HAL open science

Three numerical approaches to find mutually unbiased bases using Bell inequalities

Maria Prat Colomer, Luke Mortimer, Irénée Frérot, Máté Farkas, Antonio Acín

▶ **To cite this version:**

Maria Prat Colomer, Luke Mortimer, Irénée Frérot, Máté Farkas, Antonio Acín. Three numerical approaches to find mutually unbiased bases using Bell inequalities. Quantum, 2022, 6, pp.778. 10.22331/q-2022-08-17-778 . hal-03613073v2

HAL Id: hal-03613073

https://hal.science/hal-03613073v2

Submitted on 18 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Three numerical approaches to find mutually unbiased bases using Bell inequalities

Maria Prat Colomer[1,2,3], Luke Mortimer[1,2], Irénée Frérot[2,4], Máté Farkas[2], and Antonio Acín[2,5]

[1]These authors contributed equally to this work.

[2]ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, 08860 Castelldefels, Spain

[3]CFIS-Centre de Formació Interdisciplinària Superior, UPC-Universitat Politècnica de Catalunya, 08028 Barcelona, Spain

[4]Univ Grenoble Alpes, CNRS, Grenoble INP, Institut Néel, 38000 Grenoble, France

[5]ICREA-Institucio Catalana de Recerca i Estudis Avançats, Lluis Companys 23, 08010 Barcelona, Spain

arXiv:2203.09429v3 [quant-ph] 9 Aug 2022

Mutually unbiased bases correspond to highly useful pairs of measurements in quantum information theory. In the smallest composite dimension, six, it is known that between three and seven mutually unbiased bases exist, with a decades-old conjecture, known as Zauner's conjecture, stating that there exist at most three. Here we tackle Zauner's conjecture numerically through the construction of Bell inequalities for every pair of integers $n, d \geq 2$ that can be maximally violated in dimension d if and only if n MUBs exist in that dimension. Hence we turn Zauner's conjecture into an optimisation problem, which we address by means of three numerical methods: see-saw optimisation, non-linear semidefinite programming and Monte Carlo techniques. All three methods correctly identify the known cases in low dimensions and all suggest that there do not exist four mutually unbiased bases in dimension six, with all finding the same bases that numerically optimise the corresponding Bell inequality. Moreover, these numerical optimisers appear to coincide with the "four most distant bases" in dimension six, found through numerically optimising a distance measure in [P. Raynal, X. Lü, B.-G. Englert, *Phys. Rev. A*, 83 062303 (2011)]. Finally, the Monte Carlo results suggest that at most three MUBs exist in dimension ten.

Máté Farkas: mate.farkas@icfo.eu

1 Introduction

Mutually unbiased bases (MUBs), on the one hand, are highly symmetric bases in complex Hilbert spaces and, on the other hand, correspond to pairs of quantum measurements. The defining property of a pair of MUBs is that the overlaps between any two vectors from the two different bases is uniform. This property translates to the corresponding measurements as follows: If a measurement yields a definite outcome when measured on a quantum state, then a measurement unbiased to it will yield a uniformly random outcome on the same state. This feature makes MUBs widely useful in quantum information processing. MUBs were originally introduced in the context of optimal state determination [1], but since have been found to be useful in a variety of quantum information processing tasks, such as quantum cryptography [2, 3, 4], quantum communication tasks [5, 6], Bell inequalities [7, 8, 9] and so on (for a review, see Ref. [10]).

While MUBs have been extensively studied both in the quantum information and the mathematics community for decades, there are still open questions regarding their structure. Most notably, the maximal number of bases that are pairwise mutually unbiased is unknown for general Hilbert space dimension. A general upper bound was shown by Wootters and Fields, stating that in dimension d there exist no more than $d + 1$ MUBs [11]. In the same work, they showed that this upper bound is saturated in prime power dimensions, by providing an explicit construction. However, in composite dimensions the only known generic lower bound on the number of MUBs is $p^r + 1$, where p^r is the smallest prime power in the prime decomposition of the di-

mension (this lower bound is shown using tensor products of the Wootters–Fields construction). While in certain dimensions this lower bound has been improved [12], there exists no composite dimension in which the exact number of MUBs is known. For the smallest composite dimension, six, the number of MUBs is known to be no more than seven and no less than three, from the general bounds. However, which of the numbers in between is the exact number of MUBs in dimension six is unknown (apart from the fact that it cannot be six, following from a general result by Weiner [13]). It was first conjectured by Zauner in 1999 that there are no more than three MUBs in dimension six [14], and this conjecture has not been resolved to date, despite substantial efforts.

There are numerous works trying to prove (or disprove) Zauner's conjecture, both analytically and numerically. While not providing an exhaustive list of references here, let us note that on the analytic side, it has been shown that Zauner's conjecture is equivalent to a conjecture on orthogonal decompositions of Lie algebras [15]. Furthermore, there exist various analytic constructions of MUB triplet families (see Refs. [16, 17] and references therein), but thus far there has not been found even a single vector that is unbiased to all the vectors in any of these triplets. For notable recent developments on Zauner's conjecture see Refs. [18, 19].

On the numerical side, Bengtsson et al. introduced a distance measure of two bases that is maximised if and only if the bases are mutually unbiased [20]. This construction turns the problem of finding a set of MUBs into an optimisation problem, maximising all the pairwise distances within a set of bases. Using this approach, Raynal, Lü and Englert later constructed a two-parameter family of four bases in dimension six, such that for certain values of the parameters, the bases coincide with the numerical maximiser of the distance function [21]. These four bases are not MUBs, but based on the numerical evidence the authors refer to them as "the four most distant bases in dimension six".

Since MUBs optimise various quantum information processing tasks, it is natural to measure the closeness of a set of bases to MUBs in terms of some quantum information processing protocol. This was studied formally by Aguilar et al. [22], using the fact that MUBs optimise the success probability of a communication task called quantum random access codes (QRACs). A slight generalisation of the QRAC task is then optimised by a set of n MUBs, and finding n MUBs in dimension d corresponds to optimising the associated success probability. With this method, Aguilar et al. managed to re-prove the non-existence of $d+2$ MUBs in certain low dimensions using quantum information theoretic tools. However, the case of dimension six remains open.

In this work, we employ similar ideas to tackle Zauner's conjecture. Namely, we study a recently introduced family of Bell inequalities, known to be maximally violated by a pair of MUBs in dimension d [9]. We then extend these inequalities to new ones, maximally violated by a set of n MUBs in dimension d. Then, we apply three numerical methods for finding the maximal value of these Bell inequalities in a fixed dimension. Namely, we apply see-saw semidefinite programming (SDP), non-linear SDP, and Monte Carlo techniques. While these methods are heuristic—in the sense that there is no guarantee for finding a global maximum—they find the maximum in all the cases where the maximum is known (i.e., it is known that n MUBs exist in the given dimension d). Furthermore, when applying these techniques to dimension six and four bases, all the different numerical tools converge to the same bases, and these four bases are—numerically—very close to the "four most distant bases" of Ref. [21] (one should not expect exact equality, since the measure optimised in Ref. [21] is different from the Bell inequalities we optimise). Hence, our results support Zauner's conjecture, even though no rigorous claim can be made due to the heuristic nature of our methods. Finally, we were able to implement the Monte Carlo algorithm for $d = 10$, where—similarly to $d = 6$—we do not find more than three MUBs.

2 Preliminaries

In this section, we introduce the mathematical background and concepts necessary for turning the MUB problem into an optimisation problem. Namely, we formally introduce MUBs, Bell inequalities and the specific family of Bell inequalities tailored for MUBs.

2.1 Mutually unbiased bases

Let us take a d-dimensional Hilbert space $\mathcal{H} \cong \mathbb{C}^d$, and two orthonormal bases on it, $\{|b_j^1\rangle\}_{j=1}^d$ and $\{|b_k^2\rangle\}_{k=1}^d$. We say that these two bases are *mutually unbiased* if

$$|\langle b_j^1|b_k^2\rangle|^2 = \frac{1}{d} \quad \forall j,k \in [d], \tag{1}$$

where $[d] \equiv \{1,2,\ldots,d\}$. A simple example in dimension two is the computational basis $\{|0\rangle, |1\rangle\}$ and the Hadamard basis $\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$.

One may also associate an orthonormal basis with a quantum measurement. In general, a quantum measurement is described by a positive operator-valued measure (POVM), which, in the d-dimensional, d-outcome case corresponds to a set of d positive semidefinite operators $B_j \geq 0$ on $\mathbb{C}^d$, adding up to the identity operator $\mathbb{1}$. Given an orthonormal basis $\{|b_j\rangle\}_{j=1}^d$ on $\mathbb{C}^d$, one can define the corresponding POVM $\{B_j = |b_j\rangle\langle b_j|\}_{j=1}^d$, consisting of rank-1 projections onto the basis elements. We say that two measurements are MUBs if they correspond to a pair of orthonormal bases that are MUBs.

2.2 Bell inequalities

We look at MUBs in the context of *Bell scenarios* (see Ref. [23] for a review). Bell scenarios describe physical experiments performed by two distant parties, usually referred to as Alice and Bob. These parties share many copies of a bipartite (quantum) state, and perform local measurements on these copies. The experiment is described by the *correlation*, p, with elements $p(a,b|x,y)$ specifying the probability of Alice (Bob) observing locally the outcome a (b) upon choosing the measurement setting x (y). In quantum theory, the shared state is described by a density operator $\rho \geq 0$ with unit trace ($\operatorname{tr}\rho = 1$) on a tensor product Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$. The measurements are described by local POVMs $\{A_a^x\}$ and $\{B_b^y\}$ on the Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, respectively. For a fixed state and measurements, the correlation is given by the Born rule,

$$p(a,b|x,y) = \operatorname{tr}[\rho(A_a^x \otimes B_b^y)]. \tag{2}$$

Note that for a pure state, $\rho = |\psi\rangle\langle\psi|$, with $|\psi\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ and $\langle\psi|\psi\rangle = 1$, the Born rule reduces to

$$p(a,b|x,y) = \langle\psi|A_a^x \otimes B_b^y|\psi\rangle. \tag{3}$$

Bell functionals are linear functionals of correlations, i.e., functionals of the form

$$W(p) = \sum_{a,b,x,y} c_{abxy} p(a,b|x,y), \tag{4}$$

where c_{abxy} are real coefficients. Non-trivial *Bell inequalities* are Bell functionals for which $W(p) \leq \beta_L$ holds for every correlation of the form

$$p(a,b|x,y) = \int_\Lambda \mathrm{d}\mu(\lambda) p_A(a|x,\lambda) p_B(b|y,\lambda) \tag{5}$$

(also called a *local* correlation, considered as the notion of classicality in Bell scenarios), but for which there exists a quantum correlation p of the form (2) such that $W(p) > \beta_L$. In Eq. (5), Λ is a measurable set with a probability measure μ, $\lambda \in \Lambda$, and p_A and p_B are conditional probability distributions.

While the original interest in Bell inequalities was precisely this separation of local and quantum correlations, we will be interested in their quantum maximum (or *maximal quantum violation*), i.e., the tight upper bound $W(p) \leq \beta_Q$ satisfied by all correlations of the form (2).

2.3 Bell inequalities for mutually unbiased bases

In this work, we are interested in a family of Bell inequalities that was introduced in Ref. [9], and is parametrised by an integer $d \geq 2$. For a fixed d, Alice has d^2 measurement settings labelled as $x = x_1x_2$ with $x_1, x_2 \in [d]$. Each of these measurements has three outcomes, $a \in \{1,2,\perp\}$. Bob, on the other hand, has two measurement settings, $y \in \{1,2\}$, with d outcomes each, $b \in [d]$. The Bell inequality then reads

$$\begin{aligned} W_d(p) = &\sum_{x_1,x_2,y} [p(y,x_y|x_1x_2,y) - p(\bar{y},x_y|x_1x_2,y)] \\ &- \frac{1}{2}\sqrt{\frac{d-1}{d}} \sum_{x_1,x_2} [p_A(1|x_1x_2) + p_A(2|x_1x_2)], \end{aligned} \tag{6}$$

where $\bar{y}$ flips the value of $y \in \{1,2\}$, and $p_A(a|x) = \sum_b p(a,b|x,y)$ is the marginal probability distribution of Alice (which is independent of y).

This is a non-trivial Bell inequality with maximal quantum violation $\beta_Q = \sqrt{d(d-1)}$ [9].

The maximal violation can be achieved with the maximally entangled d-dimensional state $|\phi_d^+\rangle \equiv \frac{1}{\sqrt{d}}\sum_{j=1}^d |j\rangle\otimes|j\rangle$, and *any* pair of MUB measurements on Bob's side. Moreover, if the dimension is fixed to be d, this is the only way in which the maximal quantum violation can be achieved, up to local unitary freedom [9]. This property of the Bell inequality (6) forms the core of our numerical approaches to construct MUBs.

The above Bell inequality can be straightforwardly extended to a set of n measurements on Bob's side. The Bell inequality for n measurements is a sum of Bell inequalities of the form (6). For each pair $y, z \in [n]$ such that $y < z$ (denoted in the following as $(y, z) \in \text{Pairs}[n]$), we introduce d^2 settings for Alice, labelled as $x = (y, z)x_y x_z$ with $x_y, x_z \in [d]$, and take a copy of the Bell inequality in Eq. (6), defined as

$$\begin{aligned} W_d^{(y,z)}(p) = &\sum_{x_y,x_z,w} [p(a_w, x_w|(y,z)x_y x_z, w) \\ &-p(\bar{a}_w, x_w|(y,z)x_y x_z, w)] \\ &-\frac{1}{2}\sqrt{\frac{d-1}{d}} \sum_{x_y,x_z} [p_A(1|(y,z)x_y x_z) \\ &+p_A(2|(y,z)x_y x_z)], \end{aligned} \tag{7}$$

where $w \in \{y, z\}$, $a_y = 1$, $a_z = 2$, and $\bar{a}_w$ flips the value of a_w. The final Bell inequality then reads

$$W_{d,n}(p) = \sum_{(y,z)\in\text{Pairs}[n]} W_d^{(y,z)}(p). \tag{8}$$

It is clear that $W_{d,n}(p) \leq \binom{n}{2}\sqrt{d(d-1)}$, by applying the known bound to each individual term in the above sum. Moreover, if the dimension is d, the only way to reach this maximum (up to local unitary freedom) is by using the maximally entangled state, and if the n measurements on Bob's side correspond to MUBs. Hence, we can reformulate the MUB problem in terms of these Bell inequalities:

Proposition 1. *$W_{d,n}(p) = \binom{n}{2}\sqrt{d(d-1)}$ can be achieved in dimension d if and only if n MUBs exist in dimension d.*

2.4 The optimisation problem

According to the above proposition, finding n MUBs in dimension d can be cast as an optimisation problem, maximising $W_{d,n}(p)$ over d-dimensional quantum states and measurements. To see how to do this explicitly, let us first write out the Bell inequality W_d in Eq. (6) in terms of a quantum state $|\psi\rangle$ and measurements $\{A_a^x\}$, $\{B_b^y\}$, using Born's rule in Eq. (3):

$$\begin{aligned} W_d(|\psi\rangle, \{A_a^x\}, \{B_b^y\}) = &\sum_{x_1,x_2,y} (\langle\psi|A_y^{x_1x_2}\otimes B_{x_y}^y|\psi\rangle \\ &-\langle\psi|A_{\bar{y}}^{x_1x_2}\otimes B_{x_y}^y|\psi\rangle) \\ &-\frac{1}{2}\sqrt{\frac{d-1}{d}}\sum_{x_1,x_2}(\langle\psi|A_1^{x_1x_2}\otimes\mathbb{1}|\psi\rangle \\ &+\langle\psi|A_2^{x_1x_2}\otimes\mathbb{1}|\psi\rangle). \\ &=\sum_{j,k}\langle\psi|[(A_1^{jk}-A_2^{jk})\otimes(B_j^1-B_k^2) \\ &-\frac{1}{2}\sqrt{\frac{d-1}{d}}(A_1^{jk}+A_2^{jk})\otimes\mathbb{1}]|\psi\rangle, \end{aligned} \tag{9}$$

where in the second equality we wrote out the summation over y and switched to the notation $x_1x_2 \to jk$ with $j, k \in [d]$. Maximising the inequality (9) in terms of the state and the measurements can then be written as the optimisation problem

$$\begin{aligned} \max_{|\psi\rangle,\{A_a^x\},\{B_b^y\}} \quad & W_d(|\psi\rangle, \{A_a^x\}, \{B_b^y\}) \\ \text{s.t.} \quad & |\psi\rangle \in \mathbb{C}^d\otimes\mathbb{C}^d, \quad \langle\psi|\psi\rangle = 1 \\ & A_a^x, B_b^y \in \mathcal{L}_{\text{sa}}(\mathbb{C}^d) \quad \forall a,b,x,y \\ & A_a^x \geq 0 \quad \forall a,x, \quad B_b^y \geq 0 \quad \forall b,y \\ & \sum_a A_a^x = \mathbb{1} \quad \forall x, \quad \sum_b B_b^y = \mathbb{1} \quad \forall y, \end{aligned} \tag{10}$$

where $\mathcal{L}_{\text{sa}}(\mathbb{C}^d)$ is the set of self-adjoint linear operators on $\mathbb{C}^d$. Optimising $W_{d,n}$ can be written in a similar fashion, with:

$$\begin{aligned} W_{d,n}(|\psi\rangle, \{A_a^x\}, \{B_b^y\}) = &\sum_{(y,z)\in\text{Pairs}[n]} \Big\{ \\ &\sum_{j,k}\langle\psi|[(A_1^{(y,z)jk}-A_2^{(y,z)jk})\otimes(B_j^y-B_k^z) \\ &-\frac{1}{2}\sqrt{\frac{d-1}{d}}(A_1^{(y,z)jk}+A_2^{(y,z)jk})\otimes\mathbb{1}]|\psi\rangle\Big\}. \end{aligned} \tag{11}$$

The optimisation problem is then

$$\begin{aligned} \max_{|\psi\rangle,\{A_a^x\},\{B_b^y\}} \quad & W_{d,n}(|\psi\rangle,\{A_a^x\},\{B_b^y\}) \\ \text{s.t.} \quad & |\psi\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d, \quad \langle\psi|\psi\rangle = 1 \\ & A_a^x, B_b^y \in \mathcal{L}_{\text{sa}}(\mathbb{C}^d) \ \ \forall a,b,x,y \\ & A_a^x \geq 0 \ \ \forall a,x, \quad B_b^y \geq 0 \ \ \forall b,y \\ & \sum_a A_a^x = \mathbb{1} \ \ \forall x, \quad \sum_b B_b^y = \mathbb{1} \ \ \forall y. \end{aligned} \tag{12}$$

In this work, we consider various approaches to solve the optimisation problem (12). In particular, from Proposition 1 it follows that n MUBs exist in dimension d if and only if the solution of the above optimisation problem is $\binom{n}{2}\sqrt{d(d-1)}$. We will facilitate the problem using knowledge about the optimal realisation of the Bell inequality from Ref. [9].

First of all, we notice that the value $\binom{n}{2}\sqrt{d(d-1)}$ can only be achieved in dimension d with the maximally entangled state [9]. Without loss of generality, we therefore impose that $|\psi\rangle = |\phi_d^+\rangle = \frac{1}{\sqrt{d}}\sum_{j=1}^d |j\rangle \otimes |j\rangle$. We can then use the fact that for any two operators A and B on $\mathbb{C}^d$ we have that $\langle\phi_d^+|A \otimes B|\phi_d^+\rangle = \frac{1}{d}\operatorname{tr}(A^T B)$, where $(.)^T$ is the transposition in the basis $\{|j\rangle\}$. As a second simplification, we notice that in order to saturate the bound $\binom{n}{2}\sqrt{d(d-1)}$ in dimension d, Alice's measurement operators A_1^x and A_2^x, and all of Bob's measurement operators must be trace-1 [9]. For such operators we have that $\langle\phi_d^+|A_1^x \otimes \mathbb{1}|\phi_d^+\rangle = \langle\phi_d^+|A_2^x \otimes \mathbb{1}|\phi_d^+\rangle = \frac{1}{d}$. The second term in Eq. (11) is then a constant, $-\binom{n}{2}\sqrt{d(d-1)}$, and does not influence the optimisation problem. The simplified Bell expression finally reads

$$W_{d,n}^+(\{A_a^x\},\{B_b^y\}) = \sum_{(y,z)\in\text{Pairs}[n]} \left\{ \frac{1}{d}\sum_{j,k} \operatorname{tr}\left((A_1^{(y,z)jk} - A_2^{(y,z)jk})^T (B_j^y - B_k^z)\right) \right\}, \tag{13}$$

and its maximum quantum value $W_{d,n}^+$ satisfies:

$$W_{d,n}^+ \leq W_{\text{MUB}}(d,n) \equiv n(n-1)\sqrt{d(d-1)}\,. \tag{14}$$

The simplified optimisation problem becomes

$$\begin{aligned} \max_{\{A_a^x\},\{B_b^y\}} \quad & W_{d,n}^+(\{A_a^x\},\{B_b^y\}) \\ \text{s.t.} \quad & A_a^x, B_b^y \in \mathcal{L}_{\text{sa}}(\mathbb{C}^d) \ \ \forall a,b,x,y \\ & A_a^x \geq 0 \ \ \forall a,x, \quad B_b^y \geq 0 \ \ \forall b,y \\ & \operatorname{tr} A_a^x = 1 \ \ \forall a,x, \quad \operatorname{tr} B_b^y = 1 \ \ \forall y,b \\ & \sum_a A_a^x = \mathbb{1} \ \ \forall x, \quad \sum_b B_b^y = \mathbb{1} \ \ \forall y. \end{aligned} \tag{15}$$

The optimal value of this optimisation problem is $n(n-1)\sqrt{d(d-1)}$ if and only if n MUBs exist in dimension d.

We may further simplify the optimisation problem. From Ref. [9] we know that in the optimal realisation the B_b^y operators are rank-1 projections, $B_j^y = |b_j^y\rangle\langle b_j^y|$ (they are projections onto the basis elements of MUBs). For such operators, we have that

$$(B_j^y - B_k^z)^3 = [1 - \operatorname{tr}(B_j^y B_k^z)](B_j^y - B_k^z), \tag{16}$$

which implies that the spectrum of $B_j^y - B_k^z$ is contained in $\{0, \pm\lambda_{jk}^{yz}\}$, where $\lambda_{jk}^{yz} \equiv \sqrt{1 - \operatorname{tr}(B_j^y B_k^z)} = \sqrt{1 - |\langle b_j^y|b_k^z\rangle|^2}$. Moreover, we have that $\operatorname{tr}[(B_j^y - B_k^z)^2] = 2(\lambda_{jk}^{yz})^2$ and $\operatorname{tr}(B_j^y - B_k^z) = 0$, and therefore $B_j^y - B_k^z$ has one eigenvalue λ_{jk}^{yz}, one eigenvalue $-\lambda_{jk}^{yz}$ and the rest of the eigenvalues are 0. Furthermore, in the optimal realisation we have that $(A_1^{(y,z)jk})^T$ is the rank-1 projection onto the eigenspace of $B_j^y - B_k^z$ corresponding to λ_{jk}^{yz}, and $(A_2^{(y,z)jk})^T$ is the rank-1 projection onto the eigenspace of $B_j^y - B_k^z$ corresponding to $-\lambda_{jk}^{yz}$. With these final simplifications the Bell expression reads

$$\begin{aligned} W_{d,n}^{+B}(\{B_b^y\}) &= \frac{2}{d}\sum_{(y,z)\in\text{Pairs}[n]}\sum_{j,k}\sqrt{1 - \operatorname{tr}(B_j^y B_k^z)} \qquad (17) \\ &= \frac{2}{d}\sum_{(y,z)\in\text{Pairs}[n]}\sum_{j,k}\sqrt{1 - |\langle b_j^y|b_k^z\rangle|^2} \qquad (18) \end{aligned}$$

and the corresponding optimisation problem is

$$\begin{aligned} \max_{\{B_b^y\}} \quad & W_{d,n}^{+B}(\{B_b^y\}) \\ \text{s.t.} \quad & B_b^y \in \mathcal{L}_{\text{sa}}(\mathbb{C}^d) \ \ \forall a,b,x,y \\ & (B_b^y)^2 = B_b^y \ \ \forall b,y \\ & \operatorname{tr} B_b^y = 1 \ \ \forall b,y \\ & \sum_b B_b^y = \mathbb{1} \ \ \forall y. \end{aligned} \tag{19}$$

Note that projectivity and self-adjointness implies positive semidefiniteness, and therefore positive semidefiniteness does not need to be imposed.

The optimal value of the optimisation problem (19), denoted as $W_{\max}(d,n)$, is $n(n-1)\sqrt{d(d-1)}$ if and only if n MUBs exist in dimension d. In fact, Alice's measurements have been completely removed from the problem, and one may regard Eq. (19) as a purely geometrical problem: find n orthonormal bases, $\{\{|b_j^y\rangle \mid j \in [d]\} \mid y \in [n]\}$, maximising the "MUB-ness measure" of Eq. (18).

In the following three sections we apply three numerical methods to solve the optimisation problems (15) and (19) in order to numerically tackle Zauner's conjecture.

3 See-saw SDP

3.1 Methodology

One arrives at a relatively simple method of optimising problem (15) by noticing that the objective function is bi-linear in the A_a^x and B_b^y matrices. That is, if every A_a^x is fixed, then the problem simplifies to optimising a linear functional of the B_b^y matrices with a series of linear and positive semidefinite constraints. This is a standard SDP, for which there exist efficient solvers.

The *see-saw* optimisation technique starts with fixing the set of A_a^x matrices satisfying the constraints of the problem (15), either with random values or based on some prior knowledge. We then solve the problem for the B_b^y matrices, which is a standard SDP. Then, we fix the B_b^y matrices to the optimum found, and solve the resulting SDP for the A_a^x matrices, and so on. By repeating this process, the system eventually converges to a stable result, i.e. the value of the objective function does not change beyond a given threshold within a certain window of iterations (a change of less than 10^{-9} for 10 iterations in our implementation). Although the see-saw method has never been proven to converge to the global optimum, for our current problem it has never failed to converge within the chosen precision if given sufficient time, always to the value expected (i.e., to $W_{\mathrm{MUB}}(d,n)$ whenever it is known that n MUBs exist in dimension d). We implemented the see-saw algorithm with the help of the SDP solving library MOSEK [24].

3.2 Results

The values obtained with the see-saw algorithm are shown in Table 1 (the values displayed are $1-W_{d,n}/W_{\mathrm{MUB}}(d,n)$ for easier comparison across different n and d, where $W_{d,n}$ is the result of the optimisation). Notice that whenever n MUBs exist in dimension d, the see-saw method correctly converges to the MUB solution, and whenever it is known that n MUBs do not exist in dimension d, the method does indeed converge to a value less than $W_{\mathrm{MUB}}(d,n)$. For the unknown case of four MUBs in dimension six, the see-saw method could not find four MUBs, supporting Zauner's conjecture. Furthermore, the optimal measurements found by the see-saw method are numerically very close to the "four most distant bases" of Ref. [21] (see also Section 6.3). Note, however, that the results simply mean that the method could not find four MUBs in dimension six, but one cannot rule out the possibility that they exist.

The method also has consistent convergence, for example, for $d = 2$ and $n = 2$, all of 10000 see-saw optimisations from random starts converged to the correct value of $W_{\mathrm{MUB}}(2,2) = 2.82843$ up to five decimal places, albeit finding a different set of optimum matrices. Similarly, for 10000 optimisations for $d = 2$ and $n = 4$, all optimisations converged to the same value of 16.72616 (correctly signifying non-existence). We did not perform a similar sized convergence analysis for the larger dimensions due to the time required to optimise, however, several runs were always performed to certify a minimal level of consistency.

All of the results for this method were obtained on a desktop PC with 8GB of RAM using 4 cores, with times varying between milliseconds for the smallest problem ($d = 2$, $n = 2$) and hours for the largest one ($d = 6$, $n = 4$). Since SDP solvers are efficiently parallelised, this method offers good parallel scaling, however, the memory requirement is the highest of all of our methods, since it requires explicit storage and optimisation of the A_a^x matrices.

4 Non-linear SDP

4.1 Methodology

An alternative approach to the optimisation is to focus on problem (19), which features only the B_b^y matrices and thus contains fewer variables for

n \ d	2	3	4	5	6
2	0.00000	0.00000	0.00000	0.00000	0.00000
3	0.00000	0.00000	0.00000	0.00000	0.00000
4	**0.01440**	0.00000	0.00000	0.00000	**0.00004**
5	-	**0.00391**	0.00000	0.00000	-
6	-	-	**0.00186**	0.00000	-
7	-	-	-	**0.00091**	-

Table 1: The values $1 - W_{d,n}/W_{\mathsf{MUB}}(d,n)$, where $W_{d,n}$ is the result of the optimisation, obtained with the see-saw method after convergence for various dimensions (d) and numbers of bases (n), to 5 decimal places. The values depicted are consistently obtained by multiple runs. The values in bold indicate that the value was at least 10^{-5} away from zero, meaning that the method could not find n MUBs in dimension d. The "-" symbol indicates that we have not performed the optimisation. Note that in all the known cases, the algorithm predicts correctly existence/nonexistence, and it predicts that four MUBs do not exist in dimension six.

a reduction in the size of the search space as well as the memory required. The downside, however, is that the objective function is now non-linear (not even bi-linear) and thus many efficient solvers (i.e. for standard SDP systems) can no longer be applied. To optimise this problem we adapt a method based on the work by Yamashita et al. for optimising a non-linear SDP using a primal-dual interior point method [25]. This method, assuming a few basic conditions (discussed later), is guaranteed to converge to a Karush–Kuhn–Tucker (KKT) point, a point satisfying a series of constraints known as the KKT conditions, which are necessary for optimality [26].

These KKT conditions are only sufficient (imply a global minimum) in a subset of cases, the main of which being that the problem is convex, which is unfortunately not true in our case. An interesting property of our search-space, however, is that by taking the derivative of our objective function it can be shown that all local minima are global minima if the constraints are met, thus implying that our search space is, in fact, a series of discontinuous convex regions. This signifies that for the MUB-existence case this method will always converge to MUBs, although no claim can be made for the non-existence case.

In order to implement the method of Yamashita et al., we parametrise the measurement operators $\{B_y^x\}$ by a real vector $\mathbf{x} = (x_i)_i$. To be able to deal with real numbers instead of complex ones, we note that every self-adjoint matrix $B = B_r + \mathrm{i}B_i$ (where B_r is real symmetric and B_i is real anti-symmetric) can be mapped to the real symmetric matrix $\hat{B}$ via

$$B \mapsto \hat{B} = \begin{bmatrix} B_r & B_i \\ -B_i & B_r \end{bmatrix}. \tag{20}$$

It is easy to verify that $B \geq 0$ if and only if $\hat{B} \geq 0$, and $\operatorname{tr} B = \frac{1}{2} \operatorname{tr} \hat{B}$.

We therefore define—in line with the method of Yamasitha et al.—a matrix $X(\mathbf{x}) = \sum_i C_i x_i + D$, which is a block diagonal matrix containing the $\hat{B}_b^y$ matrices on its diagonal in such a way that the linear constraints $\sum_b B_b^y = \mathbb{1}$ and $\operatorname{tr} B_b^y = 1$ of the optimisation problem (19) are already enforced. The real parameters x_i correspond to those elements of the $\hat{B}_b^y$ matrices that are free after enforcing the linear constraints. While the constraint $B_b^y \geq 0$ is superfluous for the problem (19), we chose to include this in our optimisation problem, as this constraint is heavily used in the method of Yamashita et al. With the parametrisation above, this constraint is equivalent to $X(\mathbf{x}) \geq 0$. The last remaining constraint is projectivity, $(B_b^y)^2 = B_b^y$ for all y and b, which is equivalent to $X^2(\mathbf{x}) = X(\mathbf{x})$. We enforce this constraint through $g(\mathbf{x}) \equiv \|X^2(\mathbf{x}) - X(\mathbf{x})\|_F^2 = 0$, where $\|.\|_F$ is the Frobenius norm. Further, we denote the objective function in terms of $\mathbf{x}$ by $W(\mathbf{x})$, suppressing the d, n indices whenever it does not lead to confusion.

The method requires introducing Lagrange multipliers (dual variables) for every constraint. In our case, there is a single inequality constraint $X(\mathbf{x}) \geq 0$, to which we assign the dual variable Z, which is a matrix with the same dimensions as $X(\mathbf{x})$. Furthermore, we have a single equality constraint, $g(\mathbf{x}) = 0$, to which we assign the

dual variable y, which is a scalar. The resulting Lagrangian reads

$$L(\mathbf{x}, y, Z) = W(\mathbf{x}) - yg(\mathbf{x}) - \mathrm{tr}[Z^T X(\mathbf{x})]. \quad (21)$$

The algorithm for solving the optimisation problem is iterative, and each iteration begins with the calculation of G, the Hessian of the Lagrangian. In our case this can be quite expensive so we opt to use the alternative update method also proposed in Ref. [25] based on the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, which approximates the Hessian without requiring the full calculation of the second derivatives. This G is then used to form a series of linear equations which we solve using a stabilised bi-conjugate gradient method in order to obtain the update directions for the primal ($\mathbf{x}$) and dual (y, Z) variables. Following this, a simple line search is performed to find the optimum step size, then the variables are updated. This process is then repeated until the following barrier KKT conditions are met for some barrier parameter μ:

$$r(\mathbf{x}, y, Z, \mu) \equiv \begin{pmatrix} \nabla L(\mathbf{x}, y, Z) \\ g(\mathbf{x}) \\ X(\mathbf{x})Z - \mu I \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad (22)$$

Through repeated iterations of this method with values of μ converging towards zero, this algorithm has been proven to always converge to a KKT point of the system, assuming certain conditions. The first of these is that the functions $W(\mathbf{x})$ and $g(\mathbf{x})$ are both twice continuously differentiable, which in our case is true: $W(\mathbf{x})$ is simply a sum of square roots of polynomials of $\mathbf{x}$, whilst $g(\mathbf{x})$ is a vector of polynomials of $\mathbf{x}$. The second condition is that the vector $\mathbf{x}$ must remain within a finite set during the optimisation, which for us is true since infinite values are non-optimal for the MUB functional. The third condition is that the matrices C_i must be linearly independent, which for us is true since they serve to place a single component of $\mathbf{x}$ to one (diagonal) or two (off-diagonal) positions in the matrix $X(\mathbf{x})$.

Similarly to the the see-saw method, there exists some non-determinism to this method due to it starting from a randomised interior point. This is found by first beginning with a random vector, then performing gradient descent using the derivatives of the constraints until the vector satisfies all constraints within some precision. This vector is then used as the starting point for the iterative method. Although perhaps one could start with a "good guess" for the vector and optimise from there, we decided to begin from a random point to attempt to cover a wider search space.

4.2 Results

The values obtained through this algorithm are shown in Table 2. The optimal values, as well as the optimal bases found, agree with the result of the see-saw method up to high numerical precision. In particular, we find the same set of four bases to numerically optimise the six dimensional case (see Section 6.3 for further details).

Regarding performance, a large amount of time can be saved if certain parameters, such as the initial step size, are chosen correctly. Notably, there appears to be no universally optimal set of parameters, and many of the results are obtained through the manual tweaking of the parameters for the specific system. All of the results for this method were obtained on a standard desktop PC using 4 cores, with the times varying between milliseconds for the smallest problem and an hour for the largest one. This method has a significantly reduced memory cost compared to the see-saw method, since here the A_a^x matrices are not used, as well as in general offering faster optimisation.

5 The MUB problem as a ground state problem

5.1 Methodology

Finally, we apply an optimisation method inspired by statistical physics, namely simulated annealing [27]. Given n orthonormal bases $\{|b_j^y\rangle; j \in [d]\}(y \in [n])$ in a Hilbert space of dimension d, our goal is to maximise the expression of Eq. (18), which we rewrite here for completeness (we keep the d, n dependence implicit throughout this section, and keep the $+B$ superscript implicit in Eq. (18)):

$$W[\mathbf{x}] = \frac{2}{d} \sum_{(y,z)\in \mathrm{Pairs}[n]} \sum_{j,k=1}^{d} \sqrt{1 - |\langle b_j^y | b_k^z \rangle|^2}, \quad (23)$$

n \ d	2	3	4	5	6
2	0.00000	0.00000	0.00000	0.00000	0.00000
3	0.00000	0.00000	0.00000	0.00000	0.00000
4	**0.01440**	0.00000	0.00000	0.00000	**0.00004**
5	-	**0.00391**	0.00000	0.00000	-
6	-	-	**0.00161**	0.00000	-
7	-	-	-	**0.00091**	-

Table 2: The values $1 - W_{d,n}/W_{\mathsf{MUB}}(d,n)$, where $W_{d,n}$ is the result of the optimisation, obtained at an approximate KKT point with $\|r(\mathbf{x}, y, Z, \mu)\| \equiv \sqrt{\|\nabla L(\mathbf{x}, y, Z)\|^2 + |g(\mathbf{x})|^2 + \|X(\mathbf{x})Z - \mu I\|_F^2} \leq 10^{-5}$ for various dimensions (d) and numbers of bases (n), to 5 decimal places. The values depicted are consistently obtained by multiple runs. The values in bold indicate that the value was at least 10^{-5} away from zero, meaning that the method could not find n MUBs in dimension d. Note that in all the known cases, the algorithm predicts correctly existence/nonexistence, and it predicts that four MUBs do not exist in dimension six.

n \ d	2	3	4	5	6	7	8	9	10
2	0.	0.	0.	0.	0.	0.	0.	0.	0.
3	0.	0.	0.	0.	0.	0.	0.	0.	0.
4	**0.01440**	0.	0.	0.	**0.00004215**	0.	0.	0.	**0.000005938**
5	-	**0.003912**	0.	0.	-	0.	0.	0.	-
6	-	-	**0.001609**	0.	-	0.	0.	-	-
7	-	-	-	**0.0009073**	-	0.	0.	-	-
8	-	-	-	-	-	0.	0.	-	-

Table 3: Monte Carlo results. Relative deviation from the MUB optimum in Eq. (23) $1 - W_{\max}/W_{\mathrm{MUB}}$. "0." indicates that n MUBs have been found in dimension d up to at least 10^{-10} precision. The positive values in bold are the numerical maxima found over at least three independent Monte Carlo simulations. For $d = 10$, $n = 4$, three out of ten independent simulations found the indicated value; it remains possible that this does not correspond yet to the true optimum.

where $\mathbf{x} = \{|b_j^y\rangle\}$ defines the collection of n bases. When the n bases are mutually unbiased ($|\langle b_j^y|b_k^z\rangle|^2 = 1/d$ for all $y, z \in \mathrm{Pairs}[n]$ and all $j, k \in [d]$), we have $W = W_{\mathrm{MUB}} = n(n-1)\sqrt{d(d-1)}$. In general, the maximal value is $W_{\max} \leq W_{\mathrm{MUB}}$, with equality if and only if n MUBs exist in dimension d. Our strategy is then to optimise the bases by maximising W (or equivalently, minimising $-W$) via simulated annealing. As detailed below, this amounts to parametrising the vectors $|b_j^y\rangle$ by some parameters $\mathbf{x}$, and regarding $-W(\mathbf{x})$ as the energy of the configuration $\mathbf{x}$. The ground state (namely, the lowest-energy configuration) of $-W$ is found by sampling $\mathbf{x}$ with probability $\propto \exp[\beta W(\mathbf{x})]$, progressively ramping up the inverse temperature β. In contrast to gradient-descent-based approaches, this allows to explore a variety of local minima of $-W$ at nonzero temperature, with the hope of finally converging to the global minimum when the temperature approaches zero. As our results show, this hope is indeed confirmed by solid evidence. If n MUBs do exist, they are obtained at the end of the optimisation, saturating the value $W_{\max} = W_{\mathrm{MUB}}$. Otherwise, the algorithm converges to an optimum (presumably the global optimum) $W_{\max} < W_{\mathrm{MUB}}$, supporting that n MUBs do not exist in dimension d.

Parametrising the bases.– Each vector $|b_j^y\rangle$ is simply parametrised by its decomposition in the canonical basis: $|b_j^y\rangle = \sum_{i=1}^d U_{ji}^y |e_i\rangle$, where $\mathbf{x} := \{U^y\}_{y=1}^n$ are n complex $d \times d$ unitary matrices. The overlaps are then obtained as $\langle b_j^y|b_k^z\rangle = \sum_{i=1}^d (U_{ji}^y)^* U_{ki}^z$, where $(.)^*$ is the complex conjugation.

Simulated annealing.– The basic idea of simulated annealing [27] is to consider:

$$\langle W\rangle_\beta = \int d\mathbf{x}\; W(\mathbf{x})\; \frac{e^{\beta W(\mathbf{x})}}{Z_\beta}\;, \tag{24}$$

where $Z_\beta = \int d\mathbf{x}\; e^{\beta W(\mathbf{x})}$ is a normalisation factor. Eq. (24) corresponds to an effective thermal average, where the parameters $\mathbf{x}$ are sampled from a Gibbs distribution $e^{\beta W(\mathbf{x})}/Z_\beta$, in which $-W(\mathbf{x})$ plays the role of the energy, and $1/\beta$ is the temperature. We have that [27]:

$$\langle W\rangle_\beta \underset{\beta\to\infty}{\longrightarrow} W_{\max}\;. \tag{25}$$

In words, progressively ramping up the inverse temperature β, the thermal average Eq. (24) converges towards the global maximum of the function W. At each value of β, the Gibbs distribution in Eq. (24) is sampled via a Markov-chain Monte Carlo algorithm [28].

Monte Carlo sampling.– The Markov chain is a list of samples $\{\mathbf{x}_i\}_{i=1}^N$, generated in such a way that in the limit of infinitely many samples, the average value of $W(\{\mathbf{x}_i\})$ over the N samples converges towards the exact average value, up to $O(1/\sqrt{N})$ corrections:

$$\frac{1}{N}\sum_{i=1}^{N} W(\mathbf{x}_i) = \langle W\rangle_\beta + O(1/\sqrt{N})\;. \tag{26}$$

In order to sample $\mathbf{x}$ according to the Gibbs distribution [Eq. (24)], we implement a Metropolis algorithm [28]. That is, we start from an arbitrary initial configuration $\mathbf{x}_1$, and then iterate:

1. propose a new configuration $\mathbf{x}_{\text{new}}$ (see below)
2. compute the difference $\Delta = \beta[W(\mathbf{x}_{\text{new}}) - W(\mathbf{x}_i)]$
3. if $\Delta > 0$, accept the move
4. if $\Delta < 0$, accept the move with probability e^Δ
5. if the move is accepted, update $\mathbf{x}_{i+1} = \mathbf{x}_{\text{new}}$; otherwise $\mathbf{x}_{i+1} = \mathbf{x}_i$.

Implementation of the updates.– In our implementation, we ramp β linearly from $\beta_i = 1$ to $\beta_f \approx 10^4$–10^5 in $n_{\text{steps}} = 10^3$ steps. For each value of β, we attempt $n_{\text{attempts}} = 10^5$ Metropolis updates. Each move consists of selecting randomly one basis among n, and rotate randomly all its elements. Specifically, the moves are proposed as follows

1. choose randomly and uniformly one of the bases $y \in [n]$
2. draw $2d^2$ independent random numbers $\{(r_{jk}, s_{jk})\}_{(j,k)\in[d]^2}$, uniformly in the interval $[-\epsilon, \epsilon]$ (see below for the choice of ϵ)
3. define U^y_{new} as $U^y_{jk} + r_{jk} + is_{jk}$
4. make U^y_{new} unitary via the Gram–Schmidt procedure.

The parameter ϵ, which defines the typical amplitude of the proposed moves, is adapted throughout the algorithm in order to ensure a constant acceptance rate. Intuitively, when the temperature is very high, large moves involving a potentially large change in energy are required to efficiently explore the parameter space. Progressively ramping down the temperature, the bases start to stabilise in the vicinity of the maxima of W, and large moves become often rejected by the Metropolis rule. On the other hand, if ϵ is very small, the moves will not efficiently explore the parameter space. As a compromise, we adapt ϵ such that the acceptance rate r_{accept} of the Metropolis update (that is, for each value of β, r_{accept} is the number N_{accept} of accepted moves divided by the number N_{attempts} of attempts), is kept between $r_{\min} = 0.32$ and $r_{\max} = 0.48$. We initialise $\epsilon = 1$ at the beginning of the simulation, and whenever $r_{\text{accept}} < r_{\min}$, we change ϵ to 0.8ϵ. Similarly, whenever $r_{\text{accept}} > r_{\max}$, we change ϵ to 1.2ϵ.

Optimal bases.– As the goal of the simulation is not to accurately estimate the thermal average $\langle W\rangle_\beta$ [Eq. (24)], but only to efficiently find its global maximum, we do not carry out a detailed evaluation of the error on $\langle W\rangle_\beta$ as estimated from our samples. Throughout the simulation, we record the optimal bases $\mathbf{x}_{\text{opt}}$ encountered, corresponding to the maximal value of W_{opt} found so far. As a last step of the optimisation, we set $\beta = \infty$ and start a new simulation starting from $\mathbf{x}_{\text{opt}}$ as the initial configuration. Effectively, this amounts to only accepting the moves that increase W, in order to achieve as many digits of precision as needed for W_{opt}, as well as for the basis elements themselves.

5.2 Results

The results of this optimisation are summarised in Table 3. For a meaningful comparison of differ-

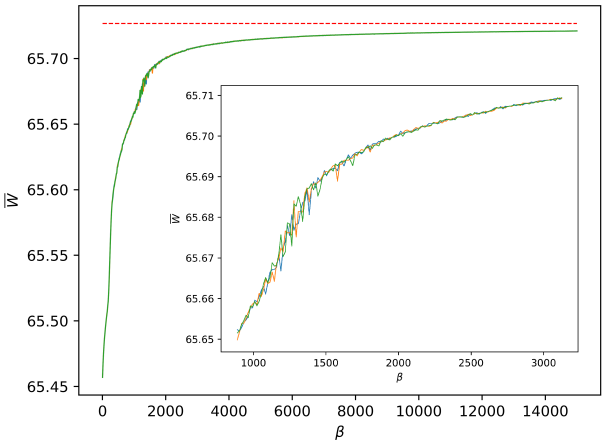

Figure 1: Monte Carlo simulation for $d = 6$, $n = 4$. As a function of the inverse temperature β, mean value of the Bell operator [cf. Eq. (24)]. Three independent simulations are plotted, and the dashed red line indicates the maximal value found throughout the simulations (see text). Inset: Zoom of the main panel for $1000 < \beta < 3000$.

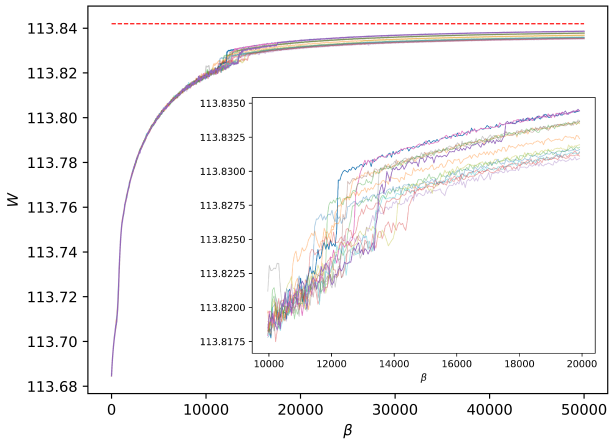

Figure 2: Monte Carlo simulation for $d = 10$, $n = 4$. As a function of the inverse temperature β, mean value of the Bell operator [cf. Eq. (24)]. Ten independent simulations are plotted, and the dashed red line indicates the maximal value found throughout the simulations. Inset: Zoom of the main panel for $10000 < \beta < 20000$.

ent dimensions d, and different number of bases n, we indicate the relative deviation to the MUB optimum: $1 - W_{\max}/W_{\mathrm{MUB}}$.

Power-of-prime dimensions.– All dimensions $2 \leq d \leq 9$ except $d = 6$ are powers of prime, for which the maximal number of MUBs is exactly $d + 1$. Our numerical simulations are consistent with this fact, and our variational optimisation systematically reconstructs n MUBs for all $n \leq d + 1$ [1]. For $d = 2, 3, 4, 5$, we also performed the optimisation for $n = d + 2$, where three independent simulations gave the same optimum within numerical accuracy, also in agreement with the see-saw and non-linear SDP methods, and with the fact that there cannot exist $d + 2$ MUBs in dimension d. The corresponding optimal solutions are analysed analytically in Section 6 for $d = 2$ and 3.

d=6.– In the case $d = 6$, we do not find more than $n = 3$ MUBs. For $n = 4$, three independent simulations gave consistently the optimum $W_{\mathrm{opt}}(6, 4) = W_{\mathrm{MUB}}(1 - 0.00004215) \approx 65.723938549\ldots$ (within numerical accuracy). The optimal bases found also coincide within numerical precision with those found by the see-saw and non-linear SDP techniques of the previous sections (also see Section 6.3 for a close analytical construction). The complete evolution of $\langle W \rangle_\beta$ is illustrated in Fig. 1 for these three simulations.

d=10.– The simulations for $d = 10$ are reaching the limits of our current implementation. We did find $n = 3$ MUBs, but not $n = 4$. As illustrated in Fig. 2 over ten independent simulations, three of them converged to the same optimum $W_{\mathrm{opt}}(10, 4) = W_{\mathrm{MUB}}(1 - 0.000005938) \approx 113.8413197\ldots$. This value is reported in Table 3, and is our best estimate for the true optimum[2]. Even though some of our simulations are trapped in local optima, our results support the conjecture that no more than three MUBs exist in dimension 10.

6 Analytic constructions

In this section we describe analytic constructions that match the best bases found numerically for the cases of four bases in dimension two, five bases in dimension three, and four bases in dimension six.

[1] For $d = 8$ we actually stopped at $n = 8$, and for $d = 9$ at $n = 5$, because of the high numerical cost of these higher-dimensional instances.

[2] Using our open source code, Markus Grassl was able to run the simulation for $d = 10$ and $n = 4$ on a computer cluster 2000 times. He reported to us that the most frequent optimum in these runs agrees with our findings, however, in two instances he found the slightly larger value of $W_{\mathrm{opt}}(10, 4) \approx 113.8414358$.

6.1 Dimension two

One can parametrise any rank-1 qubit projection B using the Bloch representation

$$B = \frac{1}{2}(\mathbb{1} + \vec{r} \cdot \vec{\sigma}), \tag{27}$$

where $\vec{r} = (x, y, z)$ is a unit vector in $\mathbb{R}^3$, $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ is a vector of the Pauli matrices, and $\vec{r}\cdot\vec{\sigma} = x\sigma_x + y\sigma_y + z\sigma_z$. Accordingly, for the case of four rank-1 projective measurements on a qubit, we use the parametrisation $B_1^y = \frac{1}{2}(\mathbb{1} + \vec{r}_y \cdot \vec{\sigma})$ for $y \in \{1, 2, 3, 4\}$, and $B_2^y = \mathbb{1} - B_1^y$. Consider then the four measurements parametrised by the four vectors

$$\begin{aligned} \vec{r}_1 &= \frac{1}{\sqrt{3}}(1, 1, 1) \\ \vec{r}_2 &= \frac{1}{\sqrt{3}}(1, -1, -1) \\ \vec{r}_3 &= \frac{1}{\sqrt{3}}(-1, 1, -1) \\ \vec{r}_4 &= \frac{1}{\sqrt{3}}(-1, -1, 1), \end{aligned} \tag{28}$$

defining a regular tetrahedron on the Bloch sphere. It is straightforward to verify that by plugging these measurements into Eq. (17) we get

$$W_{2,4}^{+B}(\{B_b^y\}) = 4(\sqrt{3} + \sqrt{6}) \approx 16.7262, \tag{29}$$

which agrees (after normalisation) with the values in Tables 1, 2, and 3 up to numerical precision.

6.2 Dimension three

For the three-dimensional case, we parametrise each basis $y \in \{1, \ldots, 5\}$ by a unitary matrix U_y, whose columns correspond to the basis vectors. Fixing the computational basis allows us to write the first basis as $U_1 = \mathbb{1}$. We then take the second basis to be the Fourier basis

$$U_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega_3 & \omega_3^2 \\ 1 & \omega_3^2 & \omega_3 \end{bmatrix}, \tag{30}$$

where $\omega_3 = \mathrm{e}^{\frac{2\pi \mathrm{i}}{3}}$. We define the remaining three bases as

$$\begin{aligned} U_3 &= \mathrm{diag}(\omega_6, \bar{\omega}_6, 1) U_2 \\ U_4 &= \mathrm{diag}(1, \omega_6, \bar{\omega}_6) U_2 \\ U_5 &= \mathrm{diag}(\bar{\omega}_6, 1, \omega_6) U_2, \end{aligned} \tag{31}$$

where $\omega_6 = \mathrm{e}^{\frac{2\pi \mathrm{i}}{6}}$, $\bar{\omega}_6$ is its complex conjugate, and $\mathrm{diag}(x, y, z)$ is a diagonal matrix of x, y, and z. Note that all the bases U_2, U_3, U_4, and U_5 are unbiased to U_1. It is straightforward to verify that by plugging the measurements B_b^y corresponding to the bases U_y into Eq. (17), we get

$$W_{3,5}^{+B}(\{B_b^y\}) = 8(\sqrt{2} + \sqrt{5} + \sqrt{6}) \approx 48.7982, \tag{32}$$

which agrees (after normalisation) with the values in Tables 1, 2, and 3 up to numerical precision.

6.3 Dimension six

All three numerical methods converged to (numerically) the same set of four bases in dimension six. Just like in the case of five bases in dimension three, this set has the property that one basis is unbiased to the other three bases. Upon closer inspection, one finds that these bases are numerically very close to the "four most distant bases" of Ref. [21], which were found via maximising the MUBness measure

$$D^2 \propto \sum_{(y,z)\in \mathrm{Pairs}[n]} \sum_{j,k} |\langle b_j^y | b_k^z \rangle|^2 (1 - |\langle b_j^y | b_k^z \rangle|^2) \ . \tag{33}$$

One can again parametrise the bases by four unitary matrices such that the first one is $U_1 = \mathbb{1}$. The rest of the bases are based on Eq. (6) of Ref. [21], in which the authors describe a family of three unitary matrices, depending on two parameters, θ_t and θ_x. The "optimal" values of these parameters (optimality here originally means maximising the distance measure of Ref. [20]) is then determined by finding the unique real solution of Eq. (20) in Ref. [21], plugging it into Eq. (19) of Ref. [21] to obtain θ_t, and plugging it into Eq. (21) of Ref. [21] to obtain θ_x. These analytic values correspond to approximately $(\theta_x, \theta_t) \approx (0.9852276, 1.0093680)$.

If we substitute the resulting bases $\{B_b^y\}$ into our optimisation problem in Eq. (19), we obtain (analytically)

$$W_{6,4}^{+B}(\{B_b^y\}) \approx 65.7239381 \tag{34}$$

Comparing with our optimum ≈ 65.7239385, we observe that: 1) the analytical optimum (for the bases of Ref. [21]) and our numerical optimum differ after the eighth significant figure; and 2) consistently, our numerical optimum is larger than the analytical one. This should not be a surprise,

for the two solutions optimise different MUBness measures, respectively Eq. (18) and Eq. (33). As a further comparison, one can look at the overlaps $\mathrm{tr}(B_j^y B_k^z) = |\langle b_j^y | b_k^z \rangle|^2$ of the bases found. We indeed find that one of the bases is unbiased to the other three, i.e.

$$\mathrm{tr}(B_j^1 B_k^y) = |\langle b_j^y | b_k^z \rangle|^2 = 0.166667 \quad \forall j, y, k. \tag{35}$$

Apart from these, there are three different values of overlaps, whose approximate values 0.124, 0.181, and 0.152 agree with our numerical findings up to two or three significant digits.

In conclusion, our three numerical methods, together with the approach of Ref. [21] all appear to converge to essentially the same set of four bases in dimension six. Small differences in the overlaps originate in the different MUBness measures which are optimised. The results of these four independent different numerical approaches may signify that Zauner's conjecture is indeed correct, although they cannot rule out the contrary.

7 Conclusions

We reformulated the existence problem of MUBs as an optimisation problem, using a recently found family of Bell inequalities. We then applied three numerical methods suitable for optimising Bell inequalities in order to tackle the existence problem: see-saw SDP optimisation, non-linear SDP, and Monte Carlo techniques. The results of all these numerical optimisations is in full accordance with the known cases in dimensions $d = 2, 3, 4, 5$, where we find $d + 1$ MUBs. Furthermore, whenever it is known that n MUBs do not exist in a given dimension d, all the different algorithms converge to the same set of bases in all cases (with a slight difference between the see-saw method and the other two methods for $d = 4$, $n = 6$), and these bases are not MUBs.

We applied our numerical techniques to the open case of four MUBs in dimension six. All three algorithms suggest that there do not exist four MUBs in dimension six, by converging to a Bell value strictly smaller than the hypothetical MUB value. Moreover, the bases found by all three algorithms are very close numerically with the "four most distant bases" in dimension six of Ref. [21]. Hence, our findings provide further numerical evidence for Zauner's conjecture. In the next composite dimension, $d = 10$, our Monte Carlo results suggest that no more than $n = 3$ MUBs exist.

It is important to point out that the numerical methods used in this work are heuristic, i.e. there is no guarantee of convergence to the global optimum. As such, heuristic numerics can never provide a rigorous proof of the non-existence of MUBs (only that of existence, by explicitly finding MUBs). To overcome this shortcoming, a plausible future direction towards a rigorous numerical proof is using a variant of the Navascués–Pironio–Acín hierarchy of SDPs [29] for maximising a Bell inequality in a fixed dimension. While such numerical optimisation is computationally significantly more expensive than those in our work, it provides certifiable upper bounds on Bell inequality violations, and therefore could in principle be used to rigorously prove the non-existence of MUBs.

Code availability. The numerical findings presented in this paper can be reproduced using the codes made available on public repositories. For the see-saw SDP, see `https://github.com/Lumorti/seesaw`. For the non-linear SDP, see `https://github.com/Lumorti/nonlinear`. For the Monte Carlo simulations, see `https://github.com/mariaprat/mubs-montecarlo`.

8 Acknowledgements

The authors thank Markus Grassl for providing a large set of data obtained by running our code on a computer cluster. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 754510, the Government of Spain (FIS2020-TRANQI, Severo Ochoa CEX2019-000910-S), Fundacio Cellex, Fundacio Mir-Puig, Generalitat de Catalunya (CERCA, AGAUR SGR 1381 and QuantumCAT), the ERC AdG CERQUTE and the AXA Chair in Quantum Information Science, the Agence Nationale de la Recherche (Qu-DICE project ANR-PRC-CES47), the John Templeton Foundation (Grant No. 61835), ICFO and SPIE under a Maria Yzuel Fellowship Award.

References

[1] I D Ivanovic. Geometrical description of quantal state determination. *Journal of Physics A: Mathematical and General*, 14(12):3241–3245, 1981. doi:10.1088/0305-4470/14/12/019.

[2] G. Brassard C. H. Bennett. Quantum cryptography: Public key distribution and coin tossing. *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing (IEEE, 1984)*, 175:8, 1984. doi:10.1016/j.tcs.2011.08.039.

[3] Artur K. Ekert. Quantum cryptography based on Bell's theorem. *Phys. Rev. Lett.*, 67:661–663, 1991. doi:10.1103/PhysRevLett.67.661.

[4] Dagmar Bruß. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.*, 81:3018–3021, 1998. doi:10.1103/PhysRevLett.81.3018.

[5] Armin Tavakoli, Alley Hameedi, Breno Marques, and Mohamed Bourennane. Quantum random access codes using single d-level systems. *Phys. Rev. Lett.*, 114:170502, 2015. doi:10.1103/PhysRevLett.114.170502.

[6] Máté Farkas and Jędrzej Kaniewski. Self-testing mutually unbiased bases in the prepare-and-measure scenario. *Phys. Rev. A*, 99:032316, 2019. doi:10.1103/PhysRevA.99.032316.

[7] H. Bechmann-Pasquinucci and N. Gisin. Bell inequality for qunits with binary measurements. *Quantum Info. Comput.*, 3(2):157–164, 2003. doi:10.26421/QIC3.2-6.

[8] Jędrzej Kaniewski, Ivan Šupić, Jordi Tura, Flavio Baccari, Alexia Salavrakos, and Remigiusz Augusiak. Maximal nonlocality from maximal entanglement and mutually unbiased bases, and self-testing of two-qutrit quantum systems. *Quantum*, 3:198, 2019. doi:10.22331/q-2019-10-24-198.

[9] Armin Tavakoli, Máté Farkas, Denis Rosset, Jean-Daniel Bancal, and Jędrzej Kaniewski. Mutually unbiased bases and symmetric informationally complete measurements in Bell experiments. *Science Advances*, 7(7):eabc3847, 2021. doi:10.1126/sciadv.abc3847.

[10] Thomas Durt, Berthold-Georg Englert, Ingemar Bengtsson, and Karol Życzkowski. On mutually unbiased bases. *International Journal of Quantum Information*, 08(04):535–640, 2010. doi:10.1142/S0219749910006502.

[11] William K Wootters and Brian D Fields. Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, 191(2):363–381, 1989. doi:10.1016/0003-4916(89)90322-9.

[12] Paweł Wocjan and Thomas Beth. New construction of mutually unbiased bases in square dimensions. *Quantum Info. Comput.*, 5(2):93–101, 2005. doi:10.26421/QIC5.2-1.

[13] Mihály Weiner. A gap for the maximum number of mutually unbiased bases. *Proc. Amer. Math. Soc.*, 141:1963–1969, 2013. doi:10.1090/S0002-9939-2013-11487-5.

[14] Gerhard Zauner. Quantendesigns: Grundzüge einer nichtkommutativen Designtheorie. *PhD thesis*, 1999.

[15] P. Oscar Boykin, Meera Sitharam, Pham Huu Tiep, and Pawel Wocjan. Mutually unbiased bases and orthogonal decompositions of Lie algebras. *Quantum Info. Comput.*, 7(4):371–382, 2007. doi:10.26421/QIC7.4-6.

[16] Stephen Brierley and Stefan Weigert. Constructing mutually unbiased bases in dimension six. *Phys. Rev. A*, 79:052316, 2009. doi:10.1103/PhysRevA.79.052316.

[17] Philippe Jaming, Máté Matolcsi, Péter Móra, Ferenc Szöllősi, and Mihály Weiner. A generalized Pauli problem and an infinite family of MUB-triplets in dimension 6. *Journal of Physics A: Mathematical and Theoretical*, 42(24):245305, may 2009. doi:10.1088/1751-8113/42/24/245305.

[18] Gary McConnell, Harry Spencer, and Afaq Tahir. Evidence for and against Zauner's MUB conjecture in $\mathbb{C}^6$. 2021. doi:10.48550/arXiv.2103.08703.

[19] Sander Gribling and Sven Polak. Mutually unbiased bases: polynomial optimization and symmetry. 2021. doi:10.48550/arXiv.2111.05698.

[20] Ingemar Bengtsson, Wojciech Bruzda, Åsa Ericsson, Jan-Åke Larsson, Wojciech Tadej, and Karol Życzkowski. Mutually unbiased bases and Hadamard matrices of order six. *Journal of Mathematical Physics*, 48(5):052106, 2007. doi:10.1063/1.2716990.

[21] Philippe Raynal, Xin Lü, and Berthold-Georg Englert. Mutually unbiased bases in six dimensions: The four most distant bases. *Phys. Rev. A*, 83:062303, 2011. doi:10.1103/PhysRevA.83.062303.

[22] Edgar A. Aguilar, Jakub J. Borkała, Piotr Mironowicz, and Marcin Pawłowski. Connections between mutually unbiased bases and quantum random access codes. *Phys. Rev. Lett.*, 121:050501, 2018. doi:10.1103/PhysRevLett.121.050501.

[23] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner. Bell nonlocality. *Rev. Mod. Phys.*, 86:419–478, 2014. doi:10.1103/RevModPhys.86.419.

[24] MOSEK ApS. *MOSEK Fusion API for C++ 9.2.49*, 2021. URL: https://docs.mosek.com/9.2/cxxfusion/index.html.

[25] Hiroshi Yamashita, Hiroshi Yabe, and Kouhei Harada. A primal–dual interior point method for nonlinear semidefinite programming. *Mathematical programming*, 135(1):89–121, 2012. doi:10.1007/s10107-011-0449-z.

[26] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. doi:10.1017/CBO9780511804441.

[27] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983. doi:10.1126/science.220.4598.671.

[28] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092, 1953. doi:10.1063/1.1699114.

[29] Miguel Navascués, Stefano Pironio, and Antonio Acín. Bounding the set of quantum correlations. *Phys. Rev. Lett.*, 98:010401, 2007. doi:10.1103/PhysRevLett.98.010401.