Double Double Toil and Trouble

Charles J. Colbourn
Arizona State University

Rijeka (virtually), July 2021

Set Systems and Steiner Systems

- (Uniform) Set system of type (v, k):
  - a set V of v points or elements
  - a collection $\mathcal{B}$ of k-subsets of V, called blocks
- Regular set system of type (v, k, r):
  - a set system of type (v, k) so that every point is in exactly r blocks.
- Steiner system $S(t, k, v)$:
  - a regular set system of type $(v, k, \frac{\binom{v-1}{t-1}}{\binom{k-1}{t-1}})$, $(V, \mathcal{B})$, for which each t-subset $T \subseteq V$ satisfies $T \subseteq B$ for exactly one $B \in \mathcal{B}$.

Set Systems and Steiner Systems

Charles J. Colbourn
Arizona State University

- (Uniform) Set system of type (v, k):
  - a set V of v points or elements
  - a collection $\mathcal{B}$ of k-subsets of V, called blocks
- Regular set system of type (v, k, r):
  - a set system of type (v, k) so that every point is in exactly r blocks.
- Steiner system $S(t, k, v)$:
  - a regular set system of type $(v, k, \frac{\binom{v-1}{t-1}}{\binom{k-1}{t-1}})$, $(V, \mathcal{B})$, for which each t-subset $T \subseteq V$ satisfies $T \subseteq B$ for exactly one $B \in \mathcal{B}$.

Set Systems and Steiner Systems

Charles J. Colbourn
Arizona State University

- Uniform set systems, regular set systems, and Steiner systems — and their duals — have been extensively applied in coding theory, communications, experimental design, etc. etc. — and in access and load balancing for storage systems.

Set Systems and Steiner Systems

Charles J. Colbourn
Arizona State University

- Uniform set systems, regular set systems, and Steiner systems — and their duals — have been extensively applied in coding theory, communications, experimental design, etc. etc. — and in access and load balancing for storage systems.

The General Problem

- Use a (regular, uniform) set system to associate data items (points) to storage units (blocks), so that accesses to storage units are balanced.
- Because in a regular set system, every point is in the same number of blocks, and every block contains the same number of points, they achieve the “balance” that we want!
- Or do they? Not all data items have the same long-term frequency of access, or popularity.

The General Problem

- Use a (regular, uniform) set system to associate data items (points) to storage units (blocks), so that accesses to storage units are balanced.
- Because in a regular set system, every point is in the same number of blocks, and every block contains the same number of points, they achieve the "balance" that we want!
- Or do they? Not all data items have the same long-term frequency of access, or popularity.

The General Problem

Popularity

Charles J. Colbourn
Arizona State University

- Dau and Milenkovic (2018) suggest ranking the data items by popularity, ordering the data items from most to least popular.
- Placing a total ordering on the points and a total ordering on the blocks leads to a unique incidence matrix for the Steiner (or set) system.

Charles J. Colbourn
Arizona State University

The Sum Metrics

Primal

- Let $A = (a_{ij})$ be a $n \times m$ incidence matrix.
- For each column $0 \leq j < m$, define the weighted column sum (block sum) σ_j to be

$$\sum_{i=0}^{n-1} i \cdot a_{ij}$$

- Then define

$$\begin{aligned} \text{MinSum}(A) &= \min(\sigma_j : 0 \leq j < m\} \\ \text{MaxSum}(A) &= \max(\sigma_j : 0 \leq j < m\} \\ \text{DiffSum}(A) &= \text{MaxSum}(A) - \text{MinSum}(A). \end{aligned}$$

The Sum Metrics

Primal

Charles J. Colbourn
Arizona State University

- Goals: Order the rows of A (i.e., the points of the underlying set system) to achieve
  - large MinSum
  - small MaxSum
  - small DiffSum

The Sum Metrics

Dual

- Let $A = (a_{ij})$ be a $n \times m$ incidence matrix.
- Transpose A and consider sum metrics for A^{T}.
- This gives point sums for A, and we focus on ordering the blocks of A.

The Sum Metrics

Dual

- Goals: Order the columns of A (i.e., the blocks of the underlying set system, or the rows of A^T) to achieve
  - large (dual) MinSum
  - small (dual) MaxSum
  - small (dual) DiffSum

Example $S(3,4,10)$ – Regular of type (10,4,12)

One incidence matrix A_1

$$
\begin{array}{l}
111111110000011110000000000000\\
111000001111110000000011100000\\
100110000100101001110001011000\\
100001010011000101010110001010\\
001011001000100010011110010001\\
001000110110001000111000100011\\
010100010001100010101100101100\\
010010101010000101100101000101\\
000101101101010001001000010110\\
000000000000011110000011111111
\end{array}
$$

$\mathrm{MinSum}(A_1) = 6$, $\mathrm{MaxSum}(A_1) = 30$, $\mathrm{DiffSum}(A_1) = 24$
$\mathrm{MinSum}(A_1^T) = 86$, $\mathrm{MaxSum}(A_1^T) = 262$, $\mathrm{DiffSum}(A_1^T) = 176$

Example $S(3, 4, 10)$

Another incidence matrix A_2


```
000111010010000011010111000100
010100010011101000100100110010
110000110001000111001000001110
010010001110000101111000100001
101100000001011010011001100001
111000001100001000010110011100
100101101000010001100010010011
001011001000110110000100101010
000010110100111000101011001000
001001100111100100000001010101
```

$\text{MinSum}(A_2) = 11$, $\text{MaxSum}(A_2) = 25$, and $\text{DiffSum}(A_2) = 14$

$\text{MinSum}(A_2^T) = \text{MaxSum}(A_2^T) = 174$, and $\text{DiffSum}(A_2^T) = 0$

Example $S(3,4,10)$

Another incidence matrix A_2

Charles J. Colbourn
Arizona State University

000111010010000011010111000100

010100010011101000100100110010

110000110001000111001000001110

010010001110000101111000100001

101100000001011010011001100001

111000001100001000010110011100

100101101000010001100010010011

001011001000110110000100101010

000010110100111000101011001000

001001100111100100000001010101

$\mathrm{MinSum}(A_2) = 11$, $\mathrm{MaxSum}(A_2) = 25$, and $\mathrm{DiffSum}(A_2) = 14$
$\mathrm{MinSum}(A_2^T) = \mathrm{MaxSum}(A_2^T) = 174$, and $\mathrm{DiffSum}(A_2^T) = 0$

The Questions

- The sum metrics for a specific incidence matrix are easily calculated.
- Can we optimize each metric over all incidence matrices for a specified set system?
- Can we optimize each metric over all incidence matrices for all (uniform, regular, or Steiner) set systems having the same parameters?

Reversal

Charles J. Colbourn
Arizona State University

- Suppose that A is a regular set system of type (v, k, r), with MinSum m and dual MinSum d.
- Reverse the ordering of the columns and the ordering of the rows to get an incidence matrix R.
- Then R has MaxSum $k(v-1) - m$ and dual MaxSum $r(b-1) - d$.
- And vice versa. So we can focus on MInSum to understand MaxSum.
- But be careful! Although DiffSum is MaxSum minus MinSum, this is for a specific incidence matrix, and it may happen that no incidence matrix with largest MinSum can also have smallest MaxSum.

Reversal

Charles J. Colbourn
Arizona State University

- Suppose that A is a regular set system of type (v, k, r), with MinSum m and dual MinSum d.
- Reverse the ordering of the columns and the ordering of the rows to get an incidence matrix R.
- Then R has MaxSum $k(v-1) - m$ and dual MaxSum $r(b-1) - d$.
- And vice versa. So we can focus on MInSum to understand MaxSum.
- But be careful! Although DiffSum is MaxSum minus MinSum, this is for a specific incidence matrix, and it may happen that no incidence matrix with largest MinSum can also have smallest MaxSum.

How bad can labellings be?

Naive bounds

- In an $S(t, k, v)$ having two disjoint blocks, labelling the points of one with $\{0, \ldots, k-1\}$ and of the second with $\{v-k, \ldots, v-1\}$, one gets MinSum $\binom{k}{2}$, MaxSum $k(v-k) + \binom{k}{2}$ and DiffSum $k(v-k)$ — and these are the worst possible.
- Similarly labelling all blocks containing a particular point with $\{0, \ldots, r-1\}$ yields dual MinSum $\binom{r}{2}$, and its reversal has dual MaxSum $r(b-r) + \binom{r}{2}$ — and these are the worst possible.
- But we cannot get both at the same time! Indeed, for $\lambda_2 = \binom{v-2}{t-2}$, the worst dual DiffSum is $(r-\lambda_2)(b-r+\lambda_2)$.

Some easy counting

Primal

- Consider a regular set system of type (v, k, r).
- It must have $b = \frac{vr}{k}$ blocks.
- The total block sum is

$$r \sum_{i=0}^{v-1} i = r \binom{v}{2}$$

- The average block sum is

$$\frac{r}{b} \binom{v}{2} = \frac{1}{2} k(v - 1)$$

- an upper bound on MinSum
- a lower bound on MaxSum

More easy counting

Dual

- Consider a regular set system of type (v, k, r).
- It must have $b = \frac{vr}{k}$ blocks.
- The total point sum is

$$k \sum_{i=0}^{b-1} i = k \binom{b}{2}$$

- The average point sum is

$$\frac{k}{v} \binom{b}{2} = \frac{1}{2} r(b-1)$$

- an upper bound on dual MinSum
- a lower bound on dual MaxSum

A primal bound

- Consider an $S(t,k,v)$. It has average block sum $\frac{1}{2}k(v-1)$.
- Now choose a set of points $X = \{x_1, \ldots, x_{t-1}\}$. Let C be the blocks that contain all points of X. Form the derived set system D with respect to X.
- D has $\frac{v-t+1}{k-t+1}$ blocks, each of size $k-t+1$, that partition $\{0,\ldots,v-1\} \setminus X$.
- The average block sum of a block in D is

$$\frac{k-t+1}{v-t+1}\left[\sum_{i=0}^{v-1} i - \sum_{j=1}^{t-1} x_i\right]$$

A primal bound
still

Charles J. Colbourn
Arizona State University

- Choose $X = \{0, \ldots, t-2\}$.
- The average block sum of a block in C is

$$\binom{t-1}{2} + \frac{k-t+1}{v-t+1}\left[(t-1)(v-t+1) + \binom{v-t+1}{2}\right]$$

- The average block sum of a block in C is

$$\binom{t-1}{2} + \frac{(k-t+1)(v+t-2)}{2}$$

The Basic Bound

Primal

Charles J. Colbourn
Arizona State University

	MinSum $\leq$	MaxSum $\geq$	DiffSum $\geq$
$t=2$	$\frac{(k-1)v}{2}$	$\frac{(k+1)v-2k}{2}$	$v-k$
$t=3$	$\frac{(k-2)v+k+4}{2}$	$\frac{(k+2)v-3k-4}{2}$	$2(v-k)-4$
$t=4$	$\frac{(k-3)v+2k+6}{2}$	$\frac{(k+3)v-4k-6}{2}$	$3(v-k)-6$
$\vdots$			

The Current Landscape

Steiner triple systems $S(2,3,v)$

- There exists an $S(2,3,v)$ with MinSum equal to v, the largest possible. (DM18).
- For v sufficiently large, some $S(2,3,v)$ has MinSum at most $c \log v$ (CCDGLLM20)
- Every $S(2,3,v)$ has DiffSum at least v when $v \geq 7$ (DM18), at least $v+1$ when $v \geq 13$ (CCDGLLM20)
- For every admissible $v \geq 13$, there exists an $S(2,3,v)$ with DiffSum at most $v+7$ (CCDGLLM20)
- For infinitely many admissible $v \geq 13$, there exists an $S(2,3,v)$ with dual DiffSum equal to 0 – an egalitarian system (C21+)

The Current Landscape

Steiner systems $S(2, 4, v)$

Charles J. Colbourn
Arizona State University

- For infinitely many admissible v, there exists an $S(2, 4, v)$ with dual DiffSum equal to 0 – an egalitarian system (Lusi-C 21)

The Topic for Today

Steiner quadruple systems $S(3, 4, v)$

- We examine the well-known doubling construction for $S(3, 4, v)$s.
- Then we label the points and the blocks to examine consequences for the sum metrics.

Steiner quadruple systems $S(3,4,v)$

Doubling

- Let $(V, \mathcal{B})$ be an $S(3,4,v)$.
- We form an $S(3,4,2v)$ on elements $V \times \{0,1\}$ — we write x_i for (x,i) with $x \in V$ and $i \in \{0,1\}$.
- The blocks are
  - Type 1: $\{a_i, b_j, c_\ell, d_m\}$ whenever $\{a,b,c,d\} \in \mathcal{B}$ and $i + j + \ell + m \equiv 1 \pmod 2$, and
  - Type 2: $\{a_0, b_0, a_1, b_1\}$ whenever $a, b \in V$ and $a \neq b$.
- The doubling construction by Hanani (1960) uses $i + j + \ell + m \equiv 0 \pmod 2$ instead. In fact you can choose this parity arbitrarily for each block of the $S(3,4,v)$.

Doubling a Point Labelling

- Let $(V, \mathcal{B})$ be an $S(3,4,v)$. It has $b = \frac{1}{24}v(v-1)(v-2)$ blocks.
- Suppose that it has block sums $\{\sigma_i : 0 \leq i < b\}$.
- Double to form an $S(3,4,2v)$ by assigning point x_0 the label x and point x_1 the label $x+v$. What are its block sums?
  - For each $B \in \mathcal{B}$ with block sum σ in the $S(3,4,v)$, doubling makes four Type 1 blocks with block sum $\sigma + v$ and four with block sum $\sigma + 3v$.
  - For each $0 \leq a < b < v$, there is a Type 2 block with block sum $2v + 2a + 2b$.

Doubling a Point Labelling

- Let $(V, \mathcal{B})$ be an $S(3,4,v)$. Recall that its MinSum is at most $v+2$ and its MaxSum at least $3v-6$.
- And the doubled $S(3,4,2v)$ must have MinSum at most $2v+2$ and MaxSum at least $6v-6$.
- The Type 2 blocks have sums between $2v+2$ and $6v-6$.
- And if the $S(3,4,v)$ has MinSum $v+\alpha$ and MaxSum $3v+\beta$, the $S(3,4,2v)$ has MinSum $2v+\alpha$ and MaxSum $3(2v)+\beta$.

Doubling a Point Labelling

- A nice consequence: When $v = 2^\ell$, there is an $S(3, 4, v)$ with MinSum $v + 2$ and MaxSum $3v - 6$.
- Earlier we saw an $S(3, 4, 10)$ with MinSum $10 + 1$ and MaxSum $3(10) - 5$. So there is an $S(3, 4, 2^\ell 5)$ with MinSum $2^\ell 5 + 1$ and MaxSum $3(2^\ell 5) - 5$ for all $\ell \geq 1$.
- Similarly there is an $S(3, 4, 14)$ with MinSum $14 + 1$ and MaxSum $3(14) - 5$, yielding another infinite class.

Doubling a Block Labelling

- Let $(V, \mathcal{B})$ be an $S(3,4,v)$.
- It has $b = \frac{1}{24}v(v-1)(v-2)$ blocks. Call them $B_0 \ldots B_{b-1}$.
- We will label the

$$d = \frac{1}{24}2v(2v-1)(2v-2) = 8\left[\frac{1}{24}v(v-1)(v-2)\right] + \binom{v}{2}$$

blocks of the $S(3,4,2v)$ from doubling.

Doubling a Block Labelling

First Positions

- First we treat the blocks as 4-tuples, arranging them so that either
  - every element appears in the first position an even number of times, or
  - all but one element appears in the first position an even number of times, and the last appears a number of times that is odd and at least 3.

Doubling a Block Labelling

First Positions

- The idea: An element is odd if it appears in the first position an odd number of times.
- Suppose there are two odd elements o_0 and o_1.
- If some block has one on the first position and the other in a later position, swap them to reduce the number of odd elements by 2.
- Otherwise define
  - A to be all elements in the first position in some block containing o_0 and o_1. Note $|A| \geq (v-2)/2$.
  - for $i \in \{0,1\}$, B_i to be all elements in the first position in some block containing o_i but not o_{1-i} or in a block containing o_i in the first position. Note $|B_i| \geq (v-2)/2$ because the 3-GDD of type $2^{(v-2)/2}$ obtained by deriving w.r.t. o_i and then deleting o_{1-i} has no independent set of size larger than $(v-2)/2$.

Doubling a Block Labelling

First Positions

Charles J. Colbourn
Arizona State University

- Choose an element a that appears in two of A, B_0, B_1.
- Swap a and an odd element in the two corresponding blocks to reduce the number of odd elements by 2.

Doubling a Block Labelling

First Positions

Charles J. Colbourn
Arizona State University

- If there is one odd element o that is in the first position in only one block (o, a, b, c), if any of a, b, or c appear in two or more blocks in the first position, swap with o in this block.
- Otherwise in the $(v-1)(v-2)/6-1$ blocks that contain o not in the first position, choose an element e that appears the most times in the first position – this is at least $\frac{(v-1)(v-2)/6-1}{v-4}$ – and swap e and o in one block. This works provided that $v \geq 14$.

Doubling a Block Labelling

Patterns

Charles J. Colbourn
Arizona State University

- Next we give two 8×4 arrays of subscripting patterns:

P_2	P_4
1000	1000
0111	1110
1110	0111
0001	0001
1011	1101
0100	1011
1101	0010
0010	0100

- P_{-2} is obtained from P_2 by interchanging 0 and 1 throughout; similarly P_{-4} from P_4.

Doubling a Block Labelling

Type 1 Blocks

- Order all blocks of an $S(3,4,v)$ as $B_0, \ldots, B_{b-1}$, forming each as a 4-tuple so that the occurrences of elements in first positions is as before.
- Choose block offsets $s_0, \ldots, s_{b-1}$, each from $\{\pm 2, \pm 4\}$ so that, for each element, the sum of all offsets of blocks containing that element in the first position is 0.
- For each $0 \le i < b$, write $B_i = (x_0, x_1, x_2, x_3)$. Then for $0 \le j < 4$,
    - when (r_0, r_1, r_2, r_3) is the jth row of P_{s_i}, define block $D_{4i+j} = \{(x_0, r_0), (x_1, r_1), (x_2, r_2), (x_3, r_3)\}$.
    - when (r_0, r_1, r_2, r_3) is the $(7-j)$th row of P_{s_i}, define $D_{8b+\binom{v}{2}-1-4i-j} = \{(x_0, r_0), (x_1, r_1), (x_2, r_2), (x_3, r_3)\}$.

Doubling a Block Labelling

Type 2 Blocks

- We have ordered all blocks $D_0, \dots, D_{4b-1}$ and $D_{4b+\binom{v}{2}}, \dots, D_{8b+\binom{v}{2}-1}$ of the doubled $S(3,4,2v)$.
- To finish up, we must order the Type 2 blocks in some order as $D_{4b}, \dots, D_{4b+\binom{v}{2}-1}$.
- It has been shown that the edges of the complete graph of order v can be ordered to get dual DiffSum 0 when $v \not\equiv 0 \pmod 4$, 1 when $v \equiv 0 \pmod 4$, provided that $v \geq 6$ (Stewart 1964, C 2021).
- Let $e_0, \dots, e_{\binom{v}{2}-1}$ be such an ordering.
- For $0 \leq j < \binom{v}{2}$, with $e_i = \{x_0, x_1\}$, define $D_{4b+j} = \{(x_0,0),(x_0,1),(x_1,0),(x_1,1)\}$.

Doubling a Block Labelling
The Conclusion

Charles J. Colbourn
Arizona State University

- Whenever $v \equiv 4, 20 \pmod{24}$, there is an $S(3, 4, v)$ with dual DiffSum 0.
- Whenever $v \equiv 8, 16 \pmod{24}$ and $v > 8$, there is an $S(3, 4, v)$ with dual DiffSum 1.

Summing Up

- Block labelling of SQSs can also be obtained by doubling other 3-wise balanced designs, but the details are more involved!
- Thanks for listening!