WHEN DOES JOINT SOURCE-CHANNEL CODING OUTPERFORM TANDEM CODING FOR SYSTEMS WITH MARKOVIAN MEMORY ?

Yangfan Zhong, Fady Alajaji and L. Lorne Campbell

Department of Mathematics and Statistics
Queen's University, Kingston, ON K7L 3N6, Canada
{yangfan, fady, campblll}@mast.queensu.ca

ABSTRACT

In this work, we explore the merits of joint source-channel coding (JSCC) versus traditional tandem coding, which consists of separately performing and concatenating source and channel coding, from the point of view of the error exponent (or reliability function). For a communication system with a discrete stationary ergodic Markov (SEM) source $\mathbf{Q}$ and a discrete channel $\mathbf{W}$ with additive SEM noise, we present sufficient conditions under which the JSCC error exponent $E_J(\mathbf{Q},\mathbf{W})$ strictly outperforms the tandem coding error exponent $E_T(\mathbf{Q},\mathbf{W})$. We observe that these conditions are satisfied by a large class of SEM source-channel pairs. Numerical examples also indicate that $E_J(\mathbf{Q},\mathbf{W})$ can be nearly twice as large as $E_T(\mathbf{Q},\mathbf{W})$ for many source-channel pairs even though $E_J(\mathbf{Q},\mathbf{W})$ is upper bounded by $2E_T(\mathbf{Q},\mathbf{W})$.

1. INTRODUCTION

In [8], we investigate the joint source-channel coding (JSCC) error exponent $E_J(\mathbf{Q},\mathbf{W})$ for a communication system with memory. Specifically, we derive an upper bound for $E_J(\mathbf{Q},\mathbf{W})$ for a system consisting of a stationary ergodic Markov (SEM) source $\mathbf{Q}$ and a channel $\mathbf{W}$ with additive SEM noise $\mathbf{P_W}$ (for the sake of brevity, we hereafter refer to this channel as the SEM channel $\mathbf{W}$). We also examine Gallager's lower bound for $E_J(\mathbf{Q},\mathbf{W})$ [3] (which is valid for arbitrary source-channel pairs with memory), when specialized to the SEM source-channel system. By comparing the upper and lower bounds, we provide the condition under which they coincide, hence exactly determining $E_J(\mathbf{Q},\mathbf{W})$. We note that this condition holds for many SEM source-channel pairs.

In this paper, we focus our interests on the comparison of the JSCC error exponent $E_J(\mathbf{Q},\mathbf{W})$ with the tandem coding error exponent $E_T(\mathbf{Q},\mathbf{W})$, the exponent resulting by separately performing and concatenating optimal source and channel coding. As in [6, 7], which consider the JSCC error exponent for discrete memoryless systems, we investigate the situation where $E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W})$ for the same SEM source-channel pair. Indeed, as pointed out in [6], this inequality, when it holds, provides a theoretical underpinning and justification for JSCC design as opposed to the widely used classical tandem or separate coding approach, since the former method provides a faster exponential rate of decay for the error probability, which often translates into substantial reductions in complexity and delay for real world applications. We herein establish sufficient conditions for which $E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W})$, and we observe via numerical examples that such conditions are satisfied by a wide class of SEM source-channel pairs.

This work was supported in part by NSERC and PREA.

2. PRELIMINARIES

Without loss of generality, we only deal with first-order Markov sources since any k-th order Markov source can be converted to a first-order Markov source by k-step blocking it. For the sake of convenience (since we will apply the following results to both the SEM source and the SEM channel), throughout this section, we use $\mathbf{P}$ to denote a first-order SEM source with finite alphabet $\mathcal{U} = \{1, 2, ..., M\}$ and transition distribution matrix $P \triangleq [p_{ij}]_{M\times M}$. The entropy rate of source $\mathbf{P}$ is given by

$$H(\mathbf{P}) = -\sum_i \sum_j \pi_i p_{ij} \log_2 p_{ij},$$

where $\pi = (\pi_1, \pi_2, ..., \pi_M)$ is the stationary distribution of the stochastic matrix P.

For any $\alpha \in [0, 1]$, we set $P(\alpha) \triangleq \left[p_{ij}^{\alpha}\right]_{M\times M}$, which is nonnegative and irreducible. The Perron-Frobenius Theorem [4] asserts that the matrix $P(\alpha)$ possesses a maximal positive eigenvalue $\lambda_\alpha(\mathbf{P})$ with positive (right) eigenvector

$$\mathbf{v}(\alpha) \triangleq (v_1(\alpha), ..., v_M(\alpha))^T$$

such that $\sum_i v_i(\alpha) = 1$. As in [5], we define the artificial Markov source $\widetilde{\mathbf{P}}_\alpha$ with respect to the original source $\mathbf{P}$ such that the transition matrix $\widetilde{P}(\alpha) \triangleq [\widetilde{p}_{ij}(\alpha)]_{M\times M}$, where

$$\widetilde{p}_{ij}(\alpha) \triangleq \frac{p_{ij}^{\alpha} v_j(\alpha)}{\lambda_\alpha(\mathbf{P}) v_i(\alpha)}.$$

(It can be easily verified that $\sum_j \widetilde{p}_{ij}(\alpha) = 1$.) We emphasize that the artificial source retains the stochastic characteristics (ergodicity) of the original source because $\widetilde{p}_{ij}(\alpha) = 0$ if and only if $p_{ij} = 0$. The entropy rate of the artificial Markov source is hence given by

$$H(\widetilde{\mathbf{P}}_\alpha) = -\sum_i \sum_j \pi_i(\alpha) \widetilde{p}_{ij}(\alpha) \log_2 \widetilde{p}_{ij}(\alpha),$$

where $\pi(\alpha) \triangleq (\pi(\alpha)_1, \pi(\alpha)_2, ..., \pi(\alpha)_M)$ is the stationary distribution of the stochastic matrix $\widetilde{P}(\alpha)$.

3. JSCC ERROR EXPONENT

Definition 1 A joint source-channel code with blocklength n for a discrete source with finite alphabet $\mathcal{S}$ described by the sequence of $k(n)$-dimensional distributions $\mathbf{Q} \triangleq \{Q^{(k(n))} : \mathcal{S}^{k(n)}\}_{k(n)=1}^{\infty}$ and a discrete channel described by the sequence of n-dimensional transition distributions $\mathbf{W} \triangleq \{W^{(n)} : \mathcal{X}^n \to \mathcal{Y}^n\}_{n=1}^{\infty}$ with common input and output alphabets $\mathcal{X} = \mathcal{Y} = \{0, 1, ..., B-1\}$ is a pair of mappings: $f_n : \mathcal{S}^{k(n)} \longrightarrow \mathcal{X}^n$ and $\varphi_n : \mathcal{Y}^n \longrightarrow \mathcal{S}^{k(n)}$.

In this work, we confine our attention to discrete channels with (modulo B) additive noise of n-dimensional distribution $\mathbf{P_W} \triangleq \{P_W^{(n)} : \mathcal{Z}^n\}_{n=1}^{\infty}$. The channels are described by

$$Y_i = X_i \oplus Z_i \quad (\text{modulo } B),$$

where Y_i, X_i and Z_i are the output, input and noise symbols at time i, and $Z_i \in \mathcal{Z} = \{0, 1, ..., B-1\}$ is independent of X_i, $i = 1, 2, ..., n$. The average probability of error of the code (f_n, φ_n) is

$$\begin{aligned} & P_e^{(n)}(Q^{(k(n))}, W^{(n)}) \\ = & \sum_{\mathbf{s}} Q^{(k(n))}(\mathbf{s}) \sum_{\mathbf{y}:\varphi_n(\mathbf{y})\neq \mathbf{s}} W^{(n)}(\mathbf{y}|f_n(\mathbf{s})). \end{aligned}$$

Since $k(n)$ source symbols are mapped to n channel symbols, $R_t \triangleq k(n)/n$ source symbols/channel use is called the code's transmission rate, which is assumed to be independent of n.

Definition 2 The JSCC error exponent $E_J(\mathbf{Q}, \mathbf{W})$ for source $\mathbf{Q}$ and channel $\mathbf{W}$ is defined as the largest number E for which there exists a sequence of joint source-channel codes (f_n, φ_n) with

$$E \leq \liminf_{n\to\infty} -\frac{1}{n} \log_2 P_e^{(n)}(Q^{(k(n))}, W^{(n)}).$$

Proposition 1 [8] For an SEM source $\mathbf{Q}$ and an SEM channel $\mathbf{W}$ with noise $\mathbf{P_W}$ such that the entropy rates of $\mathbf{Q}$ and $\mathbf{P_W}$ satisfy $R_t H(\mathbf{Q}) + H(\mathbf{P_W}) < \log_2 B$, $E_J(\mathbf{Q}, \mathbf{W})$ is positive and determined exactly by

$$\begin{aligned} E_J(\mathbf{Q}, \mathbf{W}) &= \rho^* \log_2 B - (1+\rho^*) \\ &\quad \times \log_2 \left[\lambda_{\frac{1}{1+\rho^*}}^{R_t}(\mathbf{Q}) \lambda_{\frac{1}{1+\rho^*}}(\mathbf{P_W}) \right] \end{aligned}$$

if $\rho^* \leq 1$, where ρ^* satisfies the equation

$$R_t H\left(\widetilde{\mathbf{Q}}_{\frac{1}{1+\rho^*}}\right) + H\left(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{1+\rho^*}}}\right) = \log_2 B.$$

Otherwise (if $\rho^* > 1$), the following bounds hold

$$\begin{aligned} E_J(\mathbf{Q}, \mathbf{W}) &\leq \rho^* \log_2 B - (1+\rho^*) \\ &\quad \times \log_2 [\lambda_{\frac{1}{1+\rho^*}}^{R_t}(\mathbf{Q}) \lambda_{\frac{1}{1+\rho^*}}(\mathbf{P_W})], \end{aligned}$$

and

$$E_J(\mathbf{Q}, \mathbf{W}) \geq \log_2 B - 2\log_2 [\lambda_{\frac{1}{2}}^{R_t}(\mathbf{Q}) \lambda_{\frac{1}{2}}(\mathbf{P_W})].$$

Remarks:

1. For an SEM source-channel pair ($\mathbf{Q}$,$\mathbf{W}$) with $R_t H(\mathbf{Q}) + H(\mathbf{P_W}) \geq \log_2 B$, $E_J(\mathbf{Q}, \mathbf{W}) = 0$.
2. As a byproduct, the lower and upper bounds of the error exponent for SEM sources/channels directly follow from Proposition 1, cf. [8].
3. It is also shown in [8] that the lower and upper bounds for $E_J(\mathbf{Q}, \mathbf{W})$ enjoy a form that is similar to Csiszár's bounds for discrete memoryless source-channel pairs [1], which are expressed as the minimum of the sum of the source error exponent and the lower/upper bounds of the channel error exponent.

4. TANDEM CODING ERROR EXPONENT

Definition 3 A tandem code $(f_n^*, \varphi_n^*) \triangleq (f_{cn} \circ f_{sn}, \varphi_{sn} \circ \varphi_{cn})$ for a discrete source $\mathbf{Q}$ and a discrete channel $\mathbf{W}$ is composed independently of a $(k(n), M)$ block source code (f_{sn}, φ_{sn}) defined by $f_{sn} : \mathcal{S}^{k(n)} \longrightarrow \{1, 2, ..., M\}$ and $\varphi_{sn} : \{1, 2, ..., M\} \longrightarrow \mathcal{S}^{k(n)}$ with source code rate $R_s \triangleq \log_2 M/k(n)$ source code bits/source symbol, and an (n, M) block channel code (f_{cn}, φ_{cn}) defined by $f_{cn} : \{1, 2, ..., M\} \longrightarrow \mathcal{X}^n$ and $\varphi_{cn} : \mathcal{Y}^n \longrightarrow \{1, 2, ..., M\}$ with channel code rate $R_c \triangleq \log_2 M/n$ source code bits/channel use, where "$\circ$" denotes composition and R_s and R_c are independent of n.

The transmission rate for the tandem code is $R_t = k(n)/n = R_c/R_s$ source symbols/channel use. The code's average error probability is given by

$$\begin{aligned} & P_{e^*}^{(n)}(Q^{(k(n))}, W^{(n)}) \\ = & \sum_{\mathbf{s}} Q^{(k(n))}(\mathbf{s}) \sum_{\mathbf{y}:\varphi_n^*(\mathbf{y})\neq \mathbf{s}} W^{(n)}(\mathbf{y} \mid f_n^*(\mathbf{s})). \end{aligned}$$

Definition 4 The tandem error exponent $E_T(\mathbf{Q}, \mathbf{W})$ for source $\mathbf{Q}$ and channel $\mathbf{W}$ is defined as the largest number $\widehat{E}$ for which there exists a sequence of tandem codes (f_n^*, φ_n^*) such that

$$\widehat{E} \leq \liminf_{n\to\infty} -\frac{1}{n} \log_2 P_{e^*}^{(n)}(Q^{(k(n))}, W^{(n)}).$$

Since the tandem coding exponent results from separately performing and concatenating optimal source and channel coding, it can be easily shown that

$$E_T(\mathbf{Q}, \mathbf{W}) = \sup_R \min \left\{ R_t e\left(\frac{R}{R_t}, \mathbf{Q}\right), E(R, \mathbf{W}) \right\}, \quad (1)$$

where $e(R, \mathbf{Q})$ and $E(R, \mathbf{W})$ are the source and channel error exponents, respectively. To evaluate $E_T(\mathbf{Q}, \mathbf{W})$ for an SEM source-channel pair $\mathbf{Q}$ and $\mathbf{W}$, we recall the fact that $e(R, \mathbf{Q})$ is 0 for $R < H(\mathbf{Q})$, strictly increasing in $H(\mathbf{Q}) \leq R \leq \log_2 \lambda_0(\mathbf{Q})$ and infinity for $R > \log \lambda_0(\mathbf{Q})$,[1] while $E(R, \mathbf{W})$ is decreasing in R, and vanishes at $R = C(\mathbf{W})$, where $C(\mathbf{W}) = \log_2 B - H(\mathbf{P_W})$ is the capacity of the SEM channel $\mathbf{W}$. Therefore, if $R_t e(\log_2 \lambda_0(\mathbf{Q}), \mathbf{Q}) \geq E(R_t \log_2 \lambda_0(\mathbf{Q}), \mathbf{W})$, then the graphs of $R_t e\left(\frac{R}{R_t}, \mathbf{Q}\right)$ and $E(R, \mathbf{W})$ must have exactly one intersection R_o and by (1)

$$E_T(\mathbf{Q}, \mathbf{W}) = R_t e\left(\frac{R_o}{R_t}, \mathbf{Q}\right) = E(R_o, \mathbf{W}).$$

If $R_t e(\log_2 \lambda_0(\mathbf{Q}), \mathbf{Q}) < E(R_t \log_2 \lambda_0(\mathbf{Q}), \mathbf{W})$, then there is no intersection between $R_t e\left(\frac{R}{R_t}, \mathbf{Q}\right)$ and $E(R, \mathbf{W})$. In this case, it follows by (1) that

$$E_T(\mathbf{Q}, \mathbf{W}) = E(R_t \log_2 \lambda_0(\mathbf{Q}), \mathbf{W}).$$

In general, we know that $E_J(\mathbf{Q}, \mathbf{W}) \geq E_T(\mathbf{Q}, \mathbf{W})$ since by definition tandem coding is a special case of JSCC. Meanwhile, Proposition 1 states that $E_J(\mathbf{Q}, \mathbf{W}) = E_T(\mathbf{Q}, \mathbf{W}) = 0$ if $R_t H(\mathbf{Q}) \geq C(\mathbf{W})$ for SEM source-channel pairs.

[1] It can be shown that $\log_2 \lambda_0(\mathbf{Q}) \leq \log_2 |\mathcal{S}|$ with equality iff the stochastic matrix Q is strictly positive.

We are hence interested in determining the conditions for which $E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W})$ when $R_t H(\mathbf{Q}) < C(\mathbf{W})$. Although both $E_J(\mathbf{Q},\mathbf{W})$ and $E_T(\mathbf{Q},\mathbf{W})$ are not always determined, we can still provide some sufficient conditions for which $E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W})$. Before we proceed, we first show that JSCC exponent can at most be equal to double the tandem coding exponent for SEM source-channel systems.

Theorem 1 For an SEM source $\mathbf{Q}$ and an SEM channel $\mathbf{W}$, the JSCC exponent is upper bounded by

$$E_J(\mathbf{Q},\mathbf{W}) \le 2E_T(\mathbf{Q},\mathbf{W}).$$

Proof: It can be shown as in [1] (by introducing Markov types [2]) that for an SEM source $\mathbf{Q}$ and an SEM channel $\mathbf{W}$, the JSCC error exponent satisfies

$$E_J(\mathbf{Q},\mathbf{W}) \le \min_R \left[R_t e\left(\frac{R}{R_t},\mathbf{Q}\right) + E(R,\mathbf{W})\right], \qquad (2)$$

where $e(R,\mathbf{Q})$ is the source error exponent for $\mathbf{Q}$ and $E(R,\mathbf{W})$ is the channel error exponent for $\mathbf{W}$. We know that $R_t e\left(\frac{R}{R_t},\mathbf{Q}\right)$ is an increasing function when $R \le R_t \log_2 \lambda_0(\mathbf{Q})$ and is infinity when $R > R_t \log_2 \lambda_0(\mathbf{Q})$, and $E(R,\mathbf{W})$ is a decreasing function of R. We thereby denote

$$R_o \triangleq \begin{cases} \text{the rate satisfying} \quad R_t e(\frac{R_o}{R_t},Q) = E(R_o,W), \\ \quad \text{if } R_t e\left(\log_2 \lambda_0(\mathbf{Q}),\mathbf{Q}\right) \ge E(R_t \log_2 \lambda_0(\mathbf{Q}),\mathbf{W}), \\ R_t \log_2 \lambda_0(\mathbf{Q}), \\ \quad \text{if } R_t e\left(\log_2 \lambda_0(\mathbf{Q}),\mathbf{Q}\right) < E(R_t \log_2 \lambda_0(\mathbf{Q}),\mathbf{W}), \end{cases}$$

and according to (1) we can always write that

$$E_T(\mathbf{Q},\mathbf{W}) = E(R_o,\mathbf{W}).$$

Suppose the minimum of (2) is attained at R_m,; then we have

$$\begin{aligned} E_J(\mathbf{Q},\mathbf{W}) &\overset{(a)}{\le} R_t e\left(\frac{R_m}{R_t},\mathbf{Q}\right) + E(R_m,\mathbf{W}) \\ &\overset{(b)}{\le} R_t e\left(\frac{R_o}{R_t},\mathbf{Q}\right) + E(R_o,\mathbf{W}) \\ &\overset{(c)}{\le} 2E(R_o,\mathbf{W}) \\ &= 2E_T(\mathbf{Q},\mathbf{W}), \end{aligned}$$

where equality in (a) holds if $R_m \ge R_{cr}$, where

$$R_{cr} \triangleq \log_2 B - H(\mathbf{P}_{W_{\frac{1}{2}}})$$

is the critical rate of the SEM channel. Equality in (b) holds iff $R_m = R_o$, and equality in (c) holds iff $R_t e\left(\log_2 \lambda_0(\mathbf{Q}),\mathbf{Q}\right) \ge E(R_t \log_2 \lambda_0(\mathbf{Q}),\mathbf{W})$. □

5. SUFFICIENT CONDITIONS FOR WHICH JSCC OUTPERFORMS TANDEM CODING

For an SEM source $\mathbf{Q}$, if its entropy rate is equal to $\log_2 \lambda_0(\mathbf{Q})$, then its error exponent is reduced to

$$e(R,\mathbf{Q}) = \begin{cases} 0, & \text{if } R \le \log_2 \lambda_0(\mathbf{Q}) \\ +\infty, & \text{if } R > \log_2 \lambda_0(\mathbf{Q}). \end{cases}$$

In this case, the source is incompressible and only channel coding is performed in both JSCC and tandem coding; as a result, $E_J(\mathbf{Q},\mathbf{W}) = E_T(\mathbf{Q},\mathbf{W}) = E(R_t \log_2 \lambda_0(\mathbf{Q}),\mathbf{W})$. Thus, we assume in the rest of the section that $H(\mathbf{Q}) < \log_2 \lambda_0(\mathbf{Q})$ and $R_t H(\mathbf{Q}) < C(\mathbf{W}) = \log_2 B - H(\mathbf{P_W})$. We also assume that the source and channel are both SEM.

Theorem 2 If $R_t H\left(\widetilde{\mathbf{Q}}_{\frac{1}{2}}\right) + H\left(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{2}}}\right) \ge \log_2 B$, then

$$E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W}).$$

Theorem 2 states that if $E_J(\mathbf{Q},\mathbf{W})$ is determined exactly, no matter whether $E_T(\mathbf{Q},\mathbf{W})$ is known or not, then the JSCC exponent is larger than the tandem coding exponent.

Conversely, if $E_T(\mathbf{Q},\mathbf{W})$ is determined exactly, irrespective of whether $E_J(\mathbf{Q},\mathbf{W})$ is determined or not, the strict inequality also holds, as shown by the following result.

Theorem 3 (a). If $R_t H(\mathbf{Q}) + H(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{2}}}) \ge \log B$, then

$$E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W}).$$

(b). Otherwise, if $R_t H(\mathbf{Q}) + H(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{2}}}) < \log B$ and

$$R_t \log \lambda_0(\mathbf{Q}) + H(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{2}}}) > \log B,$$

there must exist a unique ρ_1 satisfying

$$R_t H(\widetilde{\mathbf{Q}}_{\frac{1}{1+\rho_1}}) = \log B - H(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{2}}}).$$

For such ρ_1, if

$$\begin{aligned} &(1+\rho_1)R_t[H(\widetilde{\mathbf{Q}}_{\frac{1}{1+\rho_1}}) - \log \lambda_{\frac{1}{1+\rho_1}}(\mathbf{Q})] \\ &\qquad \le \log B - 2\log \lambda_{\frac{1}{2}}(\mathbf{P_W}), \end{aligned}$$

then $E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W})$.

For the case when $E_J(\mathbf{Q},\mathbf{W})$ and $E_T(\mathbf{Q},\mathbf{W})$ are both unknown, if the lower bound for $E_J(\mathbf{Q},\mathbf{W})$ is strictly bigger than the upper bound for $E_T(\mathbf{Q},\mathbf{W})$, then $E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W})$ automatically holds. This yields the following condition.

Theorem 4 Assume that (ρ_1,ρ_2) is a pair of finite numbers satisfying

$$\begin{cases} \rho_1 R_t H\left(\widetilde{\mathbf{Q}}_{\frac{1}{1+\rho_1}}\right) - R_t(1+\rho_1)\log_2 \lambda_{\frac{1}{1+\rho_1}}(\mathbf{Q}) \\ = \rho_2 H\left(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{1+\rho_2}}}\right) - (1+\rho_2)\log_2 \lambda_{\frac{1}{1+\rho_2}}(\mathbf{P_W}), \\ R_t H\left(\widetilde{\mathbf{Q}}_{\frac{1}{1+\rho_1}}\right) + H\left(\widetilde{\mathbf{P}}_{\mathbf{W}_{\frac{1}{1+\rho_2}}}\right) = \log_2 B. \end{cases}$$

If

$$\begin{aligned} &\log_2 B - 2\log_2[\lambda_{\frac{1}{2}}^{R_t}(\mathbf{Q})\lambda_{\frac{1}{2}}(\mathbf{P_W})] \\ > \;& \rho_1 R_t H\left(\widetilde{\mathbf{Q}}_{\frac{1}{1+\rho_1}}\right) - (1+\rho_1)R_t \log_2 \lambda_{\frac{1}{1+\rho_1}}(\mathbf{Q}), \end{aligned}$$

then $E_J(\mathbf{Q},\mathbf{W}) > E_T(\mathbf{Q},\mathbf{W})$.

We next illustrate Theorems 2-4 for the following simple example. Suppose the transmission rate $R_t = 1$. Consider a binary SEM source $\mathbf{Q}$ and a ternary SEM channel $\mathbf{W}$, both with symmetric transition matrices given by

$$Q = \begin{bmatrix} q & 1-q \\ 1-q & q \end{bmatrix}$$

and

$$P_W = \begin{bmatrix} p & (1-p)/2 & (1-p)/2 \\ (1-p)/2 & p & (1-p)/2 \\ (1-p)/2 & (1-p)/2 & p \end{bmatrix},$$

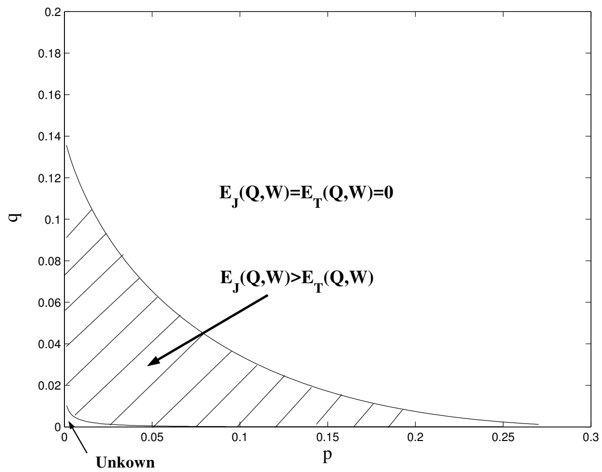

Fig. 1. When is $E_J(\mathbf{Q}, \mathbf{W}) > E_T(\mathbf{Q}, \mathbf{W})$ satisfied?

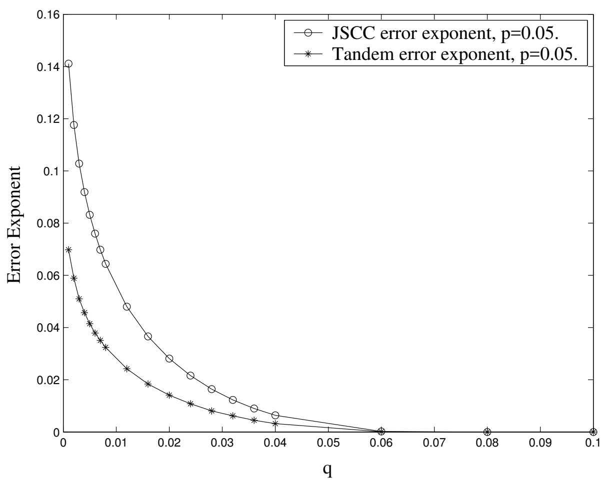

Fig. 2. Comparison of $E_J(\mathbf{Q}, \mathbf{W})$ and $E_T(\mathbf{Q}, \mathbf{W})$ for $p = 0.05$.

such that $0 < p, q < 0.5$.[2] If (q, p) satisfies any one of the conditions of Theorems 2-4, then $E_J(\mathbf{Q}, \mathbf{W}) > E_T(\mathbf{Q}, \mathbf{W})$. The range for which the inequality holds is summarized in Fig. 1.

For the channel $p = 0.05$ and $p = 0.025$, we plot the JSCC and tandem coding exponents against q whenever they are exactly determined, see Figs. 2 and 3. We see that for these source-channel pairs, $E_J(\mathbf{Q}, \mathbf{W})$ substantially outperforms $E_T(\mathbf{Q}, \mathbf{W})$ (indeed $E_J(\mathbf{Q}, \mathbf{W}) \approx 2E_T(\mathbf{Q}, \mathbf{W})$) for a large class of (q, p) pairs. For other SEM source-channel pairs (not necessarily with binary source alphabets or ternary channel alphabets) and transmission rates not equal to one, we have similar results; this indicates that the JSCC exponent is strictly better than the tandem coding exponent for a wide family of SEM systems.

6. CONCLUSION

We study the advantages of JSCC over the traditional tandem coding by providing a systematic comparison of the JSCC exponent $E_J(\mathbf{Q}, \mathbf{W})$ and the tandem coding exponent $E_T(\mathbf{Q}, \mathbf{W})$ for communication systems with Markovian memory.

We first show that $E_J(\mathbf{Q}, \mathbf{W}) \leq 2E_T(\mathbf{Q}, \mathbf{W})$ and give the conditions for equality, and we then provide sufficient conditions for which $E_J(\mathbf{Q}, \mathbf{W}) > E_T(\mathbf{Q}, \mathbf{W})$. Numerical results indicate that the inequality holds for many SEM source-channel pairs, and that $E_J(\mathbf{Q}, \mathbf{W}) \approx 2E_T(\mathbf{Q}, \mathbf{W})$ in many cases, which means that for the same error probability P_e, JSCC would require around half the delay of tandem coding, that is,

$$P_e \approx 2^{-nE_T(\mathbf{Q},\mathbf{W})} = 2^{-\frac{n}{2}E_J(\mathbf{Q},\mathbf{W})}$$

for n sufficiently large.

[2] P_W is not the channel probability transition matrix; it is the transition matrix of the SEM channel noise.

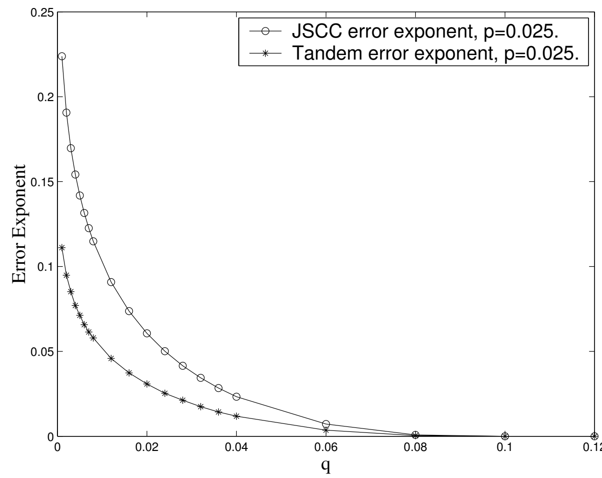

Fig. 3. Comparison of $E_J(\mathbf{Q}, \mathbf{W})$ and $E_T(\mathbf{Q}, \mathbf{W})$ for $p = 0.025$.

7. REFERENCES

[1] I. Csiszár, "Joint source-channel error exponent," *Probl. Contr. Inform. Theory*, vol. 9, pp. 315–328, 1980.

[2] L. Finesso, C. C. Liu, P. Narayan, "The optimal error exponent for Markov order estimation," *IEEE Trans. Inform. Theory*, vol. 42, pp. 1488–1497, Oct. 1996.

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

[4] E. Seneta, *Non-negative Matrices and Markov Chains*, New York: Springer-Verlag, 1981.

[5] Karol Vašek, "On the error exponent for ergodic Markov source," *Kybernetika*, vol. 16, pp. 318-329, Nov. 1980.

[6] Y. Zhong, F. Alajaji, and L. L. Campbell, "When is joint source-channel coding worthwhile: an information theoretic perspective," *Proc. 22nd Bienn. Symp. Commun.*, Canada, pp. 121-123, June 2004.

[7] Y. Zhong, F. Alajaji, and L. L. Campbell, "On the joint source-channel coding error exponent for discrete memoryless systems: computations and applications," submitted to *IEEE Trans. Inform. Theory*, Oct. 2004.

[8] Y. Zhong, F. Alajaji, and L. L. Campbell, "On the joint source-channel coding error exponent for systems with Memory," submitted to *2005 IEEE Int'l. Symp. Inform. Theory (ISIT'05)*, Jan. 2005.