Conflicts between Optimality Criteria in Incomplete-Block Designs for Microarray Experiments

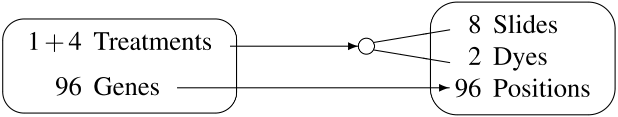
R. A. Bailey

Queen Mary University of London

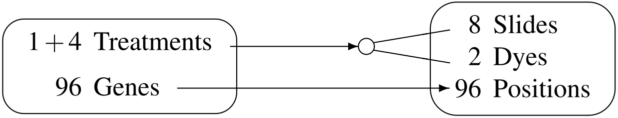
r.a.bailey@qmul.ac.uk

May 2007

A small microarray experiment

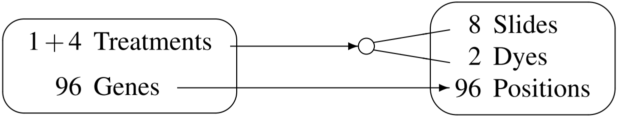


A small microarray experiment



- There is 1 ‘control’ treatment (labelled 0) and 4 other treatments.

A small microarray experiment



- There is 1 ‘control’ treatment (labelled 0) and 4 other treatments.
- ○ shows that we need to know a specific (non-orthogonal) design for the allocation of the treatments to the dye-slide combinations, such as

slides

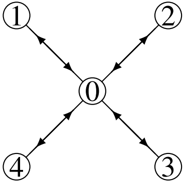
	1	2	3	4	5	6	7	8
red	0	1	0	2	0	3	0	4
green	1	0	2	0	3	0	4	0

Representation of the design as an oriented graph

Treatments are vertices; slides are edges, oriented from green to red.

slides

	1	2	3	4	5	6	7	8
red	0	1	0	2	0	3	0	4
green	1	0	2	0	3	0	4	0

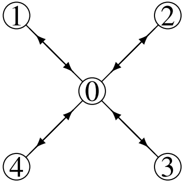


Representation of the design as an oriented graph

Treatments are vertices; slides are edges, oriented from green to red.

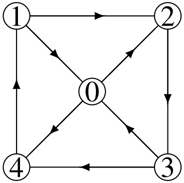
slides

	1	2	3	4	5	6	7	8
red	0	1	0	2	0	3	0	4
green	1	0	2	0	3	0	4	0



slides

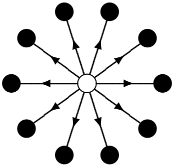
	1	2	3	4	5	6	7	8
red	0	2	0	4	2	3	4	1
green	1	0	3	0	1	2	3	4



Two applications

Large Brazilian forestry experiment (v. Julio Bueno)

10 varieties of tree are being grown in several states for 5 years.
It is proposed to compare each with control using a single-reference design with 10 slides. Is this a good use of resources?



Large number of mutations in yeast (Hughes et al., *Cell*, **102**)

300 mutant varieties of yeast were compared with wild-type yeast using a double-reference design with 600 slides. Was that the best use of resources?

Model and optimality criteria

t treatments $\quad$ b slides (call these “blocks”) $\quad$ 2 dyes

Model and optimality criteria

t treatments b slides (call these "blocks") 2 dyes

Assume that the logarithm of the intensity of treatment i coloured with dye l in block k has expected value

$$\tau_i + \beta_k + \delta_l$$

and variance σ^2, independent of all other responses.

Model and optimality criteria

t treatments b slides (call these "blocks")   2 dyes

Assume that the logarithm of the intensity of treatment i coloured with dye l in block k has expected value

$$\tau_i + \beta_k + \delta_l$$

and variance σ^2, independent of all other responses.

To estimate all the $\tau_i - \tau_j$, we need $b \geq t - 1$.

Model and optimality criteria

t treatments b slides (call these “blocks”) 2 dyes

Assume that the logarithm of the intensity of treatment i coloured with dye l in block k has expected value

$$\tau_i + \beta_k + \delta_l$$

and variance σ^2, independent of all other responses.

To estimate all the $\tau_i - \tau_j$, we need $b \geq t-1$.

A design is A-optimal if it minimizes the sum of the variances of the estimators of these differences.

Model and optimality criteria

t treatments $\quad$ b slides (call these "blocks") $\quad$ 2 dyes

Assume that the logarithm of the intensity of treatment i coloured with dye l in block k has expected value

$$\tau_i + \beta_k + \delta_l$$

and variance σ^2, independent of all other responses.

To estimate all the $\tau_i - \tau_j$, we need $b \geq t - 1$.

A design is A-optimal if it minimizes the sum of the variances of the estimators of these differences.

A design is D-optimal if it minimizes the volume of the confidence ellipsoid for the vector $(\tau_1, \ldots, \tau_t)$ subject to $\sum \tau_i = 0$.

Model and optimality criteria

t treatments b slides (call these "blocks")   2 dyes

Assume that the logarithm of the intensity of treatment i coloured with dye l in block k has expected value

$$\tau_i + \beta_k + \delta_l$$

and variance σ^2, independent of all other responses.

To estimate all the $\tau_i - \tau_j$, we need $b \geq t - 1$.

A design is A-optimal if it minimizes the sum of the variances of the estimators of these differences.

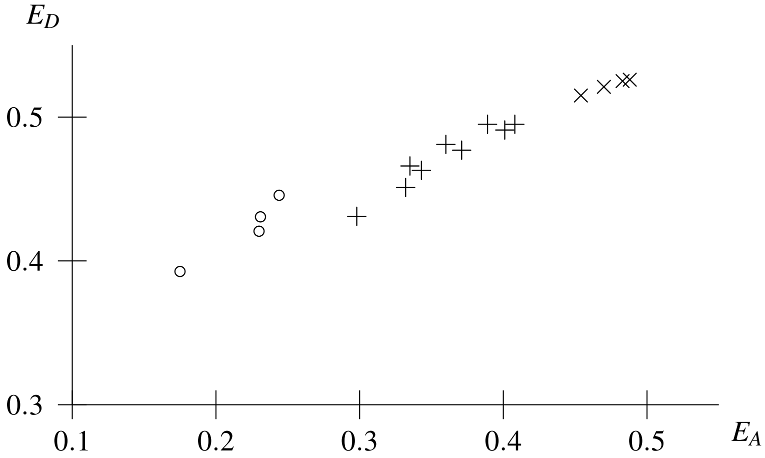
A design is D-optimal if it minimizes the volume of the confidence ellipsoid for the vector $(\tau_1, \ldots, \tau_t)$ subject to $\sum \tau_i = 0$.

If $t = 2$ then A-optimal = D-optimal.

Temporarily ignore the dyes

We will come back to them later.

Typical behaviour of the optimality criteria



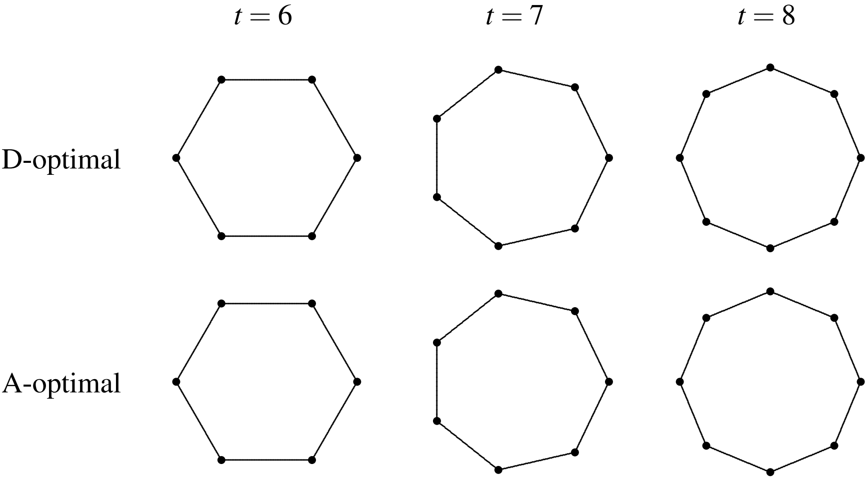
Optimality criteria for all connected equireplicate designs with 8 treatments in 12 blocks of size 2: graphs with edge-connectivity 3, 2, 1 are shown as $\times$, $+$, $\circ$ respectively

What happens when $b = t$?

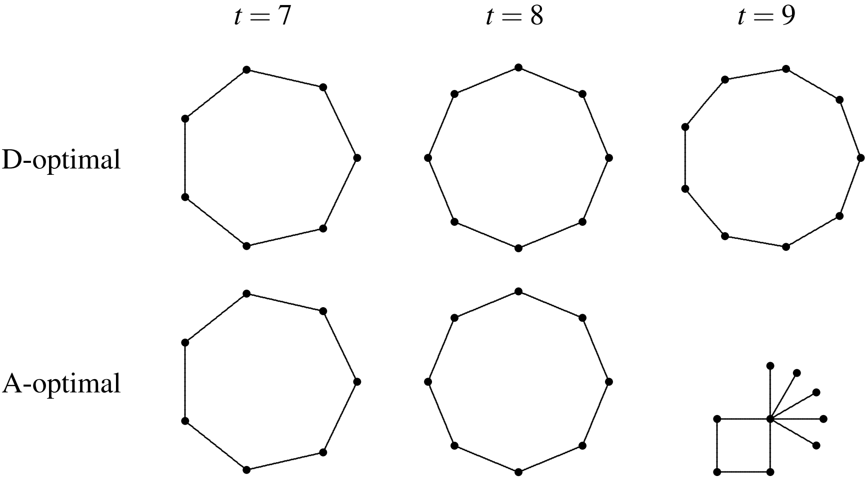
Computer investigation by

- Jones and Eccleston (1980)
- Kerr and Churchill (2001)
- Wit, Nobile and Khanin (2005)
- Ceraudo (2005).

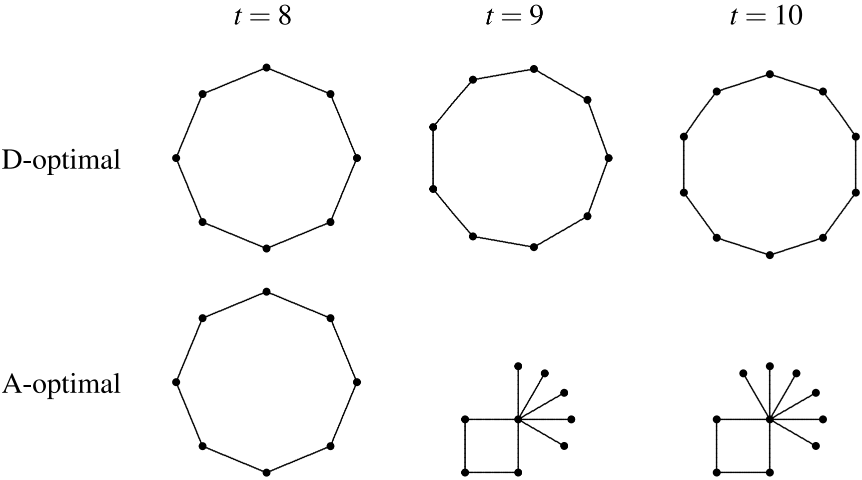
Optimal designs when $b = t$



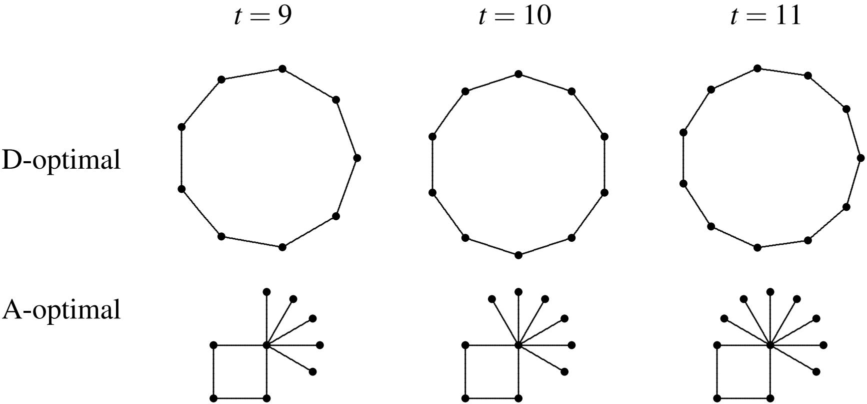
Optimal designs when $b = t$



Optimal designs when $b = t$



Optimal designs when $b = t$



D-optimality

Cheng (1978), after Gaffke (1978), after Kirchhoff (1847):

$$E_D = \frac{(t \times \text{number of spanning trees})^{1/(t-1)}}{2\bar{r}}$$

D-optimality

Cheng (1978), after Gaffke (1978), after Kirchhoff (1847):

$$E_D = \frac{(t \times \text{number of spanning trees})^{1/(t-1)}}{2\bar{r}}$$

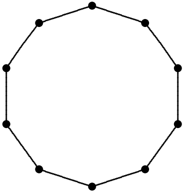
number of spanning trees $=$
number of ways of removing $b-t+1$ edges without disconnecting the graph, (which is easy to calculate by hand when $b-t$ is small)

D-optimality

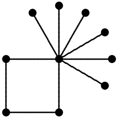
Cheng (1978), after Gaffke (1978), after Kirchhoff (1847):

$$E_D = \frac{(t \times \text{number of spanning trees})^{1/(t-1)}}{2\bar{r}}$$

number of spanning trees $=$
number of ways of removing $b-t+1$ edges without disconnecting the graph, (which is easy to calculate by hand when $b-t$ is small)



10 spanning trees

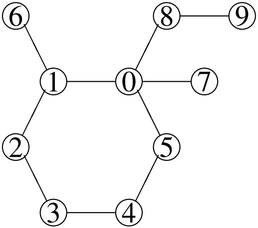


4 spanning trees

The loop design is uniquely D-optimal when $b = t$.

A-optimality

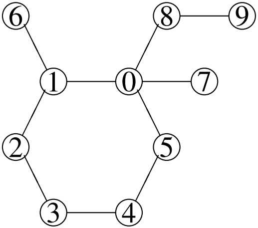
If $b = t$, the graph contains a single circuit.



A-optimality

If $b = t$, the graph contains a single circuit.

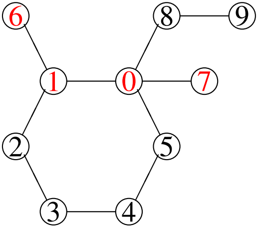
Let $V_{ij} =$ variance of estimator of $\tau_i - \tau_j$.



A-optimality

If $b = t$, the graph contains a single circuit.

Let $V_{ij} =$ variance of estimator of $\tau_i - \tau_j$.

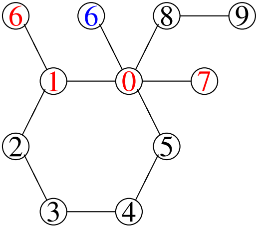


$$V_{67} = V_{61} + V_{10} + V_{07} = V_{10} + 4\sigma^2$$

A-optimality

If $b = t$, the graph contains a single circuit.

Let $V_{ij} =$ variance of estimator of $\tau_i - \tau_j$.



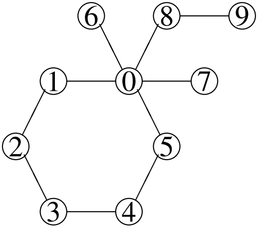
$$V_{67} = V_{61} + V_{10} + V_{07} = V_{10} + 4\sigma^2$$

$$V_{67} = V_{60} + V_{07} = 4\sigma^2$$

A-optimality

If $b = t$, the graph contains a single circuit.

Let $V_{ij} =$ variance of estimator of $\tau_i - \tau_j$.

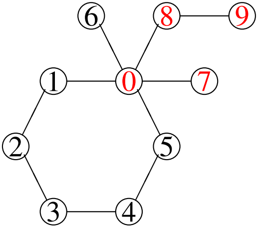


$$V_{67} = V_{60} + V_{07} = 4\sigma^2$$

A-optimality

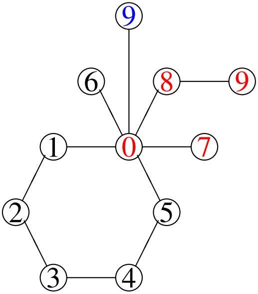
If $b = t$, the graph contains a single circuit.

Let $V_{ij} =$ variance of estimator of $\tau_i - \tau_j$.



$$V_{97} = V_{98} + V_{80} + V_{07} = V_{80} + 4\sigma^2$$

A-optimality



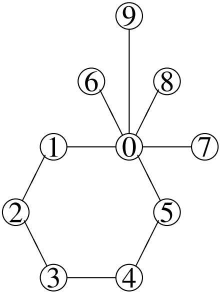
If $b = t$, the graph contains a single circuit.

Let $V_{ij} =$ variance of estimator of $\tau_i - \tau_j$.

$$V_{97} = V_{98} + V_{80} + V_{07} = V_{80} + 4\sigma^2$$

$$V_{97} = V_{90} + V_{07} = 4\sigma^2$$

A-optimality



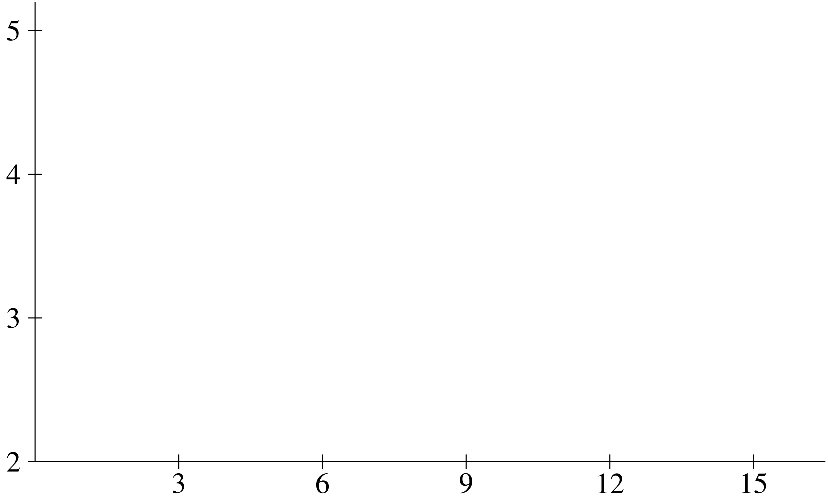
If $b = t$, the graph contains a single circuit.

Let $V_{ij} =$ variance of estimator of $\tau_i - \tau_j$.

For a given size of circuit, the total variance is minimized when everything outside the circuit is attached to the same vertex of the circuit.

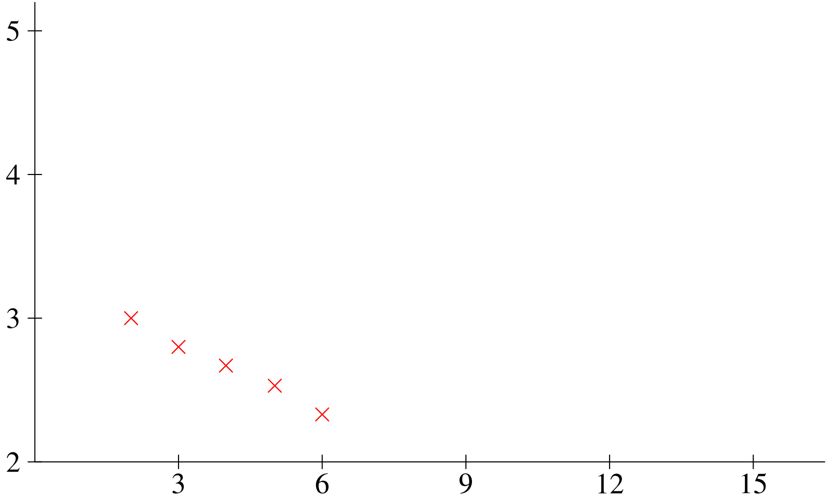
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



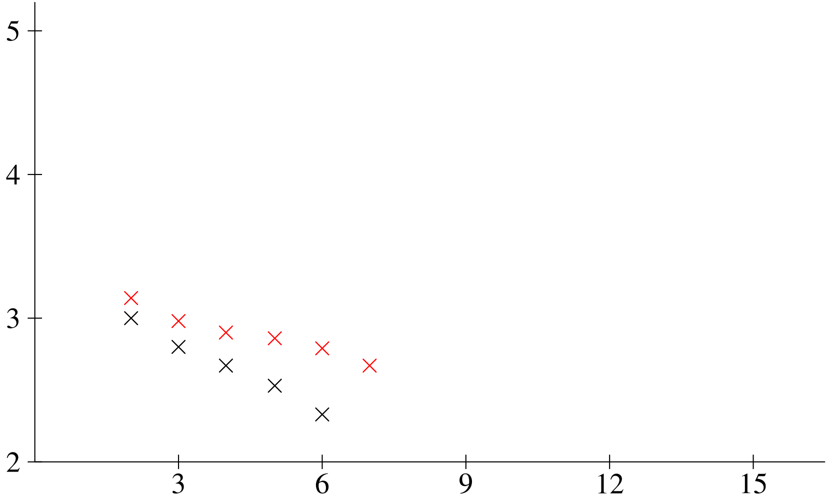
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



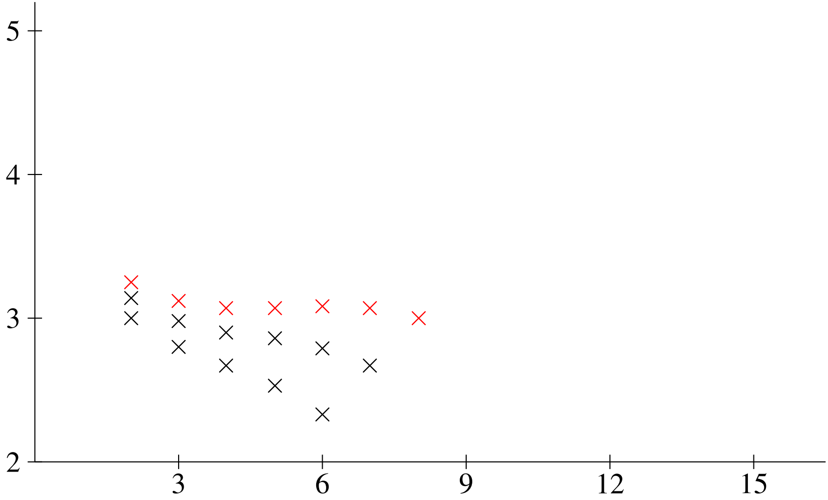
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



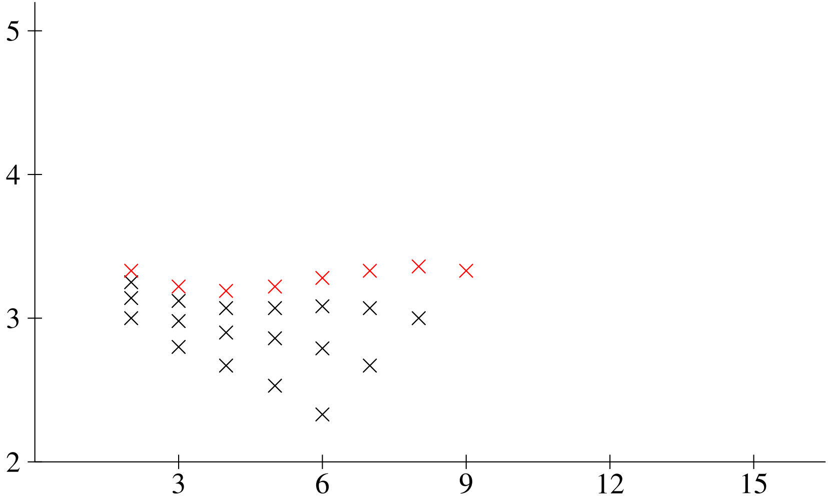
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



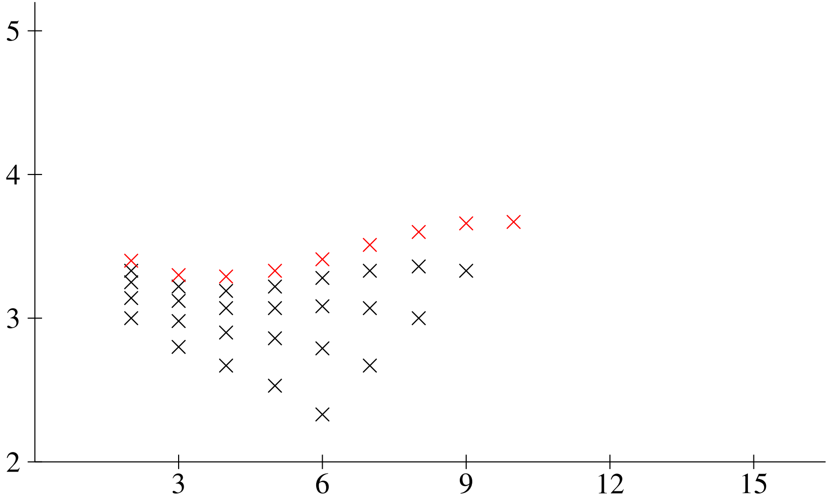
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



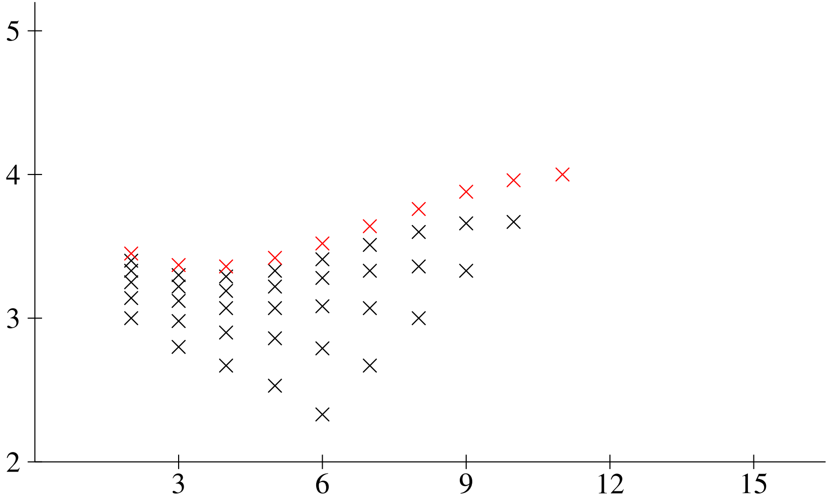
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



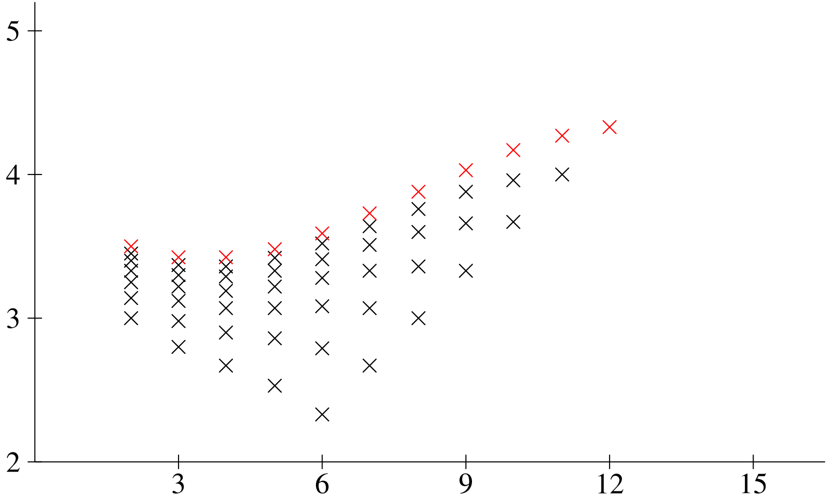
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



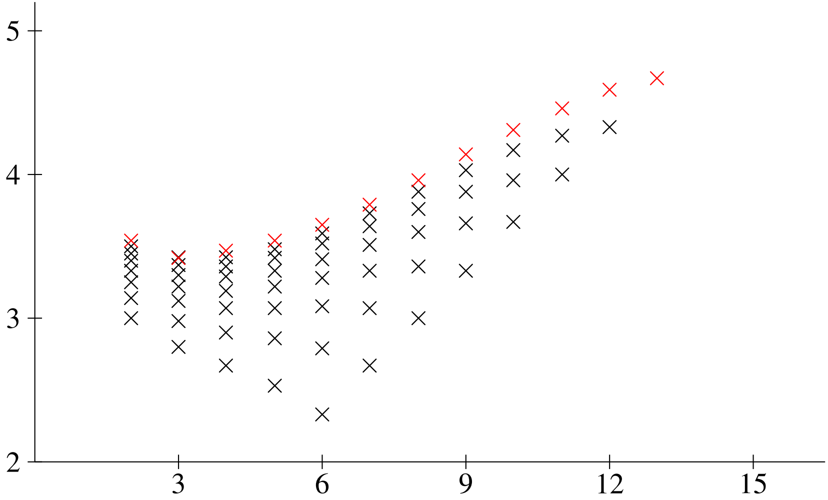
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



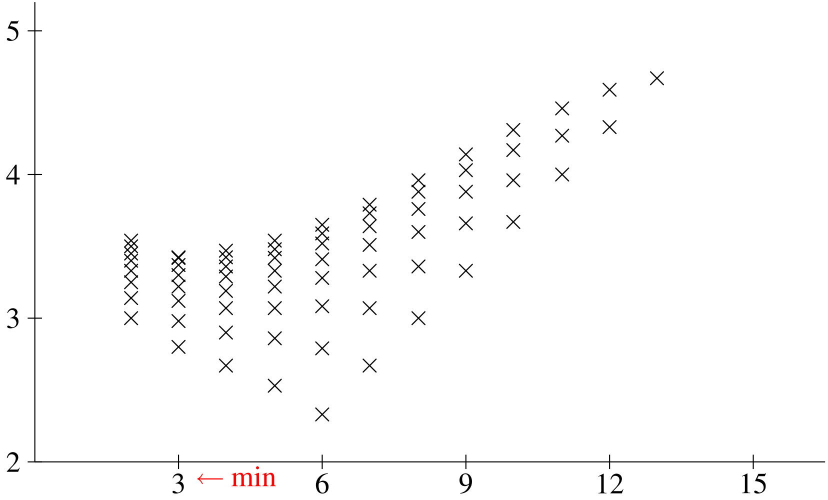
A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.

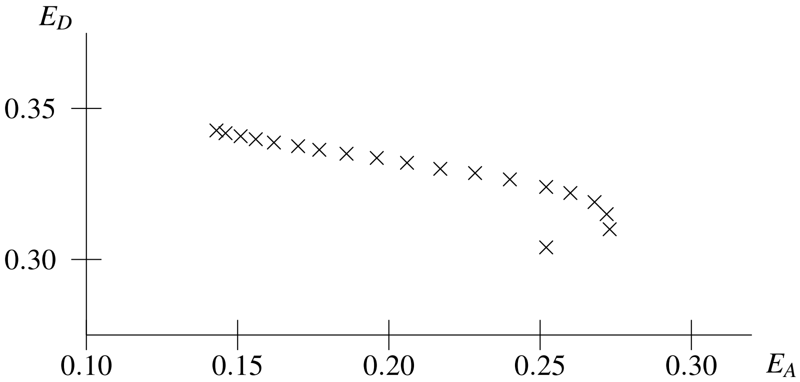


A-optimality: continued

Average pairwise variance is a cubic function of the size of the circuit.



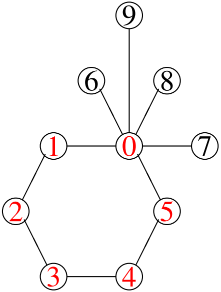
Optimality criteria for designs for 20 treatments in 20 blocks



The two criteria give essentially reverse rankings.

Assigning colours to a circuit with leaves

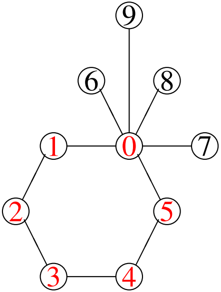
The difference between the colours can be estimated only from the circuit.



Assigning colours to a circuit with leaves

The difference between the colours can be estimated only from the circuit.

More leaves → smaller circuit → larger variance for colour difference.

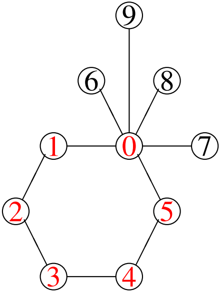


Assigning colours to a circuit with leaves

The difference between the colours can be estimated only from the circuit.

More leaves → smaller circuit → larger variance for colour difference.

Variance between circuit nodes increases unless the arrows are directed around the circuit.



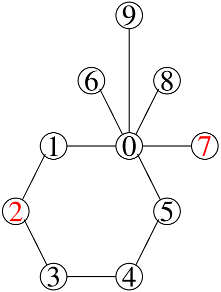
Assigning colours to a circuit with leaves

The difference between the colours can be estimated only from the circuit.

More leaves → smaller circuit → larger variance for colour difference.

Variance between circuit nodes increases unless the arrows are directed around the circuit.

Variance between a leaf and a circuit node increases because the leaf occurs with only one colour.



Assigning colours to a circuit with leaves

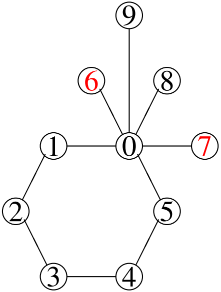
The difference between the colours can be estimated only from the circuit.

More leaves → smaller circuit → larger variance for colour difference.

Variance between circuit nodes increases unless the arrows are directed around the circuit.

Variance between a leaf and a circuit node increases because the leaf occurs with only one colour.

Variance between leaves increases unless they all have the same colour.



Assigning colours to a circuit with leaves

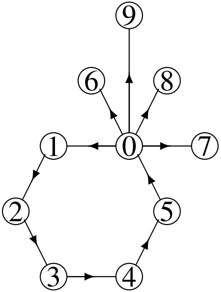
The difference between the colours can be estimated only from the circuit.

More leaves $\rightarrow$ smaller circuit $\rightarrow$ larger variance for colour difference.

Variance between circuit nodes increases unless the arrows are directed around the circuit.

Variance between a leaf and a circuit node increases because the leaf occurs with only one colour.

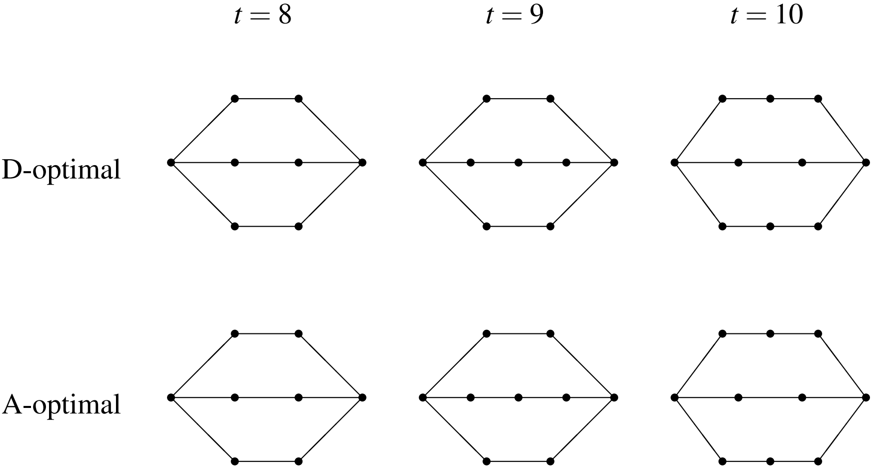
Variance between leaves increases unless they all have the same colour.



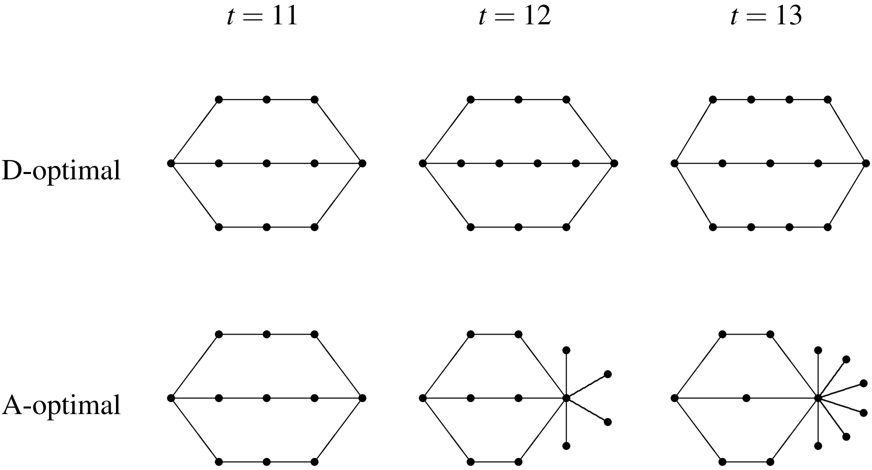
What happens when $b = t + 1$?

A similar analysis shows that the A-optimality and D-optimality criteria conflict when $t \geq 12$.

Optimal designs when $b = t + 1$



Optimal designs when $b = t + 1$



What happens for larger values of $b - t$?

Bad news theorem

Given any fixed value of $b - t$, there is a threshold T such that when $t \geq T$ the A- and D-optimality criteria conflict.

In fact, when $t \geq T$, the A-better designs have many vertices of valency 1 (leaves) attached to single vertex of some small graph, whereas the D-better designs have no leaves.

How do we resolve this conflict between A-optimality and D-optimality?

How do we resolve this conflict between A-optimality and D-optimality?

Jones and Eccleston just commented on the discrepancy.

How do we resolve this conflict between A-optimality and D-optimality?

Jones and Eccleston just commented on the discrepancy.

Kerr and Churchill recommended even replication for each treatment, so that each treatment is coloured with each dye for half its replications. This avoids designs with leaves.

How do we resolve this conflict between A-optimality and D-optimality?

Jones and Eccleston just commented on the discrepancy.

Kerr and Churchill recommended even replication for each treatment, so that each treatment is coloured with each dye for half its replications. This avoids designs with leaves.

Wit, Nobile and Khanin gave a non-standard definition of A-optimality and appeared to favour D-optimality.

How do we resolve this conflict between A-optimality and D-optimality?

Jones and Eccleston just commented on the discrepancy.

Kerr and Churchill recommended even replication for each treatment, so that each treatment is coloured with each dye for half its replications. This avoids designs with leaves.

Wit, Nobile and Khanin gave a non-standard definition of A-optimality and appeared to favour D-optimality.

RAB proposes that we need $b \geq (9/8)t$.

How do we resolve this conflict between A-optimality and D-optimality?

Jones and Eccleston just commented on the discrepancy.

Kerr and Churchill recommended even replication for each treatment, so that each treatment is coloured with each dye for half its replications. This avoids designs with leaves.

Wit, Nobile and Khanin gave a non-standard definition of A-optimality and appeared to favour D-optimality.

RAB proposes that we need $b \geq (9/8)t$.

How do we resolve this conflict between A-optimality and D-optimality?

Jones and Eccleston just commented on the discrepancy.

Kerr and Churchill recommended even replication for each treatment, so that each treatment is coloured with each dye for half its replications. This avoids designs with leaves.

Wit, Nobile and Khanin gave a non-standard definition of A-optimality and appeared to favour D-optimality.

RAB proposes that we need $b \geq (9/8)t$.

Good news theorem

Inserting 1 or 2 (or sometimes 3) vertices into the edges of a graph with no leaves gives a lower average pairwise variance than attaching the extra vertices to a single vertex of that graph.

Strategy for choosing a design when $t/8 \le b - t \le t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
2. Insert up to 2 treatments in each edge.

Strategy for choosing a design when $t/8 \leq b-t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
2. Insert up to 2 treatments in each edge.

Forestry example

$t = 10 + 1 = 11$

Strategy for choosing a design when $t/8 \leq b-t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
2. Insert up to 2 treatments in each edge.

Forestry example

$t = 10 + 1 = 11 \qquad \surd$

Strategy for choosing a design when $t/8 \leq b-t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
2. Insert up to 2 treatments in each edge.

Forestry example

$t = 10+1 = 11 \qquad \surd$

$t = 12 \Rightarrow b-t = 2$

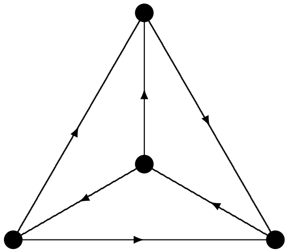
Strategy for choosing a design when $t/8 \leq b - t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
2. Insert up to 2 treatments in each edge.

Forestry example

$t = 10 + 1 = 11$ $\surd$

$t = 12 \Rightarrow b - t = 2$



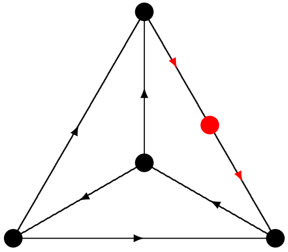
Strategy for choosing a design when $t/8 \leq b - t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
2. Insert up to 2 treatments in each edge.

Forestry example

$t = 10 + 1 = 11$ $\surd$

$t = 12 \Rightarrow b - t = 2$



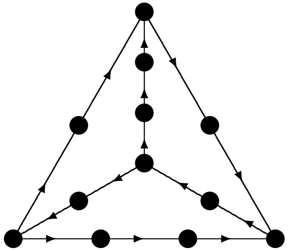
Strategy for choosing a design when $t/8 \leq b - t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
2. Insert up to 2 treatments in each edge.

Forestry example

$t = 10 + 1 = 11$ $\quad\surd$

$t = 12 \Rightarrow b - t = 2$



Strategy for choosing a design when $t/8 \leq b-t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.

2. Insert up to 2 treatments in each edge.

Strategy for choosing a design when $t/8 \leq b - t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
   1.1 Find the best equireplicate design for $b-t$ treatments in $b-t$ blocks of size 3.

2. Insert up to 2 treatments in each edge.

Strategy for choosing a design when $t/8 \leq b-t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
   1.1 Find the best equireplicate design for $b-t$ treatments in $b-t$ blocks of size 3.
   1.2 Draw its Levi graph (which is bipartite, with $2(b-t)$ vertices of valency 3).

2. Insert up to 2 treatments in each edge.

Strategy for choosing a design when $t/8 \leq b-t \leq t/2$

1. Choose the best equireplicate design for $2(b-t)$ treatments in $3(b-t)$ blocks, including dye allocation.
   1.1 Find the best equireplicate design for $b-t$ treatments in $b-t$ blocks of size 3.
   1.2 Draw its Levi graph (which is bipartite, with $2(b-t)$ vertices of valency 3).
   1.3 Using the algorithm from Hall's Marriage Theorem, (also König's Theorem) orient the edges so that each lower vertex has 2 out-edges and 1 in-edge and each upper vertex has 1 out-edge and 2 in-edges.
2. Insert up to 2 treatments in each edge.

Example with large t

$t = 216$

Example with large t

$t = 216$

Need $b - t \geq t/8 = 27$ so try $b - t = 27$.

Example with large t

$t = 216$

Need $b - t \geq t/8 = 27$ so try $b - t = 27$.

The best equireplicate design for 27 treatments in 27 blocks of size 3 is the simple cubic lattice.

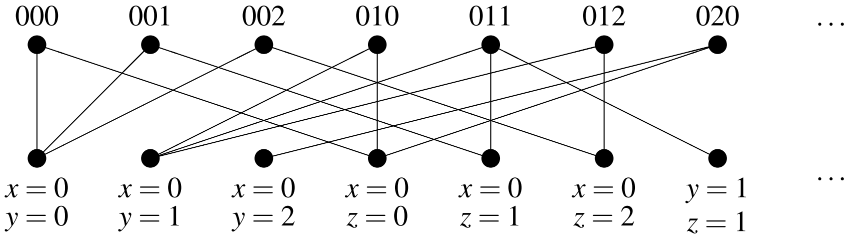
Example with large t

$t = 216$

Need $b - t \geq t/8 = 27$ so try $b - t = 27$.

The best equireplicate design for 27 treatments in 27 blocks of size 3 is the simple cubic lattice.

Its Levi graph has 54 vertices and 81 edges.



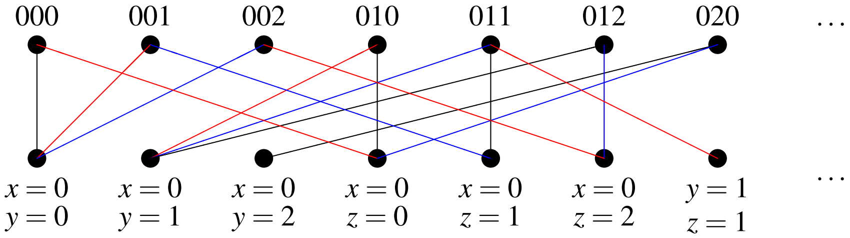
Example with large t

$t = 216$

Need $b - t \geq t/8 = 27$ so try $b - t = 27$.

The best equireplicate design for 27 treatments in 27 blocks of size 3 is the simple cubic lattice.

Its Levi graph has 54 vertices and 81 edges.



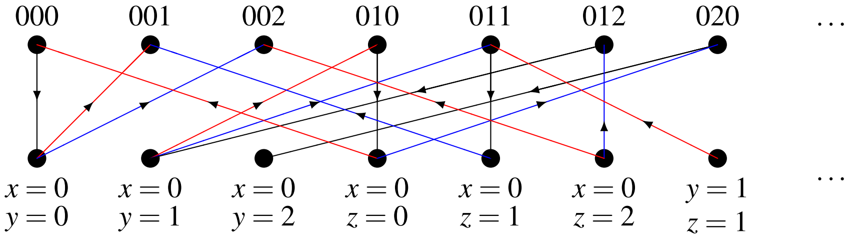
Example with large t

$t = 216$

Need $b - t \geq t/8 = 27$ so try $b - t = 27$.

The best equireplicate design for 27 treatments in 27 blocks of size 3 is the simple cubic lattice.

Its Levi graph has 54 vertices and 81 edges.



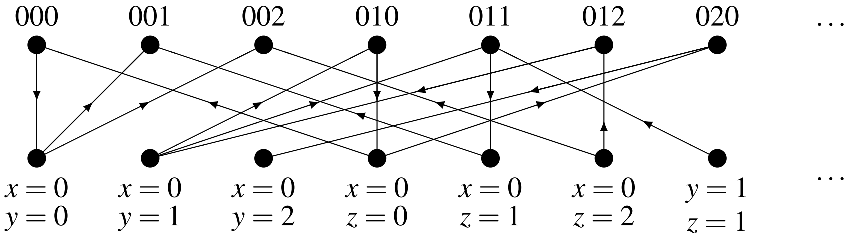
Example with large t

$t = 216$

Need $b - t \geq t/8 = 27$ so try $b - t = 27$.

The best equireplicate design for 27 treatments in 27 blocks of size 3 is the simple cubic lattice.

Its Levi graph has 54 vertices and 81 edges.



Inserting 2 vertices into each edge gives 216 vertices and 243 edges.