

Covering Arbitrary Point Patterns

Arya Mazumdar and Ligong Wang

Department of EECS / Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
{aryam,wlg}@mit.edu

Abstract—This paper considers the problem of covering an arbitrary *point pattern*—a set of λT points in the interval $[0, T]$ —with a subset of $[0, T]$ that is drawn from a predefined codebook. The subset is required to contain either all or a certain proportion of the points in the pattern, depending on the problem setting. Also, all subsets in this codebook must have Lebesgue measure not exceeding dT where $d \leq 1$ is a given constant. The problem of interest here is to find the trade-off between d and the size of the codebook. We find this trade-off asymptotically as T goes to infinity. When the subset is required to cover all the points, the answer turns out to be the same as in the case where the points were randomly generated by a Poisson process of intensity λ , the latter being obtained in an earlier work.

I. INTRODUCTION

Consider a source that generates a “point pattern” x —a finite subset of the interval $[0, T]$. After observing x , we draw a “covering set”—a subset \mathcal{X} of $[0, T]$ which must contain x —from a beforehand-prepared codebook of such subsets. Clearly there is a trade-off between the size of the codebook and the sizes (i.e., Lebesgue measures) of the covering sets in the codebook.

This trade-off was first considered in [1] in the flavor of a continuous-time rate-distortion problem. Here the *rate* R is defined as the logarithm of the size of the codebook divided by T , and the *distortion* d is defined as the expected Lebesgue measure of the covering set divided by T , provided that it covers all the points in the pattern. Unlike previous works on rate-distortion problems for point processes (mostly Poisson processes) [2]–[6], this new model makes reconstruction in a space different from that of the source. Indeed, the source is a finite subset of $[0, T]$, while the reconstruction set is in general uncountably infinite and has a positive Lebesgue measure. Despite this asymmetry, this model turns out to be very natural in that it has a simple solution, and that it is closely related to the Poisson channel [7].

The work of [1]

- provides a closed-form expression of the rate-distortion function when the source is a homogeneous Poisson process of intensity λ :¹

$$R_{\text{Pois}}(d, \lambda) = -\lambda \log d, \quad 0 < d \leq 1; \quad (1)$$

- shows that the rate-distortion function for any other random source generating no more than λ points per second is at most $R_{\text{Pois}}(d, \lambda)$;
- shows that $R_{\text{Pois}}(d, \lambda)$ is achievable in the adversarial setting where the source is arbitrary, containing no more than λ points per second, and where the codebook could be random; and
- extends all the above results to the Wyner-Ziv setting [8] where the reconstructor knows some of the points in the source.

One open question from [1] is in the adversarial setting: how many random bits are needed to generate the random codebook in order to cover adversarial point patterns of λT points? In the present work we answer this question and show that we do not need any random bits at all: a *deterministic* codebook of size $2^{R_{\text{Pois}}(d, \lambda) + \epsilon}$, with any $\epsilon > 0$, is sufficient to cover all size- λT point patterns. We formulate and prove this result in Section II. In Section II we also provide a self-contained converse proof for the adversarial setting which is entirely different from the converse proof for Poisson sources provided in [1]. This converse proof is based on a very simple nonasymptotic lower bound on the codebook size. Both the achievability and the converse proofs employ some ideas present in Goblick’s thesis [9] and Wyner’s paper on covering of the unit n -square [10], but they also contain some new elements. Our result gives further validation to the fact commented upon in [1]: *the Poisson process is the most difficult to cover*.

We also consider the problem where only a proportion of the points in the source must be covered, and derive an expression on the minimum required rate. Namely, if only $b \in (d, 1]$ proportion of the points in the pattern

This work was supported in part by AFOSR under Grant No. FA9550-11-1-0183, and by NSF under Grant No. CCF-1017772.

¹Unless otherwise stated, all logarithms in this paper are base-two.

are to be covered, then the minimum possible as well as achievable rate is $\lambda D_2(b||d)$, where

$$D_2(p||q) \triangleq p \log \frac{p}{q} + (1-p) \log \frac{1-p}{1-q}, \quad p, q \in [0, 1] \quad (2)$$

is the *relative entropy* between the two Bernoulli distributions with parameters p and q , respectively. This part is a generalization of the problem in Section II and is discussed in Section III.

II. COVERING ALL POINTS

We now formulate the covering-arbitrary-point-patterns problem. Let $\mathcal{P}(\lambda, T)$ denote the set of point patterns on $[0, T]$ that contain no more than λT points:

$$\mathcal{P}(\lambda, T) \triangleq \{\mathbf{x} \subseteq [0, T]: |\mathbf{x}| \leq \lambda T\}. \quad (3)$$

Given $d \in (0, 1]$, let the codebook \mathcal{C} consist of subsets of $[0, T]$ of Lebesgue measures not exceeding dT :

$$\mathcal{C} \subseteq \{\mathcal{X} \subseteq [0, T]: \mu(\mathcal{X}) \leq dT\}, \quad (4)$$

where $\mu(\cdot)$ denotes the Lebesgue measure on \mathbb{R} . The codebook \mathcal{C} is called λ -covering if every point pattern in $\mathcal{P}(\lambda, T)$ is covered by at least one codeword in \mathcal{C} , i.e., if

$$\forall \mathbf{x} \in \mathcal{P}(\lambda, T), \quad \exists \mathcal{X} \in \mathcal{C}: \mathbf{x} \subseteq \mathcal{X}. \quad (5)$$

Denote

$$M(\lambda, d, T) \triangleq \min\{|\mathcal{C}|: \mathcal{C} \text{ is } \lambda\text{-covering}\}. \quad (6)$$

Our first result is the following:

Theorem 1: For all $\lambda > 0$ and $d \in (0, 1]$,

$$\lim_{T \rightarrow \infty} \frac{\log M(\lambda, d, T)}{T} = -\lambda \log d. \quad (7)$$

The proof of the achievability part of Theorem 1 combines the idea of dividing the interval $[0, T]$ into small parts as in [1] with the random covering idea as in [9], [10] (see also [11]) and is given in Section II-A. The converse part of Theorem 1 can follow directly from [1, Theorem 1]. However, we give an entirely new proof to this converse which, unlike [1, Theorem 1], does not rely on any special property of the Lebesgue measure, and is hence applicable to covering point patterns on general measurable spaces. In the new proof we derive a *nonasymptotic* lower bound on the codebook size (Proposition 1) which does *not* follow from [1].

A. Proof of Theorem 1—Achievability

We show existence of a λ -covering codebook by probabilistic methods (see [11]).

Like [1], we partition the interval $[0, T]$ into T/Δ parts of equal size Δ .² The parts are namely

$$\mathcal{I}^\Delta \triangleq \{[0, \Delta), \dots, [T - \Delta, T]\}, \quad (8)$$

Since every point pattern in $\mathcal{P}(\lambda, T)$ is covered by the union of at most λT parts in \mathcal{I}^Δ , we now reduce the problem of covering elements of $\mathcal{P}(\lambda, T)$ to the problem of covering all subsets of \mathcal{I}^Δ of size λT .³

We randomly generate a size- M codebook $\mathcal{C} \subseteq \{\mathcal{X} \subseteq [0, T]: \mu(\mathcal{X}) = dT\}$ in the following way: each covering set $\mathcal{X} \in \mathcal{C}$ is the union of dT/Δ parts picked uniformly from \mathcal{I}^Δ and is chosen independently of the other covering sets in \mathcal{C} .

Fix $\mathbf{x}^\Delta \subseteq \mathcal{I}^\Delta$, $|\mathbf{x}^\Delta| = \lambda T$. The probability that a covering set randomly chosen as above covers \mathbf{x}^Δ is

$$\frac{\binom{T/\Delta - \lambda T}{dT/\Delta - \lambda T}}{\binom{T/\Delta}{dT/\Delta}}. \quad (9)$$

Since all M covering sets in \mathcal{C} are chosen independently, the probability that no element in \mathcal{C} covers \mathbf{x}^Δ is

$$\left(1 - \frac{\binom{T/\Delta - \lambda T}{dT/\Delta - \lambda T}}{\binom{T/\Delta}{dT/\Delta}}\right)^M \quad (10)$$

which is upper-bounded by

$$\exp\left(-M \frac{\binom{T/\Delta - \lambda T}{dT/\Delta - \lambda T}}{\binom{T/\Delta}{dT/\Delta}}\right), \quad (11)$$

as $1 - a \leq e^{-a}$ for all $a \in \mathbb{R}$.

We next upper-bound the probability that there exists at least one λT -subset of \mathcal{I}^Δ that is *not* covered by any element in \mathcal{C} . To this end, we use the union bound and the fact that the total number of λT -subsets of \mathcal{I}^Δ is

$$\binom{T/\Delta}{\lambda T}. \quad (12)$$

²We assume Δ to be small and ignore the edge effect that T or dT might not be divisible by Δ .

³Note that we do not need to consider subsets of \mathcal{I}^Δ of sizes smaller than λT , as every such subset is contained in a subset of size equal to λT .

These yield that the probability of interest is upper-bounded by

$$\binom{T/\Delta}{\lambda T} \exp \left(-M \frac{\binom{T/\Delta - \lambda T}{dT/\Delta - \lambda T}}{\binom{T/\Delta}{dT/\Delta}} \right). \quad (13)$$

If (13) is less than one, then there must exist a deterministic \mathcal{C} that is λ -covering. The condition (13) being less than one is equivalent to

$$M > \frac{\binom{T/\Delta}{dT/\Delta}}{\binom{T/\Delta - \lambda T}{dT/\Delta - \lambda T}} \ln \binom{T/\Delta}{\lambda T}. \quad (14)$$

We next apply the bound (see, e.g., [12])

$$\frac{2^{nH_2(k/n)}}{n+1} \leq \binom{n}{k} \leq 2^{nH_2(k/n)}, \quad (15)$$

where $H_2(\cdot)$ denotes the binary entropy function:

$$H_2(p) \triangleq -p \log p - (1-p) \log(1-p), \quad p \in [0, 1]. \quad (16)$$

We obtain that (14) holds whenever

$$\begin{aligned} \log M &> \frac{T}{\Delta} H_2(d) - \left(\frac{T}{\Delta} - \lambda T \right) H_2 \left(\frac{d - \lambda \Delta}{1 - \lambda \Delta} \right) \\ &+ \log \left(\frac{T}{\Delta} - \lambda T + 1 \right) + \log \left(\frac{T}{\Delta} H_2(\lambda \Delta) \ln 2 \right). \end{aligned} \quad (17)$$

Letting T increase to infinity we obtain that a sequence of λ -covering codebooks of size M_T exists if

$$\begin{aligned} \lim_{T \rightarrow \infty} \frac{\log M_T}{T} \\ > \frac{1}{\Delta} \left(H_2(d) - (1 - \lambda \Delta) H_2 \left(\frac{d - \lambda \Delta}{1 - \lambda \Delta} \right) \right). \end{aligned} \quad (18)$$

Since the RHS of (18) is achievable in the limit as T goes to infinity for every positive Δ , so must be its infimum over Δ . This infimum is

$$\begin{aligned} \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \left(H_2(d) - (1 - \lambda \Delta) H_2 \left(\frac{d - \lambda \Delta}{1 - \lambda \Delta} \right) \right) \\ = -\lambda \log d. \end{aligned} \quad (19)$$

This proves the achievability part of Theorem 1.

Remark 1: A stronger version of Theorem 1 is true: almost all codebooks generated by the random choice are λ -covering if we discount a little in the trade-off between rate and distortion (where the loss goes to 0 as T goes to infinity). Indeed, if we require (13) to be less than some positive constant ϵ , then with probability at least $1 - \epsilon$ the codebook \mathcal{C} is λ -covering. For any positive ϵ this adds only a term of $\frac{1}{T} \log \log(1/\epsilon)$ to the RHS of (17), which goes to 0 as T goes to infinity.

B. Proof of Theorem 1—Converse

We prove the following nonasymptotic lower bound on $|\mathcal{C}|$, and then the converse part of Theorem 1 follows immediately:

Proposition 1: Any λ -covering codebook \mathcal{C} on $[0, T]$ whose codewords have Lebesgue measures not exceeding dT must satisfy

$$|\mathcal{C}| \geq \frac{1}{d^{\lambda T}}. \quad (20)$$

Proof: We first observe the following relation between $\mathcal{P}(\lambda, T)$ and the λT -dimensional hypercube $[0, T]^{\times \{1, \dots, \lambda T\}}$: every point in $[0, T]^{\times \{1, \dots, \lambda T\}}$ can be mapped to a point pattern in $\mathcal{P}(\lambda, T)$ by collecting all its coordinates. (Note that a point with two or more identical coordinates is mapped to a point pattern with less than λT points.) We further observe that a point in $[0, T]^{\times \{1, \dots, \lambda T\}}$ is contained in $\mathcal{X}^{\times \{1, \dots, \lambda T\}}$ with $\mathcal{X} \subseteq [0, T]$ if, and only if, all its coordinates lie in \mathcal{X} , i.e., if the point pattern it maps to is covered by \mathcal{X} . The codebook \mathcal{C} being λ -covering is hence equivalent to

$$\bigcup_{\mathcal{X} \in \mathcal{C}} \mathcal{X}^{\times \{1, \dots, \lambda T\}} = [0, T]^{\times \{1, \dots, \lambda T\}}. \quad (21)$$

Since by assumption $\mu(\mathcal{X}) \leq dT$ for all $\mathcal{X} \in \mathcal{C}$, we know that the volume of each term inside the union in (21) is at most $(dT)^{\lambda T}$. On the other hand, the hypercube on the RHS of (21) has volume $T^{\lambda T}$. Hence, for (21) to hold, we need

$$|\mathcal{C}| \geq \frac{T^{\lambda T}}{(dT)^{\lambda T}} = \frac{1}{d^{\lambda T}}. \quad (22)$$

III. COVERING SOME POINTS

Let us now generalize the notion of “covering”. What if, instead of all the λT points, the codeword is only required to cover a proportion b of these points? We formulate this problem as follows.

Let the set of point patterns $\mathcal{P}(\lambda, T)$ be as in (3) and the codebook \mathcal{C} be as in (4). Fix $b \in [0, 1]$. We say a pattern \mathbf{x} is *b-covered* by a covering set \mathcal{X} if

$$|\mathbf{x} \cap \mathcal{X}^c| \leq (1 - b)\lambda T. \quad (23)$$

A codebook \mathcal{C} is said to be (λ, b) -covering if

$$\forall \mathbf{x} \in \mathcal{P}(\lambda, T), \quad \exists \mathcal{X} \in \mathcal{C}: \mathbf{x} \text{ is } b\text{-covered by } \mathcal{X}. \quad (24)$$

Denote

$$M(\lambda, b, d, T) = \min\{|\mathcal{C}|: \mathcal{C} \text{ is } (\lambda, b)\text{-covering}\}. \quad (25)$$

Generalizing Theorem 1, we have the following theorem:

Theorem 2: For any $\lambda > 0$, $b \in [0, 1]$ and $d \in (0, 1]$, at least

$$\lim_{T \rightarrow \infty} \frac{\log M(\lambda, b, d, T)}{T} = \begin{cases} \lambda D_2(b \| d), & b > d \\ 0, & b \leq d. \end{cases} \quad (26)$$

where $D_2(\cdot \| \cdot)$ is defined in (2).

Not surprisingly, when we let $b = 1$, the RHS of (26) becomes the RHS of (7). Also note that

$$D_2(b \| d) \leq b \log \frac{1}{d}, \quad b > d, \quad (27)$$

so covering a proportion b of every pattern containing λT points is easier than covering every pattern containing $b\lambda T$ points.

The proof of Theorem 2 is a generalization of that of Theorem 1, and is given below.

A. Proof of Theorem 2—Achievability

We partition the interval $[0, T]$ into parts of size Δ as in (8). We then randomly generate a size- M codebook \mathcal{C} by choosing each covering set as the union of dT/Δ parts picked uniformly in \mathcal{I}^Δ independently of the other covering sets in the codebook.

Fix a point pattern $\mathbf{x} \in \mathcal{P}(\lambda, T)$ with $|\mathbf{x}| = \lambda T$.⁴ We map \mathbf{x} to a *multiset* \mathbf{x}^Δ from \mathcal{I}^Δ in the following way: for each point in \mathbf{x} , add the (unique) part in \mathcal{I}^Δ which contains this point to \mathbf{x}^Δ . Hence each part in \mathcal{I}^Δ appears in \mathbf{x}^Δ as many times as the number of points in \mathbf{x} that lie in this part. (Note that, unlike in the covering-all-points case, now we must take into account the repetitions of parts in \mathbf{x}^Δ , namely, we must keep track of the number of points inside each occupied part. This is because missing at most k occupied parts does not imply missing at most k points in the pattern.) Now a covering set \mathcal{X} b -covers \mathbf{x} if, and only if, it b -covers \mathbf{x}^Δ , i.e., if it covers at least $b\lambda T$ elements of \mathbf{x}^Δ . Further, a codebook generated as above is b -covering if, and only if, it b -covers all size- λT multisets from \mathcal{I}^Δ .

We first consider the case where the elements of \mathbf{x}^Δ are all different from each other. In this case, the total number of covering sets that are of Lebesgue measure dT , that are unions of parts in \mathcal{I}^Δ , and that can b -cover \mathbf{x}^Δ is given by

$$\sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}. \quad (28)$$

We next show that, even if \mathbf{x} contains repetitions, the total number of different codewords that b -cover \mathbf{x}^Δ is

⁴It is sufficient to consider only those point patterns with *exactly* λT points, as a pattern with fewer points is always subset of a pattern with λT points.

$$\frac{1}{\lambda T} \sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}. \quad (29)$$

To this end, let \mathbf{y}^Δ be any size- λT subset of \mathcal{I}^Δ *without repetitions* which contains all the elements that are in \mathbf{x}^Δ . Consider any covering set $\tilde{\mathcal{X}}$ that covers i elements in \mathbf{y}^Δ . We produce another $(\lambda T - 1)$ covering sets in the following way: we fix the parts in $\tilde{\mathcal{X}}$ that lie outside \mathbf{y}^Δ , and cyclically shift the parts in $\tilde{\mathcal{X}}$ that lie inside \mathbf{y}^Δ . For example, for a covering set that contains (only) the first and third parts in \mathbf{y}^Δ , we first shift it to one which contains the second and fourth parts in \mathbf{y}^Δ , then shift it to one which contains the third and fifth parts in \mathbf{y}^Δ , etc. Each one of these λT covering sets (including the original one) covers i elements in \mathbf{y}^Δ . (Note, however, that some of these λT covering sets might be the same.) If we uniformly pick one from these λT covering sets, then the probability that each part in \mathbf{y}^Δ , as well as each part in \mathbf{x}^Δ , be covered is equal to $i/(\lambda T)$. This further implies that the expected number of covered parts in \mathbf{x}^Δ equals i . Hence at least one of these λT covering sets will cover no less than i parts in \mathbf{x}^Δ (repetitions counted). In particular, if $i \geq b\lambda T$, then this covering set b -covers \mathbf{x}^Δ . We have thus shown that at least a proportion of $1/(\lambda T)$ of all covering sets which b -cover \mathbf{y}^Δ will also b -cover \mathbf{x}^Δ . Since, by assumption, \mathbf{y}^Δ contains no repetitions, the number of covering sets that b -covers \mathbf{y}^Δ is given by (28). This implies that the number of different covering sets which cover \mathbf{x}^Δ is at least (29).

We now obtain that the probability that a uniformly chosen covering set will b -cover \mathbf{x}^Δ , where \mathbf{x}^Δ is any size- λT multiset from \mathcal{I}^Δ , is at least

$$\frac{\frac{1}{\lambda T} \sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}}{\binom{T/\Delta}{dT/\Delta}}. \quad (30)$$

Since the covering sets in \mathcal{C} are chosen independently, we further have that the probability that no element of \mathcal{C} will b -cover \mathbf{x}^Δ is at most

$$\left(1 - \frac{\frac{1}{\lambda T} \sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}}{\binom{T/\Delta}{dT/\Delta}} \right)^M$$

$$\begin{aligned}
&\leq \exp \left(-M \frac{\frac{1}{\lambda T} \sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}}{\binom{T/\Delta}{dT/\Delta}} \right) \quad (31) \\
&\leq \exp \left(-M \frac{\binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}}{\lambda T \binom{T/\Delta}{dT/\Delta}} \right), \\
&\quad i \in \{b\lambda T, \dots, \lambda T\}. \quad (32)
\end{aligned}$$

We next upper-bound the probability that there exists at least one size- λT multiset from \mathcal{I}^Δ that is *not* covered by any element of \mathcal{C} . To this end, we use the union bound and the fact that the total number of such multisets is

$$\binom{T/\Delta + \lambda T - 1}{\lambda T}. \quad (33)$$

These yield that the probability that there exists a multiset that is not covered by any element of \mathcal{C} is at most

$$\binom{T/\Delta + \lambda T - 1}{\lambda T} \exp \left(-M \frac{\binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}}{\lambda T \binom{T/\Delta}{dT/\Delta}} \right) \quad (34)$$

where i can be any integer between $b\lambda T$ and λT . If (34) is less than one, then there must exist at least one deterministic codebook that b -covers all multisets, and hence all size- λT point patterns. That (34) be less than one is equivalent to

$$M > \frac{\lambda T \binom{T/\Delta}{dT/\Delta} \ln \left(\frac{T/\Delta + \lambda T - 1}{\lambda T} \right)}{\binom{\lambda T}{i} \binom{T/\Delta - \lambda T}{dT/\Delta - i}}. \quad (35)$$

We again apply the bound (15) to obtain that (35) holds whenever

$$\begin{aligned}
\log M &> \log(\lambda T) + \frac{T}{\Delta} H_2(d) - \lambda T H_2(\eta) \\
&\quad - \left(\frac{T}{\Delta} - \lambda T \right) \cdot H_2 \left(\frac{d - \eta \lambda \Delta}{1 - \lambda \Delta} \right) \\
&\quad + \log(\lambda T + 1) + \log \left(\frac{T}{\Delta} - \lambda T + 1 \right) \\
&\quad + \log \left(H_2 \left(\frac{\lambda T}{T/\Delta + \lambda T - 1} \right) \cdot \ln 2 \right) \quad (36)
\end{aligned}$$

for some $\eta \in [b, 1]$. Letting T increase to infinity we obtain that a sequence of (λ, b) -covering codebooks of

size M_T exists if

$$\lim_{T \rightarrow \infty} \frac{\log M_T}{T} > \frac{1}{\Delta} \left(H_2(d) - \lambda \Delta H_2(\eta) - (1 - \lambda \Delta) H_2 \left(\frac{d - \eta \lambda \Delta}{1 - \lambda \Delta} \right) \right) \quad (37)$$

for some $\eta \in [b, 1]$. Since (37) is achievable for all positive Δ and all $\eta \in [b, 1]$, we conclude that a sequence of desired codebooks exists provided

$$\lim_{T \rightarrow \infty} \frac{\log M_T}{T} > \min_{\eta \in [b, 1]} \lim_{\Delta \downarrow 0} \frac{1}{\Delta} \left(H_2(d) - \lambda \Delta H_2(\eta) - (1 - \lambda \Delta) H_2 \left(\frac{d - \eta \lambda \Delta}{1 - \lambda \Delta} \right) \right) \quad (38)$$

$$= \min_{\eta \in [b, 1]} D_2(\eta \| d) \quad (39)$$

$$= \begin{cases} D_2(b \| d), & b > d, \\ 0, & b \leq d. \end{cases} \quad (40)$$

This completes the proof of the achievability part of Theorem 2.

Remark 2: Remark 1 in Section II-A applies here as well. Indeed, with probability $1 - \epsilon$ for any $\epsilon > 0$ the random choice will produce a (λ, b) -covering codebook without any loss in rate asymptotically as T goes to infinity.

Remark 3: The discrete-time problem encountered in our proof of the achievability part of Theorem 2 is similar to a problem of covering binary vectors by a constant-weight code with some Hamming-covering radius. More specifically, in the latter problem, we seek a constant-weight- dT/Δ code \mathcal{C} such that, for any weight- λT binary vector, there exists an element of \mathcal{C} within Hamming distance $(1 - 2b + d/\Delta)\lambda T$ from the vector. This problem has been considered in [13].

B. Proof of Theorem 2—Converse

Like in Section II-B, we map every point in $[0, T]^{\times \lambda T}$ to a point pattern on $[0, T]$. We consider only the points whose coordinates are all different. (The set of points with identical coordinates has Lebesgue measure zero, and can hence be ignored.) Such a point is mapped to a point pattern containing exactly λT points. If this point pattern is b -covered by \mathcal{X} , then at least $b\lambda T$ coordinates of the corresponding point must be in \mathcal{X} , which means the point must be contained in

$$\bigcup_{\substack{s \subseteq \{1, \dots, \lambda T\} : \\ |s| \geq b\lambda T}} \mathcal{X}^{\times s} \times [0, T]^{\times (\{1, \dots, \lambda T\} \setminus s)}. \quad (41)$$

The union (41) can also be written as the union of nonoverlapping regions as follows:

$$\bigcup_{i=b\lambda T}^{\lambda T} \bigcup_{\substack{\mathbf{s} \subseteq \{1, \dots, \lambda T\} \\ |\mathbf{s}|=i}} \mathcal{X}^{\times \mathbf{s}} \times ([0, T] \setminus \mathcal{X})^{\times (\{1, \dots, \lambda T\} \setminus \mathbf{s})}. \quad (42)$$

We now know that, for \mathcal{C} to be (λ, b) -covering, the union of (42) over all $\mathcal{X} \in \mathcal{C}$ must cover the whole hypercube of $[0, T]^{\times \{1, \dots, \lambda T\}}$ (except perhaps for the subset containing points with identical coordinates, whose Lebesgue measure is zero). Next note that the volume of (42) is at most

$$\sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} (dT)^i ((1-d)T)^{\lambda T-i}. \quad (43)$$

Therefore we have

$$|\mathcal{C}| \geq \frac{T^{\lambda T}}{\sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} (dT)^i ((1-d)T)^{\lambda T-i}} \quad (44)$$

$$\begin{aligned} &= \frac{1}{\sum_{i=b\lambda T}^{\lambda T} \binom{\lambda T}{i} d^i (1-d)^{\lambda T-i}} \\ &\geq \left(((1-b)\lambda T + 1) \cdot \max_{i \in \{b\lambda T, \dots, \lambda T\}} \binom{\lambda T}{i} d^i (1-d)^{\lambda T-i} \right)^{-1}. \end{aligned} \quad (45)$$

This further implies

$$\begin{aligned} \log |\mathcal{C}| &\geq -\log((1-b)\lambda T + 1) \\ &\quad + \min_{i \in \{b\lambda T, \dots, \lambda T\}} \left\{ -\log \binom{\lambda T}{i} + i \log \frac{1}{d} \right. \\ &\quad \left. + (\lambda T - i) \log \frac{1}{1-d} \right\} \end{aligned} \quad (47)$$

$$\begin{aligned} &\geq -\log((1-b)\lambda T + 1) \\ &\quad + \min_{i \in \{b\lambda T, \dots, \lambda T\}} \left\{ -\lambda T H_2 \left(\frac{i}{\lambda T} \right) + i \log \frac{1}{d} \right. \\ &\quad \left. + (\lambda T - i) \log \frac{1}{1-d} \right\} \end{aligned} \quad (48)$$

$$\begin{aligned} &= -\log((1-b)\lambda T + 1) \\ &\quad + \lambda T \cdot \min_{i \in \{b\lambda T, \dots, \lambda T\}} D_2 \left(\frac{i}{\lambda T} \parallel d \right), \end{aligned} \quad (49)$$

where in the second inequality we used (15). Next, note that

$$\begin{aligned} &\min_{i \in \{b\lambda T, \dots, \lambda T\}} D_2 \left(\frac{i}{\lambda T} \parallel d \right) \\ &\geq \min_{i \in [b\lambda T, \lambda T]} D_2 \left(\frac{i}{\lambda T} \parallel d \right) \end{aligned} \quad (50)$$

$$= \begin{cases} D_2(b \parallel d), & b > d, \\ 0, & b \leq d. \end{cases} \quad (51)$$

Combining (49) with (51) and letting T increase to infinity proves the converse part of Theorem 2.

REFERENCES

- [1] A. Lapidoth, A. Malär, and L. Wang, "Covering point patterns," in *Proc. IEEE Int. Symp. Inform. Theory*, Saint Petersburg, Russia, July 31–August 5 2011, pp. 149–153.
- [2] I. Rubin, "Information rates and data-compression schemes for Poisson processes," *IEEE Trans. Inform. Theory*, vol. 20, no. 2, pp. 200–210, Mar. 1974.
- [3] R. G. Gallager, "Basic limits on protocol information in data communication networks," *IEEE Trans. Inform. Theory*, vol. 22, no. 4, pp. 385–398, July 1976.
- [4] S. Verdú, "The exponential distribution in information theory," *Problems of Information Transmission*, vol. 31, no. 1, pp. 86–95, 1996.
- [5] A. S. Bedekar, "On the information about message arrival times required for in-order decoding," in *Proc. IEEE Int. Symp. Inform. Theory*, Washington DC, June 24–29, 2001.
- [6] T. P. Coleman, N. Kiyavash, and V. G. Subramanian, "The rate-distortion function of a Poisson process with a queueing distortion measure," in *Proceedings Data Compression Conference 2008*, Cliff Lodge, Snowbird, Utah, USA, Mar. 2008.
- [7] A. D. Wyner, "Capacity and error exponent for the direct detection photon channel — part I and II," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1462–1471, Nov. 1988.
- [8] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
- [9] T. J. Goblick, "Coding for a discrete information source with a distortion measure," Ph.D. dissertation, Massachusetts Institute of Technology, 1963.
- [10] A. D. Wyner, "Random packings and coverings of the unit n-sphere," *Bell System Techn. J.*, vol. 46, pp. 2111–2118, 1967.
- [11] T. Berger, *Rate Distortion Theory—a mathematical basis for data compression*. Prentice Hall, 1971.
- [12] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. North-Holland, Amsterdam, 1977.
- [13] T. Etzion, V. Wei, and Z. Zhang, "Bounds on the sizes of constant weight covering codes," *Design, Codes and Cryptography*, vol. 5, pp. 217–239, 1995.