Achievable Rate Region for Iterative Multi-User Detection via Low-cost Gaussian Approximation: Draft of Journal Version

Xiaojie Wang, *Student Member, IEEE,* Chulong Liang, Li Ping, *Fellow, IEEE,* and Stephan ten Brink, *Senior Member, IEEE*

***Abstract*—We establish a multiuser extrinsic information transfer (EXIT) chart area theorem for the interleave-division multiple access (IDMA) scheme, a special form of superposition coding, in multiple access channels (MACs). A low-cost multi-user detection (MUD) based on the Gaussian approximation (GA) is assumed. The evolution of mean-square errors (MSE) of the GA-based MUD during iterative processing is studied. We show that the K-dimensional tuples formed by the MSEs of K users constitute a conservative vector field. The achievable rate is a potential function of this conservative field, so it is the integral along any path in the field with value of the integral solely determined by the two path terminals. Optimized codes can be found given the integration paths in the MSE fields by matching EXIT type functions. The above findings imply that i) low-cost GA detection can provide near capapcity performance, ii) the sum-rate capacity can be achieved independently of the integration path in the MSE fields; and iii) the integration path can be an extra degree of freedom for code design.**

***Index Terms*—EXIT chart, non-orthogonal multiple access, area theorem, MAC capacity.**

I. Introduction

The multiple access channel (MAC) capacity region is determined by a tuple of rates R_k, $1 \leq k \leq K$ of the K individual users [1], [2]. The achievable rates constitute the MAC capacity region, which is bounded by $2^K - 1$ constraints.

To achieve arbitrary points of the capacity region, joint detection and decoding is required which has prohibitive complexity as the number of users grows. Theoretically, successive interference cancellation (SIC) together with time-sharing or rate-splitting can achieve the entire capacity region [3]. SIC involves subtraction of successfully detected signals. If practical forward error control (FEC) codes are used, each subtraction incurs an extra overhead in terms of either power or rate relative to an ideal capacity achieving code [4, Fig. 13.3]. Such overheads accumulate during SIC steps, moving it away from the capacity limit particularly when the number of users is large. Also, both time-sharing and rate-splitting involve segmenting a data frame of a user into several sub-frames. The reduced sub-frame length implies reduced coding gain for a practical turbo or low-density parity-check (LDPC) type code, which further worsens the losses of accumulation.

Iterative detection can alleviate the loss accumulation problem using soft cancellations instead of hard subtraction. A turbo or LDPC code involving iterative detection can be optimized by matching the so-called extrinsic information transfer (EXIT) functions of two local processors [5], [6]. In a single-user point-to-point channel, such matching can offer near capacity performance, as shown the area properties [7], [8].

Interleave-division multiple-access (IDMA) is a low-cost transmission scheme for MACs [9]. A Gaussian approximation (GA) of cross-user interference is key to a low-cost IDMA detector [10]. For comparison, consider a common *a posteriori* probability (APP) multi-user detector (MUD) [9] and let K be the number of users. The per-user complexity of a GA-based MUD remains roughly the same for all K, while that of an APP-based MUD is exponental in K.

A question naturally arises: At such low cost, what is the achievable performance of IDMA under GA-based MUD? Some partial answers to this question are available. It is shown that IDMA is capacity approaching when all users see the same channel [11]. However, for the general MAC system, the code design for IDMA becomes difficult since the performance of the iterative detection and decoding is determined by K different parameters, not a single parameter in equal-power equal-rate case, i.e., all users have the same power and code rate. Previous works on IDMA focus on the power

Part of the results is under review of the IEEE International Symposium on Information Theory 2019, Paris, France.

This is a draft of the Journal Version of the submitted ISIT 19 Paper, further providing numerical results.

X.J. Wang and S. ten Brink are with Institute of Telecommunications, Pfaffenwaldring 47, University of Stuttgart, 70569 Stuttgart, Germany (e-mail: {wang, tenbrink}@inue.uni-stuttgart.de).

C. Liang and L. Ping are with the Department of Electronic Engineering, City University of Hong Kong, Hong Kong SAR, China (e-mail: {chuliang, eeliping}@cityu.edu.hk).

optimization [12], [13] and code design for equal-power equal-rate scenario [14]. It is also known that the GA-based MUD can achieve some points in the capacity region for multiple-input multiple-output (MIMO) MACs [15]. To the best of the authors' knowledge, no previous work has shown that IDMA under the GA-based MUD can achieve the entire MAC capacity region.

This paper provides a comprehensive analysis of the achievable performance of IDMA under GA-based MUD. We approach the problem based on multi-dimensional curve matching. Let v_k be the mean-square error (MSE) (i.e., the variance) for the GA-based MUD for user k, with $v_k = 0$ indicating perfect decoding. Using the mutual information (MI) and minimum MSE (MMSE) theorem [8], [16], we show that the achievable sum-rate can be evaluated using a line integral along a valid path in the K-dimensional vector field $\boldsymbol{v} = [v_1, v_2, \cdots, v_K]^T$. Furthermore, the integral is path-independent and its value is solely determined by the two terminations. The path independence property greatly simplifies the code optimization problem. We gain some interesting insights from the discussions in this paper.

- A low-cost GA-based MUD can provide near optimal performance. In particular, it is provably capacity-achieving for Gaussian signaling.
- The achievable rates with finite modulation is also studied, we show that the loss to Gaussian capacity can be made arbitrarily small by imposing a larger number of users or data layers.
- FEC codes optimized for single-user channels may not be good choices for MACs. The FEC codes should be carefully designed to match the GA-based MUD, which facilitates iterative detection. We will provide examples for the related code design.
- A multi-user area theorem of EXIT chart is established for the code design.
- The sum-rate capacity is a potential function in the MSE field formed by $\boldsymbol{v}$, which leads to the path independence property.
- All points of the MAC capacity region are achievable using only one FEC code per user. This avoids the loss related to the frame segmentation of SIC as aforementioned.
- The above results can be extended to MIMO MAC channels straightforwardly.

We will provide simulation results to show that properly designed IDMA can approach the sum-rate MAC capacity for different decoding paths in the MSE field within $1\,\mathrm{dB}$.

This paper is structured as follows. In Sec. II, we present the multiuser iterative detection and decoding scheme in IDMA along with the matching condition. Then, we derive the achievable rates of IDMA and show its implication in code design in single antenna setup in Sec. III. The achievable rate analysis is further extend to MIMO cases in Sec. IV. Sec. V provides code design examples and numerical results verifying our theorems. Finally, Sec. VI concludes the paper.

II. Iterative IDMA Receiver

Consider a general K-user MAC system, which is described by

$$y = \sum_{i=1}^{K} \sqrt{P_i} h_i x_i + n$$

where P_i denotes the received signal strength of the ith user's signal, h_i denotes the fading coefficients of the user, x_i is the ith transmit signal and n is the additive white Gaussian noise (AWGN) with zero mean and unit variance, i.e., $\mathcal{CN}\left(0, \sigma^2 = 1\right)$.

The iterative receiver is depicted in Fig. 1. The

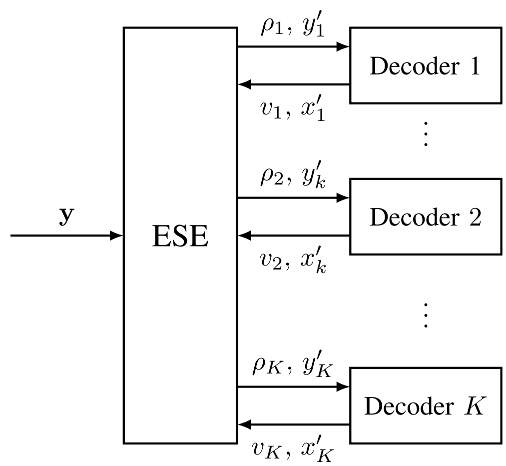

Fig. 1. The iterative multi-user detection and decoding model in IDMA.

elementary signal estimator (ESE) module has access to the channel observation y and feedbacks x'_k from all the users' decoders. It performs the so called soft interference cancellation (SoIC) and provides each decoder a "clean observation", i.e., a signal with reduced interference. The decoders fetch the corresponding signals from ESE and performs its decoding while treating the residual signals of other users as noise. Thus, we refer to the decoder as "single user decoder". Through the iterative message passing between the ESE and the single-user decoders, every user can decode its signal.

To gurantee that the signals can be perfectly recovered, properly designed FEC codes shall be applied. This can be done by, e.g., using EXIT chart based design (see [7], [17]).

A. ESE functions

The decoder feedbacks x'_k are characterized by the MSE of its estimates, denoted by v_i. The "decoder observation" after SoIC is characterized by the signal-to-noise ratio (SNR), denoted by ρ_i, assuming that the interference is Gaussian-distributed[1], i.e.,

$$y'_k = \sqrt{\rho_k} x_k + z \tag{2}$$

where z, comprised of AWGN and multi-user interference, is assumed to be Gaussian-distributed $\mathcal{CN}(0,1)$. This assumption greatly simplifies the multi-user detection, henceforth it is referred to as GA-based MUD. Therefore, the ESE transfer function for the ith user is given by

$$\rho_k = \frac{P_k |h_k|^2}{\sum\limits_{j=1, j\neq i}^{K} P_j |h_j|^2 v_j + \sigma^2}, \ \forall k = 1, 2, \ldots, K. \tag{3a}$$

We can also express (3a) in a vector form as

$$\boldsymbol{\rho} = \phi(\boldsymbol{v}) \tag{3b}$$

where $\boldsymbol{\rho} = [\rho_1, \rho_2, \cdots, \rho_K]^T$ and $\boldsymbol{v} = [v_1, v_2, \cdots, v_K]^T$. Due to the fact that the MSE is bounded by $0 \leq v_i \leq \mathrm{E}\left[|x_i|^2\right] = 1$,we obtain that the SNR is also bounded by

$$\rho_{k,\min} = \frac{P_k |h_k|^2}{\sum\limits_{j=1, j\neq i}^{K} P_j |h_j|^2 + \sigma^2} \leq \rho_k \leq \frac{P_k |h_k|^2}{\sigma^2} = \rho_{k,\max}. \tag{3c}$$

This bound implies that the single user decoder shall be able to decode its signal before its ESE input reaches the maximum, i.e., $\exists \rho'_k \leq \rho_{k,\max}, \ v_k(\rho'_k) = 0$.

As the consequence of the iterative processing, we can write the SNR and MSE vector as a function depending on a "time" or "iteration" variable t, i.e., $\boldsymbol{\rho} = \boldsymbol{\rho}(t)$, $\boldsymbol{v} = \boldsymbol{v}(t)$ and $\boldsymbol{\rho}(t) = \phi(\boldsymbol{v}(t))$.

B. DEC functions

The decoder (DEC) transfer functions can be characterized by

$$v_k = \psi_k(\rho_k), 0 \leq v_k \leq 1, \ \forall k = 1, 2, \ldots, K. \tag{4a}$$

Similar to (3b), we can write (4a) in a vector form as

$$\boldsymbol{v} = \psi(\boldsymbol{\rho}). \tag{4b}$$

Note that the boundaries for the uncooperative MAC[2] are given by

$$\boldsymbol{v} = \psi(\boldsymbol{\rho}(t=0)) = \mathbf{1} \tag{4c}$$

$$\boldsymbol{v} = \psi(\boldsymbol{\rho}(t=\infty)) = \mathbf{0} \tag{4d}$$

where t is a variable which addresses the evolution of the SNR $\boldsymbol{\rho}$ or MSE $\boldsymbol{v}$ through iterative processing. (4c) indicates that no a priori information is present to the ESE at the beginning of iterations. (4d) ensures error-free decoding at the end. The decoders are typically APP decoders, so that the MSEs $\boldsymbol{v}$ are also the conditional MMSE, i.e., $v_k = \mathrm{E}\left[|x_k - \mathrm{E}[x_k | x'_k]|^2\right]$. Moreover, it is commonly assumed that the symbol estimates after APP decoding can be modeled as an observation from the AWGN channel, i.e.,

$$x'_k = \sqrt{\rho'_k} x_k + w \tag{5}$$

where w follows a Gaussian-distribution $\mathcal{CN}(0,1)$.

C. Matching condition

The matching codes which allow error-free decoding yet with highest code rate (will be shown in Sec. III-A) shall satisfy

$$\psi(\boldsymbol{\rho}(t)) = \phi^{-1}(\boldsymbol{\rho}(t)). \tag{6}$$

In other words, it is sufficient to match the code components along a K-dimensional line, which is given by $\boldsymbol{\rho}(t)$. It is noteworthy to mention that it is not necessary to match the functions in the entire K-dimensional space, i.e., requiring $\psi(\boldsymbol{\rho}) = \phi^{-1}(\boldsymbol{\rho}), \forall \boldsymbol{\rho}$. Matching along the path given by $\boldsymbol{\rho}(t)$ is much easier and achieves the MAC capacity (see Sec. III-A).

III. Achievable rates

The fundamental relation between achievable rate and MMSE in AWGN channels $y = x + n$ is found by Guo et. al. [16] as

$$R(\mathrm{snr}) = \int_0^{\mathrm{snr}} \mathrm{mmse}(\rho)\, d\rho.$$

for any input distribution of x. Later in [8], [18], it is extended to iterative decoding scheme involving side information from a priori channel and extrinsic channel.

$$R = \int_0^{\infty} \mathrm{mmse}(\rho_{ap} + \rho_{ext})\, d\rho_{ap}.$$

The achievable rates in IDMA under the Gaussian assumptions in (2) and (5) can be written as [16], [18]

$$R_k = \int_0^{\infty} f\left(\rho_k + f^{-1}(v_k)\right) d\rho_k, \ \forall k = 1, 2, \ldots, K. \tag{7a}$$

where $f(\rho_k) = v_k$ denotes the MMSE as a function of SNR.

[1]The Gaussian assumption is valid for a large number of users with arbitrary transmit symbols x_i as the consequence of the central limit theorem or if the transmit signals x_i are Gaussian by themselves.

[2]For MAC with cooperative encoders, a more general constraint can be $\mathbf{1}^T \boldsymbol{v}(t=0) = K$ and $v_k \geqslant 0, \ \forall k = 1, 2, \ldots, K$.

A. Gaussian alphabets

We consider Gaussian signals, i.e., x_i are Gaussian distributed which can be achieved by using e.g., superposition coded modulation (SCM) [19]. Therefore, the MMSE is given by $f_G(\rho) = \frac{1}{1+\rho}$, and the achievable rates are

$$R_k = \int_0^\infty \frac{1}{\rho_k + v_k^{-1}} d\rho_k = -\int_{v_k=1}^{v_k=0} \frac{g_k}{\mathbf{g}^T \boldsymbol{v} + \sigma^2} dv_k, \quad \forall k = 1, 2, \ldots, K \tag{7b}$$

where $\mathbf{g}^T = \left[P_1 |h_1|^2, P_2 |h_2|^2, \cdots, P_K |h_K|^2\right]^T$ contains the powers of all users, $\boldsymbol{v} = [v_1, v_2, \cdots, v_K]^T$ and $g_k = P_k |h_k|^2$ denotes the kth element of vector $\mathbf{g}$. The derivation is shown in Appendix A.

The sum-rate of all users is

$$R_{\text{sum}} = \sum_{k=1}^{K} R_k = -\int_{L(t)} \frac{\mathbf{g}}{\mathbf{g}^T \boldsymbol{v}(t) + \sigma^2} \cdot d\boldsymbol{v}(t) \tag{8a}$$

where (8a) denotes a line integral defined by $L = \boldsymbol{v}(t),\ t \in [0, \infty]$. It can be further shown that the integrands constitute a gradient of a scalar field (aka, potential function), i.e., $\nabla_{\boldsymbol{v}} \log\left(\sigma^2 + \mathbf{g}^T \boldsymbol{v}\right) = \frac{\mathbf{g}}{\mathbf{g}^T \boldsymbol{v} + \sigma^2}$. Thus, the achievable sum-rate can be written as

$$\begin{aligned} R_{\text{sum}} &= -\int_{L=\boldsymbol{v}(t)} \left[\nabla \log\left(\sigma^2 + \mathbf{g}^T \boldsymbol{v}\right)\right] \boldsymbol{v}'(t)\, dt \\ &\overset{(4c),(4d)}{=} \log\left(1 + \frac{\sum_{k=1}^{K} P_k |h_k|^2}{\sigma^2}\right) \end{aligned} \tag{8b}$$

which is independent of the path taken for code matching. In other words, any path with matched DEC functions can achieve the sum-rate capacity. The matching condition given in (6) is thus also proved, since it can be easily verified that $R_k < -\int_{v_k=1}^{v_k=0} \frac{g_k}{\mathbf{g}^T \boldsymbol{v} + \sigma^2} dv_k$ and thus $R_{\text{sum}} < \log\left(1 + \frac{\sum_{k=1}^{K} P_k |h_k|^2}{\sigma^2}\right)$, if $\psi(\boldsymbol{\rho}(t)) < \phi^{-1}(\boldsymbol{\rho}(t))$. On the contrary, if $\psi(\boldsymbol{\rho}(t)) > \phi^{-1}(\boldsymbol{\rho}(t))$, error-free decoding is not possible.

This leads to the following theorem.

Theorem 1. *The achievable sum-rate in IDMA for any path $L(t) : \boldsymbol{v}_s = \mathbf{1} \to \boldsymbol{v}_e = \mathbf{0}$ (starting from $\boldsymbol{v}_s = \mathbf{1}$ to $\boldsymbol{v}_e = \mathbf{0}$) is given by*

$$\begin{aligned} R_{\text{sum}} &= -\int_{L(t)} f_G\left(\boldsymbol{\rho}(t) + f_G^{-1}(\boldsymbol{v}(t))\right) \cdot d\boldsymbol{\rho}(t) \\ &= \log\left(1 + \frac{\sum_{k=1}^{K} P_k |h_k|^2}{\sigma^2}\right) \end{aligned}$$

with matching codes defined in $\psi(\boldsymbol{\rho}(t)) = \phi^{-1}(\boldsymbol{\rho}(t))$ and assumptions in (2) *and* (5).

Proof: see above. ■

B. Finite alphabets

If the symbols $x_i \in \mathcal{S}_i$ are taken from finite alphabets $|\mathcal{S}_i| < \infty$, the capacity formula, in general, can not be expressed in closed-form. However, eq. (7a) is still valid, using the MMSE-formula for the underlying modulation format $f(\rho) = v$. It is also well known that the loss incurred by finite alphabets, compared with Gaussian, is negligible in the low-SNR regime. Besides, the Gaussian capacity can be approached by higher order modulations with shaping and/or SCM [19].

We provide in the following an achievable rate analysis for IDMA with quadrature phase shift keying (QPSK) signalling. The users are assumed to have same power P, modulation and coding scheme, for simplicity. Thus, the matching condition can be written as

$$v_k = \frac{1}{(K-1)} \left(\frac{1}{\rho_k} - \frac{\sigma^2}{P}\right), \ \forall k = 1, 2, \ldots, K.$$

and the achievable sum-rate can be written as

$$R_{\text{sum}} = K \cdot \int_0^\infty f_Q\left(\rho_k + f_Q^{-1}(v_k)\right) d\rho_k. \tag{9}$$

where $f_Q(\cdot)$ denotes the MMSE of QPSK, which is given by

$$f_Q(\rho) = 1 - \int_{-\infty}^{\infty} \tanh(\rho - y\sqrt{\rho}) \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy. \tag{10}$$

In AWGN channels, Gaussian signals are the hardest to estimate [20], i.e.,

$$f_X(\rho) \le f_G(\rho) = \frac{1}{1+\rho}.$$

for any input distribution X with the same variance. Hence, the achievable rate with distributions other than Gaussian is essentially smaller.

We compare the achievable rates using QPSK with different number of users in Fig. 2 by numerically solving the integral (9). Clearly, the loss to Gaussian capacity, due to finite modulation, can be made arbitrarily small by imposing a larger number of users or date layers (which may belong to a same user) into the system. Although we assumed equal-power and equal-rate for simplicity, the achievable rates analysis can be extended to other general cases straightforwardly.

We will provide code matching examples for a three user case based on QPSK signaling in Sec. V. The achievable rates can also be found by the density evolution (DE) method, which are very close to the Gaussian capacity.

C. Path vs rate tuples

Consider a simple two-user case, i.e., $K = 2$. Fig. 3 illustrates some special paths and their corresponding achievable rate tuple (or rate pair here). The simplest path is a straight line between the starting point

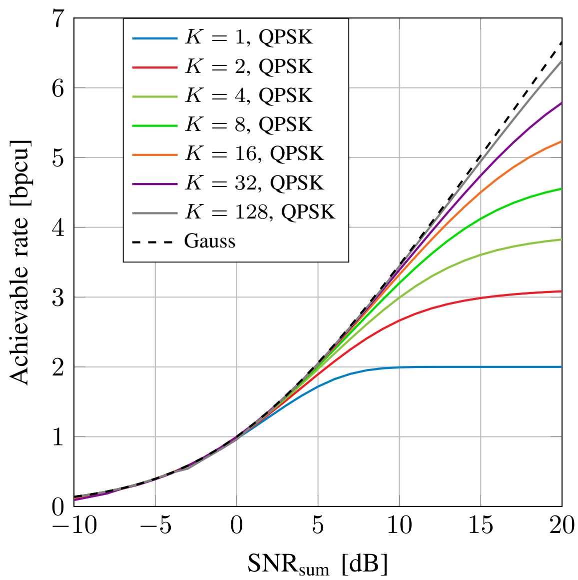

Fig. 2. Achievable rates of multiuser IDMA with matching codes and QPSK modulation; all users are assumed to have same power, modulation and coding scheme; For fair comparison, the multi-user SNR $\mathrm{SNR}_{\mathrm{sum}} = KP/\sigma^2$ shall be used as abscissa.

$\boldsymbol{v}(t=0) = \mathbf{1}$ and the stop point $\boldsymbol{v}(t=\infty) = \mathbf{0}$, denoted by *path 1*. It is straightforward to obtain

$$R_k = \frac{g_k}{\mathbf{g}^T\mathbf{1}} \log\left(1 + \frac{\mathbf{g}^T\mathbf{1}}{\sigma^2}\right), \forall k = 1, 2, \ldots, K.$$

In this case, the achievable rate of each user is proportional to the received signal power strength g_k. For the two-user case, this rate tuple coincides with the point where TDMA/FDMA achieves the sum-rate capacity (see green dot in Fig. 3). In *path 1,* it satisfies

$$v_1(t) = v_2(t) = \cdots = v_K(t) = v(t), \forall t.$$

The matching code for kth user shall have the following MSE characteristic function

$$v_k = \begin{cases} 1 & \rho_k \le \rho_{k,\min} \\ \frac{1}{\mathbf{g}^T\mathbf{1} - g_k} \cdot \left(\frac{1}{\rho_k} - \sigma^2\right) & \rho_{k,\min} \le \rho_k \le \rho_{k,\max} \\ 0 & \rho_k \ge \rho_{k,\max} \end{cases}.$$

For *path 2* and *path 3* which are comprised of K segments and each segment has merely value change (from $v_l = 1$ to $v_l = 0$) in one particular direction v_l, i.e., within the segment $\frac{dv_l}{dt} \neq 0$ and $\frac{dv_k}{dt} = 0, \forall k \neq l$. Depending on the order of the segments, there exist $K!$ such paths, which constitute the $K!$ SIC corner points of MAC capacity region. The user rate can be written as

$$R_k = \log\left(1 + \frac{g_k}{\sum\limits_{l=\pi(k)+1}^{K} g_l + \sigma^2}\right), \forall k = 1, 2, \ldots, K,$$

where $\pi(k) = k'$ denotes the permutation of user order with $1 \le \pi(k) \le K$ and $\pi(k) \neq \pi(k'), \forall k \neq k'$. The corresponding decoding functions are given by

$$v_k = \psi_k(\rho_k) = \begin{cases} 1 & \rho_k < \rho_{k,\mathrm{SIC}} \\ 0 & \rho_k \ge \rho_{k,\mathrm{SIC}} \end{cases}$$

where

$$\rho_{k,\mathrm{SIC}} = \frac{g_k}{\sum\limits_{l=\pi(k)+1}^{K} g_{\pi(l)} + \sigma^2}$$

are the decoding thresholds. The decoding functions are step functions with sharp transitions at corresponding threshold SNRs $\rho_{k,\mathrm{SIC}}$. This type of decoding functions may pose difficulties for practical code designs, compared to that with smooth transitions.

D. Achievable rate region

To achieve other points in the MAC capacity region, i.e., with maximum sum-rate but different individual user rates, other paths shall be used. In the following theorem, we show that the entire MAC capacity region can be achieved by proving the existence of paths. Examples for constructing a dedicated path achieving a feasible rate tuple are provided in *case 2* of Sec. V-A.

Theorem 2. *IDMA with GA-based MUD achieves every rate tuple in the K-user MAC capacity region $\mathcal{C}(K)$. Given a feasible target rate tuple $\mathbf{R} = [R_1, R_2, \cdots, R_K] \in \mathcal{C}(K)$, there exists at least one path defined by $L(t) = \boldsymbol{v}(t) : \boldsymbol{v}_s = \mathbf{1} \rightarrow \boldsymbol{v}_e = \mathbf{0}$ which achieves $\mathbf{R}$.*

Proof: See Appendix B. ∎

Remark: It is easy to prove that there exists a unique path for each of the $K!$ SIC corner points and the decoding functions shall be step functions. For other rate tuples, it can be verified that there exist many different paths achieving that rate tuple. The choice of the integration path poses varying degrees of difficulty for the design of matching codes. Thus, the design of an appropriate integration path could be an extra degree of freedom for code design.

IV. MU-MIMO Channel

Assume that each transmitter has $N_{t,i}$ antennas and the receiver has N_R antennas respectively; then, the received signal can be written as

$$\mathbf{y} = \sum_{k=1}^{K} \sqrt{P_k} \mathbf{H}_k \mathbf{x}_k + \mathbf{n} \tag{11}$$

where $\mathbf{H}_k$ is the channel of the kth user, $\mathbf{n}$ denotes the uncorrelated noise $\mathrm{E}\left[\mathbf{n}\mathbf{n}^H\right] = \sigma^2\mathbf{I}$. In this case, the ESE module is replaced by a linear MMSE (LMMSE)

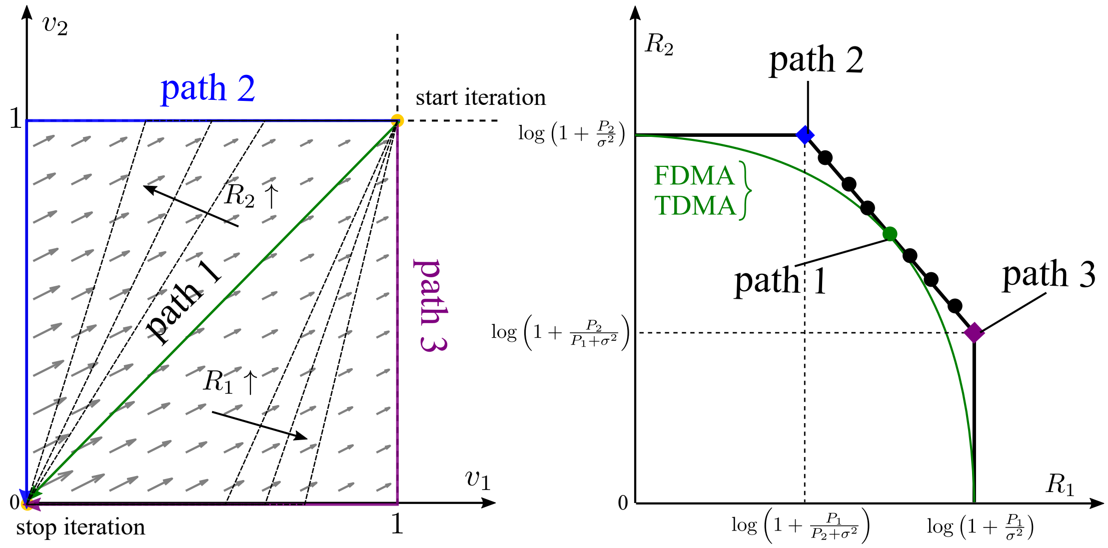

Fig. 3. Illustration (exemplary for two users) of different integration paths achieving different rate pairs (R_1, R_2); the arrows in the left figure illustrate the two-dimensional MSE vector field; the achieved rate pairs are marked in the right figure for the corresponding paths.

receiver [18]. Under the LMMSE-based ESE, the SNR of user k can be written as [21]

$$\rho_k = \frac{\sum_{i=1}^{N_{t,k}} \mathbf{h}_{k,i}^H \mathbf{R}^{-1} \mathbf{h}_{k,i}}{1 - v_k \sum_{i=1}^{N_{t,k}} \mathbf{h}_{k,i}^H \mathbf{R}^{-1} \mathbf{h}_{k,i}} \tag{12}$$

where $\mathbf{h}_{k,i}$ denotes the ith column of the kth user's channel matrix $\mathbf{H}_k$ and

$$\mathbf{R} = \sigma_n^2 \mathbf{I} + \mathbf{H}\mathbf{V}\mathbf{H}^H$$

with $\mathbf{V} = \mathrm{diag}\left(P_1 v_1, P_2 v_2, \cdots, P_K v_K\right)$ and $\mathbf{H}$ being the concatenated channels of all users. Following a similar approach in Appendix A, we obtain with the matching condition in (6) the user rate R_k as

$$R_k = \left[-\int \sum_{i=1}^{N_{t,k}} \mathbf{h}_{k,i}^H \mathbf{R}^{-1} \mathbf{h}_{k,i} dv_k \right]_{v_k=1}^{v_k=0}$$

Therefore, the sum-rate can be obtained as

$$\begin{aligned} R_{\text{sum}} &= \sum_{i=1}^{K} R_i = -\int_{\mathbf{v}=\mathbf{1}}^{\mathbf{v}=\mathbf{0}} \nabla \log\det\left[\mathbf{R}\right] d\mathbf{v} \\ &= \log\det\left[\mathbf{I} + \frac{1}{\sigma_n^2}\mathbf{H}^H \mathbf{P} \mathbf{H}\right] \end{aligned} \tag{13}$$

where $\mathbf{P} = \mathrm{diag}\left(P_1, P_2, \cdots, P_K\right)$. Path independence follows from the condition

$$\frac{\partial}{\partial v_k} \log\det\left[\mathbf{R}\right] = \mathrm{trace}\left[\mathbf{R}^{-1}\mathbf{H}_k\mathbf{H}_k^H\right] = \sum_{i=1}^{N_{t,k}} \mathbf{h}_{k,i}^H \mathbf{R}^{-1} \mathbf{h}_{k,i}$$

with Jacobi's formula.

V. Results

For simplicity, we first consider a single-input single-output (SISO) setup. We assume that the power levels $g_i = P_i|h_i|^2 (i = 1, \cdots, K)$ are constant in our code design. Thus, the multi-user SNR is defined as

$$\mathrm{SNR}_{\text{sum}} = \frac{\sum_{i=1}^{K} g_i}{\sigma^2}. \tag{14}$$

We consider $K = 3$ users with the power distribution $\boldsymbol{g} = [g_1, g_2, g_3]^T = [\frac{1}{7}, \frac{2}{7}, \frac{4}{7}]^T$ and we target the sum-rate $R_{\text{sum}} = R_1 + R_2 + R_3 = 1$ bpcu as an example. Furthermore, the capacity region (more precisely, the *dominant face* which maximizes the sum-rate) with Gaussian alphabets is given by

$$\begin{aligned} 0.1069 \le R_1 &\le \log_2\left(1 + \frac{g_1}{\sigma^2}\right) = 0.1926, \\ 0.2224 \le R_2 &\le \log_2\left(1 + \frac{g_2}{\sigma^2}\right) = 0.3626, \\ 0.4854 \le R_3 &\le \log_2\left(1 + \frac{g_3}{\sigma^2}\right) = 0.6521, \\ R_1 + R_2 &\le \log_2\left(1 + \frac{g_1 + g_2}{\sigma^2}\right) = 0.5145, \\ R_1 + R_3 &\le \log_2\left(1 + \frac{g_1 + g_3}{\sigma^2}\right) = 0.7776, \\ R_2 + R_3 &\le \log_2\left(1 + \frac{g_2 + g_3}{\sigma^2}\right) = 0.8931. \end{aligned}$$

A. ESE Functions

According to the matching condition in (6), for the design of capacity-achieving codes, the ESE transfer functions $\boldsymbol{\rho}(t) = \phi(\boldsymbol{v}(t))$ shall be determined. For this, we specify the K-dimensional decoding path $\boldsymbol{v}(t)$.

As the path independence property of Theorem 1, we can constraint $\boldsymbol{v}(t)$ to be a piecewise linear path with

n segments starting from the point $\boldsymbol{v}(t=0)=\boldsymbol{x}_0=\mathbf{1}$, acrossing the intermediate points $\boldsymbol{v}(t=i)=\boldsymbol{x}_i=[x_{i,1},\cdots,x_{i,K}]^T, i=1,2,\cdots,n-1$, and terminating at the point $\boldsymbol{v}(t=n)=\boldsymbol{x}_n=\mathbf{0}$, where $\boldsymbol{x}_i\neq\boldsymbol{x}_j,\forall i\neq j$ and for practical decoding

$$1\geq x_{1,k}\geq x_{2,k}\geq\cdots\geq x_{n-1,k}\geq 0\,\forall k \tag{15}$$

shall apply. Therefore, the path can be expressed in a vector form as

$$\boldsymbol{v}(t)=\boldsymbol{x}_i-(\boldsymbol{x}_i-\boldsymbol{x}_{i+1})\cdot(t-i), t\in[i,i+1] \tag{16}$$

for $i=0,1,2,\cdots,n-1$. With the specified path, the ESE transfer function for user k can be computed as

$$\begin{aligned}\rho_k=\phi_k(v_k)&=\frac{g_k}{\boldsymbol{g}^T\boldsymbol{v}(t)-g_kv_k(t)+\sigma^2}\\&=\frac{g_k}{\boldsymbol{g}^T\left[\boldsymbol{x}_i-(\boldsymbol{x}_i-\boldsymbol{x}_{i+1})\cdot\frac{x_{i,k}-v_k}{x_{i,k}-x_{i+1,k}}\right]-g_kv_k+\sigma^2},\\&=\frac{g_k}{\sum\limits_{k'\neq k}^{K}g_{k'}\left(\frac{x_{i,k'}-x_{i+1,k'}}{x_{i,k}-x_{i+1,k}}v_k+\frac{x_{i,k}x_{i+1,k'}-x_{i+1,k}x_{i,k'}}{x_{i,k}-x_{i+1,k}}\right)+\sigma^2},\\&\quad v_k\in[x_{i+1,k},x_{i,k}]\text{ for }i=0,1,2,\cdots,n-1.\end{aligned} \tag{17}$$

Note that when $x_{i,k}=x_{i+1,k}$, the above function is not valid. Actually, ρ_k is a vertical line from $\frac{g_k}{\boldsymbol{g}^T\boldsymbol{x}_i-g_kx_{i,k}+\sigma^2}$ to $\frac{g_k}{\boldsymbol{g}^T\boldsymbol{x}_{i+1}-g_kx_{i+1,k}+\sigma^2}$ with $v_k=x_{i,k}$. Substituting (16) into (7b), we obtain the user rate R_k

$$\begin{aligned}R_k&=-\int_{v_k(t)=1}^{v_k(t)=0}\frac{g_k}{\boldsymbol{g}^T\boldsymbol{v}(t)+\sigma^2}dv_k(t)\\&=\sum_{i=0}^{n-1}\frac{g_k(x_{i,k}-x_{i+1,k})}{\boldsymbol{g}^T(\boldsymbol{x}_i-\boldsymbol{x}_{i+1})}\log\frac{\boldsymbol{g}^T\boldsymbol{x}_i+\sigma^2}{\boldsymbol{g}^T\boldsymbol{x}_{i+1}+\sigma^2}.\end{aligned} \tag{18}$$

To achieve the target sum-rate $R_{\text{sum}}=1$ with Gaussian alphabets, it suffices to have $\sigma^2=1$. To verify the path independence property of Theorems 1 and 2, we consider three different paths for (17) and evaluate the system performance and achievable rates via density evolution and bit error rate (BER) simulations.

1) Case 1: We do not specify any intermediate point $\{\boldsymbol{x}_i\}$, i.e., $n=1$. The path is a straight line between the starting point $\boldsymbol{v}(t=0)=\mathbf{1}$ and the stop point $\boldsymbol{v}(t=\infty)=\mathbf{0}$, as discussed in Sec. III-C. The ESE function for user k is given by

$$\rho_k=\frac{g_k}{(\boldsymbol{g}^T\mathbf{1}-g_k)v_k+\sigma^2}, v_k\in[0,1]. \tag{19}$$

The rate for user k is proportional to its power g_k, i.e., $R_k=\frac{g_k}{\boldsymbol{g}^T\mathbf{1}}R_{\text{sum}}$. Thus, the corresponding rate tuple is $(R_1,R_2,R_3)=(\frac{1}{7},\frac{2}{7},\frac{4}{7})$. The transfer functions in (19) are depicted in the left most subfigure in Fig. 4 for the three users, respectively.

2) Case 2: We construct a dedicated path to achieve an arbitrarily chosen rate tuple in the MAC region, e.g., $(R_1,R_2,R_3)=(0.15,0.3,0.55)$. To find a dedicated path, we search for $\{\boldsymbol{x}_i\}$ by solving K non-linear equations given by (18). Then, the ESE transfer functions can be obtained by substituting $\{\boldsymbol{x}_i\}$ into (17). If $n>2$, there are $K(n-1)$ unknown variables $\{x_{i,k}\}$, which is larger than K. This potentially result in multiple solutions. It is noteworthy to mention that the variables $\{x_{i,k}\}$ are bounded in $[0,1]$ and shall satisfy (15). We may fix some unknown variables $\{x_{i,k}\}$ and solve the K non-linear equations given by (18) to obtain remaining unknown variables. Usually, we can fix $K(n-2)$ unknown variables and have feasible solution for the remaining K unknown variables. Here, we consider a 3-segment path having intermediate points

$$\boldsymbol{x}_1=[x_{1,1},x_{1,2},0]^T\text{ and }\boldsymbol{x}_2=[0,x_{2,2},0]^T. \tag{20}$$

Substituting (20) into (18), we obtain

$$\begin{aligned}0.15=&\frac{g_1(1-x_{1,1})}{1-g_1x_{1,1}-g_2x_{1,2}}\log_2\frac{2}{g_1x_{1,1}+g_2x_{1,2}+1}\\&+\frac{g_1x_{1,1}}{g_1x_{1,1}+g_2(x_{1,2}-x_{2,2})}\log_2\frac{g_1x_{1,1}+g_2x_{1,2}+1}{g_2x_{2,2}+1},\\0.3=&\frac{g_2(1-x_{1,2})}{1-g_1x_{1,1}-g_2x_{1,2}}\log_2\frac{2}{g_1x_{1,1}+g_2x_{1,2}+1}\\&+\frac{g_2(x_{1,2}-x_{2,2})}{g_1x_{1,1}+g_2(x_{1,2}-x_{2,2})}\log_2\frac{g_1x_{1,1}+g_2x_{1,2}+1}{g_2x_{2,2}+1}\\&+\log_2(g_2x_{2,2}+1),\\0.55=&\frac{g_3}{1-g_1x_{1,1}-g_2x_{1,2}}\log_2\frac{2}{g_1x_{1,1}+g_2x_{1,2}+1}.\end{aligned} \tag{21}$$

Solving (21), we obtain one feasible solution given by $x_{1,1}=0.2145, x_{1,2}=0.2056, x_{2,2}=0.0618$. Substituting the solution into (17), we can obtain the ESE functions. These transfer functions are depicted in the middle subfigure of Fig. 4 for the three users, respectively.

3) Case 3: We randomly choose the intermediate points $\{\boldsymbol{x}_i\}$. Then, substituting the points into (17), we obtain the ESE transfer functions and subsequently compute the rate for each user using (18).

Here, we consider a piece-wise linear path with 2 segments by specifying an intermediate point arbitrarily, e.g., $\boldsymbol{x}_1=[0.5,0.2,0.2]$. Substituting $\boldsymbol{x}_1$ into (17), we have the ESE functions as

$$\rho_1=\begin{cases}\frac{g_1}{(g_2+g_3)0.4v_1+\sigma^2}, & 0\leq v_1\leq 0.5,\\ \frac{g_1}{g_2(1.6v_1-0.6)+g_3(1.6v_1-0.6)+\sigma^2}, & 0.5\leq v_1\leq 1,\end{cases}$$

$$\rho_2=\begin{cases}\frac{g_2}{g_1 2.5v_2+g_3v_2+\sigma^2}, & 0\leq v_2\leq 0.2,\\ \frac{g_2}{g_1(0.625v_2+0.375)+g_3v_2+\sigma^2}, & 0.2\leq v_2\leq 1,\end{cases}$$

$$\rho_3=\begin{cases}\frac{g_3}{g_1 2.5v_3+g_2v_3+\sigma^2}, & 0\leq v_3\leq 0.2,\\ \frac{g_3}{g_1(0.625v_3+0.375)+g_2v_3+\sigma^2}, & 0.2\leq v_3\leq 1.\end{cases}$$

These transfer functions are depicted in the right most subfigure of Fig. 4 for the three users, respectively. The

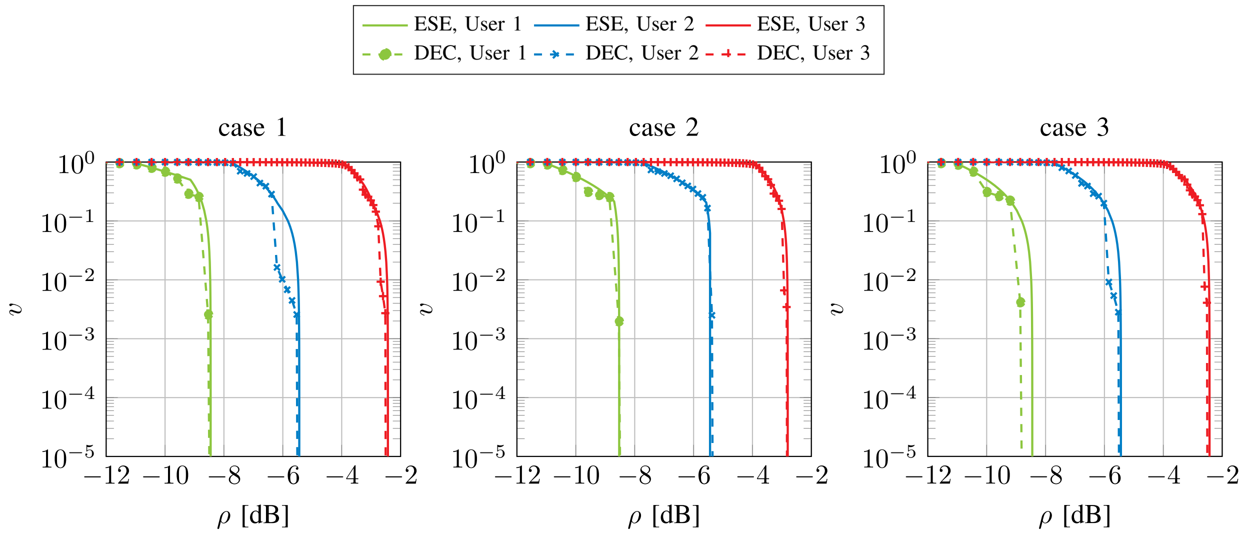

Fig. 4. ESE transfer functions and the matching LDPC code transfer functions for three different paths; three users with QPSK and the power distribution $\boldsymbol{g} = [g_1, g_2, g_3]^T = [\frac{1}{7}, \frac{2}{7}, \frac{4}{7}]^T$ are considered.

achievable rates for each user, given in (22), can be obtained by substituting $\boldsymbol{x}_1$ into (18), see Tab. I.

B. LDPC code optimization

As the ESE functions are readily available, according to the matching condition, we optimize the degree profile of LDPC codes to match the ESE functions for user k, i.e.,

$$v_k = \psi_k(\rho_k) = \begin{cases} 1, & \rho_k \leq \rho_{k,\min}, \\ \phi_k^{-1}(\rho_k), & \rho_{k,\min} \leq \rho_k \leq \rho_{k,\max}, \\ 0, & \rho_k \geq \rho_{k,\max}, \end{cases} \tag{23}$$

where $\phi_k^{-1}(\rho_k)$ is the inverse of the ESE function of user k and $\rho_{k,\min} = \frac{g_k}{\boldsymbol{g}^T\mathbf{1} - g_k + \sigma^2}$, $\rho_{k,\max} = \frac{g_k}{\sigma^2}$. Using the EXIT chart matching techniques [18], [22]–[24], the matching LDPC codes can be designed by properly choosing the degree distributions.

We basically follow the method described in Appendix 5G of [24] to design irregular LDPC codes given a target transfer function $v = \psi(\rho)$, where ρ is the *a priori* SNR and v, the decoder output, denotes the extrinsic variance. The difference is that we use mutual information instead of the mean of log likelihood ratio (LLR) to track the evolution process.

The asymptotic performance of an LDPC code ensemble can be specified by its variable node and check node edge distribution polynomials, namely

$$\lambda(x) = \sum_{i=1}^{d_{v,\max}} \lambda_i x^{i-1} \text{ and } \sum_{i=1}^{d_{c,\max}} \eta_i x^{i-1}, \tag{24}$$

where λ_i (resp., η_i) is the fraction of edges in the bipartite graph of the LDPC code connected to variable nodes (resp., check nodes) with degree i, and $d_{v,\max}$ (resp., $d_{c,\max}$) is the maximum variable node (resp., check node) degree. Moreover, we use the Gaussian approximation [25], i.e., (2) and (5), to optimize the edge distributions for the sake of simplicity.

In [23], it is shown that the decoder characteristic for an LDPC code can be computed as

$$I_{E,V} = \sum_{i=1}^{d_{v,\max}} \lambda_i \cdot J\left(\sqrt{(i-1)\left[J^{-1}\left(I_{E,C}\right)\right]^2 + 4\rho}\right), \tag{25}$$

$$I_{E,C} = 1 - \sum_{j=1}^{d_{c,\max}} \eta_j \cdot J\left(\sqrt{j-1} \cdot J^{-1}\left(1 - I_{E,V}\right)\right), \tag{26}$$

where $I_{E,V}$ (resp., $I_{E,C}$) is the extrinsic information from variable node (resp., check node) to check node (resp., variable node), ρ is the decoder input SNR, and the $J(\cdot)$ is defined by

$$J(\sigma_{ch}) = 1 - \int_{-\infty}^{\infty} \frac{e^{\frac{-(y-\sigma_{ch}^2/2)^2}{2\sigma_{ch}^2}}}{\sqrt{2\pi\sigma_{ch}^2}} \cdot \log_2\left[1 + e^{-y}\right] dy,$$

its inverse function is further denoted by $J^{-1}(\cdot)$. Substituting (26) into (25), we have (27), where the LDPC code can be charaterized by one single variable $I_{E,V}$. The degree optimization problem can be formulated in (28). The cost function in (28) is to maximize the code rate. Let $I_{E,V,\text{ini}}(\rho)$ be the initial extrinsic information given by the channel, which can be written as

$$\begin{aligned} I_{E,V,\text{ini}}(\rho) &= \sum_{i=1}^{d_{v,\max}} \lambda_i \cdot J\left(\sqrt{(i-1)\left[J^{-1}(0)\right]^2 + 4\rho}\right) \\ &= J\left(2\sqrt{\rho}\right). \end{aligned} \tag{29}$$

Let $I_{E,V,\text{fin}}(\rho)$ denote the extrinsic information upon convergence, which corresponds to an output extrinsic

$$
\begin{aligned}
R_1 = &\frac{g_1(1-0.5)}{g_1(1-0.5)+g_2(1-0.2)+g_3(1-0.2)}\log_2\left(\frac{1+1}{g_1\cdot 0.5+g_2\cdot 0.2+g_3\cdot 0.2+1}\right)\\
&+\frac{g_1(0.5-0)}{g_1(0.5-0)+g_2(0.2-0)+g_3(0.2-0)}\log_2\left(g_1\cdot 0.5+g_2\cdot 0.2+g_3\cdot 0.2+1\right)=0.157,\\
R_2 = &\frac{g_2(1-0.2)}{g_1(1-0.5)+g_2(1-0.2)+g_3(1-0.2)}\log_2\left(\frac{1+1}{g_1\cdot 0.5+g_2\cdot 0.2+g_3\cdot 0.2+1}\right)\\
&+\frac{g_2(0.2-0)}{g_1(0.5-0)+g_2(0.2-0)+g_3(0.2-0)}\log_2\left(\frac{g_1\cdot 0.5+g_2\cdot 0.2+g_3\cdot 0.2+1}{0+1}\right)=0.281,\\
R_3 = &\frac{g_3(1-0.2)}{g_1(1-0.5)+g_2(1-0.2)+g_3(1-0.2)}\log_2\left(\frac{1+1}{g_1\cdot 0.5+g_2\cdot 0.2+g_3\cdot 0.2+1}\right)\\
&+\frac{g_3(0.2-0)}{g_1(0.5-0)+g_2(0.2-0)+g_3(0.2-0)}\log_2\left(\frac{g_1\cdot 0.5+g_2\cdot 0.2+g_3\cdot 0.2+1}{0+1}\right)=0.562. \qquad (22)
\end{aligned}
$$

$$
I_{E,V} = \sum_{i=1}^{d_{v,\max}} \lambda_i \cdot J\left(\sqrt{(i-1)\left[J^{-1}\left(1-\sum_{j=1}^{d_{c,\max}}\eta_j\cdot J\left(\sqrt{j-1}\cdot J^{-1}\left(1-I_{E,V}\right)\right)\right)\right]^2+4\rho}\right), \qquad (27)
$$

$$
\begin{aligned}
&\max_{\{\lambda_i\}} \sum_{i=1}^{d_{v,\max}} \frac{\lambda_i}{i}\\
\text{s.t. } &\sum_{i=1}^{d_{v,\max}} \lambda_i = 1,\\
&\sum_{i=1}^{d_{v,\max}} \lambda_i \cdot J\left(\sqrt{(i-1)\left[J^{-1}\left(1-\sum_{j=1}^{d_{c,\max}}\eta_j\cdot J\left(\sqrt{j-1}\cdot J^{-1}\left(1-I_{E,V}\right)\right)\right)\right]^2+4\rho}\right) > I_{E,V}\\
&\text{for } \forall 0<\rho<\infty \text{ and } I_{E,V,\text{ini}}(\rho)\le I_{E,V}\le I_{E,V,\text{fin}}(\rho). \qquad (28)
\end{aligned}
$$

---

variance to the MUD $v=\psi(\rho)$, i.e., given ρ, $I_{E,V,\text{fin}}(\rho)$ should satisfy the following equation

$$
\sum_{i=1}^{d_{v,\max}} \Lambda_i \cdot f_Q\left(\frac{i\cdot\left[J^{-1}\left(I_{E,C,\text{fin}}\right)\right]^2}{4}\right)=\psi(\rho), \qquad (30)
$$

where f_Q is defined in (10), and

$$
I_{E,C,\text{fin}} = 1-\sum_{j=1}^{d_{c,\max}}\eta_j\cdot J\left(\sqrt{j-1}\cdot J^{-1}\left(1-I_{E,V,\text{fin}}\right)\right)
$$

is the converged message from check nodes to variable nodes. Furthermore, we define

$$
\Lambda_i = \frac{\lambda_i/i}{\sum_{i=1}^{d_{v,\max}}\lambda_i/i}
$$

as the fraction of variable node of degree i.

The optimization in (28) is a non-convex optimization. However, given $\{\Lambda_i\}$ and $\eta(x)$, the problem in (28) can be solved using standard linear programming. We use an iterative way to optimize the edge distribution $\lambda(x)$ with fixed $\eta(x)$ in Algorithm 1. In practise, Algorithm 1 is repeated for several check edge distributions $\eta(x)$ till a matching code is found.

C. Numerical Results

With Algorithm 1, if we use all variable degrees less than $d_{v,\max}$, the optimization can be quite slow. However, the optimized degree sequences are mostly comprised of a few small degrees. Therefore, we only use a subset of degrees less than $d_{v,\max}$ to run Algorithm 1 more efficiently. In the algorithm, we set $T = 100$ and $\epsilon = 0.001$ for all optimizations. The optimized results as well as other parameters are summarized in Table I. Fig. 4 shows the optimized LDPC DEC transfer functions (denoted by dashed lines). The DEC functions match enough well with the ESE functions.

Table I
CODE OPTIMIZATION RESULTS FOR THREE CASES

Case	Case 1			Case 2		
User	User 1	User 2	User 3	User 1	User 2	User 3
Power	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{4}{7}$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{4}{7}$
Path	$[1,1,1] \to [0,0,0]$			$[1,1,1] \to [0.2145, 0.2056, 0]$ $\to [0, 0.0618, 0] \to [0,0,0]$		
Target rate	0.1429	0.2857	0.5714	0.15	0.30	0.55
Check edge distribution	$\eta_3 = 1$	$\eta_4 = 1$	$\eta_5 = 1$	$\eta_3 = 1$	$\eta_4 = 1$	$\eta_5 = 1$
variable degree set $\{d_v\}$	{2:1:30, 35:5:50, 60:10:100}			{2:1:30, 35:5:50}		
Optimized variable edge distribution	$\lambda_2, 0.5239$ $\lambda_3, 0.2140$ $\lambda_7, 0.1627$ $\lambda_{30}, 0.0685$ $\lambda_{35}, 0.0309$	$\lambda_2, 0.3770$ $\lambda_3, 0.2168$ $\lambda_7, 0.0719$ $\lambda_8, 0.1577$ $\lambda_{40}, 0.1237$ $\lambda_{100}, 0.0529$	$\lambda_2, 0.3293$ $\lambda_3, 0.2351$ $\lambda_8, 0.2500$ $\lambda_{21}, 0.0654$ $\lambda_{22}, 0.0014$ $\lambda_{45}, 0.0258$ $\lambda_{50}, 0.0930$	$\lambda_2, 0.5234$ $\lambda_3, 0.2292$ $\lambda_7, 0.0896$ $\lambda_8, 0.0590$ $\lambda_{30}, 0.0708$ $\lambda_{35}, 0.0280$	$\lambda_2, 0.3779$ $\lambda_3, 0.2290$ $\lambda_7, 0.1431$ $\lambda_8, 0.0580$ $\lambda_{50}, 0.1920$	$\lambda_2, 0.3218$ $\lambda_3, 0.2273$ $\lambda_7, 0.0879$ $\lambda_8, 0.1654$ $\lambda_{20}, 0.0501$ $\lambda_{21}, 0.0335$ $\lambda_{50}, 0.1140$
Optimized rate	0.1467	0.3014	0.5707	0.1555	0.3154	0.5522

Case	Case 3		
User	User 1	User 2	User 3
Power	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{4}{7}$
Path	$[1,1,1] \to [0.5, 0.2, 0.2] \to [0,0,0]$		
Target rate	0.157	0.281	0.562
Check edge distribution	$\eta_3 = 1$	$\eta_4 = 1$	$\eta_5 = 1$
variable degree set $\{d_v\}$	{2:1:30, 35:5:50}		
Optimized variable edge distribution	$\lambda_2, 0.5250$ $\lambda_3, 0.2138$ $\lambda_6, 0.0978$ $\lambda_7, 0.0618$ $\lambda_{30}, 0.0604$ $\lambda_{35}, 0.0412$	$\lambda_2, 0.3788$ $\lambda_3, 0.1925$ $\lambda_6, 0.0763$ $\lambda_7, 0.1589$ $\lambda_{50}, 0.1935$	$\lambda_2, 0.3289$ $\lambda_3, 0.2277$ $\lambda_8, 0.1189$ $\lambda_9, 0.1747$ $\lambda_{45}, 0.1337$ $\lambda_{50}, 0.0161$
Optimized rate	0.1588	0.2926	0.5608

Algorithm 1 Algorithm for LDPC Code Optimization in IDMA

Input: Target decoder transfer functions $v_k = \psi_k(\rho_k)$, check edge distribution $\eta_k(x)$, maximum trial T, threshold ϵ and maximum variable degree $d_{v,\max}$.

Output: The optimized variable edge distribution $\lambda^{(T)}(x)$.

Initialize $\lambda^{(0)}(x) = x$.
for $t = 1$ to T **do**
  Solve (28) by linear programming to obtain $\lambda^{(t)}(x)$, where $I_{E,V,\text{fin}}(\rho)$ in (28) is obtained by solving (30) using $\lambda^{(t-1)}(x)$.
  if $1 - \frac{\sum_{i=1}^{d_{v,\max}} \lambda_i^{(t)} \lambda_i^{(t-1)}}{\sqrt{\left(\sum_{i=1}^{d_{v,\max}} (\lambda_i^{(t)})^2\right)\left(\sum_{i=1}^{d_{v,\max}} (\lambda_i^{(t-1)})^2\right)}} \le \epsilon$ **then**
    $\lambda^{(T)}(x) = \lambda^{(t)}(x)$.
    return $\lambda^{(T)}(x)$.
  end if
end for
return $\lambda^{(T)}(x)$.

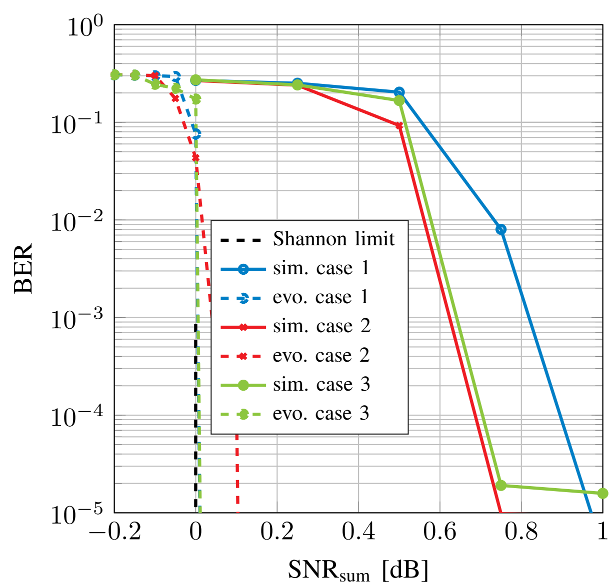

Fig. 5. BER and density evolution results for a three user MAC with matched codes; three different paths and QPSK signaling are considered.

After matching the degree distribution, we construct parity check matrices for BER simulations. The parity-check matrix of code-word length 10^5 for each user in each case is randomly generated and subsequently we remove the cycle-4 loops in the matrix by edge permutation [26]. Furthermore, we consider QPSK signaling in the simulation to verify that our Theorems work also well with finite alphabets, not only with Gaussian alphabets.

Fig. 5 shows the average BER performance for three users with matching codes along different decoding paths (denoted by solid lines, respectively), where we set a maximum iteration of 1000 between the ESE detector and the LDPC decoders. Moreover, the Shannon limit at the sum-rate $R_{\text{sum}} = 1$ along with the density evolution performance with optimized codes considering QPSK are provided. From the numerical results, we can conclude that

- the evolution thresholds for three cases with QPSK are close to the Gaussian capacity. The loss incurred by finite alphabets is negligible at the target sum-rate.
- Three cases have BER below 10^{-4} within 1 dB to the Shannon limit. The sum-rate capacity can be achieved for different paths also with QPSK signalling.
- For case 2, we computed a dedicated path to achieve an arbitrarily choosen rate tuple $(R_1, R_2, R_3) = (0.15, 0.3, 0.55)$. The numerical results also verified our path construction based on GA.

To further verify that the decoding path (or decoding trajectory $L(t)$) in the BER simulation is close to the desired path in the theory. We compare the decoding trajectories obtained via density evolution and BER simulation for three cases at the SNR $\mathrm{SNR}_{\text{sum}} = 1$ dB, where all users can decode its signal with high probability. The evolution trajectories differ from the specified paths (discussed in Sec. V-A) mainly due to the different SNRs (the specified paths assume $\mathrm{SNR}_{\text{sum}} = 0$ dB). We observe that the simulation trajectories are consistent with those of density evolution for all three cases. This further consolidates the path independence theorem, and provides numerical evidence for finite alphabet cases.

VI. Conclusion

It is proved that the simple interleave-division multiple-access (IDMA), relying on a low-cost Gaussian approximation (GA) based multi-user detector (MUD), is capacity-achieving for general Gaussian multiple access channels (GMAC) with arbitrary number of users, power distribution and with single or multiple antennas. We show that IDMA with matching codes is capacity-achieving for arbitrary decoding path in the mean-square error (MSE) vector field. This property is further used to prove that IDMA achieves not only the sum-rate capacity, but the entire GMAC capacity region. The construction of capacity-achieving codes is also provided by establishing an area theorem for multi-user extrinsic information transfer (EXIT) chart.

Appendix A
Proof of (7b)

Let $\rho_{k,\min},\ \rho_{k,\max}$ as defined in (3c). The achievable rates can be thus expressed as

$$
\begin{aligned}
R_k &= \int_{\rho_{k,\min}}^{\rho_{k,\max}} \frac{1}{\rho_k + v_k^{-1}} d\rho_k + \int_0^{\rho_{k,\min}} \frac{1}{\rho_k + 1} d\rho_k \\
&\overset{\rho_k' = \frac{d\rho_k}{dv_k}}{=} \int_{v_k=1}^{v_k=0} \frac{\rho_k'}{\rho_k + v_k^{-1}} dv_k + \int_0^{\rho_{k,\min}} \frac{1}{\rho_k + 1} d\rho_k \\
&= \int_1^0 \frac{\rho_k' - v_k^{-2} + v_k^{-2}}{\rho_k + v_k^{-1}} dv_k + \underbrace{\log\left(1 + \rho_{k,\min}\right)}_{=w_0} \\
&= \left[\log\left(\rho_k + v_k^{-1}\right) + \int_1^0 \frac{v_k^{-2}}{\rho_k + v_k^{-1}} dv_k\right]_{v_k=1}^{v_k=0} + w_0 \\
&\overset{(1a)}{=} \left[\log\left(\rho_k + v_k^{-1}\right) + \int \left(v_k^{-1} - \frac{g_k}{\mathbf{g}^T \boldsymbol{v} + \sigma^2}\right) dv_k\right]_{v_k=1}^{v_k=0} \\
&\quad + w_0 \\
&= \left[\log\left(\rho_k v_k + 1\right) - \int \frac{g_k}{\mathbf{g}^T \boldsymbol{v} + \sigma^2} dv_k\right]_{v_k=1}^{v_k=0} + w_0 \\
&= -\int_1^0 \frac{g_k}{\mathbf{g}^T \boldsymbol{v} + \sigma^2} dv_k
\end{aligned}
$$

where $g_k = P_k \left|h_k\right|^2$ is the kth element of the vector $\mathbf{g}$.

Appendix B
Proof of Theorem 2

The user rate $R_k = -\int_{v_k=1}^{v_k=0} \frac{g_k}{\mathbf{g}^T \boldsymbol{v} + \sigma^2} dv_k$ is obviously a continuous and monotone decreasing function of $\boldsymbol{v}$. If v_k are unbounded, then R_k are unbounded with the single sum-rate constraint $\sum R_k \le \log\left(\frac{\mathbf{g}^T \mathbf{1} + \sigma^2}{\sigma^2}\right)$. However, the value range of R_k is constrained by the fact that $0 \le v_l \le 1, \forall l$. Therefore, it is bounded by

$$
R_k \le -\int_{v_k=1}^{v_k=0} \frac{g_k}{g_k v_k + \sigma^2} dv_k = \log\left(\frac{g_k + \sigma^2}{\sigma^2}\right)
$$

and similarly $R_k \ge \log\left(\frac{g_k + \sigma^2}{\sum_{l \ne k} g_l + \sigma^2}\right)$. Further, the constraints on v_l leads to

$$
\begin{aligned}
R_k + R_l &\le \log\left(\frac{g_k + g_l + \sigma^2}{\sigma^2}\right), \forall k \ne l \\
R_k + R_l + R_m &\le \log\left(\frac{g_k + g_l + g_m + \sigma^2}{\sigma^2}\right), \forall k \ne l \ne m \\
&\vdots \\
\sum R_k &\le \log\left(\frac{\mathbf{g}^T \mathbf{1} + \sigma^2}{\sigma^2}\right)
\end{aligned}
$$

and these constraints constitute the MAC capacity region.

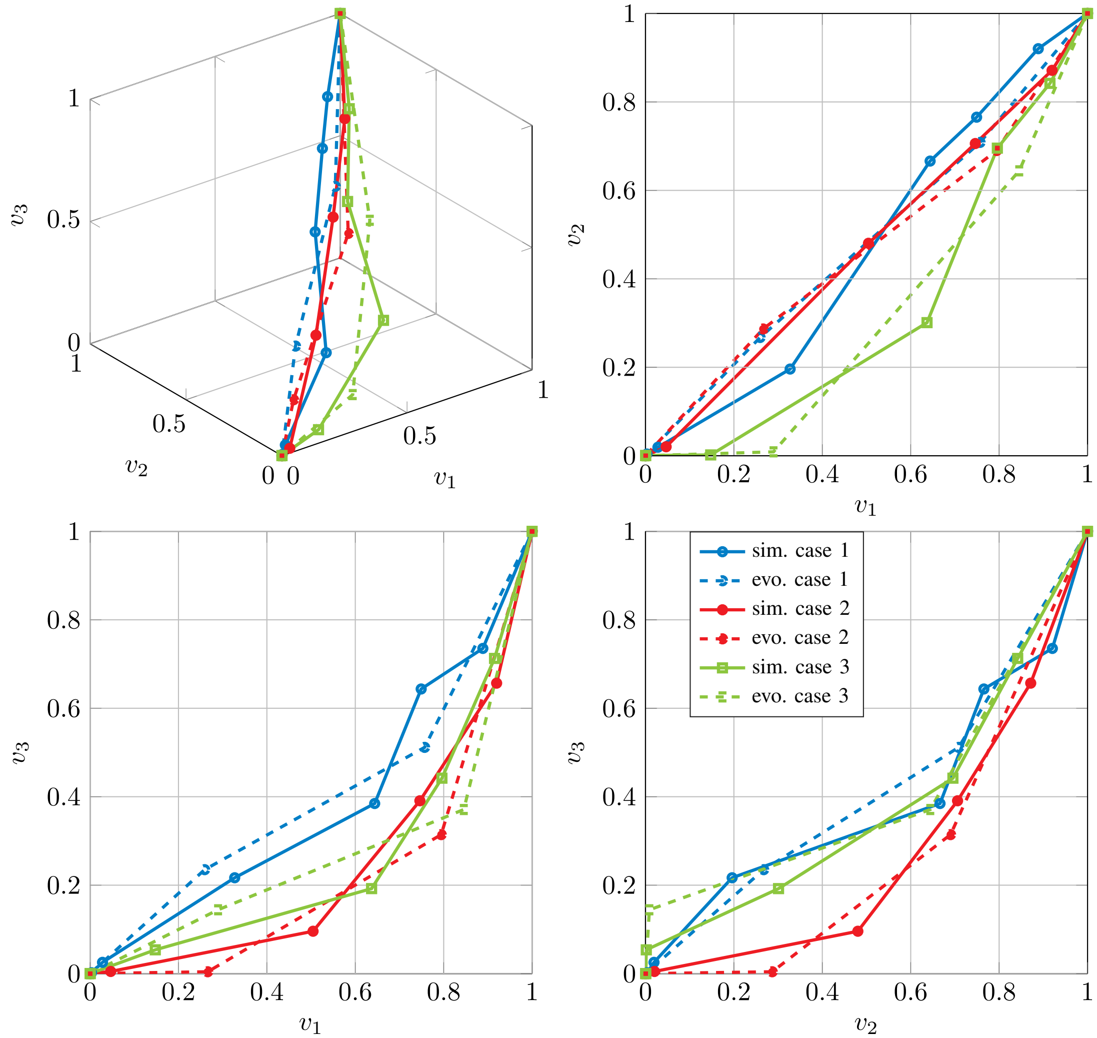

Fig. 6. Evolution and simulation trajectories for three cases at $\mathrm{SNR}_{\mathrm{sum}} = 1$ dB. Upper left: 3-D diagram for a trajectory (v_1, v_2, v_3). Upper right: Side view (v_1, v_2) of a trajectory. Lower left: Side view (v_1, v_3) of a trajectory. Lower right: Side view (v_2, v_3) of a trajectory.

REFERENCES

[1] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. John Wiley & Sons, 2006.

[2] A. E. Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[3] D. Tse and P. Viswanath, *Fundamentals of Wireless Communications*. Cambridge University Press, 2005.

[4] Y. Hu and L. Ping, "Interleave-Division Multiple Access (IDMA)," in *Multiple Access Techniques for 5G Wireless Networks and Beyond*, M. Vaezi, Z. Ding, and H. V. Poor, Eds. Springer, 2019, ch. 13, pp. 417–449.

[5] S. ten Brink, "Exploiting the Chain Rule of Mutual Information for the Design of Iterative Deoding Shemes," in *Proc. 39th Annual Allerton Conf. on Comm., Control and Computing*, 2001.

[6] T. J. Richardson, M. A. Shokrollahi, and R. L. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Transactions on Information Theory*, vol. 47, no. 2, pp. 619–637, Feb 2001.

[7] A. Ashikhmin, G. Kramer, and S. ten Brink, "Extrinsic information transfer functions: model and erasure channel properties," *IEEE Transactions on Information Theory*, vol. 50, no. 11, pp. 2657–2673, Nov 2004.

[8] K. Bhattad and K. R. Narayanan, "An MSE-Based Transfer Chart for Analyzing Iterative Decoding Schemes Using a Gaussian Approximation," *IEEE Transactions on Information Theory*, vol. 53, no. 1, pp. 22–38, Jan 2007.

[9] L. Ping, L. Liu, K. Wu, and W. K. Leung, "Interleave division multiple-access," *IEEE Transactions on Wireless Communications*, vol. 5, no. 4, pp. 938–947, April 2006.

[10] L. Ping, L. Liu, K. Y. Wu, and W. K. Leung, "Approaching the capacity of multiple access channels using interleaved low-rate codes," *IEEE Communications Letters*, vol. 8, no. 1, pp. 4–6, Jan 2004.

[11] Y. Hu, C. Liang, L. Liu, C. Yan, Y. Yuan, and L. Ping, "Interleave-division multiple access in high rate applications," *IEEE Wireless Communications Letters*, 2018.

[12] K. Li and X. Wang, "EXIT chart analysis of turbo multiuser detection," *IEEE Transactions on Wireless Communications*, vol. 4, no. 1, pp. 300–311, Jan 2005.

[13] J. Song and Y. Shu, "On Construction of Rate-Compatible Raptor-Like QC-LDPC Code for Enhanced IDMA in 5G and Beyond," in *Proc. 10th Internat. Symp. Turbo Codes*, 2018.

[14] G. Song and J. Cheng, "Low-complexity coding scheme to approach multiple-access channel capacity," in *2015 IEEE International Symposium on Information Theory (ISIT)*, June 2015, pp. 2106–2110.

[15] L. Liu, Y. Chi, C. Yuen, Y. L. Guan, and Y. Li, “Capacity-achieving iterative LMMSE detection for MIMO-NOMA systems,” *IEEE Transactions on Signal Processing*, 2019.

[16] D. Guo, S. Shamai, and S. Verdu, “Mutual information and minimum mean-square error in Gaussian channels,” *IEEE Transactions on Information Theory*, vol. 51, no. 4, pp. 1261–1282, April 2005.

[17] X. Wang, S. Cammerer, and S. ten Brink, “Near Gaussian multiple access channel capacity detection and decoding,” *arXiv preprint arXiv:1811.10938*, 2018.

[18] X. Yuan, L. Ping, C. Xu, and A. Kavcic, “Achievable Rates of MIMO Systems With Linear Precoding and Iterative LMMSE Detection,” *IEEE Transactions on Information Theory*, vol. 60, no. 11, pp. 7073–7089, Nov 2014.

[19] L. Ping, J. Tong, X. Yuan, and Q. Guo, “Superposition coded modulation and iterative linear MMSE detection,” *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 6, pp. 995–1004, August 2009.

[20] D. Guo, Y. Wu, S. S. Shitz, and S. Verdu, “Estimation in Gaussian Noise: Properties of the Minimum Mean-Square Error,” *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 2371–2385, April 2011.

[21] X. Yuan, Q. Guo, X. Wang, and L. Ping, “Evolution analysis of low-cost iterative equalization in coded linear systems with cyclic prefixes,” *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 2, pp. 301–310, February 2008.

[22] S. ten Brink, “Convergence behavior of iteratively decoded parallel concatenated codes,” vol. 49, no. 10, pp. 1727–1737, Oct. 2001.

[23] S. ten Brink, G. Kramer, and A. Ashikhmin, “Design of low-density parity-check codes for modulation and detection,” vol. 52, no. 4, pp. 670–678, Apr. 2004.

[24] X. Yuan, *Low-complexity iterative detection in coded linear systems*, City University of Hong Kong, Hong Kong, China, 2008.

[25] S.-Y. Chung, T. Richardson, and R. Urbanke, “Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation,” vol. 47, no. 2, pp. 657–670, Feb. 2001.

[26] J. A. McGowan and R. C. Williamson, “Loop removal from LDPC codes,” in *Proc. IEEE Information Theory Workshop (ITW’03)*, Paris, France, 31 Mar. - 4 Apr. 2003, pp. 230–233.