d_H}(\boldsymbol{c}_{i_l}, \boldsymbol{c}_{i_{l+1}}) = d_l\right\}\right)\bigcap\left\{\mathrm{d_H}(\boldsymbol{c}_{i_k}, \boldsymbol{c}_{i_1}) = d_k\right\}\right\} \\ &\qquad \leq 2^{-n\left(\sum_{l=1}^{k-1}(1-H(d_l/n))\right)}, \end{aligned} \tag{25}$$

where (25) follows from (23). Combining (24) and (25),

$$\begin{aligned} P_{\mathrm{RC},\sigma} &\leq \sum_{\substack{0 \leq d_l \leq n,\\ 1 \leq l \leq k}} 2^{-n\left(\left(\sum_{l=1}^{k}(d_l/n)\alpha_p\right)+\left(\sum_{l=1}^{k-1}(1-H(d_l/n))\right)\right)}, \\ &= \left(\sum_{d_k=0}^{n} 2^{-d_k\alpha_p}\right)\left(\prod_{l=1}^{k-1}\sum_{d_l=0}^{n} 2^{-n(1-H(d_l/n)+(d_l/n)\alpha_p)}\right) \\ &\leq 2^{nc_n}\left(\prod_{l=1}^{k-1}\sum_{d_l=0}^{n} 2^{-n(1-H(d_l/n)+(d_l/n)\alpha_p)}\right). \end{aligned} \tag{26}$$

If $\delta = d_l/n$ is treated as a continuous variable, then the exponent $E_1(\delta) \triangleq 1 - H(\delta) + \delta\alpha_p$ is a convex function with a unique minimum at $\delta = \tilde{\delta}_p \triangleq \frac{\sqrt{4p(1-p)}}{1+\sqrt{4p(1-p)}}$. We have $E_1(\tilde{\delta}_p) = 1 - \log(1 + \sqrt{4p(1-p)}) = R_0(p)$, and it follows from (26) that $P_{\mathrm{RC},\sigma} \leq 2^{-n((k-1)R_0(p)-kc_n)}$. As $2k/3 \leq k-1$ for $k > 2$, we have $k\lambda_p \leq (k-1)R_0(p)$,

$$P_{\mathrm{RC},\sigma} \leq 2^{-nk(\lambda_p - c_n)}, \tag{27}$$

where σ is a k-cycle with $k > 2$. A transposition is just a k-cycle with $k = 2$, and from the remark following (22), it follows that (27) holds even for $k = 2$.

D. General $\sigma \in S_m$ with $\sigma \neq \pi_0$

It is well known that any permutation $\sigma \neq \pi_0$ can be written as a product (composition) of t disjoint cycles, for $t \geq 1$ [7]. Consider a given σ which is a product of t disjoint cycles of length $k_1, \ldots, k_t$, respectively, where $k_i \geq 2$ for $1 \leq i \leq t$. Then, extending the approach employed in Sec. IV-B, and using (27), we obtain

$$P_{\mathrm{RC},\sigma} \leq 2^{-n\left(\sum_{i=1}^{t} k_i\right)(\lambda_p - c_n)}. \tag{28}$$

E. Putting it all together

Define $\Sigma_j \triangleq \{\sigma \in S_m : |\{i : \sigma(i) \neq i, 1 \leq i \leq m\}| = j\}$, and $P_{\mathrm{RC},\Sigma_j} \triangleq \sum_{\sigma \in \Sigma_j} P_{\mathrm{RC},\sigma}$, for $1 \leq j \leq m$, to equivalently express (14) as

$$D_{\mathrm{RC}}(n, R, p, \phi) \leq \sum_{j=2}^{m} P_{\mathrm{RC},\Sigma_j}. \tag{29}$$

We have $|\Sigma_2| = \binom{m}{2} \leq 2^{n(2R)}$. For all $\sigma \in \Sigma_2$, the value of $P_{\mathrm{RC},\sigma}$ is upper bounded by (19), and so

$$P_{\mathrm{RC},\Sigma_2} \leq 2^{-n(R_1(p) - c_n - 2R)}. \tag{30}$$

For a given $j > 2$, if $\sigma \in \Sigma_j$, then from (28) it follows that $P_{\mathrm{RC},\sigma} \leq 2^{-nj(\lambda_p - c_n)}$. For $j > 2$, the size of the set Σ_j satisfies $|\Sigma_j| < \prod_{i=0}^{j-1}(m-i) \leq 2^{njR}$. If we define $\beta_n \triangleq 2^{-n(\lambda_p - c_n - R)}$, then we have $P_{\mathrm{RC},\Sigma_j} \leq \beta_n^j$. Now, if $R < \lambda_p$, then because $c_n = o(1)$, there exists N such that for $n \geq N$, we have $R < \lambda_p - c_n$ and hence $\beta_n < 1$. Therefore, for $n \geq N$,

$$\sum_{j=3}^{m} P_{\mathrm{RC},\Sigma_j} \leq \sum_{j=3}^{m} \beta_n^j \leq \frac{\beta_n^3}{1 - \beta_n}. \tag{31}$$

As $\beta_n \to 0$ and $c_n \to 0$ when $n \to \infty$, it follows from (31),

$$\sum_{j=3}^{m} P_{\mathrm{RC},\Sigma_j} \leq \frac{\beta_n^3}{1 - \beta_n} \doteq \beta_n^3 \doteq 2^{-3n(\lambda_p - R)}. \tag{32}$$

Combining (29), (30), and (32), for $R < \lambda_p$,

$$D_{\mathrm{RC}}(n, R, p, \phi) \,\dot{\leq}\, 2^{-n(R_1(p) - 2R)} + 2^{-n(3\lambda_p - 3R)}. \tag{33}$$

Comparing (14) with (33), we observe that $D_{\mathrm{RC}}(n, R, p, \phi)$ is dominated by $P_{\mathrm{RC},\sigma}$ terms for σ corresponding to k-cycles with $k = 2$ and $k = 3$. The next theorem presents an explicit lower bound for $E_{D_{\mathrm{RC}}}(R, p, \phi)$ when ϕ jointly decodes all the barcodes using a maximum likelihood approach.

Theorem 2. *We have*

$$E_{D_{\mathrm{RC}}}(R, p, \phi) \geq |\eta_p(R)|^+, \tag{34}$$

where $\eta_p(R) \triangleq \min\{R_1(p) - 2R,\ 2R_0(p) - 3R\}$.

Proof: When $R < \lambda_p$ then we have $R_1(p) > 2R$. Therefore, from (33) it follows that if $R < \lambda_p$, then $E_{D_{\mathrm{RC}}}(R, p, \phi)$ is lower bounded by $\min\{R_1(p) - 2R,\ 3\lambda_p - 3R\} = \eta_p(R)$. Note that $\eta_p(R) > 0$ if and only if $R < \lambda_p$. ■

V. Bee-Identification Exponent: Upper Bound

Define the following optimum minimum distance metrics: $d^*(n, R) \triangleq \max_{C \in \mathscr{C}(n,R)} \min_{\boldsymbol{c}_i \neq \boldsymbol{c}_j} \mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_j)$, $\delta^*(n, R) \triangleq d^*(n, R)/n$, and $\delta^*(R) \triangleq \limsup_{n\to\infty} \delta^*(n, R)$. For any given codebook $C \in \mathscr{C}(n, R)$, we show that there exists a set $\mathscr{I}_C$ consisting of pairs of codeword indices (i, j), $i \neq j$, with the following three properties: (i) If $(i, j) \in \mathscr{I}_C$, then $\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_j) \leq d^*(n, R - \frac{1}{n})$, (ii) If $(i, j) \in \mathscr{I}_C$ and $(\hat{\imath}, \hat{\jmath}) \in \mathscr{I}_C$, then $\hat{\imath} \neq i, \hat{\imath} \neq j$ and $\hat{\jmath} \neq i, \hat{\jmath} \neq j$, (iii) Size of set $\mathscr{I}_C$ is equal to $m/4$. A set satisfying the above properties can be constructed iteratively as follows

- *Step 1*: For a given codebook $C \in \mathscr{C}(n, R)$, initialize $\mathscr{I}_C$ to be the empty set and let $\mathcal{T} = C$.
- *Step 2*: As $|\mathcal{T}| \geq m/2$, there exists $\boldsymbol{c}_i, \boldsymbol{c}_j \in \mathcal{T}$, with $i \neq j$, satisfying $\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_j) \leq d^*(n, R - \frac{1}{n})$. Include the pair (i, j) to $\mathscr{I}_C$, and let $\mathcal{T} = \mathcal{T} \setminus \{\boldsymbol{c}_i, \boldsymbol{c}_j\}$.
- *Step 3*: If $|\mathscr{I}_C| < m/4$, then go to *Step 2*, else stop.

Assume that the receiver employs ML decoding, and interpret each pair $(i, j) \in \mathscr{I}_C$ as a transposition $\sigma = (i\ j)$. Let $A_{(i,j)}$ denote the error event that the receiver incorrectly decodes the channel induced permutation to transposition $(i\ j)$ (instead of the identity permutation π_0), i.e. $A_{(i,j)} = \{\pi_0 \to (i, j)\}$. Then,

$$D(C, p, \phi) \geq \Pr\left\{ \bigcup_{(i,j)\in\mathscr{I}_C} A_{(i,j)} \right\}. \tag{35}$$

Using de Caen's lower bound [8], the expression on the right side in (35) can itself be lower bounded by

$$\sum_{(i,j)\in\mathscr{I}_C} \frac{\left(\Pr\{A_{(i,j)}\}\right)^2}{\Pr\{A_{(i,j)}\} + \sum\limits_{\substack{(\hat{\imath},\hat{\jmath})\in\mathscr{I}_C \\ (\hat{\imath},\hat{\jmath})\neq(i,j)}} \Pr\left\{A_{(i,j)} \cap A_{(\hat{\imath},\hat{\jmath})}\right\}},$$

$$\overset{(a)}{=} \sum_{(i,j)\in\mathscr{I}_C} \frac{\left(\Pr\{A_{(i,j)}\}\right)^2}{\Pr\{A_{(i,j)}\} + \sum\limits_{\substack{(\hat{\imath},\hat{\jmath})\in\mathscr{I}_C \\ (\hat{\imath},\hat{\jmath})\neq(i,j)}} \Pr\left\{A_{(i,j)}\right\} \Pr\left\{A_{(\hat{\imath},\hat{\jmath})}\right\}},$$

$$\geq \frac{\sum\limits_{(i,j)\in\mathscr{I}_C} \Pr\{A_{(i,j)}\}}{1 + \sum\limits_{(\hat{\imath},\hat{\jmath})\in\mathscr{I}_C} \Pr\left\{A_{(\hat{\imath},\hat{\jmath})}\right\}}, \tag{36}$$

where (a) follows because events $A_{(i,j)}$ and $A_{(\hat{\imath},\hat{\jmath})}$ are independent when $(\hat{\imath}, \hat{\jmath}) \neq (i, j)$. Now

$$\begin{aligned} \sum_{(i,j)\in\mathscr{I}_C} \Pr\{A_{(i,j)}\} &\overset{(b)}{\dot{\geq}} \sum_{(i,j)\in\mathscr{I}_C} 2^{-n\left(2\delta^*\left(n, R-\frac{1}{n}\right)\alpha_p\right)}, \\ &\overset{(c)}{\geq} 2^{-n\left(2\delta^*(n, R-\frac{1}{n})\alpha_p - \left(R - \frac{2}{n}\right)\right)}, \\ &\doteq 2^{-n(2\delta^*(R)\alpha_p - R)}, \end{aligned} \tag{37}$$

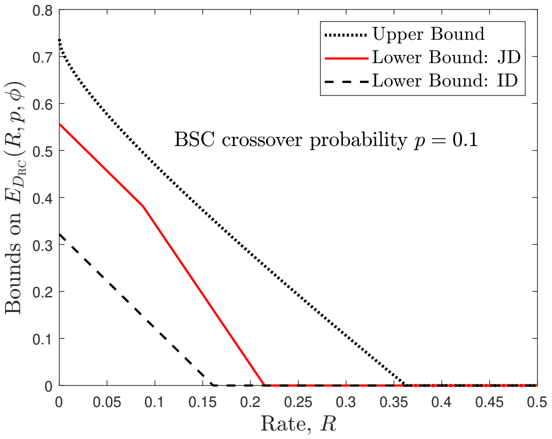

Fig. 3: Lower bounds on $E_{D_{\mathrm{RC}}}(R, p, \phi)$ using independent decoding (ID) (see (10)) and joint decoding (JD) (see (34)) of barcodes. Upper bound (41) holds for all code sequences.

where (b) follows as $\mathrm{d_H}(C_{\pi_0}, C_{(i,j)}) \le 2\, d^*(n, R - \frac{1}{n})$ for $(i,j) \in \mathscr{I}_C$, and (c) follows because $|\mathscr{I}_C| \ge m/4$. If $R_{\mathrm{UB}}(p) \triangleq \sup\{R : 2\delta^*(R)\alpha_p > R\}$, then combining (35), (36), (37), and noting that $x/(1+x)$ increases with x,

$$D(C, p, \phi) \,\dot{\ge}\, \frac{2^{-n(2\delta^*(R)\alpha_p - R)}}{1 + 2^{-n(2\delta^*(R)\alpha_p - R)}}, \\ \doteq 2^{-n(2\delta^*(R)\alpha_p - R)}, \quad 0 \le R < R_{\mathrm{UB}}(p). \tag{38}$$

As (38) is true for all $C \in \mathscr{C}(n, R)$, we have

$$D_{\mathrm{RC}}(n, R, p, \phi) \,\dot{\ge}\, 2^{-n(2\delta^*(R)\alpha_p - R)}, \quad 0 \le R < R_{\mathrm{UB}}(p). \tag{39}$$

The value $\delta^*(R)$ can be upper bounded as [9], [10]

$$\delta^*(R) \le \delta_{\mathrm{LP}}(R) \triangleq \frac{1}{2} - \sqrt{\delta_{\mathrm{GV}}(1-R)(1 - \delta_{\mathrm{GV}}(1-R))}. \tag{40}$$

Theorem 3. *We have*

$$E_{D_{\mathrm{RC}}}(R, p, \phi) \le |2\delta^*(R)\alpha_p - R|^+ \le |2\delta_{\mathrm{LP}}(R)\alpha_p - R|^+. \tag{41}$$

Proof: Follows immediately from (39) and (40). ■

Fig. 3 plots different bounds on $E_{D_{\mathrm{RC}}}(R, p, \phi)$. As (38) holds for all $C \in \mathscr{C}(n, R)$, we observe that the upper bound (41) is applicable to all possible codebook designs.

VI. Discussion

We demonstrated that joint decoding of barcodes provides much better random coding exponent compared to independent decoding. We also gave an explicit upper bound on the bee-identification exponent that is applicable to all possible codebook designs. Future work will characterize the exponent for *typical* random codes. Other extension of this work includes the scenario where the bee-identification error is flagged only when the fraction of incorrectly decoded barcodes exceeds a threshold. Another interesting scenario is when some of the m rows in codebook C are deleted, due to some bees being outside the hive when taking the picture.

Appendix A
Proof of Prop. 1

Let $\gamma_{k-1}, \tilde{\gamma}_{k-1} \in \mathbb{F}_{2^n}$, and $\Delta \triangleq \gamma_{k-1} \oplus \tilde{\gamma}_{k-1}$, where $\oplus$ denotes modulo-2 addition. Then, $\Pr\{\mathrm{d_H}(\gamma_{k-1}, \boldsymbol{c}_k) = d_{k-1}\} = \Pr\{\mathrm{d_H}(\tilde{\gamma}_{k-1}, \boldsymbol{c}_k + \Delta) = d_{k-1}\} \overset{\text{(i)}}{=} \Pr\{\mathrm{d_H}(\tilde{\gamma}_{k-1}, \boldsymbol{c}_k) = d_{k-1}\}$, where (i) follows from the fact that the distribution of $\boldsymbol{c}_k + \Delta$ is same as the distribution of $\boldsymbol{c}_k$. Thus $\Pr\{\mathrm{d_H}(\boldsymbol{c}_{k-1}, \boldsymbol{c}_k) = d_{k-1} | \boldsymbol{c}_{k-1} = \gamma_{k-1}\} \overset{\text{(ii)}}{=} \Pr\{\mathrm{d_H}(\boldsymbol{c}_{k-1}, \boldsymbol{c}_k) = d_{k-1}\}$. Then $\Pr\{\bigcap_{i=1}^{k-1}\{\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_{i+1}) = d_i\}\}$ is equivalently expressed as

$$\begin{aligned}
&\sum_{\gamma_1,\ldots,\gamma_{k-1} \in \mathbb{F}_{2^n}} \left(\Pr\left\{ \bigcap_{i=1}^{k-1} \{\boldsymbol{c}_i = \gamma_i\} \right\} \right. \\
&\qquad \left. \times \Pr\left\{ \bigcap_{i=1}^{k-1} \{\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_{i+1}) = d_i\} \,\middle|\, \bigcap_{i=1}^{k-1} \{\boldsymbol{c}_i = \gamma_i\} \right\} \right), \\
&= \sum_{\gamma_1,\ldots,\gamma_{k-1}} \left(\Pr\left\{ \bigcap_{i=1}^{k-1} \{\boldsymbol{c}_i = \gamma_i\} \right\} \mathbf{1}_{\{\bigcap_{i=1}^{k-2}\{\mathrm{d_H}(\gamma_i, \gamma_{i+1}) = d_i\}\}} \right. \\
&\qquad \left. \times \Pr\left\{ \mathrm{d_H}(\boldsymbol{c}_{k-1}, \boldsymbol{c}_k) = d_{k-1} \middle| \boldsymbol{c}_{k-1} = \gamma_{k-1} \right\} \right), \\
&\overset{\text{(iii)}}{=} \sum_{\gamma_1,\ldots,\gamma_{k-1}} \left(\Pr\left\{ \bigcap_{i=1}^{k-1} \{\boldsymbol{c}_i = \gamma_i\} \right\} \mathbf{1}_{\{\bigcap_{i=1}^{k-2}\{\mathrm{d_H}(\gamma_i, \gamma_{i+1}) = d_i\}\}} \right. \\
&\qquad \left. \times \Pr\left\{ \mathrm{d_H}(\boldsymbol{c}_{k-1}, \boldsymbol{c}_k) = d_{k-1} \right\} \right), \\
&= \Pr\left\{ \bigcap_{i=1}^{k-2} \mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_{i+1}) = d_i \right\} \Pr\left\{ \mathrm{d_H}(\boldsymbol{c}_{k-1}, \boldsymbol{c}_k) = d_{k-1} \right\},
\end{aligned} \tag{42}$$

where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function, (iii) follows from (ii), while (23) follows by recursively applying (42) and using the fact that $\Pr\{\mathrm{d_H}(\boldsymbol{c}_i, \boldsymbol{c}_{i+1}) = d_i\} \le 2^{-n(1 - H(d_i/n))}$ when $\boldsymbol{c}_i$ and $\boldsymbol{c}_{i+1}$ are uniformly distributed over $\mathbb{F}_{2^n}$ [6].

References

[1] T. Gernat *et al.*, "Automated monitoring of behavior reveals bursty interaction patterns and rapid spreading dynamics in honeybee social networks," *Proc. Nat. Acad. Sci. U.S.A.*, vol. 115, no. 7, pp. 1433–1438, Feb. 2018.

[2] S. Shahi, D. Tuninetti, and N. Devroye, "The strongly asynchronous massive access channel," Jul. 2018, arXiv:1807.09934 [cs.IT].

[3] S. Shahi, D. Tuninetti, and N. Devroye, "On identifying a massive number of distributions," in *Proc. 2018 IEEE Int. Symp. Inf. Theory*, Jun. 2018, pp. 331–335.

[4] I. Shomorony and R. Heckel, "Capacity results for the noisy shuffling channel," Feb. 2019, arXiv:1902.10832 [cs.IT].

[5] R. G. Gallager, *Information Theory and Reliable Communication*. New York: John Wiley and Sons, 1968.

[6] A. Barg and G. D. Forney, "Random codes: minimum distances and error exponents," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2568–2573, Sep. 2002.

[7] I. Herstein, *Topics In Algebra*, 2nd ed. John Wiley and Sons, New York, 1975.

[8] D. de Caen, "A lower bound on the probability of a union," *Discrete Mathematics*, vol. 169, no. 1, pp. 217 – 220, May 1997.

[9] R. McEliece, E. Rodemich, H. Rumsey, and L. Welch, "New upper bounds on the rate of a code via the Delsarte-MacWilliams inequalities," *IEEE Trans. Inform. Theory*, vol. 23, no. 2, pp. 157–166, Mar. 1977.

[10] S. Litsyn, "New upper bounds on error exponents," *IEEE Trans. Inform. Theory*, vol. 45, no. 2, pp. 385–398, Mar. 1999.