# Interleaving Schemes on Circulant Graphs with Two Offsets

Aleksandrs Slivkins* Jehoshua Bruck†

November 2002
Revised: September 2003, July 2006, July 2007, October 2008.

**Abstract**

Interleaving is used for error-correcting on a bursty noisy channel. Given a graph $G$ describing the topology of the channel, we label the vertices of $G$ so that each label-set is sufficiently sparse. The interleaving scheme corrects for any error burst of size at most $t$; it is a labeling where the distance between any two vertices in the same label-set is at least $t$.

We consider interleaving schemes on infinite circulant graphs with two offsets 1 and $d$. In such graph the vertices are integers; edge $ij$ exists if and only if $|i-j| \in \{1, d\}$. Our goal is to minimize the number of labels used.

Our constructions are covers of the graph by the minimal number of translates of some label-set $S$. We focus on minimizing the *index* of $S$, which is the inverse of its density rounded up. We establish lower bounds and prove that our constructions are optimal or almost optimal, both for the index of $S$ and for the number of labels.

**Keywords:** Error bursts, error-correcting codes, interleaving schemes, circulant graphs.

**AMS subject classifications:** 94B20 (Theory of ECC: burst-correcting codes), 94B25 (Theory of ECC: combinatorial codes), 05C15 (Graph theory: coloring of graphs and hypergraphs).

## 1 Introduction

Error-correcting codes work best when the errors are scattered. Since errors on noisy channels are often bursty, *interleaving* is used. The idea is to assign data points to a number of separate codes, so that the points assigned to the same code are less likely to be hit by the same error burst. The goal is to minimize the transmission overhead, which is proportional to the number of distinct codes. For a simple example, suppose we transmit a stream of bits using parity bits for error-correcting. Furthermore, suppose we know that error bursts are quite rare, but a single burst can damage up to three consecutive bits. So we split the bits into three sets as $\{123123\ldots\}$ and compute parity bits separately for each set.

The way we interleave the codes largely depends on the topology of a noisy channel. Many noisy channels are one-dimensional, time being the only dimension. 2D noisy channels occur in optical recording [20], charged-coupled devices, 2D barcodes (e.g. *MaxiCode* from UPS), and information hiding in digital images and video sequences. A holographic data storage system can be viewed as a 3D noisy channel [14, 15, 7].

*Microsoft Research, Mountain View, CA 94043. Email: *slivkins at microsoft dot com*. This research have been completed while the first author was a B.S. student at California Institute of Technology and a Ph.D. student at Cornell University.

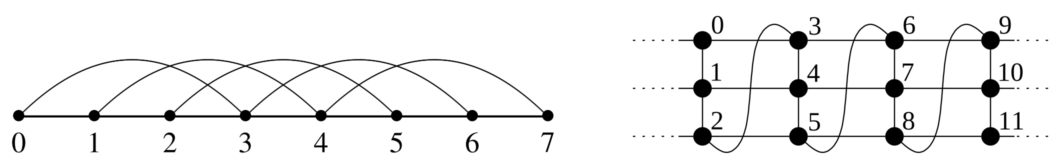
†Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125. Email: *bruck at paradise dot caltech dot edu*.

**Interleaving schemes.** Early work on interleaving concentrated on 2D rectangular-shaped error bursts [16, 17, 12, 1, 2, 8, 9]. Several other shapes have been considered as well, e.g. criss-cross errors [21, 22, 5] and circular-shaped error bursts [3]. The present paper takes after [6, 7] in that it considers *arbitrary* error bursts of a given size $t$. In other words, our goal is to make sure that no error burst of size $t$ or less contains two data points assigned to the same code.

Formally the topology of a noisy channel is given by a graph $G$ on transmitted data points, so that two data points are likely to be hit by the same error burst if and only if they are close to each other in $G$. Error bursts are then modeled as connected subgraphs of $G$. Therefore we have the following labeling problem: given a graph $G$ and an integer $t$, construct a labeling of $G$ so that no connected subgraph of size $t$ contains two vertices labeled the same, or, equivalently, the distance between any two vertices in the same label-set is at least $t$. Such a labeling is called a *$t$-interleaving scheme*, where $t$ is an *interleaving parameter*. The goal is to minimize *interleaving degree*, the number of distinct labels used. Note that for $t = 2$ it is just the graph-coloring problem.

Interleaving schemes have been introduced by Blaum et al. [6, 7]. The original paper [6] defined interleaving schemes *with repetitions*, where in any connected cluster of size $t$ any label is repeated at most $r$ times. Asymptotically optimal constructions on 2D arrays were presented for the case $r = 2$. In [7], the authors considered interleaving schemes (without repetitions) on two- and three-dimensional arrays. Their constructions are optimal for the 2D case, and optimal or nearly optimal for the 3D case, depending on $t \bmod 6$. Further work on interleaving schemes with repetitions includes [11, 23, 4]. Xu and Golomb [25] considered the inverse problem: for a given 2D array of codewords, maximize the interleaving parameter.

**Our contributions.** In this paper we extend interleaving schemes beyond arrays.[1] We consider a similar but substantially different topology: the node set is $\mathbb{Z}$, and an edge $ij$ exists if and only if $|i - j| \in \{1,\, d\}$, where $d$ is an integer parameter; we denote such graph by $G_d$ (see Figure 1a). Note that $G_d$ is essentially a 2D-array of width $d$ with a few extra edges (Figure 1b). These 'extra edges', however, break the constructions from [7], thus making our problem interesting. The graph $G_d$ belongs to the family of *circulant graphs*, which have been studied in the mathematical literature, e.g. see [10].




(a) $G_3$ as a line (b) $G_3$ as a 2D array with a few extra edges

Figure 1: $G_3$, the infinite circulant graph with two offsets $\{1,\, 3\}$.

Our main result is that for any given graph $G_d$ and integer $t$, we construct a $t$-interleaving scheme whose degree is optimal or close to optimal for these $(d, t)$ values. Our approach is to find a candidate label-set $S$ with a large density, and then to cover $\mathbb{Z}$ with a minimal number of copies of $S$. A simple lower bound on the number of copies is the inverse of the density, rounded up; we call this quantity the *index* of $S$. Most of our progress is on minimizing the index of a label-set, which is itself an interesting combinatorial problem.

Our interleaving schemes have a very simple, periodic structure. To make this more concrete, let us consider the following definition: an interleaving scheme is *periodic* if for some $p \in \mathbb{N}$ called the *period length* each integer $n$ is labeled the same as $n + p$. We distinguish three cases, with different constructions and lower bounds, depending on how $t$ compares to $d$. In all three cases, our interleaving schemes are

[1] After the initial version of this paper [24] had appeared, Jiang et al. [19, 18] considered interleaving schemes on other topologies, namely on tori [19] and on paths and cycles [18].

periodic, with period length $p = p(d,t)$. We characterize them by describing a *typical* label-set, namely the one containing 0. First, if $t$ is comparatively large, we consider a very simple interleaving scheme whose typical label-set is $p\mathbb{Z}$. This is a unique interleaving scheme of minimal degree, and the corresponding label-set is (essentially) a unique label-set of minimal index. Second, if $t$ is comparatively small, the typical label-set is of the form $\{0, t, 2t, \ldots, qt\} + p\mathbb{Z}$ for some $q, t \in \mathbb{Z}$ such that $qt < p$. Both the index and the resulting interleaving degree are nearly optimal. Third, if $d$ and $t$ are comparable then the typical label-set is of the form $\{0, a\} + p\mathbb{Z}$, for some $a < p$. This set has the minimal index, and the resulting interleaving degree is minimal in most cases and almost minimal otherwise.

**Extensions.** In the second case above, to lower bound the index of a label-set we use the lower bound derived from packing spheres on $G_d$, which is very similar to the *sphere-packing lower bound* used in [7]. Given that for interleaving schemes on 2D arrays the sphere-packing lower bound is tight [7], we investigated whether it remains tight in our setting. We give a complete characterization for odd $t$, and a partial result for even $t$; see Section 6 for further discussion.

A natural way to construct candidate label-sets (and in fact the way we did it first) is the *greedy algorithm* where we start with an empty set, and insert each consecutive number if and only if the resulting set is can be a label-set (i.e., all distances are at least $t$). We found that such algorithm often produces reasonable results, although it does not improve over Theorem 1.2. We discuss this further in Section 7.

## 1.1 Definitions and results

Throughout the paper $t$ will denote the interleaving parameter. Recall that an interleaving scheme on $G_d$ is a partition of the nodes into label-sets. A set $S \subset \mathbb{Z}$ can be a label-set if and only if the shortest-paths distance (with respect to $G_d$) between any two points of $S$ is at least $t$; we will call such sets $t$*-sparse*.

Call a set $S \subset \mathbb{Z}$ *periodic*, with a *period length* $p$, if it is that case that each integer $n$ lies in $S$ if and only if $n + p$ does. Say $S$ is $k$*-periodic* if the period $S \cap [0;p)$ consists of exactly $k$ points. We define the *density* of $S$ as $k/p$. [2] We extend this definition to non-periodic sets: we define density as $\lim_{n\to\infty} \frac{|S\cap[-n;n]\,|}{2n+1}$, whenever such limit exists. Let us say that a set is *well-formed* if the density exists; let us say that an interleaving scheme is well-formed if each of its label-sets is well-formed.

For a well-formed set $S$, a simple lower bound on the number of copies of $S$ needed to cover $\mathbb{Z}$ is given by the inverse of its density rounded up; let us call this quantity the *index* of $S$. Let $\texttt{index}(d,t)$ be the minimal index among all well-formed $t$-sparse sets on $G_d$, and let $\texttt{degree}(d,t)$ be the minimal degree among all $t$-interleaving schemes on $G_d$. Then:

**Lemma 1.1.** *For any graph $G_d$ and any $t \in \mathbb{N}$ we have* $\texttt{degree}(d,t) \geq \texttt{index}(d,t)$.

The proof is easy for well-formed interleaving schemes; the general case is somewhat more complicated, see Section 2.1 for details.[3]

If an interleaving scheme is a covering of $\mathbb{Z}$ by copies (translates) of a given $t$-sparse set $S$, we say that it is *induced* by $S$. Our interleaving schemes are induced by periodic $t$-sparse sets. Moreover, these $t$-sparse sets have a very simple structure: they are either 1- or 2-periodic, or have the property that their intersection with $[0;p)$ is $\{0, t, 2t, \ldots, qt\}$ for some $q$, where $p$ is the period; sets with this property will be called *two-offset*. Our approach is to find a $t$-sparse set with a small index, and then to cover $\mathbb{Z}$ with a minimal

[2] Note that for any two period lengths $p_1$, $p_2$ the value of density is the same. To see this, consider the interval $S \cap [0;\, p_1p_2)$.

[3] The authors wish to acknowledge that the original version of the lemma gave a lower bound of $K$ if the interleaving scheme is well-formed, and $K-1$ otherwise; the improved present version is due to the anonymous referees. We note in passing that whenever we invoke this lemma throughout this paper, the desired conclusion can also be obtained by much simpler direct arguments.

| (mod 4) | $d \equiv 0$ | $d \equiv 1$ | $d \equiv 2$ | $d \equiv 3$ |
|---|---|---|---|---|
| $t \equiv 0$ | no | no | yes* | yes* |
| $t \equiv 1$ | no | yes* | yes | yes |
| $t \equiv 2$ | yes* | yes* | no | no |
| $t \equiv 3$ | yes | yes | no | yes* |

* There exists a an optimal $t$-sparse set which is 1-periodic.

The table entry is 'yes' if and only if our interleaving scheme is optimal for the corresponding case.

Table 1: Theorem 1.2(c): our interleaving scheme is optimal in most cases.

number of copies thereof. Most of our progress is on minimizing the index of a $t$-sparse set, which is itself an interesting combinatorial problem.

Our constructions are optimal or nearly optimal. In order to state our results, let us define optimality and near-optimality, for a fixed underlying graph $G_d$. A well-formed $t$-sparse set is called *optimal* if its index is exactly $\texttt{index}(d, t)$, and *$\alpha$-approximate* if its index is at most $\alpha \times \texttt{index}(d, t)$. Similarly, a $t$-interleaving scheme is called *optimal* if its interleaving degree is exactly $\texttt{degree}(d, t)$, and *$\alpha$-approximate* if its interleaving degree is at most $\alpha \times \texttt{degree}(d, t)$.

Now we are ready to state our results. Recall that we distinguish three cases, with different constructions and lower bounds, depending on how $t$ compares to $d$.

**Theorem 1.2.** *Fix graph $G_d$ and interleaving parameter $t$. Let $\delta = \lceil d/2 \rceil$. Then:*

*(a) Suppose $t \geq d - 1$. Let $k = (t - \delta)d + \delta$. Then the set $k\,\mathbb{Z}$ is an optimal $t$-sparse set. Moreover, it is (up to translation) the only optimal $t$-sparse set that is periodic. The $t$-interleaving scheme induced by $k\,\mathbb{Z}$ is the unique optimal interleaving scheme.*

*(b) Suppose $t \leq \delta$. Then $\texttt{index}(d, t) \geq \lceil t^2/2 \rceil$. There exists a two-offset $t$-sparse set $S$ and an induced $t$-interleaving scheme which are $(1 + \frac{t}{d} + \frac{3}{t})$-approximate in general, and optimal if $t$ is even and $d \equiv \pm 1 \pmod t$. Moreover, for $d > t^3$ and even $t$ the index of $S$ is at most 1 above optimal.*

*(c) Suppose $\delta < t \leq d - 2$. Then $\texttt{index}(d, t) \geq d(3t - d)/4 + \Omega(d + t)$. There exists an optimal $t$-sparse set which is 1-periodic in many cases and 2-periodic in general (see Table 1.1). A $t$-interleaving scheme induced by this set is optimal in most cases (see Table 1.1), and $(1 + \frac{4}{d})$-approximate otherwise.*

### 1.2 Map of the paper

In Section 2 we introduce notation and prove Lemma 1.1. In Section 3 we make some observations on distances in $G_d$ and prove Theorem 1.2(a). In Section 4 we prove Theorem 1.2(c). Section 5 is on the case $t \leq \delta$, proving Theorem 1.2(b). In Section 6 we investigate when the sphere-packing lower bound is exact. In Section 7 we study the greedy algorithm for constructing $t$-sparse sets. We conclude in Section 8.

## 2 Preliminaries

An infinite circulant graph with offsets $a_1, \ldots, a_k \in \mathbb{N}$ is a graph on $\mathbb{Z}$ such that an edge $ij$ exists if and only if $|i - j| \in \{a_1, \ldots, a_k\}$. Finite circulant graphs are defined similarly: a circulant graph on $\{0, 1, \ldots, n-1\}$ with offsets $S \subset \mathbb{N}$ contains an edge $ij$ if and only if $|i - j| \bmod n \in S$. In these terms, $G_d$ is an infinite circulant graph with two offsets $\{1, d\}$. We will talk interchangeably about subgraphs of $G_d$ and subsets of $\mathbb{Z}$. We reserve $d$, $t$ for the larger offset and the interleaving parameter, respectively. Throughout the paper, most of our arguments will be slightly different depending on the parity of $d$ and $t$. To allow for a unified presentation, we let $\delta = \lceil d/2 \rceil$ and $\tau = \lceil t/2 \rceil$.

Let $\texttt{dist}(u, v)$ be the $G_d$-shortest-paths distance between points $u$, $v$, that is the number of edges in a shortest $uv$-path in $G_d$. Let $\texttt{dist}(v) = \texttt{dist}(0, v)$. Define the distance $\texttt{dist}(S, v)$ between a set $S$ and a point $v$ as the minimal distance between $v$ and $u \in S$. For an integer $r$ and a set $S$ define the *$r$-span* of $S$ as the set of points at distance less than $r$ from $S$.

Throughout the text, the notation $(a, b)$ always denotes the ordered pair, not an open interval. To distinguish the intervals from ordered pairs, we adopt the notation $[a; b]$, $(a; b)$, etc.

By default, all numbers are integers and all sets are subsets of $\mathbb{Z}$. In particular, for any two numbers $a$, $b$ an interval $[a; b]$ actually denotes the set $[a; b] \cap \mathbb{Z}$.

## 2.1 Proof of Lemma 1.1

Fix graph $G_d$ and interleaving parameter $t$, let $K = \texttt{index}(d, t)$, and denote $\beta^* = \frac{1}{K-1}$. Consider an interleaving scheme of degree $k$ with label sets $S_1, \ldots, S_k$; note that these label-sets partition $\mathbb{Z}$. If each label-set has a well-defined density, then the lemma follows trivially: the density of each set is strictly less than $\beta^*$ and the densities sum up to 1, so it must be the case that $k \geq K$.

To prove the general case, we will need to reason about the limits of infinite sequences. For each label-set $S_i$ and each $n \in \mathbb{N}$, let $\alpha_{(i,n)} = \frac{1}{2n+1} |S_i \cap [-n; n]|$. Recall that the density of $S_i$ is $\lim_{n\to\infty} \alpha_{(i,n)}$ if such limit exists.

**Claim 2.1.** *For each label-set $S_i$, if any subsequence of $\{\alpha_{(i,n)}\}_{n=0}^{\infty}$ has a limit, this limit is $\leq \beta^*$.*

*Proof.* For a fixed $i$, suppose there exists an increasing sequence $\{n_j\}_{j=0}^{\infty}$ of positive integers such that the subsequence $\{\alpha_{(i,n_j)}\}_{j=0}^{\infty}$ has a limit $\beta_i > \beta^*$. To obtain a contradiction we will construct a well-formed $t$-sparse set of index less than $K$.

By definition of a limit there exists $x > 0$ and $j_0 \in \mathbb{N}$ such that for each $j > j_0$ it is the case that $\alpha_{(i,\, n_j)} > \beta^* + x$. Take $j > j_0$ such that $n_j > \beta^* td/x$ and let $p = n_j + td$. Consider the set $S = S_i \cap [-n_j; n_j]$, and let $S^*$ be the periodic set, with period $2p + 1$, such that $S^* \cap [-p; p] = S$. Then $S^*$ consists of replicas of $S$ separated by a "padding" of length $2dt$. It follows that any two points in different replicas are at distance at least $t$ from each other, so $S^*$ is $t$-sparse. The density of $S^*$ is equal to $|S|/(2p + 1)$; with a little arithmetic one can show that it is greater than $\beta^*$, so the index of $S^*$ is less than $K$, a contradiction. Claim proved. □

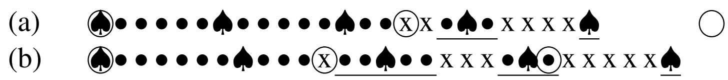
We will use the following well-known fact from calculus:

**Fact 2.2.** *Any bounded sequence $\{x_n\}_{n=0}^{\infty}$ contains a subsequence that has a limit. Moreover, this subsequence can be chosen so that its limit (a) is equal to $\limsup x_n$, (b) is equal to $\liminf x_n$.*

First, we claim that $\limsup_n \alpha_{(i,n)} \leq \beta^*$ for each label-set $S_i$. Indeed, suppose this is not the case for some $i$. Then by Fact 2.2(a) there exists a subsequence $\{\alpha_{(i,n_j)}\}_{j=0}^{\infty}$ that has a limit equal to $\limsup_n \alpha_{(i,n)} > \beta^*$, which contradicts Claim 2.1. Claim proved.

Second, we claim that $\liminf_n \alpha_{(i,n)} < \beta^*$ for each label-set $S_i$. Indeed, suppose this is not the case for some $i$. Then $\limsup_n \alpha_{(i,n)} \leq \beta^* \leq \liminf_n \alpha_{(i,n)}$, so $\lim_n \alpha_{(i,n)} = \beta^*$, and therefore $S_i$ has index $K - 1$, contradiction.

Third, it follows that $\liminf_n \alpha_{(i,n)} < \beta^*$ for all $i$. Then by Fact 2.2(b) there exists a subsequence $\{\alpha_{(1,n_j)}\}_{j=0}^{\infty}$ that has a limit equal to $\liminf_n \alpha_{(1,n)}$. Taking further subsequences (i.e. iteratively using Fact 2.2) we may further assume that $\lim_j \alpha_{(i,n_j)}$ exists for each $i$; denote it by $\beta_i$. By Claim 2.1 we have $\beta_i \leq \beta^*$ for each $i$. Since $\beta_1 < \beta^*$ and $\sum \beta_i = 1$, it follows that $k \geq K$.




We show the $t$-span of 0 for (a) $t > d-2$, (b) $t \leq d-2$. Vertices 0, $v_{\min}$, $2v_{\min}$ are encircled.

NOTATION: Recall that for $v \in \mathbb{Z}$, the $r$-span of $v$ is the set $S^*$ of all points at distance less than $r$ from $v$. We represent $S^*$ as a string where consecutive characters correspond to consecutive integers as follows: ♠ is for points $v + jd \in S^*$ ($j \in \mathbb{N}$), '•' for other elements of $S^*$, and 'x' for points not in $S^*$. Contiguous intervals of $S^*$ are underlined.

Figure 2: $v_{\min}\,\mathbb{Z}$ is $t$-sparse if and only if $t > d-2$.

# 3 Distances in $G_d$: general observations and proof of Theorem 1.2(a)

Let us make a few observations on distances in $G_d$. Recall the notation $\delta = \lceil d/2 \rceil$ and $\tau = \lceil t/2 \rceil$.

**Definition 3.1.** Fix a point $v \in \mathbb{N}$. An ordered pair $(x, y)$ of integers is a *canonical representation* of $v$ if the following conditions hold: (1) $v = xd + y$ and $-\delta < y \leq \delta$, and (2) $x$ is minimal subject to (1).

Note that in the above definition, a pair $(x, y)$ satisfying condition (1) always exists, and is unique whenever $v \not\equiv \delta \pmod{d}$. We use canonical representation to characterize the distances in $G_d$:

**Claim 3.2.** $\texttt{dist}(v) = x + |y|$ *where $(x, y)$ is the canonical representation of $v$, for any $v \in \mathbb{N}$.*

Consider a point $v \in \mathbb{Z}$. We call $v$ *remote* if $\texttt{dist}(v) \geq t$. Let $v_{\min}$ be the smallest positive remote point, and we let $v_{\max}$ be the largest non-remote point. They can be easily computed using Claim 3.2:

**Claim 3.3.** $v_{max} = d(t-1)$. *If $t > \delta$ then $v_{min} = (t-\delta)d + \delta$, else $v_{min} = t$.*

Now we are ready to prove Theorem 1.2(a):

**Proof of Theorem 1.2(a):** Suppose $t > d-2$ and let $S = v_{\min}\,\mathbb{Z}$. This set is $t$-sparse because the difference between any two elements of $S$ is either $v_{\min}$ or at least $2v_{\min} > v_{\max}$ (see Figure 2a). Since the interval between any consecutive elements of a $t$-sparse set is at least $v_{\min}$, it follows that $\texttt{index(d,t)} \geq v_{\min}$, and moreover that $S$ is a unique periodic $t$-sparse set that achieves this bound.

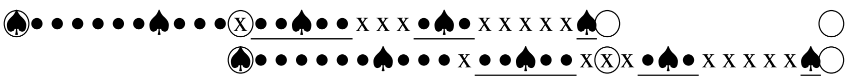
In the $t$-interleaving scheme induced by $S$ each vertex $i$ has label $(i \bmod v_{\min})$. This is a valid $t$-interleaving scheme since each label-set is a translate of $S$. The interleaving degree is $v_{\min}$, which is optimal by Lemma 1.1. It remains to prove that this is the *unique* optimal interleaving scheme, i.e. that any other interleaving scheme requires more labels. Indeed, in any other interleaving scheme there is a label-set with two consecutive vertices $u$,$v$ such that $|u - v| > v_{\min}$. Then the distance between any two points in the interval $[u+1;\, u+v_{\min}]$ is less than $t$, so all points in this interval must be labeled distinctly, not using the label of $u$ and $v$. This requires at least $v_{\min} + 1$ labels. □

Note that $v_{\min}\,\mathbb{Z}$ is $t$-sparse only if $t > d-2$ since otherwise $\texttt{dist}(2v_{\min}) < t$ (see Figure 2b).

# 4 Case $\delta < t \leq d-2$: proof of Theorem 1.2(c)

In this section we assume $\delta < t \leq d-2$. Recall that we let $\delta = \lceil d/2 \rceil$ and $\tau = \lceil t/2 \rceil$.

Note that $v_{\min} \leq v_{\max}/2$ if and only if $t \leq d-2$. We will derive our constructions and lower bounds by analyzing *triples* of consecutive elements of a $t$-sparse set. For any such triple $(v_1, v_2, v_3)$ it must be the case that the intervals $v_2 - v_1$, $v_3 - v_2$ and $v_3 - v_1$ are remote. This motivates the following definition which will be useful in the forthcoming arguments:



Using the notation in Figure 2, the two lines represent the $t$-spans of 0 and $v_{\min}$. The upper line represents the former, and the lower line represents the latter. Then $\sigma_{\min}$ is the leftmost point that is remote in both lines. Points 0, $v_{\min}$, $\sigma_{\min}$ and $\sigma_{\min} + v_{\min}$ are encircled.

Figure 3: $\sigma_{\min}$ for $(d, t) = (7, 5)$.

**Definition 4.1.** Say $(w_1, w_2)$ is a *remote pair* (with *sum* $w_1 + w_2$) if $w_1$, $w_2$ and $w_1 + w_2$ are positive and remote. Let us say that a remote pair $(w_1, w_2)$ *induces* the periodic set $\{0, w_1\} + (w_1 + w_2)\mathbb{Z}$, and the interleaving scheme where $\mathbb{Z}$ is covered by the minimal number of copies of this set.

We are especially interested in remote pairs of the form $(v_{\min} + \gamma, \cdot)$, $\gamma \in \{0, 1\}$. [4]

**Definition 4.2.** Let $\sigma_{\min}$ be the minimal sum of a remote pair of the form $(v_{\min} + \gamma, \cdot)$, $\gamma \in \{0, 1\}$ (see Figure 3). A remote pair is called *standard* if its sum is equal to $\sigma_{\min}$.

We restate Theorem 1.2(c) in the following more precise form:

**Theorem 4.3.** *Consider a pair $(G_d, t)$ such that $\delta < t \leq d - 2$.*

*(a)* $\texttt{index}(d, t) \geq \lceil \sigma_{min}/2 \rceil = d(3t - d)/4 + \Omega(d + t)$.

*(b) Any standard remote pair induces a $t$-sparse set (which is optimal by part (a));*

*(c) There exists a standard remote pair such that the induced $t$-interleaving scheme is optimal in most cases (see Table 1.1 on page 4), and $(1 + \frac{4}{d})$-approximate otherwise. In fact, in many cases one such pair is $(\sigma_{min}/2,\ \sigma_{min}/2)$.*[5]

We present the proof of Theorem 4.3 for a somewhat simpler case of odd $d$. The full proof (joint for both odd and even $d$) is in Appendix A.

### 4.1 Proof of Theorem 4.3: odd $d$

We start with a few technical claims:

**Claim 4.4.** *For any $v \in \mathbb{N}$ such that $0 < v < v_{max}$, the following are equivalent:*
*(a) $v$ is remote.*
*(b) the canonical representation of $v - v_{min}$ is a pair $(\mu_1, \mu_2)$ such that $-\mu_1 \leq \mu_2 \leq \mu_1$.*
*(c) $v = v_{min} + \mu_1 d + \mu_2$, for some $(\mu_1, \mu_2)$ such that $\mu_1 \leq 0$ and $|\mu_2| \leq \mu_1$.*

**Claim 4.5.**
*(a) $2\,v_{min} = \alpha d + 1$ where $\alpha = 2(t - \delta) + 1$.*
*(b) $\sigma_{min} = 2\,v_{min} + (\delta - \tau)(d + 1)$,*
*(c) $2\,\sigma_{min} \geq (t + 2)d$.*

[4]For odd $d$, we need to consider only pairs of the form $(v_{\min}, \cdot)$ since (as we shall see) $v_{\min} + 1$ is not remote.
[5]In this case $k\,\mathbb{Z}$, $k = \lceil \sigma_{\min}/2 \rceil$ is an optimal $t$-sparse set which induces an optimal $t$-interleaving scheme.

*Proof.* Part (a) is an easy computation which we omit.

Let us prove part (b). For each $\sigma \geq 2\,v_{\text{min}}$, let $(\alpha_{\text{c}}(\sigma),\ \beta_{\text{c}}(\sigma))$ be the canonical representation of $\sigma - 2\,v_{\text{min}}$. Let $\alpha_{\text{c}} = \alpha_{\text{c}}(\sigma_{\text{min}})$ and $\beta_{\text{c}} = \beta_{\text{c}}(\sigma_{\text{min}})$. Consider the set

$$W = \{\sigma > 2\,v_{\text{min}} \mid -\alpha_{\text{c}}(\sigma) \leq \beta_{\text{c}}(\sigma) \leq \alpha_{\text{c}}(\sigma) \text{ and } -\delta < 1 + \beta_{\text{c}}(\sigma) \leq \delta\} \tag{1}$$

First, we claim that $\sigma_{\text{min}} \in W$. Indeed, $\sigma_{\text{min}} - v_{\text{min}}$ is remote by definition of $\sigma_{\text{min}}$, thus Claim 4.4(b) says precisely that for $\sigma = \sigma_{\text{min}}$ the first condition in (1) holds. Also, by definition of canonical representation we have $-\delta < \beta_{\text{c}}$. So it remains to show that $1 + \beta_{\text{c}} \leq \delta$. Suppose this is not the case. Then $\beta_{\text{c}} = \delta$. By part (a) we have

$$\sigma_{\text{min}} = 2\,v_{\text{min}} + (\alpha_{\text{c}} d + \delta) = (\alpha + \alpha_{\text{c}})d + (\delta + 1).$$

It follows that $\sigma_{\text{min}} - 1$ is remote. Moreover, since the first condition in (1) holds for $\sigma = \sigma_{\text{min}}$, it also holds for $\sigma = \sigma_{\text{min}} - 1$, so by Claim 4.4(b) $\sigma_{\text{min}} - v_{\text{min}} - 1$ is remote. Therefore $(v_{\text{min}}, \sigma_{\text{min}} - 1)$ is a remote pair, which contradicts the minimality of $\sigma_{\text{min}}$. Claim proved.

Second, we claim that $\sigma_{\text{min}}$ is the *smallest* remote element of $W$. Indeed, assume $\sigma \in W$ for some remote $\sigma < \sigma_{\text{min}}$. Then applying Claim 4.4(b) for $v = \sigma - v_{\text{min}}$ it follows that $\sigma - v_{\text{min}}$ is remote (the condition in Claim 4.4(b) is precisely the first condition in (1)). Therefore $(v_{\text{min}}, \sigma - v_{\text{min}})$ is a remote pair, contradicting the minimality of $\sigma_{\text{min}}$. Claim proved.

Therefore

$$\alpha_{\text{c}} = \min\{x \mid \varphi(x) \geq t\} \text{ where } \varphi(x) = \max\{\texttt{dist}(\sigma) \mid \sigma \in W \text{ and } \alpha_{\text{c}}(\sigma) = x\}. \tag{2}$$

For a given $\sigma \in W$, by part (a) and the definition of $(\alpha_{\text{c}}(\sigma), \beta_{\text{c}}(\sigma))$ we have

$$\sigma = 2\,v_{\text{min}} + \alpha_{\text{c}}(\sigma) d + \beta_{\text{c}}(\sigma) = (\alpha + \alpha_{\text{c}}(\sigma))\,d + (1 + \beta_{\text{c}}(\sigma)). \tag{3}$$

By the second condition in (1), the right-hand side of (3) gives the canonical representation of $\sigma$. Therefore by Claim 3.2 it is the case that $\texttt{dist}(\sigma) = (\alpha + \alpha_{\text{c}}(\sigma)) + |1 + \beta_{\text{c}}(\sigma)|$, which is maximized, for a fixed $\alpha_{\text{c}}(\sigma)$, only if $\beta_{\text{c}}(\sigma) = \alpha_{\text{c}}(\sigma)$. It follows that $\beta_{\text{c}} = \alpha_{\text{c}}$ and moreover $\varphi(x) = 2x + \alpha + 1$. Therefore solving (2) for $\alpha_{\text{c}}$ gives $\alpha_{\text{c}} = \delta - \tau$, and part (b) follows.

Part (c) is an easy corollary of parts (ab). Plugging in the values for $\alpha$ and $\alpha_{\text{c}}$, an easy computation shows that

$$2\sigma_{\text{min}} \geq 2(\alpha + \alpha_{\text{c}})d \geq (3t - 2\delta)d \geq (t + 2)d. \quad \square$$

The following lemma extends remote pairs to interleaving schemes. We omit the proof since it is (essentially) a special case of Lemma 5.3 from the next section:

**Lemma 4.6.** *Let $S$ be the set induced by a remote pair $(w_1, w_2)$. Let $g = \gcd(w_1, w_2)$. Then the smallest number of copies of $S$ required to cover $\mathbb{Z}$ is $g\lceil (w_1 + w_2)/(2g) \rceil$, which is at most $g$ plus the index of $S$.*

Now we are ready to prove the theorem:

**Proof of Theorem 4.3: (a)** Let us denote the minimal sum of a remote pair by $\sigma$. First we claim that $\sigma = \sigma_{\text{min}}$. Indeed, let $(w_1, w_2)$, $w_1 \leq w_2$ be a remote pair with a sum $\sigma$ such that $w_1$ is minimal. If $w_1 = v_{\text{min}}$ then $\sigma = \sigma_{\text{min}}$ by definition of $\sigma_{\text{min}}$. Else we can choose $z > 0$ so that $(w_1 - z, w_2 + z)$ is a remote pair, contradicting the minimality of $w_1$. Specifically, we let $(x_1, y_1)$ be the canonical representation of $w_1$, and we choose $z \in [d-1; d+1]$ as follows. If $\texttt{dist}(w_1) > t$ let $z = d$; else we let $z = d - 1$ if $y_1 > 0$, and $z = d + 1$ otherwise. Then $w_1 - z$ and $w_2 + z$ are remote by Claim 3.2. Claim proved.

Let $S$ be a $t$-sparse set with a well-defined density $\rho$. Let $\{s_i : i \in \mathbb{Z}\}$ be an increasing enumeration of $S$. For each $i$, $(s_{i+1} - s_i,\ s_{i+2} - s_{i+1})$ is a remote pair, so its sum $s_{i+2} - s_i$ is at least $\sigma_{\text{min}}$. Then

$s_n - s_{-n} \geq n\sigma_{\min}$ for any $n > 0$, so $\rho \leq 2/\sigma_{\min}$, which gives the required lower bound on the index of $S$. By Lemma 1.1 this implies a similar lower bound on $\texttt{degree}(d,t)$.[6]

**(b)** Let $S$ be the set induced by a standard remote pair $(w_1, w_2)$. For any $u, v \in S$, either

$$|u - v| \in \{0,\ w_1,\ w_2,\ \sigma_{\min},\ \sigma_{\min} + w_1,\ \sigma_{\min} + w_2\}$$

or else $|u-v| \geq 2\sigma_{\min}$. In the latter case, since by Claim 4.5(c) we have $2\sigma_{\min} \geq (t+2)d > v_{\max}$, it follows that $\texttt{dist}(v,u) \geq t$. Therefore it remains to prove that $\sigma_{\min} + w_1$ and $\sigma_{\min} + w_2$ are remote. Indeed, by Claim 4.5(c) the canonical representation of $\sigma_{\min}$ is $(\,\cdot,\ y)$ for some $y \in [0; t/2]$. Now part (b) follows from the following general claim:

**Claim 4.7.** *If $u^*, v^* \in \mathbb{N}$ are remote, and moreover $u^* = xd + y$ so that $|y| \leq t/2$, then $u^* + v^*$ is remote.*

*Proof.* The claim is obvious if $v^* \geq v_{\max}$. Suppose $v^* < v_{\max}$. Then by Claim 4.4(c) we have

$$u^* + v^* = v_{\min} + (x + \mu_1)d + (y + \mu_2),$$

where $|\mu_2| \leq \mu_1$. Since $\texttt{dist}(u^*) = x + |y| \geq t$ it follows that $x \geq t/2 \geq |y|$, so $|y + \mu_2| \leq x + \mu_1$ and by Claim 4.4(c) $u^* + v^*$ is remote. □

**(c)** Consider a remote pair $(v_{\min}, \sigma_{\min} - v_{\min})$. We claim that for each $j \leq t - \delta$

$$(v_{\min} + j(d+1),\ \ \sigma_i - v_{\min} - j(d+1)) \tag{4}$$

is a remote pair, too. Indeed, the sum of this pair is $\sigma_{\min}$, and the first number in (4) is remote by Claim 4.4(b), so it remains to consider the second number in (4). By Claim 4.5(b) it is equal to $v_{\min} + (\alpha_c - j)d + (\alpha_c - j)$, so by Claim 4.4(b) it is remote, too. Claim proved.

By Claim 4.5(b) there exists a standard remote pair of the form $(w, w)$ if $\alpha_c = \delta - tau$ is even, and $(w, w+d+1)$ if $\alpha_c$ is odd. Let $S$ be the set induced by such a pair. If $\alpha_c$ is even, then $S = w\,Z$ induces an optimal interleaving scheme. Now suppose $\alpha_c$ is odd. With some arithmetic one can show that

$$g := \gcd(w, w+d+1) = \gcd(t, d+1).$$

By part (a) and Claim 4.5(c) we have

$$\texttt{degree}(d,t) \geq \lceil \sigma_{\min}/2 \rceil \geq (t+2)d/4 > gd/4.$$

By Lemma 4.6 set $S$ induces an interleaving scheme of degree

$$\texttt{deg}(S) = g\lceil \sigma_{\min}/2g \rceil \leq \lceil \sigma_{\min}/2 \rceil + g \leq \texttt{degree}(d,t) \times (1 + \tfrac{4}{d}),$$

so this interleaving scheme is $(1 + \frac{4}{d})$-approximate. Moreover, if both $d$ and $t$ are odd then this interleaving scheme is in fact optimal. Indeed, in this case $\sigma_{\min} = 2w + d + 1$ is even and $g$ is odd, so $2g|\sigma_{\min}$ and therefore $\texttt{deg}(S) = \lceil \sigma_{\min}/2 \rceil$, matching the lower bound from part (a). □

[6] To obtain the bound on $\texttt{degree}(d,t)$ directly, let $w_1, w_2, w_3$ be three consecutive vertices labeled the same in a $t$-interleaving scheme. Then $(w_2 - w_1, w_3 - w_2)$ is a remote pair, so its sum $w_3 - w_1$ is at least $\sigma_{\min}$. Therefore, in the interval $[0; \sigma_{\min})$ at most two vertices can be marked by each label, which requires at least $\sigma_{\min}/2$ distinct labels.

# 5 Case $t \leq \delta$: proof of Theorem 1.2(b)

Recall that we let $\delta = \lceil d/2 \rceil$ and $\tau = \lceil t/2 \rceil$. Recall that the $r$-span of a set $S$ is the set of points at $G_d$-distance less than $r$ from $S$. Following [7], we define a *$t$-sphere* as follows:

**Definition 5.1.** A *$t$-sphere* $S_t = S_t(p)$ centered at a point $p \in \mathbb{Z}$ is the $\tau$-span of $\{p\}$ if $t$ is odd, and the $\tau$-span of $\{p, p+d\}$ if $t$ is even.

To compute the size of a $t$-sphere, consider $G_d$ as a two-dimensional $d \times \infty$ mesh with "extra edges" between $(0, n)$ and $(d-1, n+1)$ for all $n$ (see Figure 1b). It is easy to see that for $t \leq \delta$ a $t$-sphere centered at $(\delta, n)$ is exactly the same in $G_d$ as in the 2D mesh, since the $t$-sphere simply does not reach the "extra edges". Therefore by [7] the size of any $t$-sphere is $\lceil t^2/2 \rceil$. Now we can state:

**Theorem 5.2** (The sphere-packing lower bound)**.**
*(a) [7] any interleaving scheme has degree at least $|S_t| = \lceil t^2/2 \rceil$,*
*(b) any well-formed $t$-sparse set has index at least $|S_t|$,*
*(c) any periodic $t$-sparse set of index $|S_t|$ induces an interleaving scheme of the same degree.*

*Proof.* Part (a) is proved in [7]; the lower bound holds because the distance between any two points in $S_t$ is less than $t$, so in any $t$-interleaving scheme all points of $S_t$ must be labeled differently. Note that part (a) also follows from part (b) in conjunction with Lemma 1.1.

To prove part (b), we claim that if $\mathtt{dist}(p, q) \geq t$ then the $t$-spheres centered at $p$ and $q$ are disjoint. Assume $S_t(p)$ and $S_t(q)$ intersect at $w$. If $t$ is odd then $\mathtt{dist}(p, w) \leq \tau - 1$ and $\mathtt{dist}(q, w) \leq \tau - 1$, so by the triangle inequality $\mathtt{dist}(p, q) < t$. Now suppose $t$ is even. Then either $\mathtt{dist}(p, w) \leq \tau - 1$ or $\mathtt{dist}(p + d, w) \leq \tau - 1$, same for $q$. Therefore by the triangle inequality $\mathtt{dist}(p, q) < t$ unless $\mathtt{dist}(p, w) = \mathtt{dist}(q, w) = \tau$. In the latter case, however, $\mathtt{dist}(p + d, w) = \mathtt{dist}(q + d, w) = \tau - 1$, so there exists a path from $p + d$ to $q + d$ with less than $t$ vertices. Shifting this path by $-d$ produces a $pq$-path of the same length. Claim proved.

Let $S$ be a well-formed $t$-sparse set of minimal index and density $\rho$. Since the sets $S_t(p)$, $p \in S$ are pairwise disjoint, the density of their union $U$ is $\rho|S_t| \leq 1$, so the index of $S$ is at least $|S_t|$, proving part (b).

For part (c), suppose the index of $S$ is exactly $|S_t|$. Then the density of $U$ is 1. Since $S$ is periodic, $U$ is periodic, too, so $U = \mathbb{Z}$. Partition $U$ as follows: let $U_i = \{v_i(p) : p \in S\}$, where $v_i(p)$ is the $i$-th vertex of $S_t(p)$ from the left. Then the sets $U_i$ are translates of $S$, hence they are $t$-sparse. Label all points of $U_i$ with $i$ to get an optimal interleaving scheme. □

In Theorem 5.2(a) we used the fact that the distance between any two points in a $t$-sphere is less than $t$. It is an open question whether there exist larger sets with this property. Finding such sets would be nice given a $\Omega(t)$ gap between (the general case of) our construction and the sphere-packing lower bound.

## 5.1 The two-offset construction

We will construct two-offset $t$-sparse sets that reach or almost reach the sphere-packing lower bound. We extend them efficiently to interleaving schemes using the following lemma:

**Lemma 5.3.** *Let $S$ be a two-offset set with a period $p$; then $S \cap [0; p) = \{0, t, 2t, \ldots, qt\}$ for some $q$. Let $\phi = p/(q+1)$ and $g = \gcd(t, p)$. Then the smallest number of copies of $S$ required to cover $\mathbb{Z}$ is $g\lceil \phi/g \rceil$, which is at most $g$ plus the index of $S$.*

*Proof.* Let's try to cover the interval $X = [0; p)$. For each integer $i$ let

$$A_i = \{(i + jt) \bmod p : j \in \mathbb{Z}\}.$$

From elementary number theory, the sets $A_0 \dots A_{g-1}$ form a disjoint partition of $X$, so the size of each $A_i$ is $p/g$. Now, each copy of $S$ intersects with exactly one $A_i$, the size of intersection being $q+1$. Therefore, one needs at least $N = \left\lceil \frac{|A_i|}{q+1} \right\rceil$ copies to cover one of the sets $A_i$, and at least $gN$ copies to cover all of them. Conversely, to cover $\mathbb{Z}$ by $gN$ copies of $S$ we can use the sets $i + j(q+1)t + S$ where $0 \le i < g$ and $0 \le j < N$. Finally, it is easy to see that $gN \le g + \lceil \phi \rceil$, where $\lceil \phi \rceil$ is the index of $S$. □

For the remainder of this section we will use the pair $(q, r)$ defined by

$$d = (q+1)\,t + r, \quad \text{where } -1 \le r \le t-2. \tag{5}$$

**Definition 5.4.** Define the *two-offset construction* as the two-offset set $S^*$ with a period

$$p^* = \begin{cases} d\tau - \tau, & \text{if } (t \text{ even and } r = -1) \text{ or } (t \text{ odd and } r = 0, 1) \\ d\tau + \tau, & \text{otherwise.} \end{cases}$$

such that $S^* \cap [0; p^*) = \{0, t, 2t\,,\ \dots\ , qt\}$.

**Lemma 5.5.** *The two-offset construction is $t$-sparse.*

*Proof.* Let $T = \{0, t, 2t\,,\ \dots\ , qt\}$, where $q$ is defined by (5). Say a node is *$T$-remote* if it is at distance at least $t$ from $T$; say a set is *$T$-remote* if all its elements are. It suffices to show that the set $jp^* + T$ is $T$-remote, for all integer $j \ge 1$.

The two right-most points of the $t$-span of $T$ are $p_1 = d(t-1) + qt$ and $p_2 = p_1 - t$; note that $p_1 < dt$. Also, note that in most cases we have $2p^* > dt$. More precisely:

$$[2p^* \le dt] \;\Rightarrow\; [t \text{ is even and } r = -1] \;\Rightarrow\; [p_2 < 2p^* < p_1 < 2p^* + t]\,.$$

Therefore the set $jp^* + T$ is $T$-remote for all integer $j \ge 2$. It remains to show that $p^* + T$ is $T$-remote.

For each pair of integers $i, j$ such that $0 \le i \le t$ and $0 \le j \le q$ let us define the interval

$$B_{ij} = v_{ij} + (i - t;\ t - i), \quad \text{where } v_{ij} = id + jt.$$

Then $B_{ij}$ is the part of the $t$-span of $v_{0j}$ that lies in $[id - t;\ id + d + t]$. It is easy to see (Figure 4) that the $t$-span of $T$ is equal to the union of the sets $B_{ij}$.

Now let $j < q$. Define the *overlap* between two integer intervals as the size of their intersection if they do intersect, and the negated number of points between them if they don't. Then $x_{ij} = t - 2i - 1$ is the overlap between $B_{ij}$ and $B_{(i,\,j+1)}$, and $y_i = t - 2i - r - 1$ is the overlap between $B_{iq}$ and $B_{(i+1,\,0)}$.

Partition the interval $[id; id + d)$ into intervals $I_{ij} = [v_{ij};\, v_{(i,\,j+1)})$ and $J_i = [v_{iq};\, v_{(i+1,\,0)})$. Say an interval is *free* if it contains some $T$-remote point. Then each interval $I_{ij}$ is free if and only if $x_{ij} < 0$, which happens if and only if $i \ge \tau$. Moreover, each interval $J_i$ is free if and only if $y_i < 0$, which happens if and only if $i \ge \lfloor \frac{t-r}{2} \rfloor$. Noting that for $p^* = d\tau \pm \tau$ we have

$$p^* + T = \{v_{\tau j} \pm \tau :\ 0 \le j \le q\},$$

it is easy to verify that all elements of $p^* + T$ lie in the intervals $I_{ij}$ and $J_i$ that are free.

Now, each interval $I_{ij}$ is free if and only if both $(v_{ij} + \tau)$ and $(v_{(i,\,j+1)} - \tau)$ are $T$-remote. So all points from $p^* + T$ that lie in some $I_{ij}$ are $T$-remote. The only point from $p^* + T$ that lies in some $J_i$ is

$$p = \begin{cases} p^* \in J_{\tau - 1}, & \text{if } p^* = d\tau - \tau, \\ p^* + qt \in J_\tau, & \text{if } p^* = d\tau + \tau. \end{cases}$$

With some easy arithmetic, we can check that $p$ is $T$-remote, too: namely, for the corresponding $J_i$ containing $p$, it suffices to check that $p$ does not lie in neither $B_{iq}$ nor $B_{(i+1,\,0)}$. Therefore, the set $p^* + T$ is $T$-remote, completing the proof. □

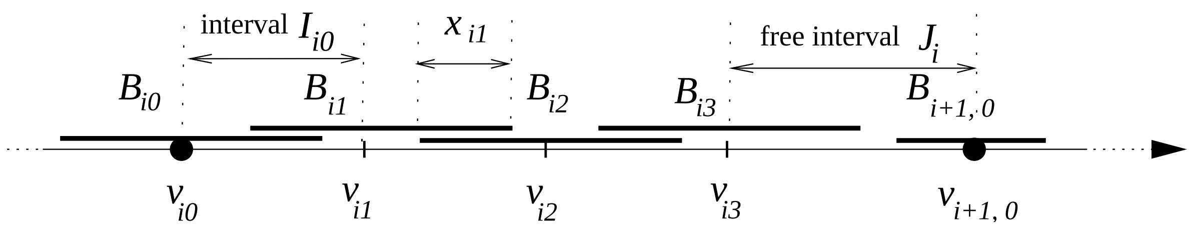



Figure 4: Span of $S = \{0, t, 2t, \ldots, qt\}$ as the union of the sets $B_{ij}$

The index of the two-offset construction is, for $p^* = d\tau \pm \tau$,

$$\left\lceil \frac{p^*}{q+1} \right\rceil = t\tau + \psi = \begin{cases} |S_t| + \tau - 1 + \psi, & t \text{ is odd} \\ |S_t| + \psi, & t \text{ is even} \end{cases} \tag{6}$$

where $\psi = \lceil \tau(r \pm 1)/(q+1) \rceil$. Thus the sphere-packing lower bound is achieved if and only if $t$ is even and $r = \pm 1$. Note that if $d > rt^2$ and $r \neq \pm 1$ then $\psi = 1$. Otherwise the index in (6) is $(1 + \frac{t}{d} + \frac{1}{t})$-approximate with respect to the sphere-packing lower bound.

We extend the two-offset construction to a $t$-interleaving scheme using Lemma 5.3. With some easy arithmetic, we can show that the resulting interleaving degree is $(1 + \frac{t}{d} + \frac{1}{t})$-approximate (with respect to the sphere-packing lower bound). By Theorem 5.2(b), whenever the two-offset construction achieves the sphere-packing lower bound, so does the induced a $t$-interleaving scheme (this can also be seen directly via Lemma 5.3). This completes the proof of Theorem 1.2(b).

# 6 More on the sphere-packing lower bound

In this section we assume that $t \leq \delta$ and investigate when the sphere-packing lower bound is exact. We solve this question for odd $t$ and give a partial result for even $t$. Recall that we let $\delta = \lceil d/2 \rceil$ and $\tau = \lceil t/2 \rceil$.

Say a set is *SLB-optimal* if it is $t$-sparse and its index reaches the sphere-packing lower bound, namely index $= |S_t|$, where $S_t$ is the $t$-sphere defined in Section 5. Define *the even construction* as the set $|S_t|\,\mathbb{Z}$. Among all sets of the form $w\mathbb{Z}$, $w \in \mathbb{Z}$ only the even construction can be SLB-optimal; the even construction is SLB-optimal if and only if it is $t$-sparse.

**Lemma 6.1.** *If $d \equiv \pm t \pmod{|S_t|}$ and $t \leq \delta$ is odd then the even construction is $t$-sparse.*

This lemma easily follows from [7]; we prove it here for the sake of completeness.

*Proof.* Recall that $|S_t| = \lceil t^2/2 \rceil$ and let $s = |S_t|$. Suppose the even construction is not $t$-sparse for some odd $t$ such that $d \equiv t \pmod{s}$. (For $d \equiv -t \pmod{s}$ the proof is similar.) Then there exist points $p > q$ such that $p \equiv q \pmod{s}$ and $\texttt{dist}(p, q) < t$. Let $(i, j)$ be the canonical representation of $p - q$, so that, in particular, $p - q = id + j$. Since $s$ divides $id + j$, it also divides $it + j$. Since by Claim 3.2 $\texttt{dist}(p, q) = i + |j| < t$, it follows that

$$it + j < it + (t - i) \leq t^2 < 2s.$$

Since $it + j$ is divisible by $s$ but is less than $2s$, it follows that it is equal to $s$.

Now we claim that $t = 2i \pm 1 = \pm(2j - 1)$. Indeed, $it + j = s = (t^2 + 1)/2$, so $2j - 1 = t(t - 2i)$. Since $|j| \leq i + |j| < t$, it follows that $2t > |2j - 1| = t\,|t - 2i|$. Therefore $|t - 2i| = 1$, and so $|2j - 1| = t$. Claim proved. It follows that $\texttt{dist}(p, q) = i + |j| = t$, contradiction. $\square$

♠ x x x x x x •♠• x x x x ••♠•• x x •••(♠)••• x x ••♠•• x x x x •♠• x x x x x x ♠

{ Left Branch } { central } { Right Branch }

♠ x x x x x x •♠• x x x x ••(♠)•• x x x ••♠•• x x x x •♠• x x x x x x ♠

{ Left Branch } { Right Branch }

$t$-spheres for $(d,t) = (8,7)$ (above) and $(d,t) = (8,6)$ (below). Centers are encircled.

NOTATION: We represent a $t$-sphere $S_t = S_t(p)$ by a string where consecutive characters correspond to consecutive numbers. We use ♠ for the points $p + kd \in S_t$, $k \in \mathbb{Z}$, • for other points of $S_t$, and 'x' for points not in $S_t$. Stations are underlined.

Figure 5: *Stations* and *branches* of a $t$-sphere.

Now we are ready to state the main theorem of this section.

**Theorem 6.2.** *For odd $t \leq \delta$, SLB-optimal constructions exist only if $d \equiv \pm t \pmod{|S_t|}$, in which case by Lemma 6.1 the even construction is SLB-optimal. For even $t \leq \delta$, the even construction is SLB-optimal only if $d \equiv \pm 1 \pmod{t}$.*

Note that for even $t$, by Theorem 6.2 and Theorem 1.2(b) it is the case that if the even construction is SLB-optimal then there exists an SLB-optimal two-offset construction. However, since the even construction is simpler, it is interesting to investigate further when exactly it is SLB-optimal. Using a computer program, for each $t \leq 42$ we have computed the 30 smallest values of $d$ when this happens. This data motivated several conjectures:

**Conjecture 6.3.** Consider the set $D_t$ of all $d \geq 2t$ such that the even construction is SLB-optimal. Then:

- If $d \equiv 1 \pmod{t}$, then it is the case that $d \in D_t$ if and only if $d - 2 \in D_t$.
- $\min(D_t) = pt - 1$, where $p$ be the smallest prime that does not divide $t/2$.
- Let $p_0, p_1, \ldots, p_n$ be the distinct prime divisors of $t/2$. Consider the sequence of intervals between consecutive elements of $D_t$. This sequence is periodic, starting from the first element, with period equal to $2 \times \prod_{j=0}^{n}(p_j - 1)$. The sum of the elements in any period is $t \times \prod_{j=0}^{n} p_j$.

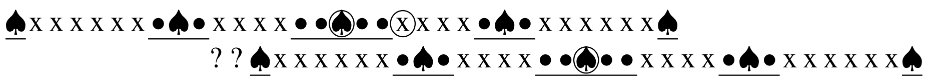
In the rest of this section we prove Theorem 6.2. We will make a heavy use of the fact that the $t$-spheres centered in any given SLB-optimal set form a partition of $\mathbb{Z}$.

For notational convenience we partition a $t$-sphere $S_t(p)$ into *stations*: contiguous clusters around the points $p + kd$, $k \in \mathbb{Z}$. We group stations into the *left branch*, the *right branch*, and (for odd $t$) the *central station*, as shown in Figure 5. In this figure, and the forthcoming figures, we represent a $t$-sphere $S_t = S_t(p)$ by a string where consecutive characters correspond to consecutive numbers. We use ♠ for the points $p + kd \in S_t$, $k \in \mathbb{Z}$, • for other points of $S_t$, and 'x' for points not in $S_t$. Stations are underlined.

## 6.1 Proof of Theorem 6.2: odd $t \leq \delta$

Let $S$ be an SLB-optimal set and let $p \in S$ be a point in $S$. Let us define $L_S = S - (\tau - 1)d$, $R_S = S + (\tau - 1)d$. It follows that $p \in L_S$ (resp. $p \in R_S$) if and only if $p$ is the leftmost (resp. rightmost) point of some $t$-sphere centered in $S$.

**Lemma 6.4.** *Suppose $S$ is an SLB-optimal set and $p \in S$. Then $p + \tau \in L_S \cup R_S$.*



- Upper row: $S_t(p)$; $p$ and $p+\tau$ are encircled.
- Lower row: $S_t(q)$; $p-d+\tau-1,\ p-d+\tau$ are labeled by '?'; $q$ is encircled.

Figure 6: For the proof of Lemma 6.4: $S_t(p)$ and $S_t(q)$ for $(d,t)=(8,5)$

*Proof.* Since the $t$-spheres centered in $S$ partition $\mathbb{Z}$, $p+\tau \in S_t(q)$ for some $q \in S$, $q \neq p$. Suppose $p+\tau$ lies in the left branch of $S_t(q)$ but is not the leftmost element thereof (see Figure 6). Then, letting $p_1 = p-d+\tau-1$, $p_2 = p_1+1$, it is easy to see that $p_1-1$ lies in $S_t(p)$, $p_2+1$ lies in $S_t(q)$, whereas $p_1$ and $p_2$ lie in neither. Thus, $p_1$ and $p_2$ are covered by some other $t$-sphere(s) centered in $S$. How can that be? In a $t$-sphere all stations except the leftmost and the rightmost ones have length $\geq 3$. Thus, $p_1$ and $p_2$ are the leftmost or the rightmost points of some $t$-spheres centered in $S$. If $p_1$ or $p_2$ is the leftmost point of such a $t$-sphere $S'$, then $S'$ intersects $S_t(p)$ at $p+\tau-1$, contradiction. So $p_1$ and $p_2$ are the rightmost elements of $t$-spheres $S_t(q_1), S_t(q_2)$ where $q_1, q_2 \in S$. Then $q_1+1=q_2$, contradiction.

So if $p+\tau$ lies in the left branch of $S_t(q)$, then $p+\tau$ must be its leftmost element, hence $p+\tau \in L_S$. Else $p+\tau$ lies in the right branch of $S_t(q)$ or in its central station. Then by a similar proof $p+\tau$ must be the rightmost element of $S_t(q)$. □

It follows that $p+t \notin S$. Indeed, if $p+t \in S$ then, since the $t$-spheres centered in $S$ are disjoint, $S_t(p+t)$ is the only $t$-sphere centered in $S$ that contains $p+\tau$. But $p+\tau$ is the inner point of $S_t(p+t)$, contradicting Lemma 6.4. Claim proved. In particular, the two-offset construction from Section 5.1 can not be SLB-optimal since it starts with $\{0, t, 2t, \ldots\}$.

**Lemma 6.5.** *Suppose $S$ is an SLB-optimal set and $p \in S$. Then:*

*(a) Exactly one of the following two statements is true (see Figure 7):*
  - *(i) $p+\tau \in L_S$ and $p-d+\tau-1 \in R_S$*
  - *(ii) $p+\tau \in R_S$ and $p+d+\tau-1 \in L_S$*

*(b) Exactly one of the following two statements is true:*
  - *(i) $p-\tau \in L_S$ and $p-d-\tau+1 \in R_S$*
  - *(ii) $p-\tau \in R_S$ and $p+d-\tau+1 \in L_S$*

*Proof.* By Lemma 6.4, either $p+\tau \in L_S$ or $p+\tau \in R_S$. Suppose $p+\tau \in L_S$ and let $p' = p-d+\tau-1$. Since $p'$ is neither in $S_t(p)$ nor in the $t$-sphere containing $p+\tau$, it is an element of some other $t$-sphere $T = S_t(q)$, $q \in S$. Since $p'$ is the leftmost element of some station of $T$, $p+\tau = p'+d+1 \in T$, unless $p'$ is the rightmost element of $T$.

Case $p+\tau \in R_S$ is solved similarly. Part (b) follows from part (a) by symmetry. □

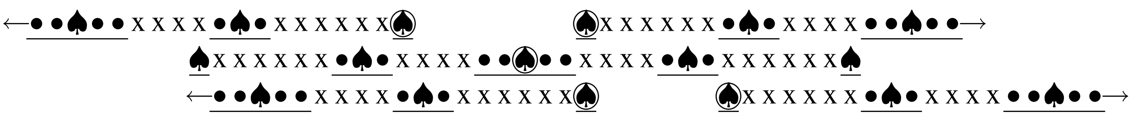
**Lemma 6.6.** *Suppose $S$ is an SLB-optimal set and $p \in S$. Then:*

*(a) exactly one of the following two statements is true (see Figure 8a):*

$$p+\tau,\ p+d-\tau+1 \in L_S \quad \text{and} \quad p-\tau,\ p-d+\tau-1 \in R_S \tag{7}$$

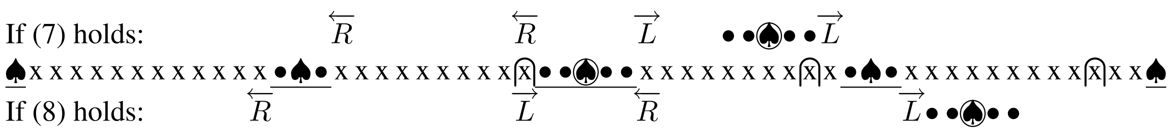
$$p+\tau,\ p-d-\tau+1 \in R_S \quad \text{and} \quad p-\tau,\ p+d+\tau-1 \in L_S \tag{8}$$

*(b) if (7) then $p+d-t \in S$, if (8) then $p+d+t \in S$ (see Figure 8b).*



- $1^{\text{st}}$ row: $p+\tau \in L_S$ and $p-d+\tau-1 \in R_S$ are labeled.
- $2^{\text{nd}}$ row: $S_t(p)$ is shown; $p$ is labeled .
- $3^{\text{rd}}$ row: $p+\tau \in R_S$ and $p+d+\tau-1 \in L_S$ are labeled.

Figure 7: The two options in Lemma 6.5, for $(d,t)=(8,5)$




(a) Middle row: $S_t(p)$ is shown, $p$ is encircled. Vertices $p \pm \tau$ and $p \pm d \pm (\tau-1)$ are labeled by $\overrightarrow{L}$ (resp. by $\overleftarrow{R}$) when they are in $L_S$ (resp. in $R_S$).

(b) Upper row: $q=p+d-t$ is encircled; the central station of $S_t(q)$ is shown. Same for $q=p+d+t$ in the lower row.

(c) Middle row: $p'-d$, $p'$ and $p'+d$ are labeled by $\bigcap$, where $p'=p+d-\tau$.

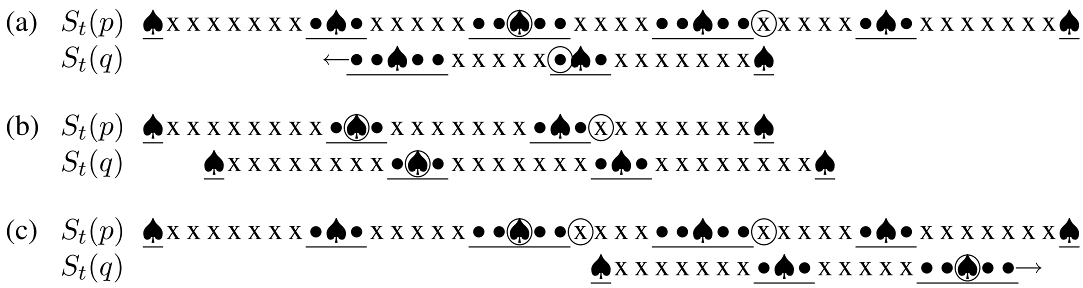
Figure 8: The two options in Lemma 6.6 for $(d,t)=(14,5)$

*Proof.* (a) By Lemma 6.5, there are four possible cases: (7), (8), $p \pm \tau \in L_S$, and $p \pm \tau \in R_S$. If $p \pm \tau \in L_S$ then by Lemma 6.5 $p_\pm \in S$, where $p_\pm = p - d\tau \pm (\tau-1)$. Since $S$ is sparse and $p_\pm \in S$, it must be the case that $\texttt{dist}(p_-, p_+) \geq t$, which contradicts the fact that $p_+ - p_- = t-1$. The case $p \pm \tau \in R_S$ is ruled out similarly.

(b) Suppose (7) holds. Let $T = S_t(q)$, $q \in S$, be the $t$-sphere containing $p' = p+d-\tau$ (Figure 8c). Say $p'$ is contained in the station $W$ of $T$. Let $W_L$, $W_R$ be the stations of $T$ immediately to the left and immediately to the right from $W$. Suppose $W$ is not the *central* station of $T$. Then either $W_L$ or $W_R$ is wider than $W$. Since $p'$ is the rightmost point of $W$, either $p-\tau \in W_L$ or $p'+d \in W_R$ (Figure 8c), so at least one of these points lie in $T$. However, we claim that both points belong to other $t$-spheres centered in $S$. Indeed, by Lemma 6.6(a) $p-\tau \in R_S$. By the same lemma $p'+1 \in L_S$ is the leftmost point of some $t$-sphere $S'$ centered in S, so $p'+d \in S'$. Claim proved. Thus, $W$ is the central station of $T$, so $q = p+d-t$.

If (8), we let $T = S_t(q)$, $q \in S$, be the $t$-sphere containing $p' = p+d+\tau$. Then by a similar argument $q = p+d+t$. □

Now we can complete the proof of the main theorem. Consider an SLB-optimal $t$-sparse set $S$ and take any $p \in S$. If (7) then by Lemma 6.6(b) $q = p+d-t \in S$. Now we apply Lemma 6.6(a) to $q$. Either (7) or (8) must hold for $q$. Since $q+\tau = p+d-\tau+1 \in L_S$, (7) does. So we apply Lemma 6.6(b) again: $q+d-t \in S$. In the same fashion, $p+k(d-t) \in S$ for any $k \in \mathbb{N}$. Since this holds for any $p \in S$, $S$ is periodic with a (not necessarily smallest) period $d-t$. Since $S$ is SLB-optimal, the density of $S$ is $w/(d-t) = 1/|S_t|$, where $w$ is the number of points of $S$ within one period. Thus, $|S_t|$ divides $d-t$. If (8) holds for $p$, then by a similar argument $|S_t|$ divides $d+t$.



(a) Labeled are: $p + \tau - 1$ in the lower row, $p$ and $p'$ in the upper row. Here $(d, t) = (9, 6)$.
(b) Labeled are: $p$, $p'$ in the upper row, $q$ in the lower row. Here $(d, t) = (10, 4)$.
(c) Labeled are: $p$, $p + \tau$, $p'$ in the upper row, $q$ in the lower row. Here $(d, t) = (9, 4)$.

Figure 9: For the proof of Lemma 6.7

### 6.2 Proof of Theorem 6.2: even $t \leq \delta$

The theorem follows from the following lemma:

**Lemma 6.7.** *Let $S$ be an SLB-optimal set containing 0. Then at least one of $\tau(d + 1)$, $-\tau(d - 1)$ is in $S$.*

*Proof.* In the lemma statement, we chose the "reference point" $0 \in S$. To clarify the proof, we write it for an arbitrary reference point $p \in S$.

The long $t$-spheres centered $S$ partition $\mathbb{Z}$. In particular, $p' = p + d + \tau$ is an element of some $t$-sphere $T = S_t(q)$, $q \in S$. Clearly, $p'$ is a leftmost element of some station of $T$. Which station? If $p'$ is in the right branch of $T$ then either $p + \tau - 1$ is in both $T$ and $S_t(p)$ (Figure 9a), or $q = p + t - 1$, which is too close to $p$ (Figure 9b).

So $p'$ lies in the left branch of $T$. Now, if $p'$ is the leftmost element of $T$ then $q = p + \tau(d + 1) \in S$, and we are done. Else $p + \tau - 1 \in S_t(p)$, $p + \tau + 1 \in T$, but $p + \tau$ is in neither $t$-sphere (Figure 9c). So $p + \tau$ must be either the leftmost or the rightmost element of some other $t$-sphere $T' = S_t(q')$, $q' \in S$. It cannot be the leftmost element since in this case $p + d + \tau - 1$ is in both $T'$ and $S_t(p)$. Thus, it is the rightmost element of $T'$, in which case $q' = p - \tau(d - 1)$. □

To prove Theorem 6.2, let $w = |S_t| = t^2/2$ and assume that the even construction $S = w\mathbb{Z}$ is SLB-optimal, i.e. that $S$ is $t$-sparse. Then by Lemma 6.7 either $(d+1)\tau \in S$ or $(d-1)\tau \in S$, so $w$ divides either $(d + 1)\tau$ or $(d - 1)\tau$, which implies $d \equiv \pm 1 \pmod t$.

## 7 Greedy approach

A natural way to construct $t$-sparse sets is the following *greedy algorithm*. We start with an empty set $S$ and $j = 0$. For each consecutive $j$, we insert $j$ into $S$ if and only if the resulting set $S \cup \{j\}$ is $t$-sparse. Since this decision depends only on the *header* $S \cap [j - dt; j]$, and (up to translation) there are only finitely many possible headers, the construction is periodic starting from some $m$ (i.e. for some $p$ and all $n \geq m$ it is the case that $n \in S$ if and only if $n + p \in S$). Therefore without loss of generality the algorithm can stop as soon as the period is detected. We define *the greedy construction* to be the set obtained by replicating the period in both directions.

Obviously, the greedy construction is $t$-sparse. In this construction, each element is as close as possible to the smaller elements, which makes one hope that it is dense enough. However, it may be the case that if we make some intervals larger, some subsequent intervals can be made shorter, thus increasing the overall density.

We found that the greedy algorithm often produces reasonable results:

**Theorem 7.1.** *Fix graph $G_d$ and $t \in \mathbb{N}$. We distinguish three cases (as in Theorem 1.2):*

*(a) Suppose $t \geq d-1$. Let $k = (t-\delta)d + \delta$. Then the greedy construction is $k\,\mathbb{Z}$, the $t$-sparse set constructed in Theorem 1.2(a).*

*(b) Suppose $t \leq \delta$. Then the greedy construction is a two-offset set if and only if $d \equiv 0, \pm 1 \pmod t$, in which case it is exactly the two-offset construction from Definition 5.4.*

*(c) Suppose $\delta < t \leq d-2$. Then the greedy construction is a 2-periodic set which is optimal when either $d$ or $t$ are odd, and near-optimal if both of them are even.*

Note that in all cases when we can prove something about the greedy algorithm, we also show that it does not improve over Theorem 1.2. Also, note that we do not have a characterization for the greedy construction in the case when $t \leq \delta$ but $d \not\equiv 0, \pm 1 \pmod t$. Indeed, computer searches show that in this case the greedy construction is quite ugly: the periods are rather long and lack apparent structure.

**Proof Sketch of Theorem 7.1:** Part (a) is trivial since by Claim 3.3 the interval between any two consecutive points in a $t$-sparse set is at least $k$, and by Theorem 1.2(a) the set $k\,\mathbb{Z}$ is $t$-sparse.

To prove part (c), let us recall Definition 4.1 and Definition 4.2. By definition of $\sigma_0$, $(v_0, \sigma_0 - v_0)$ is a remote pair. Moreover, for odd $d$ this is a standard remote pair (by definition thereof). By Theorem 4.3(b) the set $S$ induced by this pair is an optimal $t$-sparse set. Now, it is easy to see that the greedy construction is actually $S$: indeed, by definition of $\sigma_0$ the first two elements found by the greedy algorithm are $v_0$ and $\sigma_0$; the rest follows since $S$ is $t$-sparse. For even $d$, the set $S$ is $t$-sparse by an argument similar to that in Theorem 4.3(b), so the greedy construction again coincides with $S$. Now if $\sigma_0 < \sigma_1$ then (again) $(v_0, \sigma_0 - v_0)$ is a standard remote pair, so $S$ is optimal. But if $\sigma_{\text{min}} = \sigma_1 < \sigma_0$, then the above pair is not standard, in which case the index of $S$ is slightly above the lower bound of $\sigma_{\text{min}}/2$.

For part (b), we will use the notation from Section 5.1, namely the pair $(q, r)$ defined in (5), the period $p^*$ from Definition 5.4, and some notation from the proof of Lemma 5.5, in particular $T = \{0, t, 2t\,,\; \dots\;, qt\}$. The greedy algorithm starts out with an empty set $S$, then proceeds to $S = T$. Let $w$ be the next number inserted into $S$. It is easy to see that for $r \leq 1$ we have $w = p^*$, so the 'if' direction follows since the two-offset construction is $t$-sparse.

Now assume $r \geq 2$. For the converse it suffices to show that the point $w + t$ is not $T$-remote. Let $\eta = \lfloor \frac{t-r}{2} \rfloor$ and recall the definitions of $I_{ij}$ and $J_i$ from the proof of Lemma 5.5. If $\eta < \tau - 1$ then $w \in I_{\eta q}$, so the point $w + t$ is not $T$-remote since it lies in the interval $I_{(\eta+1,\,0)}$ which is not free. Else it is the case that $\eta = \tau - 1$, $r = 2$ and $t$ is even, so the point $w + t$ is again not $T$-remote because the only two $T$-remote points in $J_\eta$ are $w$ and $w + 1$, and the only $T$-remote point in $I_{(\tau,\,1)}$ is $w + t + 1$. □

Recall the greedy algorithm described above starts with an empty set $S$. This choice is quite arbitrary; instead, we can let the greedy algorithm start with any $t$-sparse set $S$ such that $\max(S) < 0$. Same as before, it can be seen that the resulting infinite set is periodic starting from some $m$, so we can define the greedy construction induced by $S$ as the set obtained by replicating the period in both directions. It is an open question how the structure (and the index) of such sets depends on $S$.

## 8 Conclusions

We consider $t$-interleaving schemes on infinite circulant graphs with two offsets $\{1, d\}$. For each pair $(d, t)$ we construct a $t$-interleaving scheme whose degree is optimal or close to optimal for this pair. Our approach is to find a candidate label-set with a large density, and then to cover $\mathbb{Z}$ with a minimal number of copies thereof. Most of our progress is on minimizing the index of a label-set (the inverse of its density, rounded up), which is itself an interesting combinatorial problem. Our interleaving schemes have a very simple, periodic structure.

Two natural directions for future research would be interleaving schemes with repetitions and interleaving schemes on general circulant graphs.

**Acknowledgments.** The authors are most grateful to the anonymous referees for their comments.

## References

[1] K.A.S. Abdel-Ghaffar, R.J. McEliece and H.C.A. van Tilborg, ”Two-dimensional burst identification codes and their use in burst correction,” *IEEE Trans. on Information Theory*, vol. IT-34, no. 3, pp. 494-504, May 1988.

[2] K.A.S. Abdel-Ghaffar, ”Achieving the Reiger bound for burst errors using two-dimensional interleaving schemes,” In Proc. *IEEE Intl. Symp. on Information Theory (ISIT)*, 1997, p. 425.

[3] C. Almeida and R. Palazzo, ”Two-dimensional interleaving using the set partition technique,” In Proc. *IEEE Intl. Symp. on Information Theory (ISIT)*, 1994, p. 505.

[4] Y. Ben-Haim and T. Etzion, ”On the optimality of coloring with a lattice,” *SIAM J. on Discrete Mathematics*, 18 (2005), pp. 844–878.

[5] M. Blaum and J. Bruck, ”MDS array codes for correcting criss-cross errors,” *IEEE Trans. on Information Theory*, vol. 46, no. 4, pp. 1068-1077, May 2000.

[6] M. Blaum, J. Bruck and P.G. Farrell, “Two-dimensional interleaving schemes with repetitions,” *IBM Research Report RJ 10047*, Oct. 1996.

[7] M. Blaum, J. Bruck and A. Vardy, “Interleaving schemes for multidimensional cluster errors,” *IEEE Trans. on Information Theory*, vol. 44, pp. 730–743, 1998.

[8] M. Blaum and P.G. Farrell, ”Array codes for cluster-error correction,” *Electron. Lett.*, vol. 30, no. 21, pp. 1752-1753, Oct. 1994.

[9] M. Breitbach, M. Bossert, V. Zyablov and V. Sidorenko, ”Array codes correcting a two-dimensional cluster of errors,” *IEEE Trans. on Information Theory*, vol. 44, no. 5, pp. 2025-2031, Sep. 1998.

[10] B. Elspas and J. Turner, “Graphs with circulant adjacency matrices,” *J. Combin. Theory*, No.9, pp. 297–307, 1970.

[11] T. Etzion and A. Vardy, “Two-dimensional interleaving schemes with repetitions: constructions and bounds,” *IEEE Trans. on Information Theory*, vol.48, pp. 428–458, 2002.

[12] P.G. Farrell, ”Array codes for correcting cluster-error patterns,” in Proc. *IEE Conf. Electronic Image Processing*, York, U.K., July 1982.

[13] E.M. Gabidulin, ”Theory of codes with maximum rank distance,” *Probl. Pered. Inform.*, vol. 21, pp. 3-16, 1985.

[14] C. Gu, J. Hong, I. McMichael, R. Saxena and F. Mok, ”Cross-talk limited storage capacity of volume holographic memory,” *J. Opt. Soc. Amer.*, vol. 9A, pp. 1978-1987, 1992.

[15] J.F. Heanie, M. C. Bashaw, and L. Hesselink, ”Volume holographic storage and retrieval of digital data,” *Science*, vol. 265, pp. 749-752, 1994.

[16] H. Imai, ”Two-dimensional Fire codes,” *IEEE Trans. on Information Theory*, vol. IT-19, no. 6, pp. 796-806, Nov. 1973.

[17] H. Imai , ”A theory of two-dimensional cyclic codes,” *Inf. Control*, vol. 34, pp. 1-21, 1977.

[18] A. Jiang and J. Bruck, ”Multicluster Interleaving on Paths and Cycles,” *IEEE Trans. on Information Theory*, vol. 51, no. 2, pp. 597–611, February 2005.

[19] A. Jiang, M. Cook and J. Bruck ”Optimal t-Interleaving on Tori,” *SIAM J. on Discrete Mathematics*, vol. 20, no. 4, pp. 841-879, 2006.

[20] M.A. Neifeld and M. McDonald, ”Error correction for increasing the usable capacity of photorefractive memories,” *Opt. Lett.*, vol. 19, pp. 1483–1485, 1994.

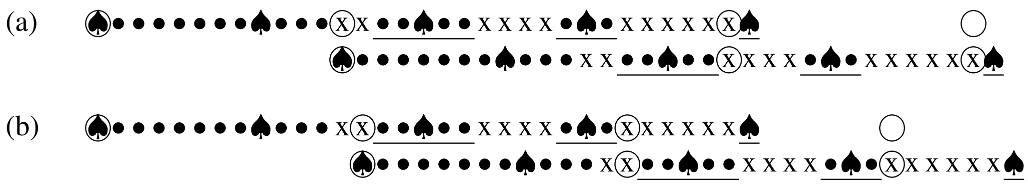
[21] R.M. Roth, ”Maximum-rank arrays codes and their application to crisscross error correction,” *IEEE Trans. on Information Theory*, vol. 37, no. 2, pp. 328-336, Mar. 1991.

[22] R.M. Roth, ”Probabilistic criss-cross error correction,” *IEEE Trans. on Information Theory*, vol. 43, no. 5, pp. 1425-1438, Sep. 1997.

[23] M. Schwartz and T. Etzion, ”Optimal 2-dimensional 3-dispersion lattices,” In Proc. *15th Intl. Symp. on Applied Algebra, Algebraic Algorithms and Error-Correcting Codes (AAECC)*, May 2003, pp. 216-225. *Lecture Notes in Computer Science*, vol. 2643, Springer-Verlag, Berlin, 2003.

[24] A. Slivkins and J. Bruck, ”Interleaving Schemes on Circulant Graphs,” Electronic Technical Report *ETR046*, Parallel and Distributed Systems Group, California Inst. of Technology, November 2002.

[25] W. Xu and S. W. Golomb, ”Optimal interleaving schemes for correcting 2-d cluster errors, In Proc. *IEEE Intl. Symp. on Information Theory (ISIT)*, 2004.



Example (a) and (b) represent $\sigma_0$ and $\sigma_1$, respectively. In each example, the upper line is the $t$-span of 0, and the lower line is the $t$-span $v_i$, both in the notation of Figure 2. Then $\sigma_i$ is the leftmost point that is remote in both lines. Points 0, $v_i$, $\sigma_i$ and $\sigma_i + v_i$ are encircled.

Figure 10: $\sigma_0$ and $\sigma_1$ for $(d, t) = (8, 5)$

## Appendix A: Full proof of Theorem 4.3

Recall that $\delta < t \leq d - 2$, $\delta = \lceil d/2 \rceil$ and $\tau = \lceil t/2 \rceil$.

Let $1_{\{\text{even } d\}}$ equal 1 if $d$ is even, and 0 otherwise. For each $i \in \{0, 1\}$ define $v_i = v_{\text{min}} + i\, 1_{\{\text{even } d\}}$, and let $\sigma_i$ be the minimal sum of a remote pair of the form $(v_i, \cdot)$; see Figures 3 and 10 for some intuition. We start with the following technical claims.

**Claim A.1.** *For any $v \in \mathbb{N}$ such that $0 < v < v_{max}$, the following are equivalent:*

*(a) $v$ is remote.*

*(b) the canonical representation of $v - v_{min}$ is a pair $(\mu_1, \mu_2)$ such that $-\mu_1 \leq \mu_2 \leq \mu_1 + 1_{\{even\ d\}}$.*

*(c) $v = v_i + \mu_1 d + \mu_2$, for some $i \in \{0, 1\}$ and $(\mu_1, \mu_2)$ such that $\mu_1 \leq 0$ and $|\mu_2| \leq \mu_1$.*

**Claim A.2.** *For each $i \in \{0, 1\}$ we have:*

*(a) $v_{min} + v_i = \alpha d + \gamma_i$ where $\alpha = 2(t - \delta) + 1$ and $\gamma_i \in \{0, 1\}$, specifically* $\gamma_i = \begin{cases} 1, & \text{if } d \text{ is odd,} \\ i, & \text{if } d \text{ is even.} \end{cases}$

*(b) $\sigma_i - v_i = v_{min} + \alpha_i(d + 1) + 1_{\{even\ d\}}$, where*

$$\alpha_i = \begin{cases} \delta - \tau - i, & \text{if both } d \text{ and } t \text{ are even,} \\ \delta - \tau & \text{otherwise.} \end{cases} \tag{9}$$

*(c) $\sigma_i = (\alpha + \alpha_i)\, d + (\alpha_i + \gamma_i)$, where $\gamma_i \leq \alpha_i + \gamma_i \leq t/2$. In particular, $2\sigma_i \geq (t + 2)d$.*

*Proof.* Part (a) is an easy computation which we omit.

Let us prove part (b). For each $i \in \{0, 1\}$ and any $\sigma \geq v_{\text{min}} + v_i$, let $(\alpha_c(\sigma),\ \beta_c(\sigma))$ be the canonical representation of $\sigma - v_{\text{min}} - v_i$. Let $\alpha_i = \alpha_i(\sigma_i)$ and $\beta_i = \beta_i(\sigma_i)$. Let

$$W_i = \{\sigma > v_{\text{min}} + v_i \mid -\alpha_c(\sigma) \leq \beta_c(\sigma) \leq \alpha_c(\sigma) + 1_{\{\text{even } d\}} \text{ and } -\delta < \gamma_i + \beta_c(\sigma) \leq \delta\} \tag{10}$$

First, we claim that $\sigma_i \in W_i$. Indeed, $\sigma_i - v_i$ is remote by definition of $\sigma_i$, thus Claim A.1(b) says precisely that for $\sigma = \sigma_i$ the first condition in (10) holds. Also, by definition of canonical representation we have $-\delta < \beta_i \leq \gamma_i + \beta_i$. So it remains to show that $\gamma_i + \beta_i \leq \delta$. Suppose not. Then $\gamma_i = 1$ and $\beta_i = \delta$. By part (a) we have

$$\sigma_i = (v_{\text{min}} + v_i) + (\alpha_i d + \delta) = (\alpha + \alpha_i)d + (\delta + 1).$$

It follows that $\sigma_i - 1$ is remote. Moreover, since the first condition in (10) holds for $\sigma = \sigma_i$, it also holds for $\sigma = \sigma_i - 1$, so by Claim A.1(b) $\sigma_i - 1 - v_i$ is remote. Therefore $(v_i, \sigma_i - 1)$ is a remote pair, which contradicts the minimality of $\sigma_i$. Claim proved.

Second, we claim that $\sigma_i$ is the *smallest* remote element of $W_i$. Indeed, assume $\sigma \in W_i$ for some remote $\sigma < \sigma_i$. Then applying Claim A.1(b) for $v = \sigma - v_i$ it follows that $\sigma - v_i$ is remote (the condition in Claim A.1(b) is precisely the first condition in (10)). Therefore $(v_i, \sigma - v_i)$ is a remote pair, contradicting the minimality of $\sigma_i$. Claim proved.

Therefore

$$\alpha_i = \min\{x \mid \varphi(x) \geq t\} \text{ where } \varphi(x) = \max\{\mathtt{dist}(\sigma) \mid \sigma \in W_i \text{ and } \alpha_c(\sigma) = x\}. \tag{11}$$

For a given $\sigma \in W_i$, by part (a) and the definition of $(\alpha_c(\sigma), \beta_c(\sigma))$ we have

$$\sigma = (v_{\min} + v_i) + \alpha_c(\sigma)d + \beta_c(\sigma) = (\alpha + \alpha_c(\sigma))\, d + (\gamma_i + \beta_c(\sigma)). \tag{12}$$

By the second condition in (10), the right-hand side of (12) gives the canonical representation of $\sigma$. Therefore by Claim 3.2 it is the case that $\mathtt{dist}(\sigma) = (\alpha + \alpha_c(\sigma)) + |\gamma_i + \beta_c(\sigma)|$, which is maximized, for a fixed $\alpha_c(\sigma)$, only if $\beta_c(\sigma) = \alpha_c(\sigma) + 1_{\{\text{even } d\}}$. It follows that $\beta_i = \alpha_i + 1_{\{\text{even } d\}}$ and moreover $\varphi(x) = 2x + \alpha + \gamma_i + 1_{\{\text{even } d\}}$. Therefore solving (11) for $\alpha_i$ gives (9), and part (b) follows.

Part (c) is an easy corollary of parts (ab). Specifically, we obtain the expression for $\sigma_i$ by plugging part (a) into part (b). We obtain the inequality for $\alpha_i + \gamma_i$ by going through all four possible parities of $(d, t)$. Finally, plugging in the values for $\alpha$ and $\alpha_i$, an easy computation shows that

$$2\sigma_i \geq 2(\alpha + \alpha_i)d \geq (3t - 2\delta)d \geq (t + 2)d. \quad \square$$

Now we are ready to prove the theorem:

**Proof of Theorem 4.3:** **(a)** Let us denote the minimal sum of a remote pair by $\sigma$. First we claim that $\sigma = \sigma_{\min}$. Indeed, Let $(w_1, w_2)$, $w_1 \leq w_2$ be a remote pair with a sum $\sigma$ such that $w_1$ is minimal. If $w_1 \in \{v_1, v_2\}$ then $\sigma = \sigma_{\min}$ by definition of $\sigma_{\min}$. Else we can choose $z > 0$ so that $(w_1 - z, w_2 + z)$ is a remote pair, contradicting the minimality of $w_1$. Specifically, we let $(x_1, y_1)$ be the canonical representation of $w_1$, and we choose $z \in [d-1; d+1]$ as follows. If $\mathtt{dist}(w_1) > t$ let $z = d$; else we let $z = d - 1$ if $y_1 > 0$, and $z = d + 1$ otherwise. Then $w_1 - z$ and $w_2 + z$ are remote by Claim **??**. Claim proved.

Let $S$ be a $t$-sparse set with a well-defined density $\rho$. Let $\{s_i : i \in \mathbb{Z}\}$ be an increasing enumeration of $S$. For each $i$, $(s_{i+1} - s_i,\ s_{i+2} - s_{i+1})$ is a remote pair, so its sum $s_{i+2} - s_i$ is at least $\sigma_{\min}$. Then $s_n - s_{-n} \geq n\sigma_{\min}$ for any $n > 0$, so $\rho \leq 2/\sigma_{\min}$, which gives the required lower bound on the index of $S$. By Lemma 1.1 this implies a similar lower bound on $\mathtt{degree}(d, t)$.

**(b)** Let $S$ be the set induced by a standard remote pair $(w_1, w_2)$. For any $u, v \in S$, either

$$|u - v| \in \{0,\ w_1,\ w_2,\ \sigma_{\min},\ \sigma_{\min} + w_1,\ \sigma_{\min} + w_2\}$$

or else $|u - v| \geq 2\sigma_{\min}$. In the latter case, since by Claim A.2(c) we have $2\sigma_{\min} \geq (t+2)d > v_{\max}$, it follows that $\mathtt{dist}(v, u) \geq t$. Therefore it remains to prove that $\sigma_{\min} + w_1$ and $\sigma_{\min} + w_2$ are remote.

We will in fact prove that $\sigma_i + w_j$ is remote for any $i, j \in \{0, 1\}$. Indeed, recall that by Claim A.2(c) the canonical representation of each $\sigma_i$ is $(\cdot,\ y)$ for some $y \in [0; t/2]$. Now part (b) follows from the following general claim:

**Claim A.3.** *If $u^*, v^* \in \mathbb{N}$ are remote, and moreover $u^* = xd + y$ so that $|y| \leq t/2$, then $u^* + v^*$ is remote.*

*Proof.* The claim is obvious if $v^* \geq v_{\max}$. Suppose $v^* < v_{\max}$. Then by Claim A.1(c) we have

$$u^* + v^* = v_i + (x + \mu_1)d + (y + \mu_2)$$

for some $i \in \{0, 1\}$ and $|\mu_2| \leq \mu_1$. Since $\mathtt{dist}(u^*) = x + |y| \geq t$ it follows that $x \geq t/2 \geq |y|$, so $|y + \mu_2| \leq x + \mu_1$ and by Claim A.1(c) $u^* + v^*$ is remote. $\square$

**(c)** Consider a remote pair $(v_i, \sigma_i - v_i)$. We claim that for each $j \leq t - \delta$

$$(v_i + j(d+1),\ \ \sigma_i - v_i - j(d+1))$$

is a remote pair, too. The sum of this pair is $\sigma_i$, so we just need to show that both numbers in this pair are remote. Indeed,

$$v_i + j(d+1) = v_{\text{min}} + jd + (j + i\,\mathtt{1}_{\{\text{even } d\}}),$$

so by Claim A.1(b) it is remote. By Claim A.2(b) we have

$$\sigma_i - v_i - j(d+1) = v_{\text{min}} + (\alpha_i - j)d + (\alpha_i - j + \mathtt{1}_{\{\text{even } d\}}),$$

so by Claim A.1(b) it is remote, too. Claim proved.

Now let us consider two cases depending on the parity of $d$ and $t$. First we consider the case when $d$ is odd or $t$ is even. Note that by Claim A.2(b) in this case we have $\sigma_1 \leq \sigma_0$, so $\sigma_{\text{min}} = \sigma_1$. Also, by Claim A.2(b) we have $\sigma_1 - 2v_1 = \alpha_1(d+1)$, where $\alpha_1$ is given by (9). Therefore there exists a standard remote pair of the form $(w, w)$ if $\alpha_1$ is even, and $(w, w+d+1)$ if $\alpha_1$ is odd. Let $S$ be the set induced by such a pair. If $\alpha_1$ is even, then $S = w\,Z$ induces an optimal interleaving scheme. Now suppose $\alpha_1$ is odd. With some arithmetic one can show that

$$g := \gcd(w, w+d+1) = \gcd(t, d+1).$$

By part (a) and Claim A.2(c) we have

$$\mathtt{degree}(d,t) \geq \lceil \sigma_{\text{min}}/2 \rceil \geq (t+2)d/4 > gd/4.$$

By Lemma 4.6 set $S$ induces an interleaving scheme of degree

$$\mathtt{deg}(S) = g\lceil \sigma_{\text{min}}/2g \rceil \leq \lceil \sigma_{\text{min}}/2 \rceil + g \leq \mathtt{degree}(d,t) \times (1 + \tfrac{4}{d}),$$

so this interleaving scheme is $(1+\frac{4}{d})$-approximate. Moreover, if both $d$ and $t$ are odd then this interleaving scheme is in fact optimal. Indeed, in this case $\sigma_{\text{min}} = 2w + d + 1$ is even and $g$ is odd, so $2g | \sigma_{\text{min}}$ and therefore $\mathtt{deg}(S) = \lceil \sigma_{\text{min}}/2 \rceil$, matching the lower bound from part (a).

Finally consider the case when $d$ is even and $t$ is odd. Then $\alpha_0 = \alpha_1$, but $\sigma_1 = \sigma_0 + 1$, so we carry out a similar argument for $(v_0, \sigma_0 - v_0)$ and prove that there is a standard remote pair of the form $(w, w+1)$ if $\alpha_0$ is even and $(w, w+d+2)$ if $\alpha_0$ is odd. By Lemma 4.6 the former case extends to an optimal interleaving scheme, whereas the latter yields a $(1+\frac{4}{d})$-approximation. $\square$