The Mismatched Multiple-Access Channel: General Alphabets

Jonathan Scarlett
University of Cambridge
jms265@cam.ac.uk

Alfonso Martinez
Universitat Pompeu Fabra
alfonso.martinez@ieee.org

Albert Guillén i Fàbregas
ICREA & Universitat Pompeu Fabra
University of Cambridge
guillen@ieee.org

***Abstract*—This paper considers channel coding for the memoryless multiple-access channel with a given (possibly suboptimal) decoding rule. Non-asymptotic bounds on the error probability are given, and a cost-constrained random-coding ensemble is used to obtain an achievable error exponent. The achievable rate region recovered by the error exponent coincides with that of Lapidoth in the discrete memoryless case, and remains valid for more general alphabets.**

I. Introduction

The problem of channel coding with a mismatched decoding rule arises in numerous settings [1]–[5]. For example, in practical systems the decoder may have imperfect knowledge of the channel, or implementation constraints may prohibit the use of an optimal decoder.

The mismatched discrete memoryless multiple-access channel (DM-MAC) was considered by Lapidoth [1], and more recently by the present authors [6]. The setup consists of a DM-MAC $W(y|x_1,x_2)$ and a decoder which maximizes the symbol-wise product of a given decoding metric $q(x_1,x_2,y)$. Lapidoth proved the achievability of the rate region given by the convex closure of the union of all rate pairs (R_1,R_2) satisfying (see Section I-B for notation)

$$R_1 \le \min_{\substack{\widetilde{P}_{X_1}=Q_1,\widetilde{P}_{X_2Y}=P_{X_2Y}\\ \mathbb{E}_{\widetilde{P}}[\log q]\ge\mathbb{E}_P[\log q]}} I_{\widetilde{P}}(X_1;X_2,Y) \tag{1}$$

$$R_2 \le \min_{\substack{\widetilde{P}_{X_2}=Q_2,\widetilde{P}_{X_1Y}=P_{X_1Y}\\ \mathbb{E}_{\widetilde{P}}[\log q]\ge\mathbb{E}_P[\log q]}} I_{\widetilde{P}}(X_2;X_1,Y) \tag{2}$$

$$R_1+R_2 \le \min_{\substack{\widetilde{P}_{X_1}=Q_1,\widetilde{P}_{X_2}=Q_2,\widetilde{P}_{Y}=P_{Y},\mathbb{E}_{\widetilde{P}}[\log q]\ge\mathbb{E}_P[\log q]\\ I_{\widetilde{P}}(X_1;Y)\le R_1,\, I_{\widetilde{P}}(X_2;Y)\le R_2}} D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times\widetilde{P}_Y) \tag{3}$$

for some input distributions Q_1 and Q_2, where $P_{X_1X_2Y}=Q_1\times Q_2\times W$, and each minimization is over all joint distributions $\widetilde{P}_{X_1X_2Y}$ satisfying the specified constraints.

The approach of [1] is based on strong typicality arguments, and [6] gives an alternative derivation based on error exponents and the method of types. In both cases, the proof is valid only in the discrete memoryless setting. In this paper, we use cost-constrained random coding [5], [7]–[9] to obtain an achievable rate region for the case of general alphabets. The region coincides with (1)–(3) in the discrete memoryless setting, with the two being related via Lagrange duality. Thus, our achievable rate region generalizes that of [1] analogously to the generalization of the single-user rate of Csiszár-Körner-Hui (see [3], [10]) given in [5].

This work has been funded in part by the European Research Council under ERC grant agreement 259663, by the European Union's 7th Framework Programme (PEOPLE-2011-CIG) under grant agreement 303633 and by the Spanish Ministry of Economy and Competitiveness under grants RYC-2011-08150 and TEC2012-38800-C03-03.

A. System Setup

We consider a 2-user memoryless MAC $W(y|x_1,x_2)$ with input alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$ and an output alphabet $\mathcal{Y}$. The decoding metric is denoted by $q(x_1,x_2,y)$. We define $W^n(\boldsymbol{y}|\boldsymbol{x}_1,\boldsymbol{x}_2)\triangleq\prod_{i=1}^n W(y_i|x_{1,i},x_{2,i})$ and $q^n(\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{y})\triangleq\prod_{i=1}^n q(x_{1,i},x_{2,i},y_i)$, where y_i is the i-th entry of $\boldsymbol{y}$ and similarly for $x_{1,i}$ and $x_{2,i}$.

The encoders and decoder operate as follows. Encoder $\nu=1,2$ selects a message m_ν equiprobably from the set $\{1,\dots,M_\nu\}$, and transmits the corresponding codeword $\boldsymbol{x}_\nu^{(m_\nu)}$ from the codebook $\mathcal{C}_\nu=\{\boldsymbol{x}_\nu^{(1)},\dots,\boldsymbol{x}_\nu^{(M_\nu)}\}$. Upon receiving $\boldsymbol{y}$ at the output of the channel, the decoder forms an estimate $(\hat{m}_1,\hat{m}_2)$ of the messages, given by

$$(\hat{m}_1,\hat{m}_2)=\operatorname*{arg\,max}_{i\in\{1,\dots,M_1\},j\in\{1,\dots,M_2\}} q^n(\boldsymbol{x}_1^{(i)},\boldsymbol{x}_2^{(j)},\boldsymbol{y}). \tag{4}$$

We assume that ties are broken at random. An error is said to have occurred if the estimate $(\hat{m}_1,\hat{m}_2)$ differs from (m_1,m_2). We distinguish between the following three types of error:

(Type 1) $\hat{m}_1\ne m_1$ and $\hat{m}_2=m_2$
(Type 2) $\hat{m}_1=m_1$ and $\hat{m}_2\ne m_2$
(Type 12) $\hat{m}_1\ne m_1$ and $\hat{m}_2\ne m_2$.

The probabilities of these events are denoted by $p_{e,1}$, $p_{e,2}$ and $p_{e,12}$ respectively, and the overall error probability is denoted by p_e. The ensemble-average error probabilities for a given random-coding ensemble are denoted by $\overline{p}_{e,1}$, $\overline{p}_{e,2}$, $\overline{p}_{e,12}$ and $\overline{p}_e$ respectively. Clearly we have

$$\max\{p_{e,1},p_{e,2},p_{e,12}\}\le p_e\le p_{e,1}+p_{e,2}+p_{e,12} \tag{5}$$

and similarly for $\overline{p}_e$.

A rate pair (R_1,R_2) is said to be achievable if, for all $\delta>0$, there exist sequences of codebooks with $M_1\ge\exp(n(R_1-\delta))$ and $M_2\ge\exp(n(R_2-\delta))$ codewords of length n for users 1 and 2 respectively such that $p_e\to 0$. We say that $E(R_1,R_2)$ is an achievable error exponent if there exist sequences of codebooks with $M_1 \geq \exp(nR_1)$ and $M_2 \geq \exp(nR_2)$ codewords of length n such that $\liminf_{n\to\infty} -\frac{1}{n}\log p_e \geq E(R_1, R_2)$.

B. Notation

The probability of an event is denoted by $\mathbb{P}[\cdot]$. The symbol $\sim$ means "distributed as". The marginals of a joint distribution $P_{XY}(x,y)$ are denoted by $P_X(x)$ and $P_Y(y)$. We write $P_X = \widetilde{P}_X$ to denote element-wise equality between two probability distributions on the same alphabet. For a distribution $P_X(x)$, expectations are denoted by $\mathbb{E}_P[\cdot]$, or simply $\mathbb{E}[\cdot]$ when the probability distribution is understood from the context. We write $\mathbb{E}_P[\log q]$ as a shorthand for $\mathbb{E}_P[\log q(X_1, X_2, Y)]$.

Given a distribution $Q(x)$ and a conditional distribution $W(y|x)$, we write $Q \times W$ to denote the joint distribution $Q(x)W(y|x)$. Mutual information with respect to a joint distribution $P_{XY}(x,y)$ is written as $I_P(X;Y)$. All logarithms have base e, and all rates are in units of nats.

II. Random-Coding Error Probability

In this section, we extend the random-coding union (RCU) bound for mismatched decoders [8], [11] to the MAC. We consider a general codeword distribution of the form

$$\Big(\{\boldsymbol{X}_1^{(i)}\}_{i=1}^{M_1}, \{\boldsymbol{X}_2^{(j)}\}_{j=1}^{M_2}\Big) \sim \prod_{i=1}^{M_1} P_{\boldsymbol{X}_1}(\boldsymbol{x}_1^{(i)}) \prod_{j=1}^{M_2} P_{\boldsymbol{X}_2}(\boldsymbol{x}_2^{(j)}). \tag{6}$$

The analysis and main results of this paper extend immediately to the case that each codeword is generated conditionally on a time-sharing sequence $\boldsymbol{u}$. However, in the mismatched setting, there are some subtle differences between the performance of this ensemble and that of explicit time-sharing, and their study is beyond the scope of this paper.

Theorem 1. *Under the random-coding distribution in* (6) *and maximum metric decoding, the ensemble average error probabilities satisfy*

$$\overline{p}_{e,\nu} \leq \mathrm{rcu}_\nu(n, M_\nu),\ \nu = 1,2 \tag{7}$$

$$\overline{p}_{e,12} \leq \min_{\nu=1,2} \mathrm{rcu}_{12,\nu}(n, M_1, M_2), \tag{8}$$

where

$$\mathrm{rcu}_1(n, M_1) \triangleq \mathbb{E}\Bigg[\min\Bigg\{1, (M_1-1)\mathbb{P}\bigg[\frac{q^n(\overline{\boldsymbol{X}}_1, \boldsymbol{X}_2, \boldsymbol{Y})}{q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})} \geq 1 \,\bigg|\, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}\bigg]\Bigg\}\Bigg] \tag{9}$$

$$\mathrm{rcu}_2(n, M_2) \triangleq \mathbb{E}\Bigg[\min\Bigg\{1, (M_2-1)\mathbb{P}\bigg[\frac{q^n(\boldsymbol{X}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{Y})}{q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})} \geq 1 \,\bigg|\, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}\bigg]\Bigg\}\Bigg] \tag{10}$$

$$\mathrm{rcu}_{12,1}(n, M_1, M_2) \triangleq \mathbb{E}\Bigg[\min\Bigg\{1, (M_1-1)\mathbb{E}\Bigg[\min\Bigg\{1, (M_2-1)\mathbb{P}\bigg[\frac{q^n(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{Y})}{q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})} \geq 1 \,\bigg|\, \overline{\boldsymbol{X}}_1\bigg]\Bigg\} \,\Bigg|\, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}\Bigg]\Bigg\}\Bigg] \tag{11}$$

$$\mathrm{rcu}_{12,2}(n, M_1, M_2) \triangleq \mathbb{E}\Bigg[\min\Bigg\{1, (M_2-1)\mathbb{E}\Bigg[\min\Bigg\{1, (M_1-1)\mathbb{P}\bigg[\frac{q^n(\overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2, \boldsymbol{Y})}{q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})} \geq 1 \,\bigg|\, \overline{\boldsymbol{X}}_2\bigg]\Bigg\} \,\Bigg|\, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}\Bigg]\Bigg\}\Bigg] \tag{12}$$

and

$$(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}, \overline{\boldsymbol{X}}_1, \overline{\boldsymbol{X}}_2) \sim P_{\boldsymbol{X}_1}(\boldsymbol{x}_1)P_{\boldsymbol{X}_2}(\boldsymbol{x}_2) \times W^n(\boldsymbol{y}|\boldsymbol{x}_1, \boldsymbol{x}_2)P_{\boldsymbol{X}_1}(\overline{\boldsymbol{x}}_1)P_{\boldsymbol{X}_2}(\overline{\boldsymbol{x}}_2). \tag{13}$$

Proof: The RCU bounds in (9)–(10) follow using the same steps as the single-user setting [11], so we focus on (11)–(12). We assume without loss of generality that message $(1,1)$ is transmitted. The type-12 error probability satisfies

$$\overline{p}_{e,12} \leq \mathbb{P}\Bigg[\bigcup_{i\neq 1, j\neq 1}\bigg\{\frac{q^n(\boldsymbol{X}_1^{(i)}, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y})}{q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})} \geq 1\bigg\}\Bigg]. \tag{14}$$

Writing the probability as an expectation given $(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})$ and applying the truncated union bound to the union over i, we obtain

$$\overline{p}_{e,12} \leq \mathbb{E}\Bigg[\min\Bigg\{1, (M_1-1) \times \mathbb{P}\Bigg[\bigcup_{j\neq 1}\bigg\{\frac{q^n(\overline{\boldsymbol{X}}_1, \boldsymbol{X}_2^{(j)}, \boldsymbol{Y})}{q^n(\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y})} \geq 1\bigg\} \,\Bigg|\, \boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{Y}\Bigg]\Bigg\}\Bigg]. \tag{15}$$

Applying the same argument again, this time to the union over j, we obtain (11). By applying the same steps with the union bounds applied in the reverse order, we obtain (12). ■

The treatment of one union at a time in (14) is crucial to the analysis. If the truncated union bound was instead applied to all $(M_1-1)(M_2-1)$ events then we would not only recover a worse bound on the error probability, but also a worse error exponent and achievable rate region in Sections III and IV respectively. Our method for refining the union bound is significantly different from that of [1]; see [6] for further discussion.

While the refined bounds in (11)–(12) do not improve the error exponent and rate region recovered under maximum-likelihood (ML) decoding, they may be of independent interest for characterizing the finite-length performance.

III. Random-Coding Error Exponents

In this section, we consider a cost-constrained ensemble with multiple cost functions, given by

$$P_{\boldsymbol{X}_\nu}(\boldsymbol{x}_\nu) = \frac{1}{\mu_{\nu,n}}\prod_{i=1}^{n} Q_\nu(x_{\nu,i})\mathbb{1}\big\{\boldsymbol{x}_\nu \in \mathcal{D}_{\nu,n}\big\}, \tag{16}$$

where for $\nu = 1, 2$, $Q_\nu \in \mathcal{P}(\mathcal{X}_\nu)$ is an input distribution, $\mu_{\nu,n}$ is a normalizing constant, and

$$\mathcal{D}_{\nu,n} \triangleq \left\{ \boldsymbol{x}_\nu \,:\, \left| \frac{1}{n} \sum_{i=1}^{n} a_{\nu,l}(x_{\nu,i}) - \phi_{\nu,l} \right| \le \frac{\delta}{n},\, l = 1, \dots, L_\nu \right\} \tag{17}$$

$$\phi_{\nu,l} \triangleq \mathbb{E}_{Q_\nu}\left[a_{\nu,l}(X_\nu)\right]. \tag{18}$$

Here $\{a_{\nu,l}\}_{l=1}^{L_\nu}$ are cost functions, and δ is a positive constant. Thus, the codewords for user ν are constrained to satisfy L_ν cost constraints in which the empirical mean of $a_{\nu,l}(\cdot)$ is close to the true mean.

Similar ensembles have been considered previously in the single-user mismatched decoding setting [5], [7]–[9]. Similarly to these works, the role of the functions $a_{\nu,l}(\cdot)$ is not to meet system constraints (e.g. power limitations), but instead to improve the performance of the random-coding ensemble itself. Thus, each cost function can be seen as a *pseudo-cost*. However, the results of this paper can easily be extended to the case that system costs are present, by handling them similarly to the pseudo-costs [8, Sec. VII].

Theorem 2. *For any pair of input distributions $\boldsymbol{Q} = (Q_1, Q_2)$, an achievable error exponent for the mismatched memoryless MAC is given by*

$$E_r(\boldsymbol{Q}, R_1, R_2) \triangleq \min\Big\{E_{r,1}(\boldsymbol{Q}, R_1), E_{r,2}(\boldsymbol{Q}, R_2), E_{r,12}(\boldsymbol{Q}, R_1, R_2)\Big\}, \tag{19}$$

where

$$E_{r,1}(\boldsymbol{Q}, R_1) \triangleq \sup_{\rho_1 \in [0,1]} E_{0,1}(\boldsymbol{Q}, \rho_1) - \rho_1 R_1 \tag{20}$$

$$E_{r,2}(\boldsymbol{Q}, R_2) \triangleq \sup_{\rho_2 \in [0,1]} E_{0,2}(\boldsymbol{Q}, \rho_2) - \rho_2 R_2 \tag{21}$$

$$E_{r,12}(\boldsymbol{Q}, R_1, R_2) \triangleq \max\Big\{E_{r,12,1}(\boldsymbol{Q}, R_1, R_2), E_{r,12,2}(\boldsymbol{Q}, R_1, R_2)\Big\} \tag{22}$$

$$E_{r,12,1}(\boldsymbol{Q}, R_1, R_2) \triangleq \sup_{\rho_1 \in [0,1], \rho_2 \in [0,1]} E_{0,12,1}(\boldsymbol{Q}, \rho_1, \rho_2) - \rho_1(R_1 + \rho_2 R_2) \tag{23}$$

$$E_{r,12,2}(\boldsymbol{Q}, R_1, R_2) \triangleq \sup_{\rho_1 \in [0,1], \rho_2 \in [0,1]} E_{0,12,2}(\boldsymbol{Q}, \rho_1, \rho_2) - \rho_2(R_2 + \rho_1 R_1), \tag{24}$$

and

$$E_{0,1}(\boldsymbol{Q}, \rho_1) \triangleq \sup_{s \ge 0, a_1(\cdot)} -\log \mathbb{E}\left[\left(\frac{\mathbb{E}\big[q(\overline{X}_1, X_2, Y)^s e^{a_1(\overline{X}_1)} \,|\, X_2, Y\big]}{q(X_1, X_2, Y)^s e^{a_1(X_1)}}\right)^{\rho_1}\right] \tag{25}$$

$$E_{0,2}(\boldsymbol{Q}, \rho_2) \triangleq \sup_{s \ge 0, a_2(\cdot)} -\log \mathbb{E}\left[\left(\frac{\mathbb{E}\big[q(X_1, \overline{X}_2, Y)^s e^{a_2(\overline{X}_2)} \,|\, X_1, Y\big]}{q(X_1, X_2, Y)^s e^{a_2(X_2)}}\right)^{\rho_2}\right] \tag{26}$$

$$E_{0,12,1}(\boldsymbol{Q}, \rho_1, \rho_2) \triangleq \sup_{s \ge 0, a_1(\cdot), a_2(\cdot)} -\log \mathbb{E}\left[\left(\mathbb{E}\left[\left(\frac{\mathbb{E}\big[q(\overline{X}_1, \overline{X}_2, Y)^s e^{a_2(\overline{X}_2)} \,\big|\, \overline{X}_1\big]}{q(X_1, X_2, Y)^s e^{a_2(X_2)}}\right)^{\rho_2} \times \frac{e^{a_1(\overline{X}_1)}}{e^{a_1(X_1)}} \,\middle|\, X_1, X_2, Y\right]\right)^{\rho_1}\right] \tag{27}$$

$$E_{0,12,2}(\boldsymbol{Q}, \rho_1, \rho_2) \triangleq \sup_{s \ge 0, a_1(\cdot), a_2(\cdot)} -\log \mathbb{E}\left[\left(\mathbb{E}\left[\left(\frac{\mathbb{E}\big[q(\overline{X}_1, \overline{X}_2, Y)^s e^{a_1(\overline{X}_1)} \,\big|\, \overline{X}_2\big]}{q(X_1, X_2, Y)^s e^{a_1(X_1)}}\right)^{\rho_1} \times \frac{e^{a_2(\overline{X}_2)}}{e^{a_2(X_2)}} \,\middle|\, X_1, X_2, Y\right]\right)^{\rho_2}\right] \tag{28}$$

$$(X_1, X_2, Y, \overline{X}_1, \overline{X}_2) \sim Q_1(x_1)Q_2(x_2) \times W(y|x_1, x_2)Q_1(\overline{x}_1)Q_2(\overline{x}_2). \tag{29}$$

For $\nu = 1, 2$, each supremum over $a_\nu(\cdot)$ is taken over all real-valued functions on $\mathcal{X}_\nu$ such that the second moment of $a_\nu(X_\nu)$ is finite.

Proof: The proof is similar for each exponent, so we focus on $E_{r,12,1}$. We define $Q_\nu^n(\boldsymbol{x}_\nu) \triangleq \prod_{i=1}^n Q_\nu(x_{\nu,i})$ and $a_{\nu,l}^n(\boldsymbol{x}_\nu) \triangleq \sum_{i=1}^n a_{\nu,l}(x_{\nu,i})$. Expanding (11) and applying Markov's inequality and $\min\{1, \alpha\} \le \alpha^\rho$ ($0 \le \rho \le 1$), we obtain[1]

$$\mathrm{rcu}_{12,1}(n, M_1) \le \frac{1}{\mu_{1,n}^{1+\rho_1}\mu_{2,n}^{1+\rho_1\rho_2}} \sum_{\boldsymbol{x}_1 \in \mathcal{D}_{1,n}, \boldsymbol{x}_2 \in \mathcal{D}_{2,n}, \boldsymbol{y}} Q_1^n(\boldsymbol{x}_1) \times Q_2^n(\boldsymbol{x}_2) W^n(\boldsymbol{y}|\boldsymbol{x}_1, \boldsymbol{x}_2) \left(M_1 \sum_{\overline{\boldsymbol{x}}_1 \in \mathcal{D}_{1,n}} Q_1^n(\overline{\boldsymbol{x}}_1) \times \left(M_2 \sum_{\overline{\boldsymbol{x}}_2 \in \mathcal{D}_{2,n}} Q_2^n(\overline{\boldsymbol{x}}_2)\left(\frac{q^n(\overline{\boldsymbol{x}}_1, \overline{\boldsymbol{x}}_2, \boldsymbol{y})}{q^n(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y})}\right)^s\right)^{\rho_2}\right)^{\rho_1} \tag{30}$$

for any $\rho_1 \in [0, 1]$, $\rho_2 \in [0, 1]$ and $s \ge 0$. From the definition of $\mathcal{D}_{\nu,n}$ in (17), we have that

$$\exp\left(\frac{a_{\nu,l}^n(\overline{\boldsymbol{x}}_\nu)}{a_{\nu,l}^n(\boldsymbol{x}_\nu)}\right)e^{2\delta} \ge 1 \tag{31}$$

for any $\boldsymbol{x}_\nu, \overline{\boldsymbol{x}}_\nu \in \mathcal{D}_{\nu,n}$. We upper-bound (30) by multiplying the $(\cdot)^s$ term by the left-hand side of (31) for $\nu = 2$ and some $l \in \{1, \dots, L_2\}$, and multiplying the $(\cdot)^{\rho_2}$ term by the

[1] In the case of continuous alphabets, summations should be replaced by integrals.

left-hand side of (31) for $\nu = 1$ and some $l \in \{1, \dots, L_1\}$. Furthermore, we replace the summations over $\mathcal{D}_{\nu,n}$ by summations over all sequences on $\mathcal{X}_\nu^n$. Expanding the resulting terms (e.g. $Q_\nu^n(\boldsymbol{x}_\nu)$) as a product from 1 to n and taking the supremum over (s, ρ_1, ρ_2) and the cost functions, we obtain a bound whose exponent is given by (23), with a prefactor of

$$\frac{e^{2\delta(\rho_1+\rho_1\rho_2)}}{\mu_{1,n}^{1+\rho_1}\mu_{2,n}^{1+\rho_1\rho_2}}. \tag{32}$$

From [8, Prop. 1] and the assumptions on the second moments of the cost functions, this prefactor is subexponential in n. By choosing $L_1 = L_2 = 3$, we can ensure that different cost functions can be used for each error type, thus allowing the suprema over $a_1(\cdot)$ and $a_2(\cdot)$ in (25)–(28) to be taken individually in each equation. It suffices to let the cost functions for $E_{0,12,1}$ and $E_{0,12,2}$ coincide, since the type-12 error exponent is the maximum of the two. ∎

Analogously to the single-user analysis of [8] and the matched MAC analysis of [9], it is possible to improve the exponents in (25)–(28) by replacing each occurrence of $e^{a_\nu(x)}$ (respectively, $e^{a_\nu(\overline{x})}$) with $e^{r_\nu(a_\nu(x)-\phi_\nu)}$ (respectively, $e^{\overline{r}_\nu(a_\nu(x)-\phi_\nu)}$) for $\nu = 1, 2$, where $\phi_\nu \triangleq \mathbb{E}_{Q_\nu}[a_\nu(X_\nu)]$ and r_ν (respectively, $\overline{r}_\nu$) is an arbitrary real number. The exponents may increase further if each equation is modified to contain multiple cost functions per user [8]. We focus on the weaker exponents in (25)–(28) for clarity of exposition, and because they yield the same achievable rate region for any given (Q_1, Q_2).

IV. Achievable Rate Region

By determining the conditions under which $E_r(\boldsymbol{Q}, R_1, R_2)$ is positive, we obtain the following achievable rate region.

Theorem 3. *An achievable rate region for the mismatched memoryless MAC is given by*

$$\mathrm{cl}\bigg(\bigcup_{Q_1,Q_2} \mathcal{R}^{\mathrm{LM}}(\boldsymbol{Q})\bigg), \tag{33}$$

where $\mathrm{cl}(\cdot)$ *denotes convex closure, and* $\mathcal{R}^{\mathrm{LM}}(\boldsymbol{Q})$ *is the set of rate pairs* (R_1, R_2) *satisfying*

$$R_1 \le \sup_{s\ge0,a_1(\cdot)} \mathbb{E}\left[\log \frac{q(X_1,X_2,Y)^s e^{a_1(X_1)}}{\mathbb{E}\big[q(\overline{X}_1,X_2,Y)^s e^{a_1(\overline{X}_1)}\,|\,X_2,Y\big]}\right] \tag{34}$$

$$R_2 \le \sup_{s\ge0,a_2(\cdot)} \mathbb{E}\left[\log \frac{q(X_1,X_2,Y)^s e^{a_2(X_2)}}{\mathbb{E}\big[q(X_1,\overline{X}_2,Y)^s e^{a_2(\overline{X}_2)}\,|\,X_1,Y\big]}\right] \tag{35}$$

and at least one of

$$R_1 \le \sup_{\rho_2\in[0,1],s\ge0,a_1(\cdot),a_2(\cdot)} -\rho_2 R_2 - \mathbb{E}\Bigg[\log\mathbb{E}\Bigg[\left(\frac{\mathbb{E}\big[q(\overline{X}_1,\overline{X}_2,Y)^s e^{a_2(\overline{X}_2)}\,\big|\,\overline{X}_1\big]}{q(X_1,X_2,Y)^s e^{a_2(X_2)}}\right)^{\rho_2} \frac{e^{a_1(\overline{X}_1)}}{e^{a_1(X_1)}}\Bigg| X_1,X_2,Y\Bigg]\Bigg] \tag{36}$$

$$R_2 \le \sup_{\rho_1\in[0,1],s\ge0,a_1(\cdot),a_2(\cdot)} -\rho_1 R_1 - \mathbb{E}\Bigg[\log\mathbb{E}\Bigg[\left(\frac{\mathbb{E}\big[q(\overline{X}_1,\overline{X}_2,Y)^s e^{a_1(\overline{X}_1)}\,\big|\,\overline{X}_2\big]}{q(X_1,X_2,Y)^s e^{a_1(X_1)}}\right)^{\rho_1} \frac{e^{a_2(\overline{X}_2)}}{e^{a_2(X_2)}}\Bigg| X_1,X_2,Y\Bigg]\Bigg] \tag{37}$$

under the joint distribution in (29)*. For* $\nu = 1, 2$*, each supremum over* $a_\nu(\cdot)$ *is taken over all real-valued functions on* $\mathcal{X}_\nu$ *such that the second moment of* $a_\nu(X_\nu)$ *is finite.*

Proof: These conditions follow from (20)–(24) using the techniques of Gallager [12]. Namely, we take $\rho_1 \to 0$ in (25) and (27), and $\rho_2 \to 0$ in (26) and (28). The convex closure operation follows using time-sharing [1]. ∎

In the remainder of this section, we outline the proof that the region described by (34)–(37) coincides with (1)–(3) for any mismatched DM-MAC. The equivalence of (34)–(35) and (1)–(2) follows using Lagrange duality, similarly to the equivalence of the primal and dual expressions for the single-user rate of Csiszár-Körner-Hui [2], [5], [8]. For the conditions in (36)–(37), we make use of the following.

Lemma 1. *For any given* (Q_1, Q_2)*, the condition* (3) *holds if and only if*

$$R_1 + R_2 \le \max\left\{ \min_{\substack{\widetilde{P}_{X_1}=Q_1,\widetilde{P}_{X_2}=Q_2,\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q]\ge\mathbb{E}_P[\log q],I_{\widetilde{P}}(X_1;Y)\le R_1}} D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times\widetilde{P}_Y), \min_{\substack{\widetilde{P}_{X_1}=Q_1,\widetilde{P}_{X_2}=Q_2,\widetilde{P}_Y=P_Y \\ \mathbb{E}_{\widetilde{P}}[\log q]\ge\mathbb{E}_P[\log q],I_{\widetilde{P}}(X_2;Y)\le R_2}} D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times\widetilde{P}_Y) \right\}. \tag{38}$$

Proof: It is easily seen that (3) holds whenever (38) holds, since each minimization in (38) is obtained by removing a constraint from the minimization in (3). It remains to show that the converse is true. To this end, we will show that the right-hand side of (3) (denoted by f_0) only exceeds that of (38) (denoted by f_0') when both (3) and (38) hold.

Assume $f_0 > f_0'$. It follows that the constraints $I_{\widetilde{P}}(X_1;Y) \le R_1$ and $I_{\widetilde{P}}(X_2;Y) \le R_2$ in (3) are active, and hence at least one of these constraints in the two minimizations of (38) are active. Let us assume that $I_{\widetilde{P}}(X_1;Y) \le R_1$ is active in the first minimization; the other case can be handled similarly. If the minimizing $\widetilde{P}_{X_1X_2Y}$ also satisfied $I_{\widetilde{P}}(X_2;Y) \le R_2$, then it would satisfy the constraints of (3), contradicting the assumption that $f_0 > f_0'$; it follows that $I_{\widetilde{P}}(X_2;Y) > R_2$. Using the identity

$$D(\widetilde{P}_{X_1X_2Y}\|Q_1\times Q_2\times\widetilde{P}_Y) = I_{\widetilde{P}}(X_1;Y) + I_{\widetilde{P}}(X_2;Y) + I_{\widetilde{P}}(X_1;X_2|Y) \tag{39}$$

and the fact that active constraints hold with equality in convex optimization problems [13], it follows that both f_0 and f_0' are greater than or equal to $R_1 + R_2$, i.e. both (3) and (38) hold. ∎

Using Lagrange duality techniques similarly to [2], [8], it can be verified that (38) holds if and only if at least one of (36)–(37) hold, thus proving the above-mentioned equivalence.

As well as being more amenable to Lagrange duality techniques, the condition in (38) simplifies the computation of the region. Specifically, instead of evaluating the minimization in (3) over a grid of (R_1, R_2) values, one can evaluate the first minimization in (38) over a range of R_1 values, and the second minimization over a range of R_2 values.

V. Application to the Single-User Setting

It was shown in [1] that one can improve on the single-user rate of Csiszár-Körner-Hui [3], [10] by treating the single-user mismatched channel as a MAC. That is, given a channel $W(y|x)$ and decoding metric $q(x, y)$, one can fix $\mathcal{X}_1$, $\mathcal{X}_2$ and $\phi : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}$ and analyze the multiple-access channel defined by $W(y|x_1, x_2) = W\big(y|\phi(x_1, x_2)\big)$ with the decoding metric $q(x_1, x_2, y) = q\big(\phi(x_1, x_2), y\big)$, and obtain better random-coding achievable rates than the standard single-user ensembles. A further improvement was obtained by expurgation, yielding an achievable rate of the form $R = R_1+R_2$ for (R_1, R_2) satisfying (1)–(3), where each minimization is further subject to $\widetilde{P}_{X_1X_2} = Q_1 \times Q_2$. As noted in [1], such expurgation is not allowed in the multiple-access setting, since it implies cooperation between the users.

By combining the techniques of [1] and the present paper, one can obtain the following theorem, whose proof is omitted due to space constraints. The key difference in the derivation compared to that of [1, Thm. 4] is that instead of keeping codewords pairs of a given joint empirical distribution, we only keep pairs such that the empirical means of L joint cost functions $a_l(x_1, x_2)$ $(l = 1, \cdots, L)$ are close to the corresponding true means, similarly to (17).

Theorem 4. *Let a single-user channel $W(y|x)$ with decoding metric $q(x, y)$ be given. Fix any alphabets $\mathcal{X}_1$ and $\mathcal{X}_2$, input distributions Q_1 and Q_2, and function $\phi : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathcal{X}$. An achievable rate for the single-user channel is given by $R = R_1 + R_2$ for any (R_1, R_2) satisfying*

$$R_1 \le \sup_{s\ge0,a(\cdot,\cdot)} \mathbb{E}\left[\log \frac{q\big(\phi(X_1, X_2), Y\big)^s e^{a(X_1,X_2)}}{\mathbb{E}\big[q\big(\phi(\overline{X}_1, X_2), Y\big)^s e^{a(\overline{X}_1,X_2)} \,\big|\, X_2, Y\big]}\right] \tag{40}$$

$$R_2 \le \sup_{s\ge0,a(\cdot,\cdot)} \mathbb{E}\left[\log \frac{q\big(\phi(X_1, X_2), Y\big)^s e^{a(X_1,X_2)}}{\mathbb{E}\big[q\big(\phi(X_1, \overline{X}_2), Y\big)^s e^{a(X_1,\overline{X}_2)} \,\big|\, X_1, Y\big]}\right] \tag{41}$$

and at least one of

$$R_1 \le \sup_{\rho_2\in[0,1],s\ge0,a(\cdot,\cdot)} -\rho_2R_2 - \mathbb{E}\Bigg[\log \mathbb{E}\Bigg[\left(\frac{\mathbb{E}\big[q\big(\phi(\overline{X}_1, \overline{X}_2), Y\big)^s e^{a(\overline{X}_1,\overline{X}_2)} \,\big|\, \overline{X}_1\big]}{q\big(\phi(X_1, X_2), Y\big)^s e^{a(X_1,X_2)}}\right)^{\rho_2} \Bigg| X_1, X_2, Y\Bigg]\Bigg] \tag{42}$$

$$R_2 \le \sup_{\rho_1\in[0,1],s\ge0,a(\cdot,\cdot)} -\rho_1R_1 - \mathbb{E}\Bigg[\log \mathbb{E}\Bigg[\left(\frac{\mathbb{E}\big[q\big(\phi(\overline{X}_1, \overline{X}_2), Y\big)^s e^{a(\overline{X}_1,\overline{X}_2)} \,\big|\, \overline{X}_2\big]}{q\big(\phi(X_1, X_2), Y\big)^s e^{a(X_1,X_2)}}\right)^{\rho_1} \Bigg| X_1, X_2, Y\Bigg]\Bigg] \tag{43}$$

under the joint distribution in (29) *with $W(y|x_1, x_2) \triangleq W(y|\phi(x_1, x_2))$. Each supremum over $a(\cdot, \cdot)$ is taken over all real-valued functions on $\mathcal{X}_1 \times \mathcal{X}_2$ such that the second moment of $a(X_1, X_2)$ is finite.*

In the discrete memoryless setting, it can be shown using Lagrange duality that Theorem 4 coincides with [1, Thm. 4]. The proof follows the same steps as those use to prove the equivalence of (34)–(37) and (1)–(3). The only difference is the additional constraint $\widetilde{P}_{X_1X_2} = Q_1 \times Q_2$ in each primal expression, and the use of joint (rather than individual) functions $a(x_1, x_2)$ in each dual expression.

VI. Conclusion

We have analyzed the memoryless MAC with general alphabets. Random-coding union (RCU) bounds on the error probabilities have been given, and error exponents for the cost-constrained ensemble have been given. The resulting achievable rate region coincides with that of Lapidoth [1] for any DM-MAC, and remains valid for more general alphabets. The application to the single-user setting has been discussed.

References

[1] A. Lapidoth, "Mismatched decoding and the multiple-access channel," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1439–1452, Sep. 1996.

[2] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai, "On information rates for mismatched decoders," *IEEE Trans. Inf. Theory*, vol. 40, no. 6, pp. 1953–1967, Nov. 1994.

[3] J. Hui, "Fundamental issues of multiple accessing," Ph.D. dissertation, MIT, 1983.

[4] I. Csiszár and P. Narayan, "Channel capacity for a given decoding metric," *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 35–43, Jan. 1995.

[5] A. Ganti, A. Lapidoth, and E. Telatar, "Mismatched decoding revisited: General alphabets, channels with memory, and the wide-band limit," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2315–2328, Nov. 2000.

[6] J. Scarlett and A. Guillén i Fàbregas, "An achievable error exponent for the mismatched multiple-access channel," in *Allerton Conf. on Comm., Control and Comp.*, Monticello, IL, Oct. 2012.

[7] S. Shamai and I. Sason, "Variations on the Gallager bounds, connections, and applications," *IEEE Trans. Inf. Theory*, vol. 48, no. 12, pp. 3029–3051, Dec. 2002.

[8] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Mismatched decoding: Finite-length bounds, error exponents and approximations," submitted to *IEEE Trans. Inf. Theory* [Online: http://arxiv.org/abs/1303.6166].

[9] J. Scarlett, A. Martinez, and A. Guillén i Fàbregas, "Cost-constrained random coding and applications," in *Inf. Theory and Apps. Workshop*, San Diego, CA, Feb. 2013.

[10] I. Csiszár and J. Körner, "Graph decomposition: A new key to coding theorems," *IEEE Trans. Inf. Theory*, vol. 27, no. 1, pp. 5–12, Jan. 1981.

[11] Y. Polyanskiy, V. Poor, and S. Verdú, "Channel coding rate in the finite blocklength regime," *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[12] R. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inf. Theory*, vol. 11, no. 1, pp. 3–18, Jan. 1965.

[13] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.