Strong Coordination over Noisy Channels: Is Separation Sufficient?

Sarah A. Obead, Badri N. Vellambi, Jörg Kliewer
Helen and John C. Hartmann Department of Electrical and Computer Engineering
New Jersey Institute of Technology
Newark, New Jersey 07102
Email: sao23@njit.edu, badri.n.vellambi@ieee.org, jkliewer@njit.edu

***Abstract*—We study the problem of strong coordination of actions of two agents X and Y that communicate over a noisy communication channel such that the actions follow a given joint probability distribution. We propose two novel schemes for this noisy strong coordination problem, and derive inner bounds for the underlying strong coordination capacity region. The first scheme is a joint coordination-channel coding scheme that utilizes the randomness provided by the communication channel to reduce the local randomness required in generating the action sequence at agent Y. The second scheme exploits separate coordination and channel coding where local randomness is extracted from the channel after decoding. Finally, we present an example in which the joint scheme is able to outperform the separate scheme in terms of coordination rate.**

I. Introduction

The problem of communication-based coordination of multi-agent systems arises in numerous applications including mobile robotic networks, smart traffic control, and distributed computing such as distributed games and grid computing [1]. Several theoretical and applied studies on multi-agent coordination have targeted questions on how agents exchange information and how their actions can be correlated to achieve a desired overall behavior. Two types of coordination have been addressed in the literature – *empirical* coordination where the histogram of induced actions is required to be close to a prescribed target distribution, and *strong* coordination, where the induced sequence of joint actions of all the agents is required to be statistically close (i.e., nearly indistinguishable) from a chosen target probability mass function (pmf).

Recently, the capacity regions of several empirical and strong coordination network problems have been established [1]–[6]. Bounds for the capacity region for the point-to-point case were obtained in [7] under the assumption that the nodes communicate in a bidirectional fashion in order to achieve coordination. A similar framework was adopted and improved in [8]. In [4], [6], [9], the authors addressed inner and outer bounds for the capacity region of a three-terminal network in the presence of a relay. The work of [4] was later extended in [5], [10] to derive a precise characterization of the strong coordination region for multi-hop networks. Starkly, the majority of the recent works on coordination have considered noise-free communication channels with the exception of two works: joint empirical coordination of the channel inputs/outputs of a noisy communication channel with source and reproduction

This work is supported by NSF grants CCF-1440014 and CCF-1439465.

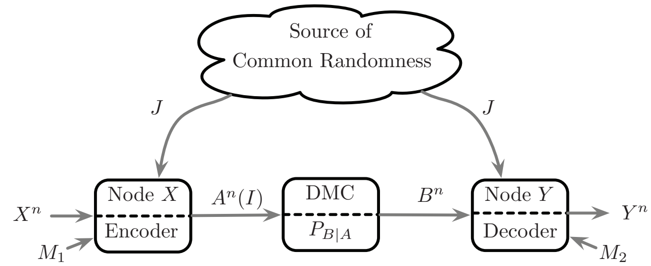

Fig. 1. Point-to-point strong coordination over a DMC.

sequences is considered in [11], and in [12], the notion of strong coordination is used to simulate a discrete memoryless channel via another channel.

In this work, we consider the point-to-point coordination setup illustrated in Fig. 1, where in contrast to [11] only source and reproduction sequences at two different nodes (X and Y) are coordinated by means of a suitable communication scheme over a discrete memoryless channel (DMC). Specifically, we propose two different novel achievable coding schemes for this noisy coordination scenario, and derive inner bounds to the underlying strong capacity region. The first scheme is a joint coordination channel coding scheme that utilizes randomness provided by the DMC to reduce the local randomness required in generating the action sequence at Node Y (see Fig. 1). The second scheme exploits separate coordination and channel coding where local randomness is extracted from the channel after decoding. Even though the proposed joint scheme is related to the scheme in [12], the presented scheme exhibits a significantly different codebook construction adapted to our coordination framework. Our scheme requires the quantification of the amount of common randomness shared by the two nodes as well as the local randomness at each of the two nodes. This is a feature that is absent from the analysis in [12]. Lastly, when the noisy channel and the correlation between X to Y are both given by binary symmetric channels (BSCs), we study the effect of the capacity of the noisy channel on the sum rate of common and local randomness. We conclude this work by showing that the joint scheme outperforms the separate scheme in terms of the coordination rate in the high-capacity regime.

II. Problem Definition

The point-to-point coordination setup we consider in this work is depicted in Fig. 1. Node X receives a sequence of actions $X^n = [X_1, X_2, \ldots, X_n] \in \mathcal{X}^n$ specified by nature where

978-1-5090-4096-4/17/$31.00 ©2017 IEEE

X^n is i.i.d. according to a pmf p_X. Both nodes have access to shared randomness J at rate R_o bits/action from a common source, and each node possesses local randomness M_k at rate ρ_k, $k = 1, 2$. Thus, in designing a *block* scheme to coordinate n actions of the nodes, we assume $J \in \{1, \ldots, 2^{nR_o}\}$, and $M_k \in \{1, \ldots, 2^{n\rho_k}\}$, $k = 1, 2$, and we wish to communicate a codeword $A^n(I)$ over the rate-limited DMC $P_{B|A}(b|a)$ to Node Y, where I denotes the (appropriately selected) coordination message. The *codeword* $A^n(I)$ is constructed based on the input action sequence X^n, the local randomness M_1 at Node X, and the common randomness J. Node Y generates a sequence of actions $Y^n \in \mathcal{Y}^n$ based on the received codeword B^n, common randomness J, and local randomness M_2. We assume that the common randomness is independent of the action specified at Node X. A tuple (R_o, ρ_1, ρ_2) is deemed *achievable* if for each $\epsilon > 0$, there exist $n \in \mathbb{N}$ and a (strong coordination) coding scheme such that the joint pmf of actions $\hat{P}_{X^n,Y^n}$ induced by this scheme and the n-fold product[1] of the desired joint pmf $P_{XY}^{\otimes n}$ are *close* in total variation, i.e.,

$$\|\hat{P}_{X^nY^n} - P_{XY}^{\otimes n}\|_{TV} < \epsilon. \tag{1}$$

We now present the two achievable coordination schemes.

III. Joint Coordination Channel Coding

This scheme follows an approach similar to those in [1], [4], [5], [10] where coordination codes are designed based on allied channel resolvability problems [13]. The structure of the allied problem pertinent to the coordination problem at hand is given in Fig. 2. The aim of the allied problem is to generate n symbols for two correlated sources X^n and Y^n whose joint statistics is close to $P_{XY}^{\otimes n}$ as defined by (1). To do so, we employ three independent and uniformly distributed messages I, K, and J and two codebooks $\mathcal{A}$ and $\mathcal{C}$ as shown in Fig. 2. To define the two codebooks, consider auxiliary random variables $A \in \mathcal{A}$ and $C \in \mathcal{C}$ jointly correlated with (X, Y) as $P_{XYABC} = P_{AC}P_{X|AC}P_{B|A}P_{Y|BC}$.

From this factorization it can be seen that the scheme consists of two *reverse test* channels $P_{X|AC}$ and $P_{Y|AC}$ used to generate the sources from the codebooks. In particular, $P_{Y|AC} = P_{B|A}P_{Y|BC}$, i.e., the randomness of the DMC contributes to the randomized generation of Y^n. Generating X^n and Y^n from I, K, J represents a complex channel resolvability problem with the following ingredients:

- Nested codebooks: Codebook $\mathcal{C}$ of size $2^{n(R_o+R_c)}$ is generated i.i.d. according to pmf P_C, i.e., $C_{ij}^n \sim P_C^{\otimes n}$ for all $(i, j) \in \mathcal{I} \times \mathcal{J}$. Codebook $\mathcal{A}$ is generated by randomly selecting $A_{ijk}^n \sim P_{A|C}^{\otimes n}(\cdot|C_{ij}^n)$ for all $(i, j, k) \in \mathcal{I} \times \mathcal{J} \times \mathcal{K}$.
- Encoding functions:
  $C^n : \{1, 2, ..., 2^{nR_c}\} \times \{1, 2, ..., 2^{nR_o}\} \to \mathcal{C}^n$,
  $A^n : \{1, ..., 2^{nR_c}\} \times \{1, ..., 2^{nR_o}\} \times \{1, ..., 2^{nR_a}\} \to \mathcal{A}^n$.
- Indices: I, J, K are independent and uniformly distributed over $\{1, ..., 2^{nR_c}\}$, $\{1, ..., 2^{nR_o}\}$, and $\{1, ..., 2^{nR_a}\}$, respectively. These indices select the pair of codewords C_{IJ}^n and A_{IJK}^n from codebooks $\mathcal{C}$ and $\mathcal{A}$.

[1]The joint pmf of n i.i.d. copies of $(X, Y) \sim p_{XY}$.

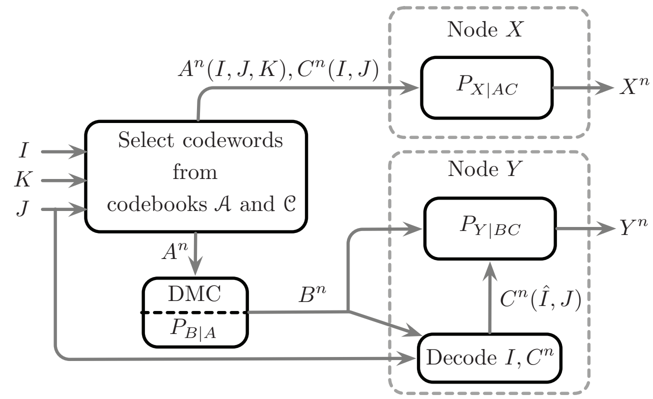

Fig. 2. A joint scheme for the allied problem.

- The selected codewords C_{IJ}^n and A_{IJK}^n are then passed through DMC $P_{X|AC}$ at Node X, while at Node Y, codeword A_{IJK}^n is sent through DMC $P_{B|A}$ whose output B^n is used to decode codeword $C_{\hat{I}J}^n$ and both are then passed through DMC $P_{Y|BC}$ to obtain Y^n.

Since the codewords are randomly chosen, the induced joint pmf of the generated actions and codeword indices in the allied problem is itself a random variable. Given a realization of the codebooks $\{a_{ijk}^n, c_{ij}^n\}$, the code-induced joint pmf of the actions and codeword indices in the allied problem is

$$\mathring{P}_{X^nY^nIJK}(x^n, y^n, i, j, k) \triangleq \frac{P_{X|AC}^{\otimes n}(x^n|a_{ijk}^n, c_{ij}^n)}{2^{n(R_c+R_o+R_a)}} \times \Big(\sum_{b^n,\hat{i}} P_{B|A}^{\otimes n}(b^n|a_{ijk}^n)\mathsf{P}_{\hat{I}|B^nJ}(\hat{i}|b^n, j)P_{Y|BC}^{\otimes n}(y^n|b^n, c_{\hat{i}j}^n)\Big), \tag{2}$$

where $\mathsf{P}_{\hat{I}|B^nJ}$ denotes the pmf induced by the operation of decoding the index I using the common randomness and the channel output at Node Y. Note that the indices for the C-codeword that generate X and Y sequences in (2) can be different since the decoding of the index I at Node Y may fail. We are done if we accomplish the following tasks: (1) identify conditions on R_o, R_c, R_a under which the code-induced pmf $\mathring{P}_{X^nY^n}$ is *close* to the design pmf $P_{XY}^{\otimes n}$ in the total variation sense; and (2) devise a strong coordination scheme by inverting the operation at Node X. This will be done in following sections by subdividing the analysis of the allied problem.

A. Resolvability constraints

Assuming that the decoding of I and the codeword C_{IJ}^n occurs perfectly at Node Y, we see that the code-induced joint pmf induced by the scheme for the allied problem for a given realization of the codebook $\{a_{ijk}^n, c_{ij}^n\}$ is

$$\check{P}_{X^nY^nIJK}(x^n, y^n, i, j, k) = \frac{P_{X|AC}^{\otimes n}(x^n|a_{ijk}^n, c_{ij}^n)}{2^{n(R_c+R_o+R_a)}} \times \Big(\sum_{b^n} P_{B|A}^{\otimes n}(b^n|a_{ijk}^n)P_{Y|BC}^{\otimes n}(y^n|b^n, c_{ij}^n)\Big). \tag{3}$$

The following result quantifies when the induced distribution in (3) is close to the n-fold product of the desired pmf P_{XY}.

Lemma 1 (Resolvability constraints). *The total variation between the code-induced pmf $\check{P}_{X^nY^n}$ in (3) and the desired pmf $P_{XY}^{\otimes n}$ vanishes, i.e., $\mathbb{E}\big[\|\check{P}_{X^nY^n} - P_{XY}^{\otimes n}\|_{TV}\big] \to 0$ as $n \to \infty$, if $R_a + R_o + R_c \geq I(XY;AC)$ and $R_o + R_c \geq I(XY;C)$.*

The proof is similar to the resolvability proof for strong coordination over multihop networks [5] and omitted due to space reasons. A full version of the proof can be found in [14].

B. Decodability constraint

Since the operation at Node Y in Fig. 2 involves the decoding of I and thus the codeword $C^n(I,J)$ using B^n and J, the induced distribution of the scheme for the allied problem will not match that of (3) unless and until we ensure that the decoding succeeds with high probability as $n \to \infty$. The following lemma quantifies the necessary rate for this decoding to succeed asymptotically almost always.

Lemma 2 (Decodability constraint). *Let $\hat{I}, C^n_{\hat{I}J}$ be the output of a typicality-based decoder that uses common randomness J to decode the index I and the codeword C^n_{IJ} from B^n. Then, $\mathbb{P}[\hat{I} \neq I] \to 0$ as $n \to \infty$, if the rate for the index I satisfies $R_c < I(B;C)$.*

The proof follows from the channel coding theorem, and is omitted due to lack of space. For a detailed proof, see [14]. Note that the above lemma also guarantees that if $R_c < I(B;C)$, then $\lim_{n\to\infty} \mathbb{E}\big[\|\check{P}_{X^nY^nIJK} - \mathring{P}_{X^nY^nIJK}\|_{TV}\big] = 0$.

C. Independence constraint

We complete modifying the allied structure to mimic the original problem with a final step. By assumption, we have a natural independence between the action sequence X^n and the common randomness J. As a result, the joint distribution over X^n and J in the original problem is a product of the marginal distributions $P_X^{\otimes n}$ and P_J. To mimic this behavior in the scheme for the allied problem, in Lemma 3 we artificially enforce independence by ensuring that the mutual information between the two vanishes.

Lemma 3 (Independence constraint). *Consider the scheme for the allied problem given in Fig. 2. Both $I(J;X^n) \to 0$ and $\mathbb{E}\big[\|\check{P}_{X^nJ} - P_X^{\otimes n} P_J\|_{TV}\big] \to 0$ as $n \to \infty$ if the code rates satisfy $R_a + R_c \geq I(X;AC)$ and $R_c \geq I(X;C)$.*

The proof of Lemma 3 builds on the results of Section III-B and the proof of Lemma 1, and is hence omitted.

In the original problem of Fig. 1, the input action sequence X^n and the index J from the common randomness source are available and the A- and C-codewords are to be selected. Now, to devise a scheme for the strong coordination problem, we proceed as follows. We let Node X choose indices I and K (and, consequently, the A- and C-codewords) from the realized X^n and J using the conditional distribution $\mathring{P}_{IK|X^nJ}$. The joint pmf of the actions and the indices is then given by

$$\hat{P}_{X^nY^nIJK} \triangleq P_X^{\otimes n} P_J \mathring{P}_{IK|X^nJ} \mathring{P}_{Y^n|IJK}. \tag{4}$$

Finally, we can argue that

$$\lim_{n\to\infty} \mathbb{E}[\|\hat{P}_{X^nY^n} - P_{XY}^{\otimes n}\|_{TV}] = 0, \tag{5}$$

since the total variation between the marginal pmf $\hat{P}_{X^nY^n}$ and the design pmf $P_{XY}^{\otimes n}$ can be bounded as

$$\begin{aligned}
\|\hat{P}_{X^nY^n} - P_{XY}^{\otimes n}\| &\leq \|\hat{P}_{X^nY^n} - \mathring{P}_{X^nY^n}\| + \|\mathring{P}_{X^nY^n} - \check{P}_{X^nY^n}\| \\
&\quad + \|\check{P}_{X^nY^n} - P_{XY}^{\otimes n}\| \\
&\leq \|\hat{P}_{X^nY^nIJK} - \check{P}_{X^nJ}\mathring{P}_{IKY^n|X^n,J}\|_{TV} \\
&\quad + \|\check{P}_{X^nY^nIJK} - \mathring{P}_{X^nY^nIJK}\|_{TV} + \|\check{P}_{X^nY^n} - P_{XY}^{\otimes n}\|_{TV} \\
&= \|P_X^{\otimes n} P_J - \check{P}_{X^nJ}\| + \|\check{P}_{X^nY^nIJK} - \mathring{P}_{X^nY^nIJK}\|_{TV} \\
&\quad + \|\check{P}_{X^nY^n} - P_{XY}^{\otimes n}\|
\end{aligned} \tag{6}$$

Since all the terms in the RHS of the above equation can be made vanishingly small provided the resolvability, decodability, and independence conditions are met, we are guaranteed that by meeting the five conditions of Lemmas 1-3 the scheme defined by (4) achieves strong coordination between Nodes X and Y by communicating over the DMC $P_{B|A}$. Note that since the operation at Nodes X and Y amount to an index selection according to $\mathring{P}_{IK|X^nJ}$, and a generation of Y^n using the DMC $P_{Y|BC}$, both operations are randomized. The last step is to derandomize the operations at Nodes X and Y by viewing the corresponding local randomness as the source of randomness in these operations. This is detailed next.

D. Local randomness rates

At Node X, local randomness is employed to randomize the selection of indices (I,K) by synthesizing the channel $\mathring{P}_{IK|X^nJ}$ whereas Node Y utilizes its local randomness to generate the action sequence Y^n by simulating the channel $P_{Y|BC}$. Using the arguments in [5], we can argue that for any given realization of J, the minimum rate of local randomness required for the probabilistic selection of indices (I,K) can be derived by quantifying the number of A and C codewords (equivalently the pair of indices I,K) jointly typical with X^n. Quantifying the list size as in [5] yields $\rho_1 \geq R_a + R_c - I(X;AC)$. At Node Y, the necessary local randomness for the generation of the action sequence is bounded by the channel simulation rate of DMC $P_{Y|BC}$ [15]. Thus, $\rho_2 \geq H(Y|BC)$.

Moreover, one can always view a part of the common randomness as local randomness, which then allows us to incorporate the rate-transfer arguments given in [5, Lemma 2]. Combining the rate-transfer argument with the constraints in Lemmas 1-3, we obtain following inner bound to the strong coordination capacity region.

Theorem 1. *A tuple (R_o, ρ_1, ρ_2) is achievable for the strong noisy coordination problem setup in Fig. 1 if for some $R_a, R_c, \delta_1, \delta_2 \geq 0$,*

$$R_a + R_o + R_c \geq I(XY;AC) + \delta_1 + \delta_2, \tag{7a}$$
$$R_o + R_c \geq I(XY;C) + \delta_1 + \delta_2, \tag{7b}$$
$$R_a + R_c \geq I(X;AC), \tag{7c}$$
$$R_c \geq I(X;C), \tag{7d}$$
$$R_c < I(B;C), \tag{7e}$$
$$\rho_1 \geq R_a + R_c - I(X;AC) - \delta_1, \tag{7f}$$
$$\rho_2 \geq H(Y|BC) - \delta_2. \tag{7g}$$

IV. Separate Coordination-Channel Coding Scheme with Randomness Extraction

As a basis for comparison, we will now introduce a separation-based scheme that involves randomness extraction. We first use a $(2^{nR_c}, 2^{nR_o}, n)$ noiseless coordination code with the codebook $\mathcal{U}$ to generate a message I of rate R_c. Such a code exists if and only if the rates R_o, R_c satisfy [1]

$$R_c + R_o \geq I(XY;U),\ R_c \geq I(X;U).$$

This coordination message I is then conveyed over the noisy channel using a rate-R_a channel codebook $\mathcal{A}$ over m channel uses. Hence, $R_c = \lambda R_a$, where $\lambda = m/n$. The decoding error probability can be made vanishingly small if $R_a < I(A;B)$. Then, from the decoder output $\hat{I}$ and the common randomness J we reconstruct the coordination sequence U^n and pass it though the test channel $P_{Y|U}$ to generate the action sequence at Node Y. Note that this separation scheme is constructed as a special case of the joint coordination-channel scheme of Fig. 2 by choosing $C = U$ and $P_{AC} = P_A P_U$.

In the following, we restrict ourselves to additive-noise DMCs, i.e., $B^m = A^m + Z^m$, where Z is the noise random variable drawn from some finite field $\mathcal{Z}$, and "+" is the native addition operation in the field. To extract randomness, we exploit the additive nature of the channel to recover the realization of the channel noise from the decoded codeword. Thus, at the channel decoder output we obtain $\hat{Z}^m = B^m + A^m(\hat{I})$, where B^m is the channel output and $A^m(\hat{I})$ the corresponding decoded channel codeword. We can then utilize a randomness extractor on $\hat{Z}^m$ to supplement the local randomness available at Node Y. The following lemma provides some guarantees with respect to the randomness extraction stage.

Lemma 4. *Consider the separation based scheme over a finite-field additive DMC. If $R_a < I(A;B)$ and we let $m, n \to \infty$ with $\frac{m}{n} = \lambda$, the following hold: (a) $\mathbb{P}[Z^m \neq \hat{Z}^m] \to 0$; (b) $\frac{1}{m} H(\hat{Z}^m) \to H(Z)$; and (c) $I(\hat{Z}^m; I, \hat{I}) \to 0$.*

The proof is omitted due to space constraints. A full version of the proof can be found in [14].

Finally, similar to the joint scheme, we can quantify the local randomness at both nodes, apply the rate transfer lemma [5, Lemma 2], and set $\lambda = 1$ to facilitate comparison with the joint scheme from Section III. The following theorem then describes an inner bound to the strong coordination region using the separate-based scheme with randomness extraction.

Theorem 2. *There exists an achievable separation based coordination-channel coding scheme for the strong setup in Fig 1 such that (1) is satisfied for $\delta_1 \geq 0, \delta_2 \geq 0$ if*

$$R_c + R_o \geq I(XY;U) + \delta_1 + \delta_2, \tag{8a}$$
$$R_c \geq I(X;U), \tag{8b}$$
$$R_c < I(A;B), \tag{8c}$$
$$\rho_1 \geq R_c - I(X;U) - \delta_1, \tag{8d}$$
$$\rho_2 \geq \max\left(0, H(Y|U) - H(Z)\right) - \delta_2. \tag{8e}$$

The proof exploits the results of Lemma 4 and is similar to the proof of Theorem 1, and is therefore omitted.

V. Example

In this section, we compare the performance of the joint and separation-based schemes in Sections III and IV, respectively, using a simple example. Specifically, we let X be a Bernoulli-$\frac{1}{2}$ source, the communication channel $P_{B|A}$ be a binary symmetric channel with crossover probability p_o (BSC(p_o)), and the conditional distribution $P_{Y|X}$ be BSC(p).

A. Basic separation scheme with randomness extraction

To derive the rate constraints for the basic separation scheme, we consider $X - U - Y$ with $U \sim$ Bernoulli $- \frac{1}{2}$ (which is known to be optimal [3]), $P_{U|X} =$ BSC(p_1), and $P_{Y|U} =$ BSC(p_2), $p_2 \in [0, p]$, $p_1 = \dfrac{p - p_2}{1 - 2p_2}$. Using this to obtain the mutual information terms in Theorem 2, we get

$$I(X;U) = 1 - h_2(p_1),\ I(A;B) = 1 - h_2(p_o), \tag{9a}$$
$$I(XY;U) = 1 + h_2(p) - h_2(p_1) - h_2(p_2), \tag{9b}$$
$$\text{and } H(Y|U) = h_2(p_2), \tag{9c}$$

where $h_2(\cdot)$ is the binary entropy function. After a round of Fourier-Motzkin elimination and upon using (9a)-(9c) in Theorem 2, we obtain the following region that is achievable using the separation-based scheme with randomness extraction:

$$R_o + \rho_1 + \rho_2 \geq h_2(p) - \min\left(h_2(p_2), h_2(p_o)\right), \tag{10a}$$
$$h_2(p_1) \geq h_2(p_o) \tag{10b}$$
$$R_c \geq 1 - h_2(p_1). \tag{10c}$$

Note that (10a) presents the achievable sum rate constraint for the required total randomness in the system.

B. Joint scheme

The rate constraints for the joint scheme are constructed in two stages. First, we derive the scheme for the codebook cardinalities $|\mathcal{A}| = 2$ and $|\mathcal{C}| = 2$, an extension to larger $|\mathcal{C}|$ is straightforward but more tedious (see Figs. 3 and 4)[2]. The joint scheme correlates the codebooks while ensuring that the decodability constraint (7e) is satisfied. To get the best tradeoff, we find the joint distribution P_{AC} that maximizes $I(B;C)$. For $|\mathcal{C}| = 2$ this is simply given by $P_{A|C}(a|c) = \delta_{ac}$, where δ_{ac} denotes the Kronecker delta. Then, the distribution $P_X(x) P_{CA|X}(c,a|x) P_{B|A}(b|a) P_{Y|BC}(y|b,c)$ that produces the boundary of the strong coordination region for the joint scheme is formed by cascading two BSCs and another symmetric channel, yielding the Markov chain $X - (C,A) - (C,B) - Y$, with the channel transition matrices $P_{CA|X} = \left[\begin{smallmatrix} 1-p_1 & 0 & 0 & p_1 \\ p_1 & 0 & 0 & 1-p_1 \end{smallmatrix}\right]$, $P_{CB|CA} = \left[\begin{smallmatrix} 1-p_o & p_o & 0 & 0 \\ 0 & 0 & p_o & 1-p_o \end{smallmatrix}\right]$, $P_{Y|CB} = \left[\begin{smallmatrix} 1-\alpha & 1-\beta & \beta & \alpha \\ \alpha & \beta & 1-\beta & 1-\alpha \end{smallmatrix}\right]^T$, with $\alpha, \beta \in [0,1]$.
Then, the mutual information terms in Theorem 1 can be expressed with $p_2 \triangleq (1-p_o)\alpha + p_o\beta$ as

$$I(X;AC) = I(X;C) = 1 - h_2(p_1),$$
$$I(XY;AC) = I(XY;C)$$
$$= 1 + h_2(p) - h_2(p_1) - h_2(p_2),$$

[2] Note that these cardinalities are not optimal. They are, however, analytically feasible and provide a good intuition about the performance of the scheme.

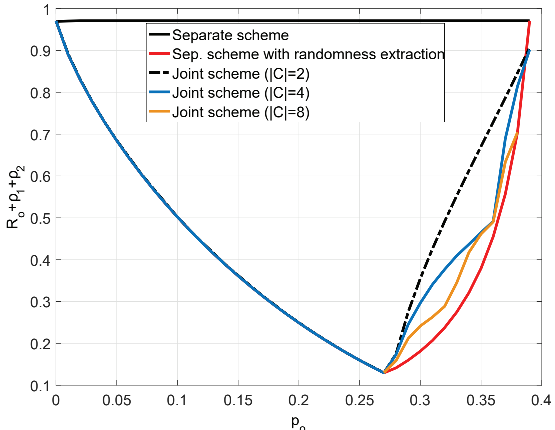

Fig. 3. Randomness sum rate vs. BSC crossover probability p_0.

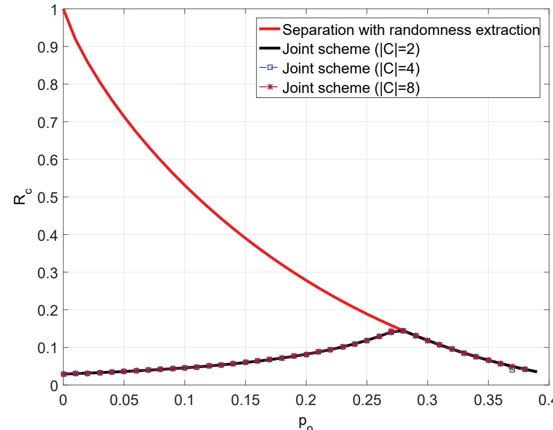

Fig. 4. Communication rate vs. BSC crossover probability p_0.

$$I(B;C) = 1 - h_2(p_o), \text{ and}$$
$$H(Y|BC) = p_o h_2(\beta) + (1-p_o) h_2(\alpha).$$

To find the minimum achievable sum rate we first perform Fourier-Motzkin elimination on the rate constraints in Theorem 1 and then minimize the information terms with respect to the parameters p_2, α, and β as follows:

$$R_o+\rho_1+\rho_2 = \min_{p_2,\alpha,\beta}\big(h_2(p)-h_2(p_2)+(1-p_o)h_2(\alpha)+p_o h_2(\beta)\big)$$
$$\text{subject to} \quad \begin{aligned} h_2(p_1) &> h_2(p_o), \\ R_c &\geq 1-h_2(p_1), \\ p &= p_1 - 2p_1p_2 + p_2. \end{aligned} \tag{11}$$

C. Numerical results

Fig. 3 compares the minimum sum rate of randomness $R_o + \rho_1 + \rho_2$ required to achieve coordination using the joint and the separate scheme with randomness extraction when the communication channel is given by BSC(p_o). The target distribution is set as $p_{Y|X} = \text{BSC}(0.4)$. The rates for the joint scheme are obtained by solving the optimization problem in (11). Similar results are obtained for the joint scheme with $|\mathcal{C}| > 2$. For the separate scheme we choose p_2 such that $h_2(p_1) = h_2(p_0)$ to maximize the amount of extracted randomness. We also include the performance of the separate scheme without randomness extraction. As can be seen from Fig. 3, both the joint scheme and the separate scheme with randomness extraction provide the same sum rate $R_o+\rho_1+\rho_2$ for $p_o \leq p'_o$ where $p'_o \triangleq \frac{1-\sqrt{1-2p}}{2}$. We also observe that for noisy channels the joint scheme approaches the performance of the separate scheme when the cardinality of C is increased. In this regime, we let $p_2 = p_0$ such that $h_2(p_2) = h_2(p_0)$ in order to maximize the amount of extracted randomness. This is done by selecting $\alpha = 0$ and $\beta = 1$ associated with $P_{Y|BC}$. However, it can be easily shown that for $p_0 > p'_0$ this does not ensure a target distribution of $P_{XY}^{\otimes n}$ anymore. Therefore, the optimization over the parameters α and β now results in a larger sum rate $R_o + \rho_1 + \rho_2$ as can be seen from Fig. 3. As p_o increases further, the required total randomness of the joint scheme approaches the one for the basic separate scheme again.

Fig. 4 provides a comparison of the communication rate for both schemes. Note that the joint scheme provides significantly smaller rates than the separation scheme with randomness extraction for $p_o \leq p'_o$, independent of the cardinality of $|\mathcal{C}|$. Thus, in this regime joint coordination-channel coding provides an advantage in terms of communication cost and outperforms the separation-based scheme for the same amount of randomness injected into the system.

References

[1] P. W. Cuff, H. H. Permuter, and T. M. Cover, "Coordination capacity," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4181–4206, 2010.

[2] E. Soljanin, "Compressing quantum mixed-state sources by sending classical information," *IEEE Trans. Inf. Theory*, vol. 48, no. 8, pp. 2263–2275, Aug. 2002.

[3] P. Cuff, "Distributed channel synthesis," *IEEE Trans. Inf. Theory*, vol. 59, no. 1, pp. 7071–7096, Nov. 2013.

[4] M. R. Bloch and J. Kliewer, "Strong coordination over a three-terminal relay network," in *Proc. IEEE Information Theory Workshop*, Hobart, Australia, Nov. 2014, pp. 646–650.

[5] B. N. Vellambi, J. Kliewer, and M. R. Bloch, "Strong coordination over multi-hop line networks," 2016. [Online]. Available: http://arxiv.org/abs/1602.09001

[6] A. Bereyhi, M. Bahrami, M. Mirmohseni, and M. R. Aref, "Empirical coordination in a triangular multi-terminal network," in *Proc. IEEE Int. Sympos. on Inform. Theory*, Istanbul, Turkey, 2013, pp. 2149–2153.

[7] A. A. Gohari and V. Anantharam, "Generating dependent random variables over networks," in *Proc. IEEE Information Theory Workshop*, Paraty, Brazil, Oct. 2011, pp. 698–702.

[8] M. H. Yassaee, A. Gohari, and M. R. Aref, "Channel simulation via interactive communications," *IEEE Trans. Inf. Theory*, vol. 61, no. 6, pp. 2964–2982, 2015.

[9] F. Haddadpour, M. H. Yassaee, A. Gohari, and M. R. Aref, "Coordination via a relay," in *Proc. IEEE Int. Sympos. on Inform. Theory*, Cambridge, MA, USA, Jul. 2012, pp. 3048–3052.

[10] M. R. Bloch and J. Kliewer, "Strong coordination over a line network," in *Proc. IEEE Int. Sympos. on Inform. Theory*, Istanbul, Turkey, Jul. 2013, pp. 2319–2323.

[11] P. Cuff and C. Schieler, "Hybrid codes needed for coordination over the point-to-point channel," in *Proc. Fourty-Nineth Annual Allerton Conf. on Commun., Control, and Comp.*, Monticello, IL, Sep. 2011, pp. 235–239.

[12] F. Haddadpour, M. H. Yassaee, S. Beig, A. Gohari, and M. R. Aref, "When is it possible to simulate a DMC channel from another?" 2013. [Online]. Available: http://arxiv.org/abs/1305.5901

[13] T. S. Han and S. Verdú, "Approximation theory of output statistics," *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 752–772, May 1993.

[14] S. A. Obead, B. N. Vellambi, and J. Kliewer, "Strong coordination over noisy channels: Is separation sufficient?" [Online]. Available: http://arxiv.org/abs/1704.08771

[15] Y. Steinberg and S. Verdú, "Channel simulation and coding with side information," *IEEE Trans. Inf. Theory*, vol. 40, no. 3, pp. 634–646, 1994.