Fractional Sudoku

A coherent configuration and fractional completion threshold

Peter Dukes* and Kate Nimegeers

University of Victoria Math & Stats

Table of contents

Introduction

Partial latin squares

A *latin square* of order n is an $n \times n$ array with entries from $[n] := \{1, 2, \ldots, n\}$ such that each symbol appears exactly once in every row and every column.

A *partial latin square* of order n is an $n \times n$ array whose cells are either empty or filled with one of n symbols in such a way that each symbol appears at most once in every row and every column.

A *completion* of a partial latin square P is a latin square which contains every entry of P.

Examples

Here is a partial latin square and a completion of it to a latin square of order 5.

1				
	2	4		
			3	
			1	

1	3	5	2	4
3	2	4	5	1
4	1	2	3	5
2	5	1	4	3
5	4	3	1	2

We can't use any row, column, or symbol very often if we want a completion to exist:

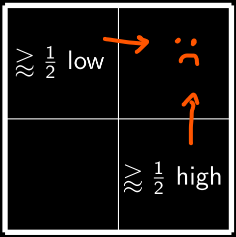
1	2	3	4	
				5

1				
2				
3				
4				
	5			

1				
	1			
		1		
			1	
				2

Barriers

We can't use any row, column, or symbol very often if we want a completion to exist:



Completion threshold

Let us say that a partial latin square is *ϵ-dense* if no row, column, or symbol is used more than ϵn times.

Completion threshold

Let us say that a partial latin square is ϵ-*dense* if no row, column, or symbol is used more than ϵn times.

Theorem

For sufficiently large n, every ϵ-dense partial latin square of order n has a completion.

- Prehistory: $\epsilon \to 0$
- 1991: $\epsilon = 10^{-7}$ (Gustavsson)
- 2013: $\epsilon \approx 10^{-4}$ (Bartlett)
- 2019: $\epsilon \approx 0.04$ (Bowditch & D.)

Completion threshold

Let us say that a partial latin square is ϵ-*dense* if no row, column, or symbol is used more than ϵn times.

Theorem

For sufficiently large n, every ϵ-dense partial latin square of order n has a completion.

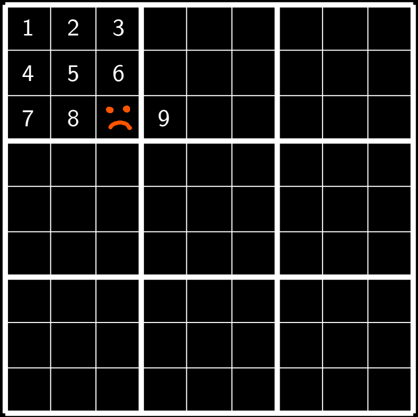
- Prehistory: $\epsilon \to 0$
- 1991: $\epsilon = 10^{-7}$ (Gustavsson)
- 2013: $\epsilon \approx 10^{-4}$ (Bartlett)
- 2019: $\epsilon \approx 0.04$ (Bowditch & D.)
- Conjecture: $\epsilon = 0.25$

A *Sudoku latin square* of type (h, w) is a latin square of order $n = hw$ divided into a $w \times h$ pattern of $h \times w$ sub-arrays (boxes), each of which contains every symbol exactly once.

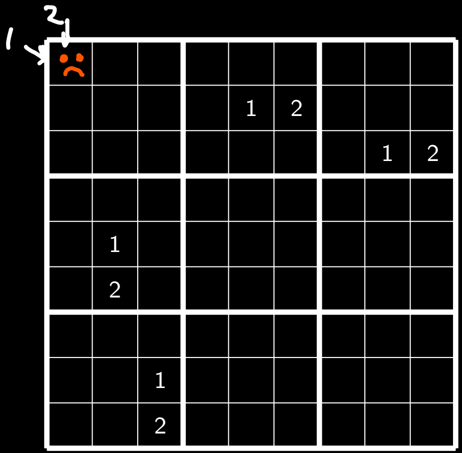
A partial Sudoku is a partial latin square in which each symbol appears at most once in every box.

Completion is defined analogously as for partial latin squares.

Stronger sparseness assumptions



Stronger sparseness assumptions



Stronger sparseness assumptions

Definition

Let us say that a Sudoku of type (h, w) is *ϵ-dense* if:

- each row, column, and box has at most ϵn filled cells; AND
- each symbol occurs at most ϵh times in any bundle of h rows corresponding to the box partition, and likewise at most ϵw times in any bundle of w columns.

Stronger sparseness assumptions

Definition

Let us say that a Sudoku of type (h, w) is ϵ-*dense* if:

- each row, column, and box has at most ϵn filled cells; AND
- each symbol occurs at most ϵh times in any bundle of h rows corresponding to the box partition, and likewise at most ϵw times in any bundle of w columns.

Question: Can ϵ-dense Sudoku always be completed for some $\epsilon > 0$?

The linear system

Linear equations

Consider an $n \times n$ Sudoku. Let the rows, columns, symbols and boxes be denoted r_i, c_j, s_k, b_ℓ, respectively, where $i, j, k, \ell \in [n]$.

Let x_{ijk} denote the number/fraction of symbols s_k placed in cell (i, j). Sudoku constraints correspond to linear equations on these variables:

- every cell has exactly one symbol: $\sum_k x_{ijk} = 1$ for each $(i, j) \in [n]^2$.
- every row has every symbol once: $\sum_j x_{ijk} = 1$ for each $(i, k) \in [n]^2$.
- " column " " : $\sum_i x_{ijk} = 1$ for each $(j, k) \in [n]^2$.

Linear equations

Consider an $n \times n$ Sudoku. Let the rows, columns, symbols and boxes be denoted r_i, c_j, s_k, b_ℓ, respectively, where $i, j, k, \ell \in [n]$.

Let x_{ijk} denote the number/fraction of symbols s_k placed in cell (i, j). Sudoku constraints correspond to linear equations on these variables:

- every cell has exactly one symbol: $\sum_k x_{ijk} = 1$ for each $(i, j) \in [n]^2$.
- every row has every symbol once: $\sum_j x_{ijk} = 1$ for each $(i, k) \in [n]^2$.
- " column " " : $\sum_i x_{ijk} = 1$ for each $(j, k) \in [n]^2$.
- every box contains every symbol once:

  $$\sum_{(i,j)\in \mathrm{box}(\ell)} x_{ijk} = 1$$

  for each $(k, \ell) \in [n]^2$.

Linear system

This results in a $4n^2 \times n^3$ linear system

$$W\vec{\mathbf{x}} = \vec{\mathbf{1}} \quad \text{(all-ones vector)}.$$

There is a naïve solution: $\vec{\mathbf{x}} = \frac{1}{n}\vec{\mathbf{1}}$.

If some entries have been pre-filled, a similar linear system can be used. We can either adjust the right side, or delete those variables x_{ijk} which are unavailable.

If we get a $\{0, 1\}$-valued solution $\vec{\mathbf{x}}$, this leads to a completion of the corresponding partial Sudoku.

Linear system

This results in a $4n^2 \times n^3$ linear system

$$W\vec{\mathbf{x}} = \vec{\mathbf{1}} \quad \text{(all-ones vector).}$$

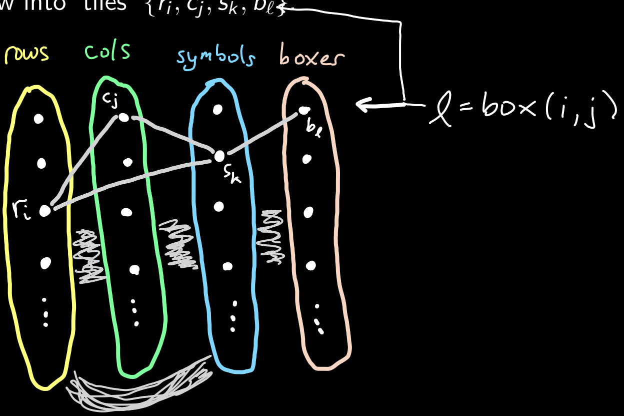
There is a naïve solution: $\vec{\mathbf{x}} = \frac{1}{n}\vec{\mathbf{1}}$.

If some entries have been pre-filled, a similar linear system can be used. We can either adjust the right side, or delete those variables x_{ijk} which are unavailable.

If we get a $[0,1]$-valued solution $\vec{\mathbf{x}}$, this leads to a *fractional* completion of the corresponding partial Sudoku.

Graph decomposition model

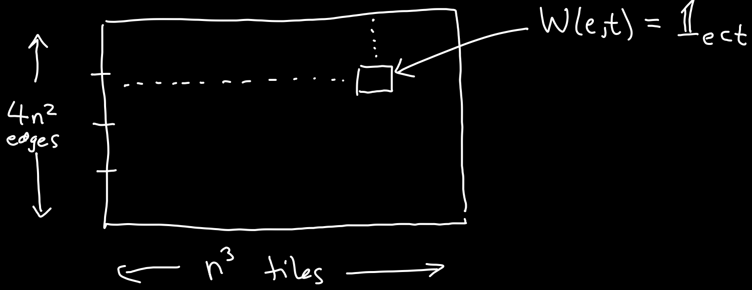
We can think of a Sudoku as an edge-decomposition of the 4-partite graph shown below into 'tiles' $\{r_i, c_j, s_k, b_\ell\}$.



Graph decomposition model

We can think of a Sudoku as an edge-decomposition of the 4-partite graph shown below into 'tiles' $\{r_i, c_j, s_k, b_\ell\}$.

W is the $\{0, 1\}$ inclusion matrix of edges versus tiles.



The normal system

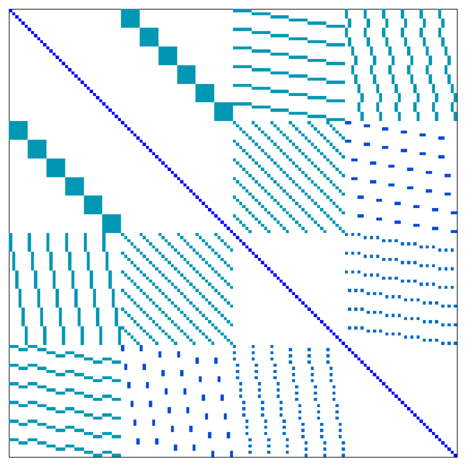
Let $M = WW^\top$ and consider instead the normal system $M\vec{\mathbf{x}} = \vec{\mathbf{1}}$.

Rows and columns of M are indexed by edges $r_i c_j, r_i s_k, c_j s_k, b_\ell s_k$. Entries tell us how many tiles contain two given edges.

That is,

$$M(e,f) = \begin{cases} n & \text{if } e = f \\ h & \text{if } e \cup f = \{c_j, s_k, b_\ell\} \text{ where } c_j \text{ meets } b_\ell \\ w & \text{if } e \cup f = \{r_i, s_k, b_\ell\} \text{ where } r_i \text{ meets } b_\ell \\ 1 & \text{if } e \cup f \text{ has exactly one of each } r_i, c_j, s_k \\ 0 & \text{otherwise.} \end{cases}$$

Coefficient matrix

$M =$ 

Coherent configuration

Using the symmetries present in Sudoku, we can express M in an algebra of fixed dimension, independent of h and w.

- Rows can be equal, unequal in the same bundle, or in different bundles. Same with columns.
- Boxes can be equal, row-adjacent, column-adjacent, or neither.
- Symbols can be equal or unequal.

These lead to 69 relations on edges.

Relations


row-col
row-symbol
col-symbol
box-symbol
row-col
row-symbol
col-symbol
box-symbol

Structure constants

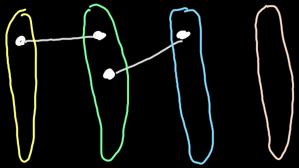
We found and stored **symbolic** structure constants for this coherent configuration using the following procedure:

- argue that they are all polynomials of degree ≤ 2 in each of h, w;
- directly compute all structure constants for the nine cases $2 \leq h, w \leq 4$;
- interpolate to arrive at symbolic expressions.

Possible values

Fix two edges e, f which are related in some way. In picking a third edge having prescribed relations with each of e, f, we multiply two of:

- row choices $\in \{0, 1, h, h-1, h-2, n-h, n-2h\}$;
- column choices $\in \{0, 1, w, w-1, w-2, n-w, n-2w\}$;
- symbol choices $\in \{0, 1, n, n-1, n-2\}$;
- box choices $\in \{0, 1, h, h-1, h-2\} * \{0, 1, w, w-1, w-2\}$.

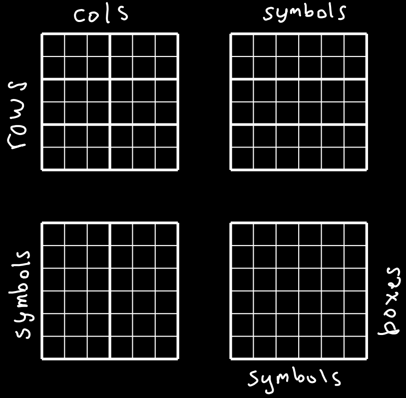


Spectral decomposition

Summary

Proposition

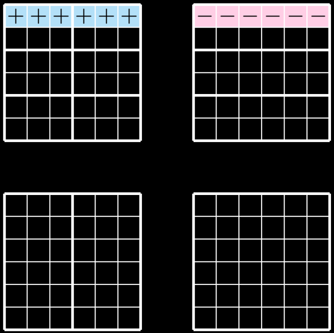
The eigenvalues of M are $\theta_j = jn$, $j = 0, 1, \ldots, 4$. *Each eigenspace has a basis of eigenvectors consisting of vectors with entries in* $\{0, \pm 1\}$.



Kernel of M

- $\theta_0 = 0$; kernel dimension $3n + (h + w)(n - 1)$

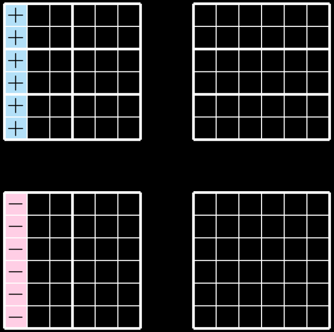
(A)



Kernel of M

- $\theta_0 = 0$; kernel dimension $3n + (h + w)(n - 1)$

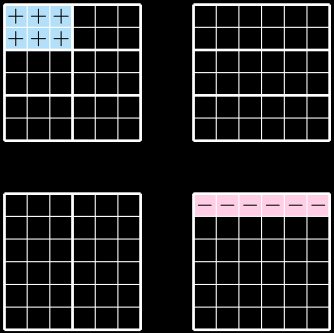
(A')



Kernel of M

- $\theta_0 = 0$; kernel dimension $3n + (h+w)(n-1)$

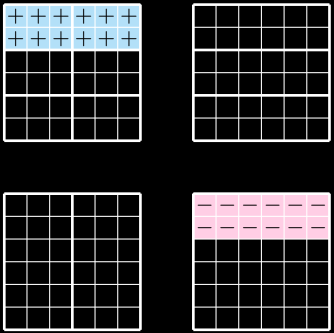
(B)



Kernel of M

- $\theta_0 = 0$; kernel dimension $3n + (h + w)(n - 1)$

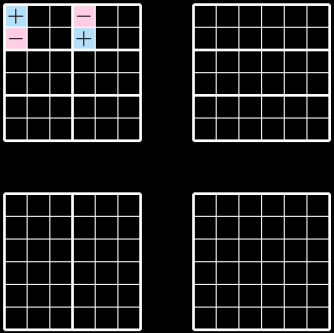
(C)



Other eigenvalues and eigenvectors

- $\theta_1 = n$; eigenspace dimension $4n^2 - (2n-3)(h+w) - 5n - 1$

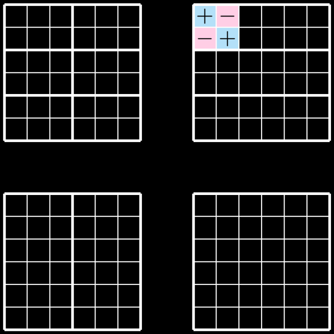
(A)



Other eigenvalues and eigenvectors

- $\theta_1 = n$; eigenspace dimension $4n^2 - (2n-3)(h+w) - 5n - 1$

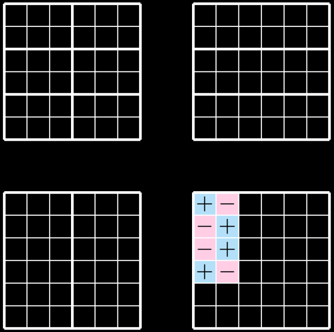
(B)



Other eigenvalues and eigenvectors

- $\theta_1 = n$; eigenspace dimension $4n^2 - (2n-3)(h+w) - 5n - 1$

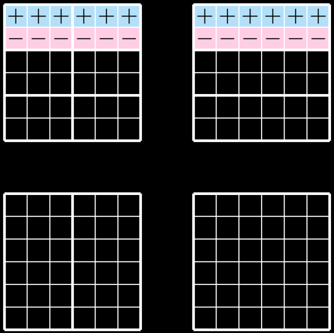
(C)



Other eigenvalues and eigenvectors

- $\theta_2 = 2n$; eigenspace dimension $(n-3)(h+w-1)+2n$

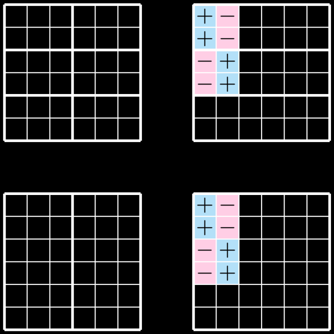
(A)



Other eigenvalues and eigenvectors

- $\theta_2 = 2n$; eigenspace dimension $(n-3)(h+w-1)+2n$

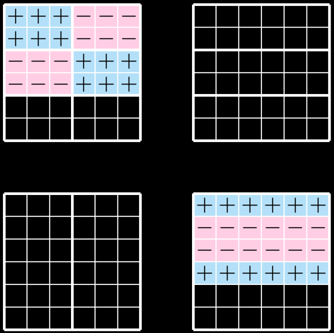
(B)



Other eigenvalues and eigenvectors

- $\theta_2 = 2n$; eigenspace dimension $(n-3)(h+w-1)+2n$

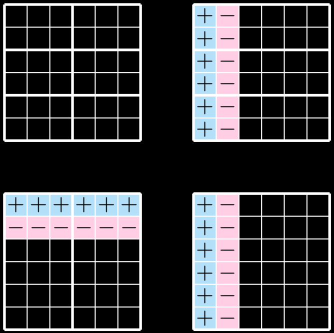
(C)



Other eigenvalues and eigenvectors

- $\theta_3 = 3n$; eigenspace dimension $n + h + w - 3$

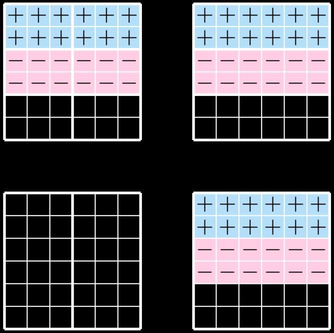
(A)



Other eigenvalues and eigenvectors

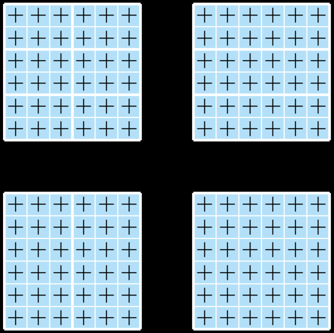
- $\theta_3 = 3n$; eigenspace dimension $n + h + w - 3$

(B)



Other eigenvalues and eigenvectors

- $\theta_4 = 4n$; eigenspace dimension 1



Projectors

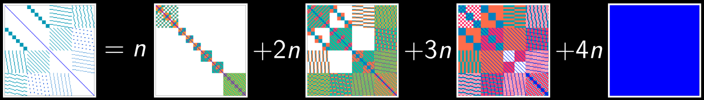
Using orthogonal projections onto the eigenspaces, we can write

$$M = nE_1 + 2nE_2 + 3nE_3 + 4nE_4.$$

Projectors

Using orthogonal projections onto the eigenspaces, we can write

$$M = nE_1 + 2nE_2 + 3nE_3 + 4nE_4.$$



A generalized inverse

Let K denote projection onto the kernel. For $x \in \mathbb{R}$, we can invert the additive shift $A = M + \frac{n}{x}K$ as

$$A^{-1} = \frac{1}{n}\left(xK + \sum_{j=1}^{4} \frac{1}{j}E_j\right).$$

With the help of computer, the choice $x = 3/2$ minimizes

$$\|A^{-1}\|_\infty = \frac{15}{4n} - \frac{7(h+w)}{8n^2} - \frac{4}{9n^2} + \frac{31(h+w) - 21}{72n^3} < \frac{15}{4n}.$$

Perturbation

Perturbed linear systems

Lemma

Let A be an $N \times N$ invertible matrix over the reals. Suppose $A - \Delta A$ is a perturbation. Then

- *$A - \Delta A$ is invertible provided $\|A^{-1}\Delta A\|_\infty < 1$; and*
- *the solution $\vec{\mathbf{x}}$ to $(A - \Delta A)\vec{\mathbf{x}} = A\vec{\mathbf{1}}$ is entrywise nonnegative provided $\|A^{-1}\Delta A\|_\infty \leq \frac{1}{2}$.*

Perturbed linear systems

Lemma

Let A be an $N \times N$ invertible matrix over the reals. Suppose $A - \Delta A$ is a perturbation. Then

- *$A - \Delta A$ is invertible provided $\|A^{-1}\Delta A\|_\infty < 1$; and*
- *the solution $\vec{\mathbf{x}}$ to $(A - \Delta A)\vec{\mathbf{x}} = A\vec{\mathbf{1}}$ is entrywise nonnegative provided $\|A^{-1}\Delta A\|_\infty \leq \frac{1}{2}$.*

Proof idea.

Use the series expansion $(A - \Delta A)^{-1} = \sum_{k=0}^{\infty}(A^{-1}\Delta A)^k A^{-1}$. □

A perturbation of M

Let S be an ϵ-dense partial Sudoku of type (h, w), where $hw = n$.

Define M_S similarly to M, so that $M_S(e, f)$ records the number of available tiles $\{r_i, c_j, s_k, b_\ell\}$ containing $e \cup f$.

A perturbation of M

Let S be an ϵ-dense partial Sudoku of type (h, w), where $hw = n$.

Define M_S similarly to M, so that $M_S(e, f)$ records the number of available tiles $\{r_i, c_j, s_k, b_\ell\}$ containing $e \cup f$.

Here is a white lie*, but morally true:

Proposition

$\|M - M_S\|_\infty < 12\epsilon n.$

*: we need to use a border to make M_S and M have the same dimensions

Main result

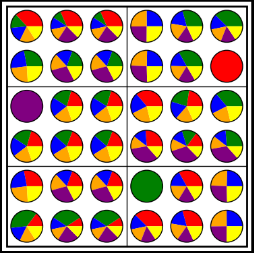
Theorem

Let $\epsilon < 1/101$. For sufficiently large h and w, every ϵ-dense partial Sudoku of type (h, w) has a **fractional** *completion, that is, an assignment of positive rational frequencies to symbols in unfilled cells so that the Sudoku conditions hold.*

Main result

Theorem

Let $\epsilon < 1/101$. For sufficiently large h and w, every ϵ-dense partial Sudoku of type (h, w) has a **fractional** *completion, that is, an assignment of positive rational frequencies to symbols in unfilled cells so that the Sudoku conditions hold.*



Proof sketch

We show $M_S\vec{\mathbf{x}} = \vec{\mathbf{1}}$ has an entrywise nonnegative solution $\vec{\mathbf{x}}$.

We shift the coefficient matrix by a multiple of K and view it as a perturbation of A (the 'empty' Sudoku).

Letting ΔA be this perturbation, we succeed when

$$\|A^{-1}\Delta A\|_\infty < \frac{15}{4n} \times 12\epsilon n + \frac{11}{2}\epsilon + o(1) = \frac{101}{2}\epsilon + o(1) < \frac{1}{2}.$$

Wrap-up

Concluding remarks

- Some details on bordering our matrices have been suppressed for clarity.

Concluding remarks

- Some details on bordering our matrices have been suppressed for clarity.
- If the boxes are asymptotically thin (say w fixed and $n = hw$ large), then the bundle condition can be dropped.

Concluding remarks

- Some details on bordering our matrices have been suppressed for clarity.
- If the boxes are asymptotically thin (say w fixed and $n = hw$ large), then the bundle condition can be dropped.
- A structure of wiggly but 'near-rectangular' boxes can be handled via a secondary perturbation.

Concluding remarks

- Some details on bordering our matrices have been suppressed for clarity.
- If the boxes are asymptotically thin (say w fixed and $n = hw$ large), then the bundle condition can be dropped.
- A structure of wiggly but ‘near-rectangular’ boxes can be handled via a secondary perturbation.
- Can this sparsity threshold for fractional completion be converted into something for actual completion?

Thank you!

References

- P.J. Dukes and K.I. Nimegeers, The linear system for Sudoku and a fractional completion threshold, `https://arxiv.org/abs/2310.15279`.
- R.A. Horn and C.R. Johnson, Matrix analysis, 2nd ed. Cambridge University Press, Cambridge, 2013.
- SageMath, the Sage Mathematics Software System (Version 9.3), The Sage Developers, 2021, `http://www.sagemath.org`.
- T. Sander, Sudoku graphs are integral. *Electron. J. Comb.*, 16 (2009), research paper N25, 7 pp.