2023 IEEE International Symposium on Information Theory (ISIT) | 978-1-6654-7554-9/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISIT54713.2023.10206920

Bounds on the Essential Covering Radius of Constrained Systems

Dor Elimelech
Electrical and Computer Engineering
Ben-Gurion University of the Negev,
Beer Sheva 8410501, Israel
doreli@post.bgu.ac.il

Tom Meyerovitch
Department of Mathematics
Ben-Gurion University of the Negev,
Beer Sheva 8410501, Israel
mtom@bgu.ac.il

Moshe Schwartz
Electrical and Computer Engineering
Ben-Gurion University of the Negev,
Beer Sheva 8410501, Israel
schwartz@ee.bgu.ac.il

***Abstract*—Motivated by applications for error-correcting constrained codes, we study the essential covering radius of constrained systems. In a recent work, the essential covering radius was suggested as new fundamental parameter of constrained systems that characterizes the error-correction capabilities of the quantized-constraint concatenation (QCC) scheme. We provide general efficiently computable upper-bounds on the essential covering radius using Markov chains and sliding-block codes, which in some cases, we show to be tight.**

I. Introduction

The *essential covering radius* was introduced in [6] as a fundamental parameter of constrained systems that reflects the error-correction capabilities of a new coding scheme for constrained systems, called *quantized-constraint concatenation* (QCC). Constrained error correction has a long history, and interest in it has been recently rekindled, motivated by applications for DNA-storage (see for example [4], [5], [7]–[9], [11]). The QCC scheme was suggested in [6] as a general framework for implementing error-correction in constrained codes, capable of correcting a $\Theta(n)$ channel errors (where n is the code length), improving upon the $O(\sqrt{n})$ possible by known schemes.

The main idea of QCC is to employ a process of embedding the codewords of an error-correcting code in a constrained system as a (noisy, irreversible) quantization process. This is in contrast to traditional methods, such as concatenation and reverse concatenation, where the encoding into the constrained system is reversible. The error-correction capabilities of QCC are therefore determined by the amount of quantization noise introduced in the embedding process of the codewords in the constrained system. The amount of noise is upper bounded by the minimal number of coordinates that need to be changed in an arbitrary word in order to transform it into a constrained word. In coding-theoretic terminology, this quantity is the covering radius of the code composed of all constrained words of a fixed length. The essential covering radius is a parameter generalizing this concept of quantization noise to the setting where the words to be quantized are generated with respect to a probabilistic model and the block length increases to infinity. A precise expression for the essential covering radius was found in [6], which we cite here in Theorem 7. This expression, however, does not immediately give an efficient computation procedure.

The goal of this paper is to efficiently compute upper bounds on the essential covering radius of constrained systems. We first propose a general efficiently computable upper bound via a solution of a linear program. We show a non-trivial example in which our bound is in fact tight. We also provide bounds using sliding-block codes. We show that in the primitive case, these bounds asymptotically attain the essential covering radius. Due to space limitations, the proofs are omitted, and may be found in [6].

II. Constrained systems and The Essential Covering Radius

Throughout this paper we shall use lower-case letters, x, to denote scalars and symbols, overlined lower-case letters, $\overline{x}$, to denote finite-length words, and bold lower-case letters, $\mathbf{x}$, to denote bi-infinite sequences. We use upper-case letters, X, for constrained systems. For a bi-infinite sequence $\mathbf{x} = \dots, \mathbf{x}_{-1}, \mathbf{x}_0, \mathbf{x}_1, \dots$ and $n \leqslant m$ we denote the subword $\mathbf{x}_n^m \triangleq \mathbf{x}_n, \dots, \mathbf{x}_m$ (and similarly $\overline{x}_n^m$ for finite words). We use Σ to denote a finite alphabet, and $[n] \triangleq \{0, 1, \dots, n-1\}$.

The set of words of length n over Σ is denoted by Σ^n. If $\overline{u} \in \Sigma^n$, we shall index its letters by $[n]$, i.e., $\overline{u} = \overline{u}_0, \overline{u}_1, \dots, \overline{u}_{n-1}$. For any $\overline{v}, \overline{u} \in \Sigma^n$, we define the Hamming distance as $d(\overline{u}, \overline{v}) \triangleq |\{i \in [n] \mid \overline{u}_i \neq \overline{v}_i\}|$. The ball of radius r (with respect to the Hamming distance) centered in $\overline{x}$ is denoted by $B_r(\overline{x})$. The covering radius of a code $C \subseteq \Sigma^n$ is the minimal integer r such that the union of balls of radius r, centered at the codewords of C, covers the whole space. That is,

$$R(C) \triangleq \min\left\{ r \in \mathbb{N} \cup \{0\} \;\middle|\; \bigcup_{\overline{c} \in C} B_r(\overline{c}) = \Sigma^n \right\}.$$

We begin by discussing constrained systems. These are often studied in the framework of symbolic dynamics (see for example [10], [12]). In a typical (one dimensional) setting we have a finite alphabet Σ, and the space of bi-infinite sequences of Σ, denoted $\Sigma^{\mathbb{Z}}$, is considered as a compact metrizable topological space, equipped with the product topology (where Σ has the discrete topology). The dynamics on the system $\Sigma^{\mathbb{Z}}$ are realized by the shift transformation, $T : \Sigma^{\mathbb{Z}} \to \Sigma^{\mathbb{Z}}$, defined by $(T\mathbf{x})_n \triangleq \mathbf{x}_{n+1}$. For a finite word $\overline{x} \in \Sigma^n$ we let $[\overline{x}]$ denote the cylinder set defined by $\overline{x}$, which is

$$[\overline{x}] \triangleq \left\{ \mathbf{x} \in \Sigma^{\mathbb{Z}} \;\middle|\; \mathbf{x}_0^{n-1} = \overline{x} \right\}. \tag{1}$$

A subshift (or shift space) $X \subseteq \Sigma^{\mathbb{Z}}$ is a compact subspace, which is invariant under the shift transformation. For a subshift X, the language of X is the set of all finite words that appear as subwords of some element in X. That is

$$\mathcal{B}(X) \triangleq \left\{ \overline{x} = (x_0 \dots x_k) \,\middle|\, \substack{\exists \mathbf{x} \in X, n \in \mathbb{Z} \text{ such that} \\ \mathbf{x}_n^{n+k} = \overline{x}, k \in \mathbb{N} \cup \{0\}} \right\}.$$

The set of words of length n in the language is denoted by $\mathcal{B}_n(X) \triangleq \mathcal{B}(X) \cap \Sigma^n$. In our setting, constrained systems are those shift spaces which can be realized by walks on some labeled graph.

Definition 1 *A shift space $X \subseteq \Sigma^{\mathbb{Z}}$ is called a constrained system (or a sofic shift) if there exists a finite directed graph $G = (V, E)$ and a labeling function $L : E \to \Sigma$ such that*

$$X = X_G \triangleq \left\{ (L(e_i))_{i \in \mathbb{Z}} \,\middle|\, \substack{(e_i)_{i \in \mathbb{Z}} \text{ is a bi-infinite} \\ \text{directed path in } G} \right\}.$$

A labeled graph $G = (V, E, L)$ is called irreducible if any two vertices are connected by a directed path. An irreducible graph is called primitive if the greatest common divisor of all cycle lengths is 1. It is well known (e.g., see [10, Theorem 4.5.8]) that an irreducible graph is primitive if and only if there exists $n \in \mathbb{N}$ such that for any two vertices $v, v' \in V$ there exists a directed path of length n from v to v'.

Definition 2 *A constrained system $X \subseteq \Sigma^{\mathbb{Z}}$ is called irreducible (respectively: primitive), if there exists an irreducible (respectively: primitive) labeled graph G such that $X = X_G$.*

We now turn to discuss the essential covering radius. Motivated by the QCC scheme, we are interested in the following question: Given an arbitrarily small $\varepsilon > 0$, what is the the smallest r such that $(1-\varepsilon)$-fraction the words in the space can be quantized to the constraint system by changing at most r symbols? Roughly speaking, the asymptotic answer to this questions is the essential covering radius of the system. Aiming toward an exact definition we state some basic definitions and well known results from ergodic theory. For any finite alphabet Σ, we consider $\Sigma^{\mathbb{Z}}$ as a measurable space, together with the Borel Σ-algebra induced by the product topology on $\Sigma^{\mathbb{Z}}$. Similarly, any subshift $Y \subseteq \Sigma^{\mathbb{Z}}$ is considered as a measurable space.

Definition 3 (Invariant and ergodic measures) *Let $Y \subseteq \Sigma^{\mathbb{Z}}$ be a subshift. A probability measure μ on Y is called* shift invariant *if $\mu(T^{-1}B) = \mu(B)$ for any measurable set B. A shift-invariant measure μ is further said to be* ergodic *if $T^{-1}B = B$ implies $\mu(B) = 0$ or $\mu(Y \setminus B) = 0$. The set of shift-invariant probability measures on Y is denoted by $M(Y)$, and the set of ergodic measures in $M(Y)$ is denoted by $M_{\mathcal{E}}(Y)$.*

For a measure $\mu \in M(Y)$ we denote by μ_n the marginal measure of μ on the first n coordinates, which is a probability measure on Σ^n. To avoid cumbersome notation, throughout this work we shall use $\mathbb{P}_\mu[A]$ in order to denote the measure $\mu(A)$, and $\mathbf{Y}$ for a random bi-infinite sequence on Y. Throughout this article we use bold upper-case letters for bi-infinite sequences of random variables, not to be confused with non-bold capital letters used to denote constrained systems.

Definition 4 *For any real $\varepsilon > 0$, two sets $A, C \subseteq \Sigma^n$, and η, a probability measure on A, we define $R_\varepsilon(C, A, \eta)$ to be*

$$\min \left\{ r \in \mathbb{N} \cup \{0\} \,\middle|\, \eta \left(A \cap \left(\bigcup_{\overline{x} \in C} B_r(\overline{x}) \right) \right) \geqslant 1 - \varepsilon \right\}.$$

We remark that when η is the uniform measure on A, $R_\varepsilon(C, A, \eta)$ is the ε-covering radius of A, namely the smallest r such that at least $(1-\epsilon)$-fraction of the words in C are at distance at most r from A.

Definition 5 *Let $X, Y \subseteq \Sigma^{\mathbb{Z}}$ be constrained systems, and $\mu \in M_{\mathcal{E}}(Y)$ be an ergodic measure. We define the ε-*covering radius *of X with respect to (Y, μ) by*

$$R_\varepsilon(X, Y, \mu) \triangleq \liminf_{n \to \infty} \frac{R(\mathcal{B}_n(X), \mathcal{B}_n(Y), \mu_n)}{n},$$

and the essential covering radius *by*

$$R_0(X, Y, \mu) \triangleq \lim_{\varepsilon \to 0} R_\varepsilon(X, Y, \mu),$$

where the limit exists due to the monotonicity of $R_\varepsilon(X, Y, \mu)$ in ε.

We remark that a typical case of particular interest is where $\mu = \mu^{\mathsf{u}}$ is the uniform Bernoulli measure, where μ_n is the uniform measure on Σ^n. In [6] an ergodic-theoretic characterization of the essential covering radius was established. This characterization uses an object from ergodic theory called an invariant extension.

Definition 6 *Let $X, Y \subseteq \Sigma^{\mathbb{Z}}$ be shift spaces, we consider $X \times Y$ as shift space, with the left shift acting by $T(\mathbf{x}, \mathbf{y}) = (T\mathbf{x}, T\mathbf{y})$. For an ergodic measure $\mu \in M_{\mathcal{E}}(Y)$, an extension of μ to $X \times Y$ is a shift-invariant measure $\nu \in M(X \times Y)$, whose Y-marginal is μ. Namely, ν satisfies that for any measurable $A \subseteq Y$, $\nu(X \times A) = \mu(A)$. An ergodic extension of μ to $X \times Y$ is an extension ν of μ that is furthermore ergodic, that is $\nu \in in M_{\mathcal{E}}(X \times Y)$. We let $M(X, Y, \mu)$ denote the set of all extensions of μ, and $M_{\mathcal{E}}(X, Y, \mu)$ denote the set of all ergodic extensions.*

The equivalent characterization of the essential covering radius is given by a minimization problem over invariant extensions.

Theorem 7 ([6]) *Let $X, Y \subseteq \Sigma^{\mathbb{Z}}$ be constrained systems, and let $\mu \in M_{\mathcal{E}}(Y)$ be an ergodic measure. Then*

$$\begin{aligned} R_0(X, Y, \mu) &= \inf\{\mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] \,|\, \nu \in M_{\mathcal{E}}(X, Y, \mu)\} \\ &= \inf\{\mathbb{P}_\nu[\mathbf{X}_0 \neq \mathbf{Y}_0] \,|\, \nu \in M(X, Y, \mu)\}. \end{aligned}$$

III. Upper Bounds On The Essential Covering Radius

While Theorem 7 gives an exact expression for the essential covering radius, the minimization problem involved is hard to solve. In general, by Theorem 7, any extension in $M(X, Y, \mu)$ induces an upper-bound on the essential covering radius. In this section we shall present two different approaches for constructing extensions for general constrained systems, thus providing upper bounds on the essential covering radius. The first approach, using Markov chains, provides an upper bound which is efficiently computable as the solution of a linear program. An alternative method for constructing extensions is by sliding-block codes. In that case, we prove that if X is primitive, the essential covering radius can be approximated by increasing the block size in such functions.

A. Markov Chains

We consider the scenario where X and Y are constrained systems generated by labeled graphs $G_X = (V_X, E_X, L_X)$ and $G_Y = (V_Y, E_Y, L_Y)$ respectively.

For an edge $e \in E_X$ or $e \in E_Y$ we denote by $\sigma(e)$ and $\tau(v)$ the source and target of the edge respectively. For simplicity, we generally assume there are no parallel edges with the same label. We focus on the case where the measure $\mu \in M_{\mathcal{E}}(Y)$ is generated by some Markov chain on the graph G_Y. This framework includes the special case where $Y = \Sigma^{\mathbb{Z}}$ and $\mu = \mu^{\mathsf{u}}$ is the uniform Bernoulli measure.

Definition 8 *Let $G = (V, E)$ be a finite directed graph. A stationary Markov chain on G is a pair (π, Q), where π is a probability measure on V and Q is a function from V to the space of probability measures on E that sends $v \in V$ to a probability measure $Q(\cdot|v)$ on E such that for every $v \in V$,*

$$\sum_{\substack{e\in E\\ \sigma(e)=v}} Q(e|v) = 1,$$

and so that for every $v \in V$ we have:

$$\pi(v) = \sum_{\substack{e\in E\\ \tau(e)=v}} \pi(\sigma(e))Q(e|\sigma(e)).$$

Note that for any Markov chain on (π, Q) on $G = (V, E)$, $Q(e|v) > 0$ implies that $\sigma(e) = v$ so we can conveniently write $Q(e)$ as an abbreviation for $Q(e|\sigma(e))$. In the case where G is a simple graph (i.e., without parallel edges), for any edge $e = (u, v) \in E$ we use the notation $Q(v|u)$ for $Q(e)$. Also, when G is a simple graph, Q may be identified with a $|V| \times |V|$ stochastic matrix (often called the transition matrix), for which π is a left eigenvector with eigenvalue 1.

There is a one-to-one correspondence between Markov chains on $G = (V, E)$ and probability measures on E that satisfy the condition

$$\sum_{\substack{e\in E\\ \sigma(e)=v}} P(e) = \sum_{\substack{e\in E\\ \tau(e)=v}} P(e).$$

Indeed, such a probability measure P corresponds to a stationary Markov chain (π, Q), where

$$\pi(v) \triangleq \sum_{\substack{e\in E\\ \sigma(e)=v}} P(e) = \sum_{\substack{e\in E\\ \tau(e)=v}} P(e),$$

and

$$Q(e|\sigma(v)) \triangleq \frac{P(e)}{\pi(\sigma v)}.$$

By abuse of notation, we denote $P = (\pi, Q)$. We assume the Markov chain does not contain degenerate vertices, i.e., $\pi(v) > 0$ for all $v \in V$. Any stationary Markov chain P induces an invariant measure $\hat{P}$ on the space of bi-infinite paths on G by

$$\hat{P}([(e_0, e_1, \dots, e_{n-1})]) = P(e_0) \prod_{i=1}^{n-1} Q(e_i).$$

for any cylinder set $[(e_0, \dots, e_{n-1})]$ corresponding to a finite path $(e_0, \dots, e_{n-1})$. We call $\hat{P}$ the stationary Markov process on G, induced by P.

If $G = G_Y$ generates the constrained system Y by the labeling function L_Y, then $P = (\pi, Q)$ induces an invariant probability measure on Y, which is the pushforward measure of $\hat{P}$ via the labeling function, i.e., for a cylinder set $[\overline{y}]$,

$$\mu_P([\overline{y}]) = \sum_{\substack{\overline{e} \text{ path in } G\\ L(\overline{e})=\overline{y}}} \pi(\sigma(e_0)) \prod_{i=0}^{|\overline{y}|-1} Q(e_i).$$

We note that $\mathbf{Y}$ is a hidden Markov process with respect to μ_P, and refer to the measure μ_P as above as *the hidden Markov measure* induced by P via the labeling function L.

Assume that $X, Y \subseteq \Sigma^{\mathbb{Z}}$ are irreducible constrained systems given by labeled graphs G_X and G_Y respectively, and assume that $\mu = \mu_{P_Y} \in M_{\mathcal{E}}(Y)$ is a measure on Y, induced by P_Y, a stationary Markov Chain on G_Y. We consider the strong product graph of G_X and G_Y given by $G_{X\times Y} = (V_{X\times Y}, E_{X\times Y}, (L_X, L_Y))$ where $V_{X\times Y} \triangleq V_X \times V_Y$, $E_{X\times Y} \triangleq E_X \times E_Y$ with $\sigma(e_x, e_y) = (\sigma(e_x), \sigma(e_y))$, $\tau(e_x, e_y) = (\tau(e_x), \tau(e_y))$ and labeling function $L_{X\times Y}$ given by:

$$L_{X\times Y}(e_x, e_y) = (L_X(e_x), L_Y(e_y)).$$

We note that a stationary Markov chain P on $G_{X\times Y}$ naturally defines a stationary Markov process $\hat{P}$ on $G_{X\times Y}$, which induces the hidden Markov measure ν_P on $X \times Y$ by the labeling function $L_{X\times Y}$.

We now give an upper bound on $R_0(X, Y, \mu)$, formulated as an optimization problem over stationary Markov chains on the product graph $G_{X\times Y}$.

Theorem 9 *Let $X, Y \subseteq \Sigma^{\mathbb{Z}}$ be shift spaces defined by the labeled graphs G_X and G_Y respectively, and let μ be the hidden Markov measure on Y induced by a the stationary Markov chain on P_Y on G_Y. Then $R_0(X, Y, \mu)$ is bounded* *above by the solution to the following linear-programming problem, denoted by $MB(G_X, G_Y, P_Y)$:*

$$\underset{P \in \mathbb{R}^{E_{X\times Y}}}{\text{minimize}} \quad \sum_{\substack{e\in E_{X\times Y}\\ L_X(e)\neq L_Y(e)}} P(e)$$

subject to

$$P(e) \geqslant 0, \qquad \forall e \in E_{X\times Y},$$

$$\sum_{e\in E_{X\times Y}} P(e) = 1,$$

$$\sum_{\substack{e'\in E_{X\times Y}\\ e'_y=e}} P(e') = P_Y(e), \qquad \forall e \in E_Y,$$

$$\sum_{\substack{e\in E_{X\times Y}\\ \sigma(e)=v}} P(e) = \sum_{\substack{e\in E_{X\times Y}\\ \tau(e)=v}} P(e), \qquad \forall v \in V_{X\times Y},$$

$$\sum_{\substack{e'\in E_{X\times Y}\\ e'_y=e_y\\ \sigma(e'_x)=\sigma(e_x)}} P(e') = Q_Y(e_y) \sum_{\substack{e'\in E_{X\times Y}\\ \sigma(e'_y)=\sigma(e_y)\\ \sigma(e'_x)=\sigma(e_x)}} P(e') \quad \forall e \in E_{X\times Y}.$$

The main tool in the proof of Theorem 9 is the following:

Proposition 10 *Any $P \in \mathbb{R}^{E_{X\times Y}}$ that satisfies the linear constraints given in the statement of Theorem 9 corresponds to a Markov chain on $G_{X\times Y}$ that induces an extension in $M(X, Y, \mu_{P_Y})$ via the pushforward map given by labeling function $L_{X\times Y}$.*

The proof of Theorem 9 now follows immediately: For any Markov chain P on $G_{X\times Y}$ satisfying the constraints of the presented optimization problem let ν_P denote the induced measure on $X \times Y$. By Proposition 10, $\nu_P \in M(X, Y, \mu_{P_Y})$. Using Theorem 7 we have

$$R_0(X, Y, \mu_{P_Y}) \leqslant \sum_{\substack{e\in E_{X\times Y}\\ L_X(e)\neq L_Y(e)}} P(e).$$

Example 1 *Consider $Y = [2]^{\mathbb{Z}}$ and $X = X_{0,k}$, the $(0,k)$-RLL system (no $k+1$ run of consecutive zeros constraint), generated by the labeled graphs G_X and G_Y shown in Figure 1. Let μ^{u} be the uniform Bernoulli measure on Y, which is generated by the stationary Markov chain P_Y, the uniform measure on E_Y. The product graph, $G_{X\times Y}$ (shown in Figure 2), is therefore a "doubled" version of the graph G_X.*

We consider the Markov measure P, defined by the edge probabilities given in Figure 2. For an appropriate choice of α, P is indeed a stationary Markov chain satisfying the conditions of Theorem 9. First, in order to get a probability measure on edges we require

$$1 = \sum_{e\in G_{X\times Y}} P(e) = 2\alpha(2^k + 2^{k-1} \cdots + 1) = 2\alpha(2^{k+1} - 1),$$

which implies that $\alpha = (2(2^{k+1} - 1))^{-1}$. We observe that for the j-th state in $V_{X\times Y}$,

$$\sum_{\sigma(e)=j} P(e) = \sum_{\tau(e)=j} P(e) = \alpha \cdot 2^{k-j},$$

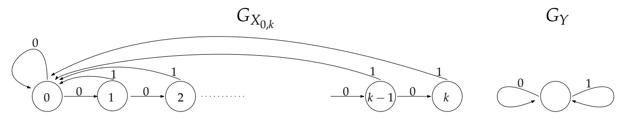

Fig. 1. Labeled graphs generating the shift spaces $X = X_{0,k}$ (the $(0,k)$-RLL shift), and $Y = [2]^{\mathbb{Z}}$.

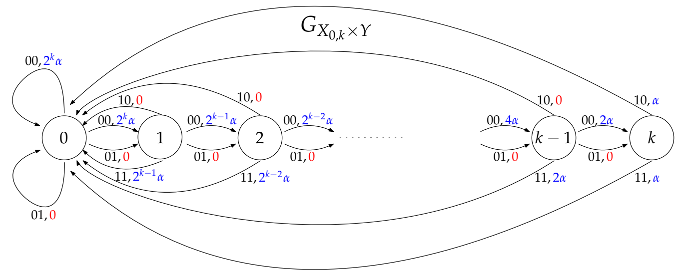

Fig. 2. The product graph $G_{X\times Y}$ for the graphs G_X and G_Y from Figure 1. Each edge is given a two-bit label, xy, corresponding to the label x from G_X and the label y from G_Y. A stationary Markov chain achieving the bound $MB(G_X, G_Y, P_Y)$ is shown by writing $P(e)$ after the label on each edge. Edges with $P(e) = 0$ are marked in red, and positive probabilities $P(e) > 0$ are written in blue.

which implies that P is indeed stationary. We also observe that for any edge with $L_Y(e) = 0$ there is a corresponding edge e' with $L_Y(e') = 1$ such that $P(e) = P(e')$. This shows that the marginal of P on G_Y is indeed $P_Y = \mu^{\mathsf{u}}$. We further observe that for any edge $e \in E_{X\times Y}$, if $\sigma(e_x) = j \in V_X$ we have

$$\sum_{\substack{e'\in E_{X\times Y}\\ e'_y=e_y\\ \sigma(e'_x)=j}} P(e') = \alpha 2^{k-j-1} = \frac{1}{2}\alpha 2^{k-j} = Q_Y(e_y) \sum_{\substack{e'\in E_{X\times Y}\\ \sigma(e'_y)=\sigma(e_y)\\ \sigma(e'_x)=j}} P(e').$$

We now get,

$$R_0(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \leqslant \sum_{\substack{e\in G_{X\times Y}\\ L_X(e)\neq L_Y(e)}} P(e) = \alpha = \frac{1}{2(2^{k+1} - 1)}.$$

In [6] it was shown that $R_0(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) = \frac{1}{2(2^{k+1}-1)}$, which implies that P attains the minimal value for the linear-programming problem given in the statement of Theorem 9, and in particular in this case, the upper bound from Theorem 9 is tight.

Example 2 *Let $X = X_{d,\infty}$ be the (d,∞)-RLL system, defined by the constraint of having a run of at least d zeroes between any two consecutive ones. Equivalently, $X_{d,\infty}$ is defined by G_X presented in Figure 3. Let $Y = [2]^{\mathbb{Z}}$ and μ^{u} be as in Example 1. The product graph $G_{X\times Y}$ is shown in Figure 4. We consider the Markov measure P, defined by the edge probabilities given in Figure 4. For $\alpha = (2(d+2))^{-1}$ we have*

$$1 = \sum_{e\in G_{X\times Y}} P(e) = 2\alpha(d+2),$$

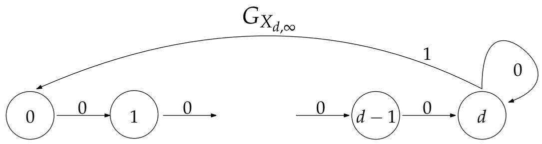

Fig. 3. A labeled graph generating the constrained system $X = X_{d,\infty}$ (the (d, ∞)-RLL shift).

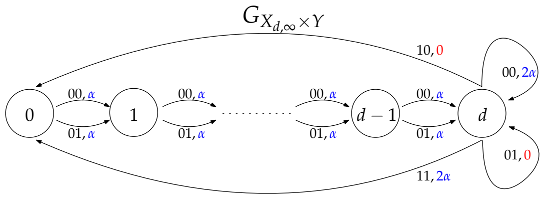

Fig. 4. The product graph $G_{X \times Y}$ for the graphs $G_{X_{d,\infty}}$ and G_Y from Figure 3 and Figure 1 respectively. Each edge is given a two-bit label, xy, corresponding to the label x from $G_{X_{d,\infty}}$ and the label y from G_Y. A stationary Markov chain achieving the bound $MB(G_X, G_Y, P_Y)$ is shown by writing $P(e)$ after the label on each edge. Edges with $P(e) = 0$ are marked in red, and positive probabilities $P(e) > 0$ are written in blue.

and so P is indeed a stationary Markov chain satisfying the conditions of Theorem 9. We now compute the upper-bound:

$$R_0(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \leqslant \sum_{\substack{e \in G_{X\times Y} \\ L_X(e) \neq L_Y(e)}} P(e) = \alpha \cdot d = \frac{d}{2(d+2)}. \quad (2)$$

For a lower bound, $R_0(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \geqslant \frac{1}{2} - \frac{1}{d+1}$, with proof in the appendix. Combining the lower and upper bounds we get,

$$\frac{1}{2} - \frac{1}{d+1} \leqslant R_0(X_{d,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) \leqslant \frac{1}{2} - \frac{1}{d+2}.$$

We note that when $d = 1$, the system $X_{1,\infty}$ is isomorphic to $X_{0,1}$, by complementing all the bits. Thus

$$R_0(X_{1,\infty}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) = R_0(X_{0,1}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}) = \frac{1}{6},$$

and in particular, the bound (2) is tight.

B. Sliding Block Codes

We now present an alternative approach for constructing extensions by using sliding-block codes. For $X, Y \subseteq \Sigma^{\mathbb{Z}}$ and $\mu \in M_{\mathcal{E}}(Y)$, given a measurable function $g : Y \to X$ which commutes with the shift transformation, the map $(g, \mathrm{Id}) : Y \to X \times Y$ defines an extension ν_g in $M(X, Y, \mu)$ by the pushforward of μ via (g, Id). That is $\nu_g(A_X \times A_Y) \triangleq \mu(A_Y \cap g^{-1}(A_X))$. We call such a function a *stationary coding function* from Y to X. By Theorem 7, for any stationary coding function g we have $R_0(X, Y, \mu) \leqslant \mathbb{P}_{\nu_g}[\mathbf{X}_0 \neq \mathbf{Y}_0]$.

Sliding-block codes are an important family of stationary coding functions. These are of particular interest to us since they provide a rich family of functions, easily described by a local rule. The properties and constructions of sliding-block codes have been extensively studied in the literature (for example, see [1]–[3]). The goal of this section is to explicitly describe the bound obtained from a sliding-block code function and to give sufficient conditions under which the essential covering radius can be approximated using extensions constructed by sliding-block codes.

Definition 11 *Let $X, Y \subseteq \Sigma^{\mathbb{Z}}$ be shift spaces. A function $\hat{f} : Y \to X$ is called a* sliding-block code *if there exist $N \in \mathbb{N}$ and a function $f : \mathcal{B}_{2N+1}(Y) \to \Sigma$ such that for all $\mathbf{y} \in Y$ and all $i \in \mathbb{Z}$, $\hat{f}(\mathbf{y})_i = f(\mathbf{y}_{i-N}^{i+N})$. In that case, $\hat{f}$ is said to be a sliding-block code of block length N.*

Let $\hat{f} : Y \to X$ be a sliding-block code function defined by a local function $f : \mathcal{B}_{2N+1}(Y) \to \Sigma$, and let $\mu \in M_{\mathcal{E}}(Y)$ be an ergodic measure. We denote the extension obtained from $\hat{f}$ by ν_f. The quantity $\mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0]$ is now easily computable:

$$\mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0] = \mathbb{P}_{\mu}\left[f(\mathbf{Y}_{-N}^{N}) \neq \mathbf{Y}_0\right] = \sum_{\substack{\overline{y} \in \mathcal{B}_{2N+1}(Y) \\ f(\overline{y}) \neq \overline{y}_N}} \mu([\overline{y}]).$$

The main result of the section is the following:

Theorem 12 *Let $X, Y \subseteq \Sigma^{\mathbb{Z}}$ be constrained systems such that X is primitive and $\mu \in M_{\mathcal{E}}(Y)$ is an aperiodic ergodic measure. Then for any $\varepsilon > 0$ there exists a sufficiently large N and a sliding-block code $\hat{f}$ of length N such that*

$$\mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0] - \varepsilon \leqslant R_0(X, Y, \mu) \leqslant \mathbb{P}_{\nu_f}[\mathbf{X}_0 \neq \mathbf{Y}_0].$$

Example 3 *Let $X_{0,k} \subseteq [2]^{\mathbb{Z}}$ denote the $(0, k)$-RLL shift as in Example 1. Let $\overline{y} \in [2]^n$ be a finite binary word. We define $c(\overline{y})$ to be the length of longest zero suffix of $\overline{y}$, i.e.,*

$$c(\overline{y}) \triangleq \max\left\{ i \,\middle|\, \overline{y} = \overline{y}_0^{n-i-1} \overline{0}_i \right\}.$$

We fix $N \in \mathbb{N}$ and consider the map $f^{(N)} : [2]^{\mathbb{Z}} \to X_{0,k}$

$$f^{(N)}(\mathbf{y})_m = \begin{cases} 1 & c\left(\mathbf{y}_{m-(N(k+1)-1)}^{m-1}\right) \equiv k \pmod{k+1}, \\ \mathbf{y}_m & \text{otherwise.} \end{cases}$$

Clearly, $\mathrm{Im}(f) \subseteq X_{0,k}$ since no run of $k+1$ zeroes may appear in $f^{(N)}(\mathbf{y})$. We note that the map $(f^{(N)}, \mathrm{Id}) : [2]^{\mathbb{Z}} \to X_{0,k} \times [2]^{\mathbb{Z}}$ is a sliding-block code. Let μ^{u} be the uniform measure over $[2]^{\mathbb{Z}}$. We note that

$$\begin{aligned}
&\mathbb{P}_{\nu_{f^{(N)}}}[\mathbf{X}_0 \neq \mathbf{Y}_0] \\
&= \mathbb{P}_{\mu^{\mathsf{u}}}\left[c\left(\mathbf{Y}_{-(N(k+1)-1)}^{-1}\right) \equiv k \pmod{k+1} \text{ and } \mathbf{Y}_0 = 0\right] \\
&= \sum_{i=0}^{N-1} \mathbb{P}_{\mu^{\mathsf{u}}}\left[c\left(\mathbf{Y}_{-(N(k+1)-1)}^{-1}\right) = i(k+1)+k \text{ and } \mathbf{Y}_0 = 0\right] \\
&= \mathbb{P}_{\mu^{\mathsf{u}}}\left[\mathbf{Y}_{-(N(k+1)-1)}^{0} = \overline{0}\right] + \sum_{i=1}^{N-1} \mathbb{P}_{\mu^{\mathsf{u}}}\left[\mathbf{Y}_{-i(k+1)}^{0} = 1\overline{0}\right] \\
&= \frac{1}{2^{N(k+1)}} + \frac{1}{2}\sum_{i=1}^{N-1} \frac{1}{2^{i(k+1)}}.
\end{aligned}$$

Taking $N \to \infty$ we indeed approach the known essential covering radius,

$$\lim_{N\to\infty} \mathbb{P}_{\nu_{f^{(N)}}}[\mathbf{X}_0 \neq \mathbf{Y}_0] = \frac{1}{2(2^{k+1}-1)} = R_0(X_{0,k}, [2]^{\mathbb{Z}}, \mu^{\mathsf{u}}).$$

References

[1] R. Adler, D. Coppersmith, and M. Hassner, "Algorithms for sliding block codes-an application of symbolic dynamics to information theory," *IEEE Transactions on Information Theory*, vol. 29, no. 1, pp. 5–22, 1983.

[2] J. Ashley and B. Marcus, "Canonical encoders for sliding block decoders," *SIAM Journal on Discrete Mathematics*, vol. 8, no. 4, pp. 555–605, 1995.

[3] J. J. Ashley, R. Karabed, and P. H. Siegel, "Complexity and sliding-block decodability," *IEEE Transactions on Information Theory*, vol. 42, no. 6, pp. 1925–1947, 1996.

[4] W. Bliss, "Circuitry for performing error correction calculations on baseband encoded data to eliminate error propagation," *IBM Tech. Discl. Bul.*, vol. 23, pp. 4633–4634, 1981.

[5] Y. M. Chee, H. M. Kiah, and H. Wei, "Efficient and explicit balanced primer codes," *IEEE Transactions on Information Theory*, vol. 66, no. 9, pp. 5344–5357, 2020.

[6] D. Elimelech, T. Meyerovitch, and M. Schwartz, "Quantized-constraint concatenation and the covering radius of constrained systems," *arXiv preprint arXiv:2302.02363*, 2023.

[7] J. L. Fan and A. R. Calderbank, "A modified concatenated coding scheme, with applications to magnetic data storage," *IEEE Transactions on Information Theory*, vol. 44, no. 4, pp. 1565–1574, 1998.

[8] R. Gabrys, P. H. Siegel, and E. Yaakobi, "Segmented reverse concatenation: a new approach to constrained ecc," in *2020 International Symposium on Information Theory and Its Applications (ISITA)*. IEEE, 2020, pp. 254–258.

[9] K. A. S. Immink, "A practical method for approaching the channel capacity of constrained channels," *IEEE Transactions on Information Theory*, vol. 43, no. 5, pp. 1389–1399, 1997.

[10] D. Lind and B. Marcus, *An introduction to symbolic dynamics and coding*. Cambridge university press, 2021.

[11] M. Mansuripur, "Enumerative modulation coding with arbitrary constraints and postmodulation error correction coding for data storage systems," in *Optical Data Storage'91*, vol. 1499. SPIE, 1991, pp. 72–86.

[12] B. H. Marcus, R. M. Roth, and P. H. Siegel, "An introduction to coding for constrained systems," *Lecture notes*, 2001.