

# Projections, Dissections and Bandwidth Expansion Mappings

Antonio Campello  
 Institute of Mathematics, Statistics  
 and Computer Science,  
 University of Campinas, São Paulo  
 13083-859, Brazil  
 campello@ime.unicamp.br

Vinay A. Vaishampayan  
 AT&T Labs-Research  
 Shannon Laboratory  
 Florham Park, NJ 07932  
 vinay@research.att.com

Sueli I. R. Costa  
 Institute of Mathematics, Statistics  
 and Computer Science,  
 University of Campinas, São Paulo  
 13083-859, Brazil  
 sueli@ime.unicamp.br

**Abstract**—We address the problem of constructing explicit mappings from a  $k$ -dimensional continuous alphabet source to an  $n$ -dimensional Gaussian channel. The source is assumed to be uniformly distributed on the unit cube  $[0, 1]^k$ . The scheme considered is based on a family of piecewise linear mappings and its performance is shown to be related to specific projected lattices of  $\mathbb{Z}^n$ . We study sufficient conditions for the mean squared error of such mappings to scale optimally with the signal-to-noise ratio of the channel and present an explicit construction for the case  $k = n - 1$ . However, in some other cases our scheme requires the source to be uniformly distributed over a fundamental region of a specific lattice, that may be not congruent to  $[0, 1]^k$ . A dissection technique is presented in order to overcome the source support mismatch and the MSE degradation of such a transformation is analyzed. An example construction of a  $2 : n$  expansion mapping using the dissection technique is presented and is shown to exhibit optimal scaling of the MSE with the channel SNR.

## I. INTRODUCTION

A source vector  $\mathbf{x}$  drawn uniformly and independently on  $[0, 1]^k$  is to be transmitted over an  $n$ -dimensional Gaussian channel and the objective is to minimize the mean squared error (MSE) between  $\mathbf{x}$  and the estimated vector  $\hat{\mathbf{x}}$ , as in Figure 1. Shannon’s separation principle and rate-distortion theory tells us that the best possible distortion is given by  $D \geq \frac{1}{2\pi e}(1 + \text{SNR})^{-n/k}$ , which means that the MSE should decay no faster than  $O(\text{SNR}^{-n/k})$ . This paper studies analog bandwidth expansion codes (i.e., when  $k < n$ ) exhibiting optimal decay  $O(\text{SNR}^{-n/k})$ .

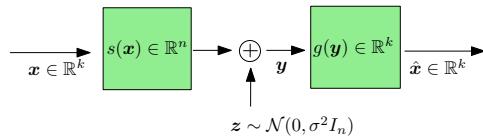


Fig. 1. Communication framework

The codes analyzed are generated by the mod-1 mapping

$$s_1(\mathbf{x}) = (A(\mathbf{x}))_1 := A\mathbf{x} \pmod{1} = A\mathbf{x} - \lfloor A\mathbf{x} \rfloor, \quad (1)$$

where  $A$  is an integer  $n \times k$  matrix. This function maps a cube in  $\mathbb{R}^k$  into a set of plane segments within a box in  $\mathbb{R}^n$

Work partially supported by FAPESP under grants 2007/56052-8 and 2012/09167-2, and CNPq under grant 309561/2009-4

(Figure 3). To achieve good performance, we need a family of mappings with matrices  $A_w$  indexed by a design parameter  $w \in \mathbb{N}$ . The specific code will be chosen based on the SNR.

The rest of the paper is organized as follows. In Section II we briefly discuss the communication model. In Section III, we deal with the geometry of the shift-map and its relation to projections of lattices. A sufficient condition to construct mappings exhibiting optimal scaling of the MSE with the channel SNR and a proof of optimality for the case  $k = n - 1$  are provided in Section IV. This condition relies on finding  $A_w$  in such a way that the lattices obtained by projecting  $\mathbb{Z}^n$  onto  $A_w^\perp$  have good packing density and such that  $A_w$  also satisfies a certain “trace-determinant” property. In Section V we observe that the “trace-determinant” condition can be overcome by a QR factorization, but this forces a constraint on the source support (the alternative communication scheme is shown in Figure 2 (bottom)). In Section VI we show how dissections of polyhedra can be employed to construct a bijection between the original and the constrained supports.

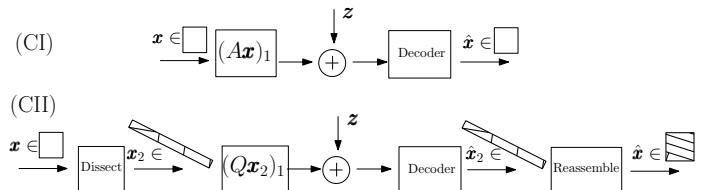


Fig. 2. A diagram of the schemes CI and CII considered here

This work can be viewed as a generalization of some previous constructions [9], [7], where the authors consider  $1 : n$  mappings. Most of the results derived here, when applied to the case  $k = 1$ , reduces to the constructions [9], [7]. The generalization, however, is not straightforward. For instance, the “trace-determinant” condition for  $k : n$  mappings is trivially true if  $k = 1$ , in which case the only issue is the construction of good projection lattice sequences. Other explicit analog codes and some asymptotic analyses are provided for example in [1], [4] and [10]. In [1] the authors provide examples of analog codes ( $2 : 1$  and  $3 : 1$ ) for sources distributed over a Voronoi region of certain lattices, while in

[4] optimality is discussed asymptotically for codes that are closely related to the mod-1 map.

From an information theoretical perspective, this paper proposes the study of analog codes in a systematic way through relating it to mathematical objects such as projections of lattices. From a mathematical point of view, it reinforces the role of projections of the cubic lattice in modulation systems started in [9], and exhibits a new interesting application to the theory of dissections of polyhedra (another application of dissections to Information Theory can be found in [8]).

## II. PRELIMINARIES

Consider the communication scheme in Figure 1 with average power constraint  $E[\|s(\mathbf{x})\|^2] \leq P$ , where the expectation is taken over a uniform source on  $[0, 1]^k$ . The Crámer-Rao bound on the MSE for this model can be evaluated as [5]:

$$\frac{1}{k} E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right] \geq \frac{\sigma^2}{k} \int_{[0,1]^k} \text{tr}((J(\mathbf{x})^t J(\mathbf{x}))^{-1}) d\mathbf{x}, \quad (2)$$

where  $J(\mathbf{x})$  is the Jacobian of  $s(\mathbf{x})$ . It can be shown [5] that if the SNR of the channel goes to infinity, the modulation function approaches the RHS of (2), which we refer to as the *low noise regime*. However, analog bandwidth expansion codes suffer from a threshold effect. Since linear modulation does not scale well as the SNR increases, in order to respect the power constraint the function  $s(\mathbf{x})$  needs to be non-linear and the signal locus will be twisted and folded. If the distance between the “folds” of the signal locus is too small, the low noise approximation will only be valid for very high SNR (the performance will be compromised due to big errors when the noise causes a “jump” to the wrong fold of the signal locus). Therefore, in the design of good codes, a geometrical parameter (the distance between “folds” of the signal locus) has to be taken into account.

Throughout the paper we use standard lattice notation as in [2]. An important definition for our purposes is the one of a primitive set of vectors. We say that a set of vectors  $\mathbf{a}_1, \dots, \mathbf{a}_k \in \mathbb{Z}^n$  is primitive if there exist  $\mathbf{a}_{k+1}, \dots, \mathbf{a}_n \in \mathbb{Z}^n$  such that  $\{\mathbf{a}_1, \dots, \mathbf{a}_n\}$  is a basis for the entire lattice  $\mathbb{Z}^n$ .

## III. GEOMETRY OF THE MAP $(A\mathbf{x})_1$

Consider the mod-1 map (or shift-map) given by Equation (1) and illustrated in Figure 3. Clearly  $s_1(\mathbf{x}) \in [-1/2, 1/2]^2$  for all  $\mathbf{x}$  and the signal locus  $s_1([0, 1]^k)$  will be a set of parallel segments of affine spaces within the box  $[-1/2, 1/2]^n$ . The Jacobian  $J(\mathbf{x})$  of this function is constant and equals  $A$ . The following proposition is an extension of the properties derived in [9] for the  $1 : n$  case and gives conditions such that the mapping (1) is well defined as an encoding function.

**Proposition 1.** *The mapping  $s_1 : [0, 1]^k \rightarrow \mathbb{R}^n$  defined by Equation (1) with  $A \in \mathbb{Z}^{n \times k}$  is injective if and only if  $\text{rank}(A) = k$  and the columns of  $A$  form a primitive set of vectors in  $\mathbb{Z}^n$ . Furthermore, the expected energy of  $s(\mathbf{x})$  for  $\mathbf{x}$  uniformly distributed over  $[0, 1]^k$  is given by  $E[s_1(\mathbf{x})^t s_1(\mathbf{x})] = n/12$ .*

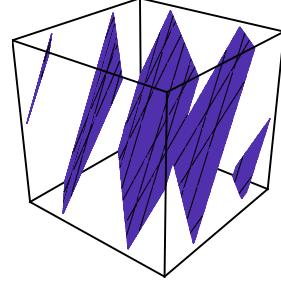


Fig. 3. Image of  $s_1(\mathbf{x})$  for  $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 2 & 2 \end{pmatrix}$

This proposition shows that the mapping  $s(\mathbf{x}) = \alpha s_1(\mathbf{x})$ , where  $\alpha = 2\sqrt{3P}/\sqrt{n}$  satisfies the power constraint. Due to periodicity and symmetry, the distance  $\delta$  between two folds of the locus can be evaluated as  $\alpha$  times the minimum distance  $d$  between two affine spaces in  $s_1(\mathbb{R}^k)$  given by

$$\delta = \min_{\mathbf{n} \in \mathbb{R}^n, \mathbf{n} \notin A^\perp} \min_{\mathbf{x} \in \mathbb{Z}^k} \|A\mathbf{x} - \mathbf{n}\|,$$

where  $A^\perp$  is the orthogonal complement of the column space of  $A$ . The inner minimization is solved by projecting  $n$  onto  $A^\perp$ , so that  $\min_{\mathbf{x} \in \mathbb{R}^k} \|A\mathbf{x} - \mathbf{n}\| = \|P_{A^\perp}(\mathbf{n})\|$ . From the fact that  $n \notin A^\perp$  iff  $P_{A^\perp}(\mathbf{n}) \neq \mathbf{0}$ , we have:

$$\delta = \min_{\mathbf{y} \neq \mathbf{y} \in P_{A^\perp}(\mathbb{Z}^n)} \|\mathbf{y}\|. \quad (3)$$

Equation (3) above shows the desired connection between the mod-1 map and projections of the cubic lattice. The minimum distance of the signal locus is exactly the norm of the shortest vector in the projection of  $\mathbb{Z}^n$  onto  $A^\perp$ . If the columns of  $A \in \mathbb{Z}^{n \times k}$  are a primitive set of vectors, then  $P_{A^\perp}(\mathbb{Z}^n)$  is a  $(n-k)$ -dimensional lattice with fundamental volume equal to  $\det P_{A^\perp}(\mathbb{Z}^n) = \sqrt{\det(A^t A)^{-1}}$  [2, Ch. 6 Thm. 4]. This means that the *center density* (i.e., the density divided by the volume of a sphere with radius 1) of the projection lattice is given by  $\Delta = (\delta/2)^{n-k} \sqrt{\det A^t A}$ , from where we get  $\delta = 2\Delta^{1/(n-k)} \det(A^t A)^{-1/2(n-k)}$ . After applying the scaling factor  $\alpha$ , the packing radius (half of the minimum distance)  $\rho$  of the signal locus is given by:

$$\rho = \frac{\alpha \delta}{2} = \frac{2\sqrt{3P}\Delta^{1/(n-k)}}{\sqrt{n} \det(A^t A)^{1/2(n-k)}}, \quad (4)$$

i.e., there is a tradeoff between  $\rho$  and  $\det(A^t A)$ .

**Example 1.** Consider the matrices

$$A_w = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ w & 1 & 0 & \cdots & 0 \\ 0 & w & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & w \end{pmatrix}, \quad (5)$$

for  $w \in \mathbb{N}$ . One can easily verify that the columns of  $A_w$  are a primitive set and  $\det(A_w^t A_w) = \sum_{i=1}^{n-1} w^{2i}$ . Since  $A_w$

generates a  $(n - 1)$ -dimensional space in  $\mathbb{R}^n$ , its projection is a 1-dimensional lattice ( $\Delta = 1/2$ ). Thus:

$$\delta = \frac{1}{\sqrt{\det(A_w^t A_w)}} = \frac{1}{(\sum_{i=1}^{n-1} w^{2i})^{1/2}}. \quad (6)$$

**Example 2.** Let  $A_w$  be the sequence of matrices given by

$$(A_w)_{ij} = \begin{cases} w^{\lfloor (i-1)/2 \rfloor} & \text{if } (i, j) = (\text{odd}, 1) \text{ or } (\text{even}, 2) \\ 0 & \text{otherwise} \end{cases}$$

and  $w \in \mathbb{N}$ . For instance, for odd  $n = 2m + 1$ , the image of  $A_w$  is generated by the vectors  $(1, 0, w, 0, w^2, \dots, 0, w^m)$  and  $(0, 1, 0, w, \dots, w^{m-1}, 0)$ . Arguing as in [7, Sec. VI], we can show that the projections of  $\mathbb{Z}^n$  onto  $A_w^\perp$  are, up to equivalence, arbitrarily close to  $\mathbb{Z}^{n-2}$ . Figure 2 displays the center density of the projection lattices as  $w$  increases when  $n = 5$ . From this and (4) we can recover the minimum distance.

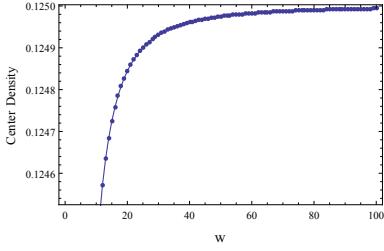


Fig. 4. Center density of  $P_{A_w^\perp}(\mathbb{Z}^5)$  as a function of  $w$ . The center density of  $\mathbb{Z}^3$  (0.125) is approached as  $w \rightarrow \infty$ .

#### IV. ANALYSIS OF THE MAP $(Ax)_1$ : COMMUNICATION SCHEME CI

We proceed now to develop a sufficient condition for a family of mappings of the form (1) to achieve the correct MSE decay order  $O(\text{SNR}^{-n/k})$ . Consider the family of mappings  $s_w(\mathbf{x}) = \alpha(A_w \mathbf{x} \pmod{1})$ , where  $w \in \mathbb{N}$  is a design parameter that will be chosen according to the SNR, and  $\alpha = 3\sqrt{2P}/\sqrt{n}$ , so that the average transmission power is  $P$ . As previously shown, the minimum distance between folds of the signal locus (after scaling) is given by  $\rho_w$  as in (4). Let  $\mathbf{y} = s_w(\mathbf{x}) + \mathbf{z}$  be the received vector, where  $\mathbf{z} \sim \mathcal{N}(0, \sigma^2 I)$  and  $\sigma$  is fixed. The analysis from now on is very similar to [9, Sec. VI B]. First we bound the MSE as

$$\text{MSE} \leq \frac{1}{k} E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \mid \|\mathbf{z}\| < \rho_w \right] + P(\|\mathbf{z}\| \geq \rho_w). \quad (7)$$

The first term of (7) is the MSE provided that no large error occurs (the value is decoded to the right fold of the curve) and approaches the RHS of (2) as  $P \rightarrow \infty$ . The second term is the probability that a normal random vector has norm greater than  $\rho_w$ . From this:

$$\text{MSE} \leq \frac{\sigma^2 n \text{tr}(A_w^t A_w)^{-1}}{12kP} + e^{-\rho_w^2/2\sigma^2} \sum_{i=0}^{n-1} \frac{(\rho_w/2\sigma^2)^i}{i!}. \quad (8)$$

To make the first term achieve the right decay order  $O(1/P^{n/k})$  it is necessary (and sufficient) that  $\text{tr}(A_w^t A_w)^{-1} = O(1/P^{(n-k)/k})$ . If the density of the projections is bounded away from zero and  $P$  grows with order  $\det(A_w^t A_w)^{1/(n-k)+\mu}$  for some  $\mu > 0$ , then due to (4) the second term is exponentially small and the first term dominates the RHS of the inequality above. This implies that we must have  $\text{tr}(A_w^t A_w)^{-1} = O(\det(A_w^t A_w)^{-1/k})$ . If the condition above is met, then the MSE of family of modulation systems indexed by  $w$  decays with order  $O(P^{n/(k+\mu)})$ , for any  $\mu$  arbitrarily close to 0. Summarizing, in order to guarantee the asymptotic analysis we must have a sequence of matrices  $A_w$  satisfying

- 1) (Injectivity) The columns of  $A_w$  are a primitive set of vectors in  $\mathbb{Z}^n$
- 2) (Minimum distance) The density of the projections of  $\mathbb{Z}^n$  onto  $A_w^\perp$  is bounded away from zero as  $w$  increases
- 3) (MSE Exponent)

$$\text{tr}(A_w^t A_w)^{-1} = O(\det(A_w^t A_w)^{-1/k}) \quad (9)$$

From an arithmetic-geometric-mean inequality argument, it is possible to show that  $\text{tr}(A^t A)^{-1} \geq k \det(A^t A)^{-1/k}$  for any full-rank  $n \times k$  matrix, with equality holding if and only if the columns of  $A$  are orthogonal and have the same norm, hence the third condition is very tight. If the matrix  $A_w$  is orthogonal, however, it is trivially satisfied and we only need to find a sequence of projections of  $\mathbb{Z}^n$  with good density.

For instance, let  $A_w$  be the matrices in Example 2. For even  $n$  the columns of  $A_w$  are orthogonal and have the same norm, so the construction clearly meets Condition (9). However a direct calculation shows that this is *not* true for odd  $n = 2m+1$  and the best possible achievable exponent from the analysis here is  $P^{-m} = P^{-(n-1)/2}$  (the same exponent provided by “giving away” one dimension).

**Example 3** (Optimality of the  $(n-1) : n$  mapping). *Let  $A_w$  be the matrices in Example 1 and  $s_w(\mathbf{x})$  the associated mappings. As shown in the example, the mapping is injective and satisfies the density condition. Furthermore  $[(A_w^t A_w)^{-1}]_{ii} = O(1/w^2)$  for all  $i = 1, \dots, n-1$ , hence the “trace-determinant” condition for the MSE exponent is achieved. From the previous analysis, if we choose the parameter  $w$  such that the power  $P = O(w^{2n-2+\mu})$ , then the non-vanishing minimum distance is also satisfied, proving that, as  $\mu \rightarrow 0$ , the MSE for this family approaches optimal decay order with the power.*

Other examples for specific values of  $k$  and  $n$  are possible to construct, however a general construction that directly satisfies the sufficient conditions may require new techniques and is left for future investigation.

#### V. ANALYSIS OF THE MAP $(Qx_2)_1$ : COMMUNICATION SCHEME CII

The matrix  $A$  determines the modulation signal set, but there are other possible mappings that produce the same signal locus. Although the determinant is inherent to the

locus, the trace is a property of the specific mapping chosen from the source support to the signal space. We provide next an alternative mapping that satisfies the “trace-determinant” condition.

Given a sequence  $A_w$  it is always possible to orthogonalize its columns (via QR factorization for instance) such that  $A_w = Q'_w R'_w$ , where  $Q'_w$  is an orthogonal matrix and  $R'_w$  is upper triangular. After scaling, we obtain  $A_w = Q_w R_w$  where  $Q_w = \beta_w Q'_w$ ,  $R_w = (1/\beta_w) R'_w$  and  $\beta_w = (\det(A_w^t A_w)^{1/2k})$ , so that  $\det R_w = 1$ . Hence, the image by  $(A_w \mathbf{x})_1$  of the unit cube  $[0, 1]^k$  is the same as the image by  $(Q_w)_1$  of the unit area parallelepiped  $\mathcal{S}_w = R_w [0, 1]^k$ , with the additional feature that  $Q_w^t Q_w = \beta_w^2 I_k$ , ensuring the trace-determinant condition. Since  $R_w$  is a volume-preserving mapping, for a source uniformly distributed over  $\mathcal{S}_w$  essentially the same asymptotic analysis developed in the previous sections hold, substituting  $A_w$  by  $Q_w$ . Thus if  $A_w$  meets the minimum distance/injectivity conditions, then the mapping

$$s_Q : \mathcal{S}_w \rightarrow \mathbb{R}^n$$

$$s_Q(\mathbf{x}_2) = \alpha(Q_w \mathbf{x}_2 \pmod{1}). \quad (10)$$

will also meet both conditions *and* the “trace-determinant”, showing that it has optimal scaling with the power. However note that the source has support  $\mathcal{S}_w \neq [0, 1]^k$ . The support is now a fundamental region of the lattice with generator matrix  $R_w$ . A method to map  $\mathcal{S}_w$  into  $[0, 1]^k$  without increasing the MSE needs to act like an isometry. We show next that dissections have this desired property.

## VI. TRANSFORMING THE SOURCE SUPPORT: DISSECTIONS

Our objective is to construct a bijection between  $[0, 1]^k$  and  $\mathcal{S}_w$  as defined in Sec. V. In order to do so we:

- i. Dissect  $[0, 1]^k$  into  $m$  non-overlapping polyhedra  $T_1, T_2, \dots, T_m$  i.e.,  $[0, 1]^k = \bigcup_{i=1}^m T_i$ ;
- ii. Dissect  $\mathcal{S}_w$  into  $m$  non-overlapping polyhedra  $\tilde{T}_1, \tilde{T}_2, \dots, \tilde{T}_m$ ,  $\mathcal{S}_w = \bigcup_{i=1}^m \tilde{T}_i$ ,

in such a way that  $T_i$  and  $\tilde{T}_i$  are related by a rigid motion  $\phi_i$  (a composition of translations, rotations and reflections) i.e.,  $\tilde{T}_i = \phi_i(T_i)$ . This enables us to establish the bijection  $\Phi$  where  $\mathbf{x}_2 = \Phi(\mathbf{x}) := \phi_i(\mathbf{x})$ ,  $\mathbf{x} \in T_i$ ,  $i = 1, 2, \dots, m$ .

Suppose that  $\mathbf{x}_2 \in \tilde{T}_i$  is transmitted. It is possible that a small noise event results in a decoded vector which lies in  $\tilde{T}_j$ ,  $j \neq i$ . The discontinuities in  $\Phi$  will then cause an increase of the MSE between  $\mathbf{x}$  and  $\hat{\mathbf{x}}$ , see Figure 6a). We control the degradation of the MSE by modifying the bijection so as to shrink each piece of the dissection. This reduces the probability of decoding to an incorrect piece of the dissection. The modified bijection is defined next.

**Encoding:** The encoder finds  $i$  such that  $\mathbf{x} \in T_i$  and applies  $s(\mathbf{x}) = s_Q((1-\varepsilon)\phi_i(\mathbf{x}) + \mathbf{t}_i)$  where  $\varepsilon > 0$  and  $\mathbf{t}_i$  are such that each piece of  $\mathcal{S}_w$  is scaled of a factor  $(1 - \varepsilon)$  and separated apart, as in Figure 6b).

**Decoding:** Given  $\mathbf{y} = s(\mathbf{x}) + \mathbf{z} = s_Q(\mathbf{x}_2) + \mathbf{z}$ , where  $\mathbf{x}_2 = (1 - \varepsilon)\phi_i(\mathbf{x}) + \mathbf{t}_i$ , the receiver first decodes  $\mathbf{y}$  to recover  $\hat{\mathbf{x}}_2 \in \mathcal{S}_w$  and then computes  $\hat{\mathbf{x}} = \phi_i^{-1}(\hat{\mathbf{x}}_2 - \mathbf{t}_i)/(1 - \varepsilon)$ .

*Sketch of the MSE Analysis:* We give a sketch of the MSE analysis, due to page constraints. A careful analysis will be given in a full version of this paper. Let  $\rho$  be half of the distance between the folds of  $s_Q(\mathcal{S}_w)$ . If we define “jump” as the event when  $\mathbf{x} \in T_i$  and  $\hat{\mathbf{x}} \in T_j$  for  $i \neq j$  i.e., the estimate  $\hat{\mathbf{x}}$  “jumps” to the wrong piece of the dissection, then the MSE can be bounded as

$$\begin{aligned} \frac{1}{k} E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \right] &\leq \frac{1}{k} E \left[ \|\mathbf{x} - \hat{\mathbf{x}}\|^2 \mid \|\mathbf{z}\| < \rho, \text{“no jump”} \right] \\ &+ P(\text{“jump”} \mid \|\mathbf{z}\| \leq \rho) + P(\|\mathbf{z}\| > \rho) = \\ &\frac{1}{k(1-\varepsilon)^2} E \left[ \|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|^2 \mid \|\mathbf{z}\| < \rho, \text{“no jump”} \right] \\ &+ P(\text{“jump”} \mid \|\mathbf{z}\| \leq \rho) + P(\|\mathbf{z}\| > \rho) \end{aligned}$$

If the noise is smaller than the distance between folds of the locus, then  $\hat{\mathbf{x}}_2$  will be given by the projection onto the corresponding fold. In this case, the error will be  $\hat{\mathbf{x}}_2 - \mathbf{x}_2 = Q_w^t \mathbf{z} / (\alpha \beta^2) = \mathbf{z}' \sim \mathcal{N}(0, \sigma^2 / \alpha \beta_w I_k)$ . Suppose now that the translation vector  $\mathbf{t}_i$  and  $\varepsilon$  where chosen such that each piece is separated of at least  $d$  from its neighbors. Then, by analyzing the projection of  $\mathbf{z}'$  onto the plane that determines a cut, we can show that  $P(\hat{\mathbf{x}} \in T_j \mid \mathbf{x} \in T_i) \leq Q(d\alpha\beta_w/\sigma)$ , where  $Q(x)$  is tail probability of a standard normal distribution, therefore the following holds:

**Proposition 2.** *If each piece  $\tilde{T}_i$  of the dissection has at most  $r$  neighbors then:*

$$P(\text{“jump”} \mid \|\mathbf{z}\| < \rho) \leq rQ(d\alpha\beta/\sigma). \quad (11)$$

Thus, if we can choose  $\varepsilon$  in such a way that  $\varepsilon \rightarrow 0$  and  $d\alpha\beta_w$  grows with order  $P^\mu$  for some  $\mu > 0$ , the “jump” events will have exponentially small probability and essentially the same asymptotic analysis will hold for  $s$  and  $s_Q$ . In the next Section we show that Montucla’s rectangle-to-square dissection meets these requirements.

### A. A construction for $2 : n$

Let  $A_w$  be the construction described in Example 2. As shown in Section IV, this construction does not exhibit optimal scaling for odd  $n = 2m + 1$ , for it does not meet the “trace-determinant” condition. We will show that a dissection technique can be employed to achieve the correct order  $P^{-n/2}$ .

First note that  $\det(A_w^t A_w) = O(w^{4m-2})$ . Using the notation of Section V, we can write  $A_w = Q_w R_w$  where  $Q_w$  is an orthogonal  $n \times 2$  matrix and  $R_w$  is a determinant one diagonal matrix  $\text{Diag}(l_w, 1/l_w)$ , with  $l_w = O(\sqrt{w})$ . Also  $\beta_w = O(w^{m-1/2})$ . From Section IV, if we set  $w$  such that  $P$  has at least order  $\det(A_w^t A_w)^{1/(n-2)+\mu} = w^{2+\mu(4m-2)}$ , large errors will occur with exponentially small probability. For the source uniformly distributed over  $\mathcal{S}_w = [l_w, 0] \times [0, 1/l_w]$ , the MSE provided that no large error occurs is approximately the RHS of (2). This implies that

$$\begin{aligned} \text{MSE} &\approx \frac{\sigma^2}{\alpha^2 k} \int_{R_w[0,1]^2} \text{tr}((Q_w^t Q_w))^{-1} dx \\ &= \frac{n\sigma^2}{12P\beta_w^2} = O\left(\frac{1}{Pw^{2m-1}}\right) \approx O\left(P^{-\frac{n}{2}}\right). \end{aligned} \quad (12)$$

Therefore the mapping  $s_Q$  from  $\mathcal{S}_w$  (i.e., after dissecting and before reassembling the pieces) has correct decay exponent. To go back to the uniform source over the cube  $[0,1]^2$  we will dissect the thin rectangle  $\mathcal{S}_w$  into  $[0,1]^2$ .

### B. Montucla's rectangle-to-square dissection

A rectangle-to-rectangle dissection that, when applied to a unit square and a rectangle with dimensions  $l \times 1/l$ , requires less than  $\lceil l \rceil + 2$  pieces is shown in [3, p. 222]. We give a formal description of that dissection in terms of the “Two-tiles” theorem [6]. The two tiles theorem asserts that if two polyhedra tile some subset of  $\mathbb{R}^n$  under the same group of isometries, then they are equidissectable. Let  $[l, 0] \times [0, 1/l]$  be the rectangle with lengths  $l \times 1/l$ . We consider a rotated version of it, namely  $\mathcal{R} = \{\alpha v_1 + \beta v_2; 0 \leq \alpha, \beta \leq 1\}$  where  $v_1 = (-\sqrt{l^2 - 1}, 1)$  and  $v_2 = (1/l^2, \sqrt{l^2 - 1}/l^2)$ . Let  $\Lambda$  be the group of translations  $\Lambda = \{x_1(1, 0) + x_2(-\sqrt{l^2 - 1}, 1) : x_1, x_2 \in \mathbb{Z}\}$ . It can be shown that both the square  $[0, 1]^2$  and the rectangle  $\mathcal{R}$  are  $\Lambda$ -tiles, hence they are equidissectable. An illustration of the dissection is shown in Figure 5. The pieces can be described as the intersection of the rectangle  $\mathcal{R}$  with translations of the square by  $\Lambda$  (or vice-versa). This way, one can count the number of pieces as the number of translation vectors in  $t \in \Lambda$  such that  $([0, 1]^2 + t) \cap \mathcal{R} \neq \emptyset$  to show that only  $\lceil l \rceil + 2$  pieces are needed.

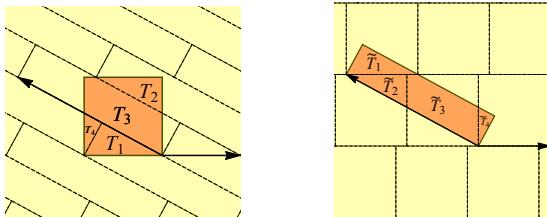


Fig. 5. Two tiles illustration

Montucla's dissection provides  $M \leq \lceil l \rceil + 2$  pieces, using  $M - 2$  vertical cuts and 1 horizontal cut. It is clear that we can separate the half-planes generated by the vertical cuts translating one of the edges of the corresponding piece to the origin, applying the scaling factor  $(1 - \varepsilon)$ , and then translating back to the original edge position, as in Figure 6b). In this case, from triangle congruence, each piece is at least a distance  $d = \frac{\varepsilon}{2}$  to its neighbor. The triangle provided by the unique “horizontal” cut can be translated such that the edge corresponding to the rectangle’s edge becomes the origin, scaled, and then translated back. This gives the distance to its neighboring pieces  $\bar{d} \approx \varepsilon/l_w \leq d$ . From the previous subsection,  $\alpha\beta_w = O(w^{m+1/2})$ , so if we choose  $\varepsilon = w^{-m+\mu}$

for any small  $\mu > 0$  the jump probability will be exponentially small in  $w$  (and thus in  $P$ ) and  $\varepsilon \rightarrow 0$  as  $w \rightarrow \infty$ .

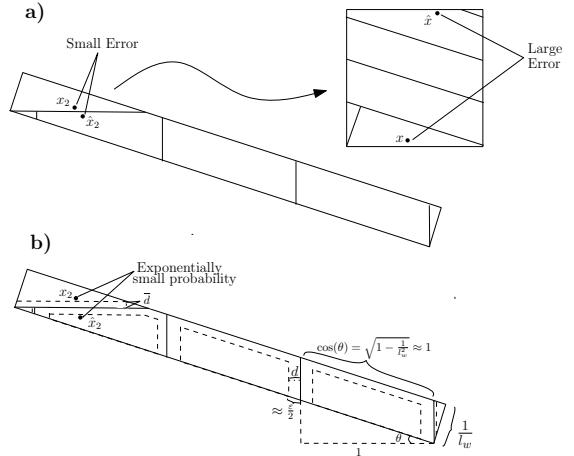


Fig. 6. a) A large error due to reassembling the pieces. b) An illustration of the technique of shrinking/separating the pieces

Summarizing, since all the pieces have at most  $r = 3$  neighbors, and each of them is separated by at least a distance  $\bar{d}$ , we have:

$$\begin{aligned} \text{MSE} &\leq \frac{1}{k(1-\varepsilon)^2} \overbrace{E[\|\mathbf{x}_2 - \hat{\mathbf{x}}_2\|^2 | \text{“no jump”, } \|\mathbf{z}\| \leq \rho]}^{O(P^{-n/2}) \text{ from (12)}} + \\ &+ \underbrace{P(\|\mathbf{z}\| \geq \rho)}_{\text{exponentially small from (8)}} + \underbrace{P(\text{“jump”} | \|\mathbf{z}\| < \rho)}_{\text{exponentially small from Prop. 2}} = O(P^{-n/2}). \end{aligned} \quad (13)$$

## REFERENCES

- [1] S. Chung. *On the construction of some capacity-approaching coding schemes*. Ph.D. dissertation, Massachusetts Institute of Technology, 2000.
- [2] J. H. Conway and N. J. A. Sloane. *Sphere-packings, lattices, and groups*. Springer-Verlag, New York, NY, USA, 1998.
- [3] G. N. Frederickson. *Dissections: Plane and Fancy*. Cambridge University Press, 2003.
- [4] M. Kleiner and B. Rimoldi. A tight bound on the performance of a minimal-delay joint source-channel coding scheme. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 136–140, 2010.
- [5] D. J. Sakrison. *Notes on analog communications*. New York: Van Nostrand Reinhold, 1970.
- [6] N. J. A. Sloane and V. A. Vaishampayan. Generalizations of schbis tetrahedral dissection. *Discrete & Computational Geometry*, 41(2):232–248, 2009.
- [7] N. J. A. Sloane, V. A. Vaishampayan, and S. I. R. Costa. The lifting construction: A general solution for the fat strut problem. In *IEEE International Symposium on Information Theory Proceedings (ISIT)*, pages 1037–1041, 2010.
- [8] C. Tian, V. A. Vaishampayan, and N. J. A. Sloane. A coding algorithm for constant weight vectors: A geometric approach based on dissections. *IEEE Transactions on Information Theory*, 55(3):1051–1060, 2009.
- [9] V. A. Vaishampayan and S. I. R. Costa. Curves on a sphere, shift-map dynamics, and error control for continuous alphabet sources. *IEEE Transactions on Information Theory*, 49:1658–1672, 2003.
- [10] N. Wernersson, M. Skoglund, and T. Ramstad. Polynomial based analog source-channel codes. *IEEE Transactions on Communications*, 57(9):2600–2606, september 2009.