

# Lattices from Linear Codes and Fine Quantization: General Continuous Sources and Channels

Farhad Shirani

New York University, New York, USA

S. Sandeep Pradhan

University of Michigan, Ann Arbor, USA

**Abstract**—In this paper we consider the information-theoretic characterization of performance limits of a broad class of multi-terminal communication problems with general continuous-valued sources and channels. In particular, we consider point-to-point source coding and channel coding with side information, distributed source coding with distortion constraints and function reconstruction problems (two-help-one). We develop an approach that uses fine quantization of the source and the channel variables followed by random coding with unstructured as well as structured (linear) code ensembles. This approach leads to lattice-like codes for general sources and channels.

## I. INTRODUCTION

Multi-terminal communication problems involving continuous-valued sources and channels have been studied extensively in the literature. For certain problems, such as point to point source coding and channel coding, and the multiple-access channel, performance limits have been derived using techniques based on weak typicality. For more complex problems such as distributed source coding, and broadcast channels, one needs a stronger technique because, for example, the Markov lemma [3] (a crucial step in the derivation of achievable rate regions) is not valid for weakly typical sequences. To address this, Wyner [2] proposed a technique for the problem of rate-distortion with side information using the technique of fine quantization where the source, the side information and the auxiliary variables are quantized to create a finite-alphabet problem. Then, the achievability results for the finite-alphabet problem are used to derive performance limits for the original problem using the convergence properties of mutual information. This problem has also been addressed using weak-\* typicality in [4] where the Markov lemma has been extended to continuous sources and side information. These are based on unstructured random code ensembles. Another technique that has been considered for linear quadratic Gaussian (LQG) sources and channels is to use subtractive dithered lattice codes [5], [6]. In this technique, the codes constructed have certain algebraic structures that can be exploited to obtain performance that is superior to those achievable using the unstructured code ensembles [7]. The drawback of these lattice codes is that (a) they are very specific to the LQG nature of the problem, and hence not amenable to non-Gaussian and nonlinear problems, (b) they are based on the point-to-point communication perspective, and hence not general enough to implement all of the multiuser techniques such as joint quantization as seen in multiple-description coding, and joint source-channel

mapping as seen in transmission of correlated sources over multiple-access channels.

In this paper, we develop a unified framework for achieving performance limits of general continuous-valued sources and channels in general multi-terminal communication setups. This is based on the fine quantization technique that can be used either with unstructured random code ensembles or structured code ensembles.

## II. POINT-TO-POINT COMMUNICATION WITH SIDE INFORMATION

We derive the optimal rate-distortion function for source coding and optimal capacity-cost function for channel coding by first discretizing the associated random variables, and then using transmission systems designed for discrete sources and channels. This approach is described in the following.

For any integer  $n > 0$ , consider the discrete set  $\frac{1}{2^n}\mathbb{Z}$ . Define the following quantization function  $Q_n : \mathbb{R} \rightarrow \mathbb{Z}_n$  as, for any  $s \in \mathbb{R}$ ,  $Q_n(s) = \arg \min_{a \in \frac{1}{2^n}\mathbb{Z}} (s - a)^2$ . For any two real numbers  $l > 0$  and  $u > 0$ , define the clipping function  $C_{l,u} : \mathbb{R} \rightarrow \mathbb{R}$ ,  $C_{l,u}(s) = \max\{\min\{u, s\}, -l\}$ . Associated with this quantization, and the clipping function, define the discrete set  $\mathbb{Z}_{l,n} \triangleq [-l, u] \cap \frac{1}{2^n}\mathbb{Z}$ , and the associated quantization cells  $A_{l,n}(i)$  for  $i = 0, 1, 2, \dots, (\lfloor u/2^n \rfloor - \lceil -l/2^n \rceil - 1)$ . Let  $\zeta_{l,u,n}(i)$  denote the quantization reconstruction of the  $i$ th cell. To reduce clutter, we denote  $Q_n(C_{l,u}(S))$  as  $S_{l,u,n}$  when the subscript is clear from the context. Moreover, we also denote  $S_{l,u} = C_{l,u}(S)$ . For the channel coding and source coding problems that will be studied in this section, we consider cost and distortion functions as follows. We assume that the cost function  $\kappa : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ , and the distortion function  $d : \mathbb{R}^3 \rightarrow \mathbb{R}^+$  are jointly continuous (for  $n$ -length vectors, we use additive cost  $\kappa_n$  and distortion  $d_n$  functions).

### A. Source Coding With Side Information at Decoder

Consider a memoryless source  $X$  with side information  $Y$  given by  $(P_{XY}, d)$  comprising of a probability measure  $P_{XY}$  on  $\mathbb{R}^2$ , with reconstruction alphabet  $\mathbb{R}$ , and a jointly continuous distortion function  $d$ .

**Definition II.1.** An  $(n, \Theta)$  transmission system is a pair of mappings  $e : \mathbb{R}^n \rightarrow \{1, 2, \dots, \Theta\}$ ,  $f : \{1, 2, \dots, \Theta\} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ . A rate distortion pair  $(R, D)$  is said to be achievable if there exists a sequence of  $(n, \Theta_n)$  transmission systems such that

$$\lim_{n \rightarrow \infty} \frac{\log \Theta_n}{n} \leq R, \quad \lim_{n \rightarrow \infty} \mathbb{E} d_n(X^n, Y^n, f(e(X^n), Y^n)) \leq D.$$

Let the operational rate-distortion function  $R_{WZ}(D)$  denote the infimum of rates  $R$  such that  $(R, D)$  is achievable.

We prove the following theorem for a general probability measure  $P_{XY}$  and continuous distortion function  $d$ .

**Theorem II.1.** For a given source  $(P_{XY}, d)$ , we have  $R_{WZ}(D) \leq \alpha(D)$ , where

$$\alpha(D) \triangleq \min_{\{P_{U|X}, g(\cdot, \cdot)\}} [I(U; X) - I(U; Y)],$$

and the minimization is carried out over all transition probability  $P_{U|X}$ , and continuous functions  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that  $Y \rightarrow X \rightarrow U$ , and  $\mathbb{E}[d(X, Y, g(U, Y))] \leq D$ .

**Proof:** Consider a joint probability measure  $P_{XYU}$  and  $g(\cdot, \cdot)$  that satisfies the conditions given in the theorem. Define an induced distortion function  $\tilde{d} : \mathbb{R}^3 \rightarrow \mathbb{R}^+$  as  $\tilde{d}(x, y, u) = d(x, y, g(y, u))$ . Note that  $\tilde{d}$  is continuous everywhere, and  $\mathbb{E}\tilde{d}(X, Y, u)$  is finite whenever  $u$  is finite. We quantize  $X$  with parameters  $l_1, u_1$  and  $n_1$ , quantize  $U$  with parameters  $l_2, u_2, n_2$  and quantize  $Y$  with parameters  $l_3, u_3, n_3$ . However, the quantized triple  $(X_{l_1, u_1, n_1}, Y_{l_3, u_3, n_3}, U_{l_2, u_2, n_2})$  may not satisfy Markov chain  $Y_{l_3, u_3, n_3} - X_{l_1, u_1, n_1} - U_{l_2, u_2, n_2}$ . To address this we consider the following approach.

We are given the source  $(X, Y)$  and an auxiliary variable  $U$  with joint distribution  $P_{XYU}$  such that  $Y - X - U$ . We consider a series of single-letter transformations of these random variables as follows.

**Step 1:** Let  $Z$  and  $W$  be two random variables that are independent of the source  $(X, Y)$  such that  $Z \in [-l_1, u_1]$  with probability 1, and the distribution  $P_{ZW}$  is given by

$$P_{ZW}(A \times B) = \frac{P_{XY}(A \cap [-l_1, u_1] \times B)}{P(X \in [-l_1, u_1])}$$

for all events  $A$  and  $B$  in Borel sigma algebra. Define

$$(\tilde{X}, \tilde{Y}) = \begin{cases} (X, Y) & \text{if } X \in [-l_1, u_1] \\ (Z, W) & \text{otherwise} \end{cases}.$$

Let  $\tilde{U}$  be a random variable that is correlated with  $(X, Y, Z, W)$  such that the distribution of  $(\tilde{X}, \tilde{U})$  is given by

$$P_{\tilde{X}, \tilde{U}}(A \times B) = \frac{P_{XU}(A \cap [-l_1, u_1] \times B)}{P(X \in [-l_1, u_1])}$$

for all event  $A$  and  $B$ , and  $(X, Y, \tilde{Y}) - \tilde{X} - \tilde{U}$ . It should be noted that  $(\tilde{X}, \tilde{Y}, \tilde{U})$  depends on  $l_1, u_1$ , however, this is not made explicit to keep the notation simple. Next we show that  $I(\tilde{X}; \tilde{U}) \approx I(X; U)$ ,  $I(\tilde{Y}; \tilde{U}) \approx I(Y; U)$ , and  $\mathbb{E}d(\tilde{X}, \tilde{Y}, \tilde{U}) \approx \mathbb{E}d(X, Y, U)$  for sufficiently large  $l_1, u_1$ . Fix an  $\epsilon > 0$ . Consider

$$\begin{aligned} I(\tilde{X}; \tilde{U}) &= \int_{-l_1}^{u_1} \int \frac{dP_{XU}(\tilde{x}, \tilde{u})}{P_X([-l_1, u_1])} \log \frac{dP_{XU}}{d(P_{\tilde{U}}P_X)}(\tilde{x}, \tilde{u}) \\ &\stackrel{(a)}{=} \int_{-l_1}^{u_1} \int \frac{dP_{XU}(\tilde{x}, \tilde{u})}{P_X([-l_1, u_1])} \log \frac{dP_{XU}}{d(P_U P_X)}(\tilde{x}, \tilde{u}) - D(P_{\tilde{U}} \| P_U) \\ &\leq \frac{1}{P_X([-l_1, u_1])} \int \int dP_{XU}(\tilde{x}, \tilde{u}) \log \frac{dP_{XU}}{d(P_U P_X)}(\tilde{x}, \tilde{u}) \\ &\rightarrow I(X; U), \end{aligned}$$

as  $l_1, u_1 \rightarrow \infty$ , where in (a) we note that  $P_{\tilde{U}} \ll P_U$ . Next observe that

$$\begin{aligned} P_{\tilde{Y}\tilde{U}}(A \times B) &= P_{\tilde{X}\tilde{Y}\tilde{U}}([-l_1, u_1] \times A \times B) \\ &= \frac{P_{XYU}([-l_1, u_1] \times A \times B)}{P_X([-l_1, u_1])} \rightarrow P_{YU}(A \times B) \end{aligned}$$

as  $l_1, u_1 \rightarrow \infty$ , and hence we have  $\lim_{l_1, u_1 \rightarrow \infty} I(\tilde{Y}; \tilde{U}) \geq I(Y; U)$ . Moreover,

$$\begin{aligned} \mathbb{E}d(\tilde{X}, \tilde{Y}, g(\tilde{Y}, \tilde{U})) &= \int_{-l_1}^{u_1} \int \int d(\tilde{x}, \tilde{y}, g(\tilde{y}, \tilde{u})) \frac{dP_{XYU}}{P_X([-l_1, u_1])}(\tilde{x}, \tilde{y}, \tilde{u}) \\ &\rightarrow \mathbb{E}d(X, Y, g(Y, U)) \end{aligned}$$

as  $l_1, u_1 \rightarrow \infty$ .

**Step 2:** Clip  $\tilde{Y}$  and  $\tilde{U}$  with parameters  $l_3, u_3$  and  $l_2, u_2$ , respectively. We still have  $\tilde{Y}_{l_3, u_3} - \tilde{X} - \tilde{U}_{l_2, u_2}$ . One can see that there exists sequence of lengths  $l_{2n}, l_{3n}$  and  $u_{2m}, u_{3m}$  such that

$$\begin{aligned} &\lim_{n, m \rightarrow \infty} \mathbb{E}d(\tilde{X}, \tilde{Y}_{l_{3n}, u_{3n}}, g(\tilde{Y}_{l_{3n}, u_{3n}}, \tilde{U}_{l_{2m}, u_{2m}})) \\ &= \lim_{n, m \rightarrow \infty} \mathbb{E}\tilde{d}(\tilde{X}, \tilde{Y}_{l_{3n}, u_{3n}}, \tilde{U}_{l_{2m}, u_{2m}}) = \mathbb{E}\tilde{d}(\tilde{X}, \tilde{Y}, \tilde{U}) = \mathbb{E}d(\tilde{X}, \tilde{Y}, g(\tilde{Y}, \tilde{U})). \end{aligned}$$

Moreover,

$$\lim_{l_2, u_2 \rightarrow \infty} I(\tilde{X}; \tilde{U}_{l_2, u_2}) = I(\tilde{X}; \tilde{U}), \quad \lim_{l_2, u_2, l_3, u_3 \rightarrow \infty} I(\tilde{Y}_{l_3, u_3}; \tilde{U}_{l_2, u_2}) = I(\tilde{Y}; \tilde{U})$$

**Step 3:** Next we quantize  $\tilde{X}$  into  $\tilde{X}_{n_1}$  and enforce the Markov chain. Before we proceed further, let us note that all random variables with  $\sim$  on top depend on  $l_1$  and  $u_1$ , and this dependence is not made explicit. Now using

$$I(\tilde{X}\tilde{Y}_{l_3, u_3}; \tilde{U}_{l_2, u_2} | \tilde{X}_{n_1}) = I(\tilde{X}, \tilde{Y}_{l_3, u_3}; \tilde{U}_{l_2, u_2}) - I(\tilde{X}_{n_1}; \tilde{U}_{l_2, u_2}),$$

we get

$$\lim_{n_1 \rightarrow \infty} I(\tilde{Y}_{l_3, u_3}; \tilde{X}; \tilde{U}_{l_2, u_2} | \tilde{X}_{n_1}) = I(\tilde{Y}_{l_3, u_3}; \tilde{U}_{l_2, u_2} | \tilde{X}) = 0. \quad (1)$$

Define  $\hat{U}_{l_2, u_2}$  as a random variable having the same alphabet as  $\tilde{U}_{l_2, u_2}$ , and that is jointly correlated with  $(\tilde{X}_{n_1}, \tilde{X}, \tilde{Y}_{l_3, u_3})$  according to the probability distribution that satisfies the Markov chain  $\tilde{Y}_{l_3, u_3} - \tilde{X} - \tilde{X}_{n_1} - \hat{U}_{l_2, u_2}$ , and the pair  $(\tilde{X}_{n_1}, \hat{U}_{l_2, u_2})$  has the same distribution as the pair  $(\tilde{X}_{n_1}, \tilde{U}_{l_2, u_2})$ . First, let us consider the following lemma.

**Lemma II.1.** For any triple of random variables  $S, T, W$  with joint distribution  $P_{STW}$ , consider a random variable  $\hat{W}$  that is correlated with  $(S, T)$  such that  $P_{TW} = P_{T\hat{W}}$  and  $S - T - \hat{W}$ , then

$$I(S; W | T) \geq D(P_{SW} \| P_{S\hat{W}}) \geq \frac{1}{2 \ln 2} V^2(P_{SW}, P_{S\hat{W}}).$$

Using (1), and Lemma II.1 with the triple  $(\tilde{X}, \tilde{Y}_{l_3, u_3}, \tilde{X}_{n_1})$ , and  $\tilde{U}_{l_2, u_2}$ , we have

$$\lim_{n_1 \rightarrow \infty} V(P_{\tilde{X}, \tilde{Y}_{l_3, u_3}, \tilde{U}_{l_2, u_2}}, P_{\tilde{X}, \tilde{Y}_{l_3, u_3}, \hat{U}_{l_2, u_2}}) = 0,$$

and hence using lower semicontinuity of mutual information in variational distance, we have

$$\lim_{n_1 \rightarrow \infty} I(\tilde{X}_{n_1}; \hat{U}_{l_2, u_2}) = I(\tilde{X}; \tilde{U}_{l_2, u_2}),$$

$$\lim_{n_1 \rightarrow \infty} I(\tilde{Y}_{l_3, u_3}; \hat{U}_{l_2, u_2}) \geq I(\tilde{Y}_{l_3, u_3}; \tilde{U}_{l_2, u_2}),$$

and since convergence in variational distance implies convergence in distribution, we have

$$\lim_{n_1 \rightarrow \infty} \mathbb{E}d(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3}, g(\tilde{Y}_{l_3, u_3}, \hat{U}_{l_2, u_2})) = \mathbb{E}d(\tilde{X}, \tilde{Y}_{l_3, u_3}, g(\tilde{Y}_{l_3, u_3}, \tilde{U}_{l_2, u_2})). \quad (2)$$

**Step 4:** Next we quantize  $Y_{l_3, u_3}$  and  $\hat{U}_{l_2, u_2}$  into  $Y_{l_3, u_3, n_3}$  and  $\hat{U}_{l_2, u_2, n_2}$ . We can show the convergence of mutual information and expected distortions as  $n_2$  and  $n_3$  become large. Let us summarize the approach used till now. We are given three infinite-alphabet random variables  $(X, Y, U)$  with joint distribution  $P_{XYU}$  and satisfying the Markov chain  $Y - X - U$ . We create three finite alphabet random variables  $(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3, n_3}, \tilde{U}_{l_2, u_2, n_2})$  satisfying the Markov chain  $(\tilde{Y}_{l_3, u_3, n_3} - \tilde{X}_{n_1} - \tilde{U}_{l_2, u_2, n_2})$ . The pair  $(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3, n_3})$  is obtained by a random transformation of the pair  $(X, Y)$ . In particular, we note that,  $(\tilde{X}, \tilde{Y}) = (X, Y)$  if  $X \in [-l_1, u_1]$ , and  $(\tilde{X}, \tilde{Y}) = (Z, W)$  otherwise. Further,  $\tilde{X}_{n_1}$  is the quantized version of  $\tilde{X}$ , and  $\tilde{Y}_{l_3, u_3, n_3}$  is the clipped and quantized version of  $\tilde{Y}$ . By choosing nine parameters  $l_1, l_2, l_3, u_1, u_2, u_3$  and  $n_1, n_2, n_3$ , we can make  $I(\tilde{X}_{n_1}; \hat{U}_{l_2, u_2, n_2})$ ,  $I(\tilde{Y}_{l_3, u_3, n_3}; \hat{U}_{l_2, u_2, n_2})$  and  $\mathbb{E}d(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3, n_3}, g(\tilde{Y}_{l_3, u_3, n_3}, \hat{U}_{l_2, u_2, n_2}))$  approach  $I(X; U)$ ,  $I(Y; U)$  and  $\mathbb{E}d(X, Y, g(Y, U))$  arbitrarily closely. In other words, for every  $\epsilon > 0$ , there exist infinitely many  $l_1, u_1, l_2, u_2, l_3, u_3$  such that for all sufficiently large  $n_1, n_2, n_3$ , we have

$$I(\tilde{X}_{n_1}; \hat{U}_{l_2, u_2, n_2}) - I(\tilde{Y}_{l_3, u_3, n_3}; \hat{U}_{l_2, u_2, n_2}) \leq I(X; U) - I(Y; U) + 2\epsilon,$$

$$\mathbb{E}d(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3, n_3}, g(\tilde{Y}_{l_3, u_3, n_3}, \hat{U}_{l_2, u_2, n_2})) \leq \mathbb{E}d(X, Y, g(Y, U)) + \epsilon.$$

**Step 5:** Now we can use Wyner-Ziv rate-distortion theorem [1] for finite alphabets to show the existence of a transmission system. In particular, one can show that the rate-distortion pair given by  $(I(\tilde{X}_{n_1}; \hat{U}_{l_2, u_2, n_2}) - I(\tilde{Y}_{l_3, u_3, n_3}; \hat{U}_{l_2, u_2, n_2}), \mathbb{E}d(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3, n_3}, g(\tilde{Y}_{l_3, u_3, n_3}, \hat{U}_{l_2, u_2, n_2})))$  is achievable for the finite-alphabet source  $(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3, n_3}, d)$ . In other words, for all  $\epsilon > 0$ , and for all sufficiently large  $m$ , there exists a transmission system  $\text{TS}_d$  with parameter  $(m, \Theta)$  for compressing the finite-alphabet source such that

$$\frac{1}{m} \log \Theta \leq I(\tilde{X}_{n_1}; \hat{U}_{l_2, u_2, n_2}) - I(\tilde{Y}_{l_3, u_3, n_3}; \hat{U}_{l_2, u_2, n_2}) + \epsilon, \quad (3)$$

$$\mathbb{E}d_m(\tilde{X}_{n_1}^m, \tilde{Y}_{l_3, u_3, n_3}^m, \tilde{X}^m) \leq \mathbb{E}d(\tilde{X}_{n_1}, \tilde{Y}_{l_3, u_3, n_3}, g(\tilde{Y}_{l_3, u_3, n_3}, \hat{U}_{l_2, u_2, n_2})) + \epsilon,$$

where  $\tilde{X}^m = f(e(\tilde{X}_{n_1}^m), \tilde{Y}_{l_3, u_3, n_3}^m)$ . This completes the desired proof.

### B. Channel Coding with Side Information at Transmitter

Consider a channel with state  $(P_{Y|XS}, P_S, \kappa)$  comprising of a transition probability  $P_{Y|XS} : \mathbb{R}^2 \times \mathcal{B} \rightarrow \mathbb{R}$ , a state random variable  $S$  with a probability measure  $P_S$ , and a cost function  $\kappa : \mathbb{R}^2 \rightarrow \mathbb{R}^+$ . We assume that the channel state is observable at the encoder noncausally. For any transition probability  $P_{X|S} : \mathbb{R} \times \mathcal{B} \rightarrow \mathbb{R}^+$ , we define the joint probability measure  $P_{XSY}$

on the measurable space  $(\mathbb{R}^3, \mathcal{B}^3)$  as the unique extension of the measure on product sets

$$P_{XSY}(A \times B \times C) = \int_A P_S(ds) \int_B P_{X|S}(s, dx) \int_C P_{Y|XS}(x, s, dy)$$

We assume that the state is IID, and the channel is stationary, memoryless and used without feedback. One can define the capacity-cost function [8] with side information  $C_{GP}(\tau)$  (we skip a formal definition due to lack of space).

**Theorem II.2.** For a given channel with state  $(P_{Y|XS}, P_S, \kappa)$ , we have  $C_{GP}(\tau) \geq \alpha(\tau)$ , where

$$\alpha(\tau) \triangleq \max_{\{P_{U|S}, g(\cdot, \cdot)\}} [I(U; X) - I(U; Y)],$$

and the maximization is carried out over all transition probability  $P_{U|S}$ , and continuous functions  $g : \mathbb{R}^2 \rightarrow \mathbb{R}$  such that  $U \rightarrow (X, S) \rightarrow Y$ , and  $\mathbb{E}[\kappa(g(U, S), S)] \leq \tau$ .

We skip the proof due to lack of space.

### III. DISTRIBUTED SOURCE CODING

Next we consider distributed source coding problem consisting of two correlated and memoryless continuous-valued sources  $X$  and  $Y$ , characterized by a probability measure  $P_{XY}$  which needs to be compressed distributively into bits to be sent to a joint decoder. The joint decoder wishes to reconstruct the sources with respect to two separable distortion measures  $d_x : \mathbb{R} \rightarrow \mathbb{R}^+$  and  $d_y : \mathbb{R} \rightarrow \mathbb{R}^+$ . This is a well-studied problem [3] and we skip the formal definition for conciseness. Let  $\mathcal{RD}$  denote the set of all achievable rate and distortion tuples  $(R_1, R_2, D_1, D_2)$ . We denote  $\mathcal{R}(D_1, D_2)$  as the set of all rates  $(R_1, R_2)$  such that  $(R_1, R_2, D_1, D_2)$  is achievable.

**Definition III.1.** Let  $\mathcal{P}(D_1, D_2)$  denote the collection of pairs of transition probabilities  $P_{U|X}$  and  $P_{V|Y}$ , and pairs of continuous functions  $g_i : \mathbb{R}^2 \rightarrow \mathbb{R}$  for  $i = 1, 2$ , such that  $\mathbb{E}d_x(X, \hat{X}) \leq D_1$ ,  $\mathbb{E}d_y(Y, \hat{Y}) \leq D_2$ , where the expectations are evaluated with the joint measure  $P_{XY}P_{U|X}P_{V|Y}$ , i.e.,  $U - X - Y - V$  form a Markov chain. For a  $(P_{U|X}, P_{V|Y}, g_1, g_2) \in \mathcal{P}(D_1, D_2)$ , let  $\alpha(P_{U|X}, P_{V|Y}, g_1, g_2)$  denote the set of rate pairs  $(R_1, R_2) \in [0, \infty)^2$  that satisfy

$$R_1 \geq I(X; U|V), \quad R_2 \geq I(Y; V|U), \quad R_1 + R_2 \geq I(XY; UV).$$

Let the information rate region be defined as

$$\alpha(D_1, D_2) = cl \left( \bigcup_{(P_{U|X}, P_{V|Y}, g_1, g_2) \in \mathcal{P}(D_1, D_2)} \alpha(P_{U|X}, P_{V|Y}, g_1, g_2) \right)$$

**Theorem III.1.** For a given source  $(P_{XY}, d_1, d_2)$ , we have  $\alpha(D_1, D_2) \subseteq \mathcal{R}(D_1, D_2)$ .

We skip the proof due to lack of space.

### IV. RECONSTRUCTION OF THE SUM

In this section, we develop the framework that addresses structured code ensembles. Consider a pair of jointly continuous random variables  $U, V$  with a joint PDF  $f_{UV}$ . We denote the joint probability measure as  $P_{UV}$ . We assume that (i)  $f_{UV}$  has

support over the compact set  $[0, M]^2$  for some finite  $M$ , and (ii) for all  $(u, v) \in [0, M]^2$ , the joint PDF satisfies  $l \leq f_{UV}(u, v) \leq L$ , for some  $L \geq l > 0$ . Consider quantizing  $U$  and  $V$  with a uniform quantizer with step size  $\Delta_n \triangleq M/n$  for some integer  $n$ . We denote

$$U_n \triangleq \Delta_n \left\lfloor \frac{U}{\Delta_n} \right\rfloor + \frac{\Delta_n}{2}, \text{ and } V_n \triangleq \Delta_n \left\lfloor \frac{V}{\Delta_n} \right\rfloor + \frac{\Delta_n}{2}.$$

We have the following theorem.

**Theorem IV.1.** *There exists a sequence  $\xi_n > 0$  such that*

$$I(U_n; U_n + V_n) \leq I(U; U + V) + \xi_n,$$

and  $\lim_{n \rightarrow \infty} \xi_n = 0$ .

**Proof:** Let us denote the  $i$ th cell of the quantizer as  $E_i \triangleq [(i-1)\Delta_n, i\Delta_n]$ , and its midpoint as  $e_i = i\Delta_n - \frac{\Delta_n}{2}$ , for  $i = 1, 2, \dots, n$ . Consider the following sequence of inequalities.

$$\begin{aligned} & H(U_n + V_n) \\ & \stackrel{(a)}{=} - \sum_{i,j=1}^n P_{UV}(E_i \times E_j) \log P_{UV}(G_{i+j}) \\ & = - \sum_{i+j \leq n} P_{UV}(E_i \times E_j) \log \sum_{k=1}^{i+j-1} P_{UV}(E_k \times E_{i+j-k}) \\ & \quad - \sum_{i+j > n} P_{UV}(E_i \times E_j) \log \sum_{k=i+j-n}^n P_{UV}(E_k \times E_{i+j-k}), \\ & \stackrel{(b)}{\leq} - \sum_{i+j \leq n} P_{UV}(E_i \times E_j) \log \sum_{k=1}^{i+j-1} \Delta_n^2 (f_{UV}(e_k, e_{i+j-k}) - \delta_n) \\ & \quad - \sum_{i+j > n} P_{UV}(E_i \times E_j) \log \sum_{k=i+j-n}^n \Delta_n^2 (f_{UV}(e_k, e_{i+j-k}) - \delta_n), \\ & \stackrel{(c)}{\leq} - \sum_{i+j \leq n} P_{UV}(E_i \times E_j) \log \left[ \Delta_n \left( \int_0^{(i+j-1)\Delta_n} f_{UV}(u, (i+j-1)\Delta_n - u) du - \zeta_n \right) - n\delta_n \Delta_n^2 \right] \\ & \quad - \sum_{i+j > n} P_{UV}(E_i \times E_j) \log \left[ \Delta_n \left( \int_{(i+j-1)\Delta_n}^{n\Delta_n} f_{UV}(u, (i+j-1)\Delta_n - u) du - \zeta_n \right) - n\delta_n \Delta_n^2 \right] \\ & \stackrel{(d)}{\leq} - \sum_{i,j} P_{UV}(E_i \times E_j) \log [f_{U+V}(e_i + e_j)] + (\eta_n - \log \Delta_n) \\ & \stackrel{(e)}{\leq} - \sum_{i,j} \Delta_n^2 f_{UV}(e_i, e_j) \log [f_{U+V}(e_i + e_j)] \\ & \quad + \sum_{i,j} \Delta_n^2 \delta_n |\log(f_{U+V}(e_i + e_j))| + (\eta_n - \log \Delta_n) \\ & \stackrel{(f)}{\leq} - \sum_{i,j} \Delta_n^2 f_{UV}(e_i, e_j) \log [f_{U+V}(e_i + e_j)] \\ & \quad + M^2 \delta_n \max\{\log(2ML), -\log(l)\} + (\eta_n - \log \Delta_n) \\ & \stackrel{(g)}{\leq} - \int f_{UV}(u, v) \log f_{U+V}(u + v) du dv + \alpha_n \\ & \quad + M^2 \delta_n \max\{\log(2ML), -\log(l)\} + (\eta_n - \log \Delta_n) \end{aligned}$$

where we have used the following arguments. In (a) we have defined for all  $k = 2, \dots, 2n$ ,

$$G_k = \{(u, v) : \exists m \text{ such that } (u, v) \in (E_m \times E_{k-m})\}.$$

In (b), we have a sequence  $\delta_n$  such that  $\lim_{n \rightarrow \infty} \delta_n = 0$ , existence of which follows from the uniform continuity of  $f_{UV}$  over the support set  $[0, M]^2$ . In (c), we have a sequence  $\zeta_n$  such that  $\lim_{n \rightarrow \infty} \zeta_n = 0$ , existence of which follows from the uniform Riemann integrability of  $f_{UV}$ . This is proved in the following lemma. In (d) we have a sequence  $\eta_n$  such that  $\lim_{n \rightarrow \infty} \eta_n = 0$ , existence of which follows from the uniform continuity of log function over the support set  $[l/2, 2ML]$ . In (e), we again have the sequence  $\delta_n$  from the uniform continuity of  $f_{UV}$  over the support set  $[0, M]^2$ . In (f) we have used the lower bound  $l$ , and upper bound  $2ML$  on  $f_{U+V}$ . In (g), we have a sequence  $\alpha_n$  such that  $\lim_{n \rightarrow \infty} \alpha_n = 0$ , existence of which follows from the Riemann integrability of the function  $f_{UV}(u, v) \log f_{U+V}(u + v)$ . Hence we see that  $H(U_n + V_n) + \log \Delta_n \approx h(U + V)$ . Using similar arguments, we can show that

$$H(U_n, U_n + V_n) + \log \Delta_n^2 \approx h(U, V) = h(U, U + V)$$

and  $H(U_n) + \log \Delta_n \approx h(U)$ . Combining these we get the desired result.

We have the next theorem saying something similar. For any quadruple of random variables  $(U, V, X, Y)$ , we have the following result.

**Theorem IV.2.** *There exists a sequence  $\xi_{n,m,l} > 0$  such that*

$$\lim_{n,m,l \rightarrow \infty} I(U_n + V_n, X_m; Y_l) \geq I(U + V, X; Y) - \xi_{n,m,l},$$

and  $\lim_{n,m,l \rightarrow \infty} \xi_{n,m,l} = 0$ .

**Proof:** Fix  $\epsilon > 0$ . We see that there exists  $m_0$  and  $l_0$  such that for all  $m > m_0$  and  $l > l_0$ , have  $I(U + V; X_m; Y_l) \geq I(U + V; X; Y) - \epsilon$ . Choose an  $m > m_0$  and an  $l > l_0$ . Note that  $X_m$  and  $Y_l$  are finite-valued random variables. Applying Theorem IV.1 on  $(U, V)$  for different realizations of  $(X_m, Y_l)$ , we see that there exists a sequence  $\xi_{n,m,l} > 0$  such that

$$\lim_{n \rightarrow \infty} I(U_n + V_n, X_m; Y_l) \geq I(U + V, X_m; Y_l) - \xi_{n,m,l},$$

and  $\lim_n \xi_{n,m,l} = 0$ . From this we get the desired result.

We next have the following lemma which we promised to prove in the proof of Theorem IV.1.

**Lemma IV.1.** *For all  $m = 1, 2, \dots, (2n-1)$ ,*

$$|A(m\Delta_n) - A_n(m\Delta_n)| \leq 2M\delta_n,$$

where

$$\begin{aligned} A(m\Delta_n) &= \int_{\max\{0, m\Delta_n - M\}}^{\min\{m\Delta_n, M\}} f_{UV}(u, m\Delta_n - u) du, \\ A_n(m\Delta_n) &= \sum_{k=\max\{1, m+1-n\}}^{\min\{m, n\}} \Delta_n f_{UV}(e_k, m\Delta_n - e_k). \end{aligned}$$

**Proof:** For all  $i, j \in \{1, 2, \dots, n\}$ , let

$$\bar{e}_{i,j} = \arg \max_{\tilde{u}: [\tilde{u}/\Delta_n] = i} \arg \max_{\tilde{v}: [\tilde{v}/\Delta_n] = j} f(\tilde{u}, \tilde{v}).$$

$$\underline{e}_{i,j} = \arg \min_{\tilde{u}: [\tilde{u}/\Delta_n] = i} \arg \min_{\tilde{v}: [\tilde{v}/\Delta_n] = j} f(\tilde{u}, \tilde{v}).$$

First consider the case when  $m = 1, 2, \dots, n$ . Then we have

$$A(m\Delta_n) = \sum_{k=1}^m \int_{E_k} f_{UV}(u, m\Delta_n - u) du \leq \sum_{k=1}^m \Delta_n f_{UV}(\bar{e}_{k,m-k+1}),$$

where the inequality follows because as  $u$  ranges over the set  $E_k$ , we have  $(m\Delta_n - u)$  range over the set  $E_{m-k+1}$ . Similarly, we have

$$A(m\Delta_n) \geq \sum_{k=1}^m \Delta_n f_{UV}(\underline{e}_{k,m-k+1}).$$

Moreover since  $(m\Delta_n - e_k) = e_{m-k+1}$ , we see that

$$\sum_{k=1}^m \Delta_n f_{UV}(\underline{e}_{k,m-k+1}) \leq A_n(m\Delta_n) \leq \sum_{k=1}^m \Delta_n f_{UV}(\bar{e}_{k,m-k+1}).$$

Now making the observation that

$$\left| \sum_{k=1}^m \Delta_n f_{UV}(\bar{e}_{k,m-k+1}) - \sum_{k=1}^m \Delta_n f_{UV}(\underline{e}_{k,m-k+1}) \right| \leq 2M\delta_n, \quad (4)$$

we get the desired result. The case when  $m = (n+1), \dots, (2n-1)$  can be handled similarly.

## V. LOSSY TWO-HELP-ONE PROBLEM

Next we consider a coding theorem for continuous sources for the two-help-one problem. Consider a triple of memoryless continuous-valued sources  $(X, Y, Z)$  characterized by a probability measure  $P_{XYZ}$ . Let  $d: \mathbb{R}^2 \rightarrow \mathbb{R}^+$  be a jointly continuous distortion function. The sources  $X$  and  $Y$  act as helpers for the third source  $Z$ . The sources need to be compressed distributively with rates  $R_1, R_2$  and  $R_3$ , respectively, into bits to be sent to a joint decoder. For simplicity we let  $R_3 = 0$ . The joint decoder wishes to reconstruct the source  $Z$  with respect to distortion function  $d$ .

**Definition V.1.** An  $(n, \Theta_1, \Theta_2)$  transmission system consists of mappings  $e_i: \mathbb{R}^n \rightarrow \{1, 2, \dots, \Theta_i\}$ , for  $i = 1, 2$ , and  $f: \{1, 2, \dots, \Theta_1\} \times \{1, 2, \dots, \Theta_2\} \rightarrow \mathbb{R}^n$ . A triple  $(R_1, R_2, D)$  is said to be achievable if there exists a sequence of  $(n, \Theta_{1n}, \Theta_{2n})$  transmission systems such that for  $i = 1, 2$ ,

$$\lim_{n \rightarrow \infty} \frac{\log \Theta_i}{n} \leq R_i, \quad \lim_{n \rightarrow \infty} \mathbb{E} d_n(Z^n, f(e_1(X^n), e_2(Y^n))) \leq D.$$

Let  $\mathbf{R}(D)$  denote the set of rates  $(R_1, R_2)$  such that  $(R_1, R_2, D)$  is achievable.

**Definition V.2.** Let  $\mathcal{P}(D)$  denote the collection of transition probabilities  $P_{QU_1V_1UV\hat{Z}|XY}$  such that (i)  $(UU_1) - (XQ) - (YQ) - (VV_1)$  form a Markov chain, (ii)  $Q$  is independent of  $(X, Y)$ , (iii)  $\hat{Z} = g(U_1, V_1, U + V)$  for some function  $g$ , and (iv)  $\mathbb{E} d(Z, \hat{Z}) \leq D_1$ , where the expectations are evaluated with distribution  $P_{XYZ}P_{QU_1V_1UV\hat{Z}|XY}$ . For a  $P_{QU_1V_1UV\hat{Z}|XY} \in \mathcal{P}(D)$ , let

$\alpha(P_{QU_1V_1UV\hat{Z}|XY})$  denote the set of rate pairs  $(R_1, R_2) \in [0, \infty)^2$  that satisfy

$$\begin{aligned} R_1 &\geq I(X; UU_1|QV_1) + I(U + V; V|QU_1V_1) - I(U; V|QU_1V_1), \\ R_2 &\geq I(Y; VV_1|QU_1) + I(U + V; U|QU_1V_1) - I(U; V|QU_1V_1) \\ R_1 + R_2 &\geq I(XY; UVU_1V_1|Q) + I(U + V; V|QU_1V_1) \\ &\quad + I(U + V; U|QU_1V_1) - I(U; V|QU_1V_1) \end{aligned}$$

where the information terms are evaluated with  $P_{XY}P_{QU_1V_1UV\hat{Z}|XY}$ . Let the rate region be defined as

$$\alpha(D) = cl \left( \bigcup_{P_{QU_1V_1UV\hat{Z}|XY} \in \mathcal{P}(D)} \alpha(P_{QU_1V_1UV\hat{Z}|XY}) \right)$$

**Theorem V.1.** For a source  $(P_{XYZ}, d)$  we have  $\alpha(D) \subseteq \mathbf{R}(D)$ .

We propose a coding scheme involving two layers to prove the theorem. The first is the Berger-Tung unstructured coding layer. The second is the structured coding layer that uses nested linear codes. Although we do not have space to give a complete proof, the key steps in the proof are as follows. First, we quantize the sources and the auxiliary variables, and apply the technique developed in source coding with side information to come up with a discrete version of the problem at hand. The Berger-Tung unstructured coding is accomplished in a straightforward way. The structured coding is accomplished using nested linear codes. The rates associated with this layer can be understood as follows, for example, assuming  $U_1, V_1$  and  $Q$  to be trivial.  $R_1 \geq I(X; U) + H(U + V) - H(U) = I(X; U) - I(U; V) + I(V; U + V)$ , and similarly  $R_2 \geq I(Y; V) - I(U; V) + I(U; U + V)$ . These rates can be achieved using nested linear codes (over arbitrarily large prime fields) along with joint-typical encoding and decoding. Finally, we use the properties of mutual information to show convergence.

## ACKNOWLEDGMENT

This work was supported by NSF grant CCF-1717299.

## REFERENCES

- [1] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 1–10, January 1976.
- [2] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder-II: General sources," *Inform. Contr.*, vol. 38, pp. 60–80, 1978.
- [3] T. Berger, "Multiterminal source coding," in *Lectures presented at CISM summer school on the Inform. Theory approach to communications*, July 1977.
- [4] P. Mitran, "On a Markov Lemma and Typical Sequences for Polish Alphabets," in *IEEE Trans. Inform. Theory*, October, 2015.
- [5] U. Erez and R. Zamir, "Achieving  $1/2 \log(1+\text{SNR})$  on the AWGN channel with lattice encoding and decoding," *IEEE Trans. Inform. Theory*, vol. IT-50, pp. 2293–2314, October 2004.
- [6] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Trans. Inform. Theory*, vol. IT-48, pp. 1250–1276, June 2002.
- [7] J. Korner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. on Inform. Theory*, vol. IT-25, pp. 219–221, March 1979.
- [8] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9, pp. 19–31, 1980.