Access Balancing in Storage Systems by Labeling Partial Steiner Systems

Y.M. Chee*, C.J. Colbourn†, H. Dau‡, R. Gabrys§, A.C.H. Ling¶, D. Lusi†, O. Milenkovic§
*National University of Singapore, pvocym@nus.edu.sg
†Arizona State University, {colbourn,dlusi}@asu.edu
‡RMIT University, sonhoang.dau@rmit.edu.au
§University of Illinois at Urbana-Champaign (UIUC), {gabrys, milenkovic}@uiuc.edu
¶University of Vermont, aling@emba.uvm.edu

***Abstract*—Storage architectures ranging from minimum bandwidth regenerating encoded distributed storage systems to declustered-parity RAIDs can employ dense partial Steiner systems to support fast reads, writes, and recovery of failed storage units. To enhance performance, popularities of the data items should be taken into account to make frequencies of accesses to storage units as uniform as possible. A combinatorial model ranks items by popularity and assigns data items to elements in a dense partial Steiner system so that the sums of ranks of the elements in each block are as equal as possible. By developing necessary conditions in terms of independent sets, we demonstrate that certain Steiner systems must have a much larger difference between the largest and smallest block sums than is dictated by an elementary lower bound. In contrast, we also show that certain dense partial $S(t, t+1, v)$ designs can be labeled to realize the elementary lower bound. Furthermore, we prove that for every admissible order v, there is a Steiner triple system ($S(2, 3, v)$) whose largest difference in block sums is within an additive constant of the lower bound.**

A full version [1] *of this paper is accessible at:* `https://arxiv.org/abs/1906.12073`.

I. Introduction

Distributed storage systems [2], [3], systems for batch coding [4], and multiserver private information retrieval systems [5] have each employed combinatorial designs for data placement, so that elements of the design are associated with data items and blocks with storage units. In these contexts, the most common types of designs employed are t-designs and t-packings. A t-(v, k, λ) *packing* is a pair $(X, \mathcal{B})$, where X, the point set, is a v-set and and $\mathcal{B}$ is a collection of k-subsets (*blocks*) of X such that every t-subset of X is contained in at most λ blocks. The packing is a t-(v, k, λ) *design* when every t-subset of X is a subset of exactly λ blocks. A t-$(v, k, 1)$ design is a *Steiner system*, denoted by $\mathrm{S}(t, k, v)$. A 2-$(v, 3, 1)$ design is a *Steiner triple system of order v*, denoted by $\mathrm{STS}(v)$. When $\lambda = 1$, a t-$(v, k, 1)$ packing is also referred to as a *partial* $S(t, k, v)$ or *partial Steiner system*.

When data items are of the same size, and data is placed on storage units using a t-design, placement of data is uniform across the storage units. Indeed in t-(v, k, λ) design, every point appears in exactly $r = \frac{\lambda\binom{v-1}{t-1}}{\binom{k-1}{t-1}}$ blocks; this is the *replication number* of the design. In order to understand why Steiner systems can be employed in data placement, we outline some examples. Large-scale distributed storage systems (DSS) must address potential loss of storage units, while not losing data. One solution is to replicate each data item and distribute these replicas among multiple storage nodes; systems such as the Hadoop Distributed File System and the Google File System employ this strategy [6]. One can further mitigate information loss by sensibly organizing the data. One example is the exact Minimum Bandwidth Regenerating (MBR) codes [2], consisting of two subcodes of which one is a Steiner system.

Steiner systems also prove useful for applications needing both high data availability and throughput, such as transaction processing. The storage systems underlying these applications require uninterrupted operation, satisfying user requests for data even in the event of disk failure and repairing these failed disks, on-line, in parallel. Continuous operation alone is not sufficient, because such systems cannot afford to suffer significant loss of performance during disk failures. *Declustered-parity RAIDS* (DPRAIDs) are designed to satisfy these requirements [7]–[9]. One can represent a DPRAID as a t-(v, k, λ) design $(X, \mathcal{B})$, with X ($|X| = v$) being the set of disks in the array, and $\mathcal{B}$ being the set of all parity stripes, each of size k. Then each disk occurs in the same number c of parity stripes, distributing the reconstruction effort evenly.

Although designs arise naturally in balancing data placement, little attention has been paid to the relative popularity of the data items. Dau and Milenkovic [10] formulate a number of problems to address access balancing, by labeling the points of the underlying design. In order to introduce their problems and results, we present more definitions and known results concerning designs.

Although storage systems handle "hot" (frequently accessed) and "cold" (infrequently accessed) data categories differently, typically they do not take the long-term popularity of the data items within each category into account, which may result in unbalanced access frequencies to the storage units. Access balancing can be achieved in part by selecting an appropriate packing or design, and by appropriate association of data items with elements of the packing or design. Dau and Milenkovic [10] propose a combinatorial model that ranks data items by popularity, and then strives to ensure that the sums of the ranks of the data elements in each block are not too small, not too large, or not too different from block to block. In Section II we summarize their model, state elementary bounds on various block sums, and provide a small but important improvement in the lower bound on the smallest possible difference among the block sums in a Steiner triple system. In Section III we establish a close connection between such block sums and the size of a maximum independent set of elements in the packing or design. For certain designs, this connection can be used to show that, no matter how data items are associated with the elements of the design, the block sums must be far from the values dictated by the elementary bounds from Section II. Indeed, in order to approach the elementary bounds, one must select designs or packings with very specific properties; we pursue this in Section IV. Our results indicate the need to find specific $S(t,k,v)$ designs, or at least 'dense' t-$(v,k,1)$ packings, to match the elementary bounds more closely. In Section V, we explore a construction of t-$(v,t+1,1)$ packings that asymptotically match the bounds and contain almost the same number of blocks as the full Steiner system $S(t,t+1,v)$. Completion of the dense t-$(v,t+1,1)$ packings to a Steiner system $S(t,t+1,v)$ appears problematic for general t; doing so without dramatically changing the block sums appears to be even more challenging. Nevertheless, in Section VI, we pursue this to establish, for every admissible order v, the existence of a Steiner triple system of order v whose difference in block sums is at most an additive constant more than the elementary lower bound.

II. Point Labelings and Block Sums

Let $D = (V, \mathcal{B})$ be a t-(v,k,λ) packing. A *point labeling* of D is a bijection $\mathsf{rk} : V \mapsto \{0,\dots,v-1\}$; our interpretation is that rk maps an element to its rank by popularity. The *reverse* $\overline{\mathsf{rk}}$ of a point labeling rk has $\overline{\mathsf{rk}}(i) = v-1-\mathsf{rk}(i)$ for each $i \in \{0,\dots,v-1\}$; the *reversal* of a point-labeled packing is one having the reverse of the point labeling. With respect to a specific point labeling rk, define $\mathsf{sum}(B,\mathsf{rk}) = \sum_{x\in B}\mathsf{rk}(x)$ when $B \in \mathcal{B}$. Then define

$$\begin{aligned}\mathsf{MinSum}(D,\mathsf{rk}) &= \min(\mathsf{sum}(B,\mathsf{rk}) : B \in \mathcal{B});\\ \mathsf{MaxSum}(D,\mathsf{rk}) &= \max(\mathsf{sum}(B,\mathsf{rk}) : B \in \mathcal{B});\\ \mathsf{DiffSum}(D,\mathsf{rk}) &= \mathsf{MaxSum}(D,\mathsf{rk}) - \mathsf{MinSum}(D,\mathsf{rk});\\ \mathsf{RatioSum}(D,\mathsf{rk}) &= \mathsf{MaxSum}(D,\mathsf{rk})/\mathsf{MinSum}(D,\mathsf{rk}).\end{aligned}$$

Following [10], one primary objective is to choose point labelings to maximize the MinSum and/or to minimize one of the other three. Each of these metrics is concerned with the worst case; to treat the average case, Yu *et al.* [11] study the minimum *variance* of the point sums. Access balancing is concerned primarily with minimizing the $\mathsf{DiffSum}$ or $\mathsf{RatioSum}$; because of the similarity between these two entities we often focus on the $\mathsf{DiffSum}$. Let $\mathcal{R}_D$ denote the set of all point labelings of D. Noting that $\mathsf{MaxSum}(D,\mathsf{rk}) = k(v-1) - \mathsf{MinSum}(D,\overline{\mathsf{rk}})$, we define

$$\begin{aligned}\mathsf{MinSum}(D) &= \max(\mathsf{MinSum}(D,\mathsf{rk}) : \mathsf{rk} \in \mathcal{R}_D);\\ \mathsf{MaxSum}(D) &= k(v-1) - \mathsf{MinSum}(D);\\ \mathsf{DiffSum}(D) &= \min(\mathsf{DiffSum}(D,\mathsf{rk}) : \mathsf{rk} \in \mathcal{R}_D);\\ \mathsf{RatioSum}(D) &= \min(\mathsf{RatioSum}(D,\mathsf{rk}) : \mathsf{rk} \in \mathcal{R}_D).\end{aligned}$$

If the storage system dictates the data layout and data items have the same size, we are free to permute the data items; this is captured by the selection of the point labeling rk. If we are also free to choose the t-$(v,k,1)$ packing that underlies the data layout, we may select a packing to improve the sum metrics defined. In order to capture this, let $\mathcal{D}_{t,k,v,b}$ denote the set of all t-$(v,k,1)$ packings having exactly b blocks. Then define

$$\begin{aligned}\mathsf{MinSum}(t,k,v,b) &= \max(\mathsf{MinSum}(D) : D \in \mathcal{D}_{t,k,v,b});\\ \mathsf{MaxSum}(t,k,v,b) &= k(v-1) - \mathsf{MinSum}(t,k,v,b);\\ \mathsf{DiffSum}(t,k,v,b) &= \min(\mathsf{DiffSum}(D) : D \in \mathcal{D}_{t,k,v,b});\\ \mathsf{RatioSum}(t,k,v,b) &= \min(\mathsf{RatioSum}(D) : D \in \mathcal{D}_{t,k,v,b}).\end{aligned}$$

When $b = \frac{\binom{v}{t}}{\binom{k}{t}}$, the packing is a Steiner system $S(t,k,v)$; in these cases we omit b from the notation to get $\mathsf{MinSum}(t,k,v)$ and similarly for all other entities.

Theorem 1. [10]

$$\begin{aligned}\mathsf{MinSum}(t,k,v) &\le \tfrac{1}{2}(v(k-t+1)+k(t-2));\\ \mathsf{MaxSum}(t,k,v) &\ge \tfrac{1}{2}(v(k+t-1)-kt);\\ \mathsf{DiffSum}(t,k,v) &\ge (v-k)(t-1);\\ \mathsf{RatioSum}(t,k,v) &\ge \tfrac{v(k+t-1)-kt}{v(k-t+1)+k(t-2)}.\end{aligned}$$

When D is an $S(t,t+1,v)$, $\mathsf{MinSum}(D) \le (v-1)+\binom{t}{2}$, $\mathsf{MaxSum}(D) \ge t(v-1)-\binom{t}{2}$, $\mathsf{DiffSum}(D) \ge (t-1)(v-t-1)$, *and* $\mathsf{RatioSum}(D) \ge \mathsf{RatioSum}(t,t+1,v) \ge \frac{t(v-1)-\binom{t}{2}}{(v-1)+\binom{t}{2}}$.

When in addition $t = 2$ (D is a Steiner triple system), the stronger bounds $\mathsf{DiffSum}(D) \ge v$ *and* $\mathsf{RatioSum}(D) \ge 2$ *hold.*

Theorem 1 provides bounds on the metrics across all Steiner systems $S(t,k,v)$ and all point labelings of them. In previous work, the focus has been on the MinSum (or equivalently, by reversal, the MaxSum). Dau and Milenkovic [10] use the Bose [12] and Skolem [13], [14] constructions of Steiner triple systems to establish the existence of an $\mathrm{STS}(v)$ D with $\mathsf{MinSum}(D) = v$, the largest possible by Theorem 1 (Brummond [15] establishes a similar result for Kirkman triple systems). They accomplish this by specifying a particular point labeling that meets the MinSum bound, but unfortunately the labeling chosen yields a MaxSum near $\frac{8}{3}v$, a $\mathsf{DiffSum}$ near $\frac{5}{3}v$, and a $\mathsf{RatioSum}$ near $\frac{8}{3}$, far from the bounds of $2v$, v, and 2, respectively. The reversal of this labeling yields a MinSum far from optimal, the same $\mathsf{DiffSum}$, and a larger $\mathsf{RatioSum}$.

One might hope to improve the $\mathsf{DiffSum}$ and $\mathsf{RatioSum}$ by choosing a different labeling or by choosing a different Steiner system $S(t,k,v)$. In Section III, we show that certain $S(t,k,v)$s cannot meet *any* of the bounds in Theorem 1.

A. Improved bounds for STSs

There is an $\mathrm{STS}(7)$ with $\mathsf{MinSum} = 6$ and $\mathsf{MaxSum} = 13$ with blocks 016, 024, 035, 123, 145, 256, and 346 (here we write abc for $\{a,b,c\}$). There is an $\mathrm{STS}(9)$ with $\mathsf{MinSum} = 9$ and $\mathsf{MaxSum} = 18$ with blocks 018, 027, 036, 045, 126, 135, 147, 234, 258, 378, 468, and 567. However, these are the *only* two Steiner triple systems with $\mathsf{DiffSum} = v$, and indeed the *only* $\mathrm{STS}(v)$ with $\mathsf{RatioSum} = 2$ is the $\mathrm{STS}(9)$.

Theorem 2. *Let D be a Steiner triple system of order $v \geq 13$. Then* $\mathsf{DiffSum}(D) \geq v+1$ *and* $\mathsf{RatioSum}(D) > 2$.

Proof. See the full version [1]. ■

III. Independent Sets

Let $D = (V, \mathcal{B})$ be a t-(v,k,λ) packing. An *independent set* in D is a subset $X \subseteq V$ such that there is no $B \in \mathcal{B}$ with $B \subseteq X$. An independent set X is *maximal* if there is no independent set Y with $X \subset Y$, and *maximum* if there is no independent set Y such that $|Y| > |X|$. The *independence number* of D, denoted $\alpha(D)$, is the size of a maximum independent set. There is a close connection between the independence number of a packing and the quality of any of its labelings.

Lemma 1. *A t-(v,k,λ) packing D has* MinSum *at most $k\alpha(D) - \binom{k}{2}$,* MaxSum *at least $k(v-1-\alpha(D)) + \binom{k}{2}$, and* $\mathsf{DiffSum}$ *at least $k(v+k-2-2\alpha(D))$.*

Proof. It suffices to prove the statement for MinSum. No matter how D is given a point labeling, on elements with ranks in $\{0,\ldots,\alpha(D)\}$, there is a block. The sum of this block is at most $\sum_{i=1}^{k}(\alpha(D)-(i-1))$. ■

Corollary 1. *Meeting the bound on* MinSum *in Theorem 1 for a t-$(v,k,1)$ packing D requires that*

$$\alpha(D) \geq \frac{v(k-t+1)}{2k} + \frac{k+t-3}{2}.$$

For example, Corollary 1 states that a necessary condition for a partial Steiner triple system D to have MinSum equal to v is that $\alpha(D) \geq \frac{v}{3}+1$.

We refine this bound by using a second disjoint independent set. Suppose that a t-(v,k,λ) packing D contains two *disjoint* independent sets of sizes γ_D and δ_D, respectively, with $\gamma_D \geq \delta_D$; two disjoint independent sets form an *independent pair*. Set

$$\gamma'_D = \min\left(\gamma_D, \tfrac{v(k-t+1)}{2k} + \tfrac{k+t}{2} - 1\right),$$
$$\delta'_D = \min\left(\delta_D, \tfrac{v(k-t+1)}{2k} + \tfrac{k+t}{2} - 1\right).$$

Two independent sets form a *maximum independent pair* when $\gamma'_D + \delta'_D$ is as large as possible.

Lemma 2. *A t-(v,k,λ) packing D with a maximum independent pair of sizes (γ_D, δ_D) has* $\mathsf{DiffSum}$ *at least $k(v+k-2-\delta'_D-\gamma'_D)$.*

Proof. See the full version [1]. ■

Corollary 2. *Meeting the bound on* $\mathsf{DiffSum}$ *in Theorem 1 for a t-$(v,k,1)$ packing D requires that D have a independent pair of sizes $(\frac{v(k-t+1)}{2k} + \frac{k+t}{2} - 1, \frac{v(k-t+1)}{2k} + \frac{k+t}{2} - 1)$.*

The independent pair in Corollary 2 does not require that either be maximum, nor that their combined size be as large as possible. For a Steiner triple system, for example, Corollary 2 asks only for two disjoint independent sets, each of size at least $\frac{v}{3}+1$, for a combined size of $\frac{2v}{3}+2$. Applying the $2v+1$ construction [16] twice to an $\mathrm{STS}(v)$, we form an $\mathrm{STS}(4v+3)$ having a maximum independent pair of sizes $(2v+2, v+1)$; although the combined size is over $\frac{3}{4}$ of the size of the STS, such a pair could not lead to the smallest $\mathsf{DiffSum}$ because the second largest of the pair is too small.

Corollary 2 gives a necessary condition, not a sufficient one. Nevertheless, some bounds can be stated.

Lemma 3. *When a t-$(v,k,1)$ packing D has two disjoint independent sets of sizes α and β, there is a point labeling with* $\mathsf{MinSum}(D) \geq \alpha + \binom{k}{2}$ *and (for the same labeling)* $\mathsf{MaxSum}(D) \leq k(v-1) - \beta - 1 - \binom{k}{2}$*, so* $\mathsf{DiffSum}(D) \leq k(v-k) - \beta - \alpha - 1$.

Proof. Any point labeling assigning labels $\{0,\ldots,\alpha-1\}$ to the points of the independent set of size α, labels $\{v-\beta,\ldots,v-1\}$ to the points of the independent set of size β, and labels $\{\alpha,\ldots,v-\beta-1\}$ to the remaining points, meets the stated bounds. ■

Labeling for access balancing must focus on Steiner triple systems, and on t-$(v,k,1)$ packings in general, having large sizes in maximum independent pairs. This choice is important, because not all such systems have even a single large independent set, as we explain next.

IV. Small Maximum Independent Sets

Can one choose an arbitrary t-$(v,k,1)$ packing, and by cleverly choosing a point labeling optimize one or more of the sum metrics? If not, how far from the bound of Theorem 1 can the best point labeling be? In order to discuss these questions, define

$$\alpha_{min}(t,k,v) = \min\{\alpha(D) : D \text{ is a } t\text{-}(v,k,1) \text{ packing}\},$$
$$\alpha^{\star}_{min}(t,k,v) = \min\{\alpha(D) : D \text{ is an } S(t,k,v)\}.$$

When an $S(t,k,v)$ exists, $\alpha_{min}(t,k,v) \leq \alpha^{\star}_{min}(t,k,v)$.

Erdős and Hajnal [17] establish that $\alpha_{min}(2,3,v) \geq \lfloor\sqrt{2v}\rfloor$; indeed a simple greedy algorithm produces an independent set of this size. Spencer [18] establishes that $\alpha_{min}(2,3,v) \geq c \cdot v\sqrt{2}/\sqrt{v-1}$, a small improvement on the Erdős-Hajnal result. State-of-the-art lower bounds [19]–[22] all differ only by constant factors, and all rely heavily on a theorem about "uncrowded" hypergraphs from [23]. This leads to the lower bound stated next; the upper bound is established using the Lovász Local Lemma.

Theorem 3. [24], [25] *For fixed k and t, there are absolute constants c and d for which*

$$cv^{\frac{k-t}{k-1}}(\log v)^{\frac{1}{k-1}} \leq \alpha_{min}(t,k,v) \leq dv^{\frac{k-t}{k-1}}(\log v)^{\frac{1}{k-1}}.$$

Phelps and Rödl [26] establish that the bounds of Theorem 3 apply to Steiner triple systems, not just partial ones; that is, $c\sqrt{v\ln v} \leq \alpha^{\star}_{min}(2,3,v) \leq d\sqrt{v\ln v}$ for absolute constants c and d. Grable, Phelps and Rödl [27] establish similar statements when $t \in \{2,3\}$ for all $k > t$. Combining Lemma 1 and the results in [26], [27], some Steiner triple systems only have point labelings far from the bounds of Theorem 1:

Theorem 4. *For infinitely many orders v, there exists an STS(v) D with* $\mathsf{MinSum}(D) \leq 3c\sqrt{v\ln v} - 3$ *and* $\mathsf{MaxSum}(D) \geq 3v - 3c\sqrt{v\ln v}$*, and hence* $\mathsf{DiffSum}(D) \geq 3v - 6c\sqrt{v\ln v} + 3$.

We must focus on specific Steiner systems or packings, if we are to obtain sum metrics at or near the basic bounds.

V. Dense t-$(v,t+1,1)$ Packings

We establish next that one can obtain metrics close to the optimal when $k=t+1$ for packings that contain all but a vanishingly small fraction of the blocks of an $S(t,t+1,v)$ as $v\to\infty$. The independent set requirements indicate that we must have a maximum independent pair having large sizes. To accomplish this, we partition all $(t+1)$-subsets of $\mathbb{Z}_v$ according to their sum modulo v, and choose one class of the partition to form the blocks of the packing. The basic strategy dates back at least a century to Bussey [28], and perhaps earlier. This is not a mere theoretical curiosity, because declustered-parity RAIDs need not have their loads perfectly balanced [7].

Theorem 5. *Let t and v be integers with $v>\binom{t+2}{2}+\binom{t+1}{2}$ so that v and $t+1$ are relatively prime. For each of the following statements, there exists a t-$(v,t+1,1)$ packing D on elements $\mathbb{Z}_v$ having $\frac{\binom{v}{t+1}}{v}=\frac{v-t}{v}\frac{\binom{v}{t}}{\binom{t+1}{t}}$ blocks.*

(1) $\mathsf{MinSum}(D)=v+\sigma$ *and* $\mathsf{MaxSum}(D)=tv+\sigma$ *whenever* $-\binom{t+2}{2}+1\le\sigma<\binom{t+1}{2}$.
(2) $\mathsf{MinSum}(D)=v+\binom{t+1}{2}-1$.
(3) $\mathsf{MaxSum}(D)=tv-\binom{t+2}{2}+1$.
(4) $\mathsf{DiffSum}(D)=(t-1)v$.
(5) $\mathsf{RatioSum}(D)=\frac{tv+\binom{t+1}{2}-1}{v+\binom{t+1}{2}-1}$.

Proof. Partition all $(t+1)$-subsets of $\mathbb{Z}_v$ into v classes $\{\mathcal{B}_\sigma : 0\le\sigma<v\}$ by placing set $S=\{x_1,\dots,x_{t+1}\}$ in class $\mathcal{B}_\sigma$ if and only if $\sigma\equiv\Sigma_{i=1}^{t+1}x_i \pmod v$. Because for any t-subset T of $\mathbb{Z}_v$ and each σ with $0\le\sigma<v$ there is a unique element s for which $\sigma\equiv s+\Sigma_{x\in T}x \pmod v$, each $\mathcal{B}_\sigma$ is a $t-(v,t+1,1)$ packing. The verification is completed in the full version [1] of the paper. ■

The packings so produced contain large independent sets. For example, when $\sigma=0$, $\{0,\dots,\lfloor\frac{v}{t+1}\rfloor\}$ forms an independent set. Theorem 5 yields packings that are dense in the following sense. When an $S(t,t+1,v)$ exists, it has $\frac{\binom{v}{t}}{\binom{t+1}{t}}$ blocks; the packings considered have a $\frac{v-t}{v}$ fraction of this number. Hence for fixed t the fraction of t-sets left uncovered by the packing approaches 0 as $v\to\infty$. Moreover, the bounds established for dense t-$(v,t+1,1)$ packings on MinSum and MaxSum match the values from Theorem 1 (which are best for Steiner systems). On the other hand, as $v\to\infty$ and t is fixed, the ratio of $\mathsf{DiffSum}$ of the packing to the bound approaches 1, and the $\mathsf{RatioSum}$ approaches its bound of $t-1$. By generalizing to partial systems, Theorem 5 applies to all parameters that are large enough, whether or not an $S(t,t+1,v)$ exists.

Although Theorem 5 establishes a $\mathsf{DiffSum}$ of $(t-1)v$ for certain dense t-$(v,t+1,1)$ packings, this may not be the best possible, as Theorem 1 ensures only that $(v-k)(t-1)$ is a lower bound on the $\mathsf{DiffSum}$. Theorem 6 gives evidence that the bound may not be the best possible, by producing a packing that achieves a smaller $\mathsf{DiffSum}$ than that of Theorem 5 when $t=3$, but is nearly as dense.

Theorem 6. *When $v>18$ is even, there is a* 3-$(v,4,1)$ *packing D with $\frac{v-4}{v-1}\frac{\binom{v}{3}}{\binom{4}{3}}$ blocks, having* $\mathsf{MinSum}(D)=v+2$, $\mathsf{MaxSum}(D)=3v-6$, *and hence* $\mathsf{DiffSum}(D)=2v-8$.

Proof. See the full version [1]. ■

VI. Sums and Steiner Triple Systems

For the intended applications in storage systems, it remains desirable to employ a Steiner system, rather than a dense packing, when possible. In what follows, we extend Theorem 5 to produce Steiner triple systems in which the sum metrics are close to optimal.

Building on the construction in Theorem 5, Schreiber [29] and Wilson [30] demonstrate that for certain values of v, the packing can be completed to an STS(v). To treat the labeling and block sums, we employ a technical lemma, whose easy proof is in the full version [1].

Lemma 4. *Let $n\equiv 1,5 \pmod 6$. Every pair in $\{\{a,b\} : a,b\in\mathbb{Z}_n\setminus\{0\}, a\equiv -2b \pmod n\}$ has $(n+1)/2\le a+b\le(n-1)/2+n$.*

Now we examine the block sums in the Schreiber-Wilson result. (In [29], [30], the $STS(v)$ is constructed, but the point labelling is not.)

Theorem 7. *Suppose that $v\equiv 1,3 \pmod 6$ and for every prime p dividing $v-2$, the order of $-2 \bmod p$ is singly even. Then there is an STS(v), D, with* $\mathsf{MinSum}(D)\ge v-2$, $\mathsf{MaxSum}(D)\le 2v+2$, *and hence* $\mathsf{DiffSum}(D)\le v+4$ *and* $\mathsf{RatioSum}(D)\le\frac{2v+2}{v-2}$.

Proof. The proof, given in the full version [1], parallels that of Theorem 8. ■

Unlike the point labelings in [10], the labeling for the Schreiber-Wilson construction in Theorem 7 need not achieve the largest MinSum or smallest MaxSum. Nevertheless it yields a substantial improvement with respect to the $\mathsf{DiffSum}$ and $\mathsf{RatioSum}$, within an additive constant of the best bound possible for the $\mathsf{DiffSum}$. Unfortunately, Theorem 7 requires that the order of $-2 \bmod p$ be singly even, and so applies to an infinite set of orders but not all admissible ones. We remedy this next, using a result from [31], but obtaining slightly weaker bounds.

Theorem 8. *Whenever $v\equiv 1,3 \pmod 6$, there is an STS(v), D, with* $\mathsf{MinSum}(D)\ge v-5$, $\mathsf{MaxSum}(D)\le 2v+2$, *and hence* $\mathsf{DiffSum}(D)\le v+7$ *and* $\mathsf{RatioSum}(D)\le\frac{2v+2}{v-5}$.

Proof. Let $n=v-2$. Using the proof of Theorem 5, construct a 2-$(v,3,1)$ packing $\mathcal{B}_0$ on $\mathbb{Z}_{v-2}$ (points $v-2$ and $v-1$ appear in no triples). Remove element 0 as well as all triples $\{\{0,x,v-2-x\} : 1\le x\le(v-3)/2\}$ to form $\mathcal{D}_0$. Let $\mathcal{E}_0$ be the set of pairs on $\mathbb{Z}_{v-2}\setminus\{0\}$ not covered by a triple of $\mathcal{D}_0$. Each pair in $\mathcal{E}_0$ has sum between $(v-1)/2$ and $(v-3)/2+v-2$ by Lemma 4. The pairs in $\mathcal{E}_0$ form a 3-regular graph G on $\mathbb{Z}_{v-2}\setminus\{0\}$. By [31, Lemma 9], G can be partitioned into three 1-factors, F_1, F_2, and F_3.

To form the STS(v) on $\mathbb{Z}_v$ with block set $\mathcal{C}$, we employ the mapping $\psi:\mathbb{Z}_{v-2}\setminus\{0\}\mapsto\mathbb{Z}_v\setminus\{(v-3)/2,(v-1)/2,(v+1)/2\}$ defined by $\psi(x)=x-1$ when $1\le x\le(v-3)/2$ and $\psi(x)=x+2$ when $(v-1)/2\le x<v-2$. Then $\mathcal{C}$ is formed as follows.

(1) When $\{x, y, z\} \in \mathcal{D}_0$, place $\{\psi(x), \psi(y), \psi(z)\}$ in $\mathcal{C}$;
(2) For $i = 1, 2, 3$, when $\{x, y\} \in F_i$, place $\{(v - 5 + 2i)/2, \psi(x), \psi(y)\}$ in $\mathcal{C}$;
(3) Place $\{(v - 3)/2, (v - 1)/2, (v + 1)/2\}$ in $\mathcal{C}$.

Triples of $\mathcal{B}_0$ have sum $v - 2$ or $2v - 4$, so triples of type (1) in $\mathcal{C}$ have sum between $v - 5$ and $v - 2$, or between $2v - 1$ and $2v + 2$. By Lemma 4, a pair $\{x, y\} \in \mathcal{E}_0$ has $(v-1)/2 \leq x+y \leq (v-1)/2+(v-3)$. Applying ψ, we have $(v-1)/2-2 \leq \psi(x)+\psi(y) \leq (v-1)/2+(v+1)$. Hence each triple of type (2) in $\mathcal{C}$ has sum at least $v-4$ and at most $2v + 1$. The block of type (3) has sum $\frac{3v-3}{2}$. ■

Although the bounds are slightly weaker, Theorem 8 applies to all admissible orders for Steiner triple systems. In conjunction with Theorem 2, for all $v \equiv 1, 3 \pmod 6$ with $v \geq 13$ one has $v + 1 \leq \mathsf{DiffSum}(2, 3, v) \leq v + 7$ and $2 + \frac{1}{v} \leq \mathsf{RatioSum}(2, 3, v) \leq 2 + \frac{12}{v-5}$. Based on computations described in the full version [1], it appears plausible that $\mathsf{DiffSum}(2, 3, v) = v+1$ when $v \geq 13$. It also appears plausible that $\mathsf{RatioSum}(2, 3, v) \in \{2 + \frac{1}{v}, 2 + \frac{2}{v}\}$ for every $v \neq 9$, but there is insufficient data to speculate on when it takes the larger value and when the smaller.

VII. Concluding Remarks

Because Theorem 5 achieves a $\mathsf{DiffSum}$ of $(t - 1)v$ for dense t-$(v, t+1, 1)$ packings, one might hope that this difference can be realized for $S(t, t+1, v)$ Steiner systems. However, Theorem 2 establishes that this does not happen when $t = 2$ unless $v \in \{7, 9\}$, although Theorem 8 is within an additive constant. The situation when $t = 3$ appears to be quite different. There is an $S(3, 4, 8)$ having MinSum 10 and MaxSum 18. Adapting the construction in [32], [33], one can produce an $S(3, 4, v)$ with MinSum $v+2$, MaxSum $3v-6$, and hence $\mathsf{DiffSum}$ $2v-8$ whenever v is a power of 2. In these cases, the upper bound on the MinSum and the lower bound on the MaxSum from Theorem 1 are met simultaneously. We do not expect this to happen for all orders, because the smallest $\mathsf{DiffSum}$ for an $S(3, 4, v)$ when $v \in \{10, 14\}$ appears to arise from systems with MinSum $v + 1$ and MaxSum $3v - 5$. It may happen that for every admissible v, an $S(3, 4, v)$ with $\mathsf{DiffSum}$ strictly smaller than $2v$ exists. If so, completing the packing from Theorem 5 could not yield the smallest $\mathsf{DiffSum}$. Nevertheless, the structure of independent sets must underlie appropriate constructions.

Acknowledgment

The work was supported by NSF grants CCF 1816913 (CJC) and 1814298 (OM).

References

[1] Y. M. Chee, C. J. Colbourn, H. Dau, R. Gabrys, A. C. H. Ling, D. Lusi, and O. Milenkovic, "Access balancing in storage systems by labeling partial Steiner systems," arXiv:1906.12073, 2019.
[2] S. El Rouayheb and K. Ramchandran, "Fractional repetition codes for repair in distributed storage systems," in *48th Ann. Allerton Conf. Comm. Control Comput. (Allerton)*, pp. 1510–1517, 2010.
[3] N. Silberstein and T. Etzion, "Optimal fractional repetition codes based on graphs and designs," *IEEE Trans. Inf. Theory*, vol. 61, no. 8, pp. 4164–4180, 2015.
[4] N. Silberstein and A. Gál, "Optimal combinatorial batch codes based on block designs," *Designs, Codes and Cryptography*, vol. 78, no. 2, pp. 409–424, 2016.
[5] A. Fazeli, A. Vardy, and E. Yaakobi, "Codes for distributed PIR with low storage overhead," in *IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 2852–2856, 2015.
[6] A. Cidon, S. M. Rumble, R. Stutsman, S. Katti, J. K. Ousterhout, and M. Rosenblum, "Copysets: Reducing the frequency of data loss in cloud storage." in *Usenix Ann. Tech. Conf.*, pp. 37–48, 2013.
[7] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, "RAID: High-performance, reliable secondary storage," *ACM Comput. Surveys (CSUR)*, vol. 26, no. 2, pp. 145–185, 1994.
[8] M. Holland and G. A. Gibson, "Parity declustering for continuous operation in redundant disk arrays," *SIGPLAN Not.*, vol. 27, no. 9, pp. 23–35, Sep. 1992.
[9] H. Dau, Y. Jia, C. Jin, W. Xi, K. S. Chan, "Parity declustering for fault-tolerant storage systems via t-designs," *IEEE Int. Conf. Big Data*, pp. 7-14, 2014.
[10] H. Dau and O. Milenkovic, "MaxMinSum Steiner systems for access balancing in distributed storage," *SIAM J. Discrete Math.*, vol. 32, no. 3, pp. 1644–1671, 2018.
[11] W. Yu, X. Zhang, and G. Ge, "Optimal fraction repetition codes for access-balancing in distributed storage," arXiv:1912.07779, 2019.
[12] R. C. Bose, "On the construction of balanced incomplete block designs," *Annals of Eugenics*, vol. 9, no. 4, pp. 353–399, 1939.
[13] H. Hanani, "A note on Steiner triple systems," *Mathematica Scandinavica*, vol. 8, no. 1, pp. 154–156, 1960.
[14] T. Skolem, "Some remarks on the triple systems of Steiner," *Mathematica Scandinavica*, pp. 273–280, 1959.
[15] W. M. Brummond, "Kirkman systems that attain the upper bound on the minimum block sum, for access balancing in distributed storage," arXiv:1906.02157, 2019.
[16] C. J. Colbourn and A. Rosa, *Triple Systems*, ser. Oxford mathematical monographs. Clarendon Press, 1999.
[17] P. Erdős and A. Hajnal, "On chromatic number of graphs and set-systems," *Acta Math. Acad. Sci. Hungar*, vol. 17, pp. 61–99, 1966.
[18] J. Spencer, "Turán's theorem for k-graphs," *Discrete Math.*, vol. 2, pp. 183–186, 1972.
[19] A. Eustis and J. Verstraëte, "On the independence number of Steiner systems," *Combin. Probab. Comput.*, vol. 22, no. 2, pp. 241–252, 2013.
[20] A. Kostochka, D. Mubayi, V. Rödl, and P. Tetali, "On the chromatic number of set systems," *Random Structures & Algorithms*, vol. 19, no. 2, pp. 87–98, 2001.
[21] A. Kostochka, D. Mubayi, and J. Verstraëte, "On independent sets in hypergraphs," *Random Structures & Algorithms*, vol. 44, no. 2, pp. 224–239, 2014.
[22] F. Tian and Z.-L. Liu, "Bounding the independence number in some (n, k, ℓ, λ)-hypergraphs," *Graphs Combin.*, vol. 34, no. 5, pp. 845–861, 2018.
[23] M. Ajtai, J. Komlós, J. Pintz, J. Spencer, and E. Szemerédi, "Extremal uncrowded hypergraphs," *J. Combin. Theory Ser. A*, vol. 32, no. 3, pp. 321–335, 1982.
[24] R. A. Duke, H. Lefmann, and V. Rödl, "On uncrowded hypergraphs," *Random Structures & Algorithms*, vol. 6, no. 2-3, pp. 209–212, 1995.
[25] V. Rödl and E. Šiňajová, "Note on independent sets in Steiner systems," *Random Structures & Algorithms*, vol. 5, no. 1, pp. 183–190, 1994.
[26] K. T. Phelps and V. Rödl, "Steiner triple systems with minimum independence number," *Ars Combin.*, vol. 21, pp. 167–172, 1986.
[27] D. A. Grable, K. T. Phelps, and V. Rödl, "The minimum independence number for designs," *Combinatorica*, vol. 15, no. 2, pp. 175–185, 1995.
[28] W. H. Bussey, "The Tactical Problem of Steiner," *Amer. Math. Monthly*, vol. 21, no. 1, pp. 3–12, 1914.
[29] S. Schreiber, "Covering all triples on n marks by disjoint Steiner systems," *J. Combin. Theory Ser. A*, vol. 15, no. 3, pp. 347–350, 1973.
[30] R. M. Wilson, "Some partitions of all triples into Steiner triple systems," in *Hypergraph seminar*. Springer, 1974, pp. 267–277.
[31] D. Bryant, C. J. Colbourn, D. Horsley, and I. M. Wanless, "Steiner triple systems with high chromatic index," *SIAM J. Discrete Math.*, vol. 31, no. 4, pp. 2603–2611, 2017.
[32] J. Doyen and M. Vandensavel, "Non isomorphic Steiner quadruple systems," *Bull. Soc. Math. Belg.*, vol. 23, pp. 393–410, 1971.
[33] K. T. Phelps and A. Rosa, "2-chromatic Steiner quadruple systems," *European J. Combin.*, vol. 1, no. 3, pp. 253–258, 1980.