Perturbation of the numerical range of unitary matrices

Ryszard Kukulski, Paulina Lewandowska, and Łukasz Pawela

Institute of Theoretical and Applied Informatics, Polish Academy of Sciences,
ul. Bałtycka 5, 44-100 Gliwice, Poland
`rkukulski@iitis.pl`
`https://www.iitis.pl/en/person/rkukulski`

Abstract. In this work we show how to approach the problem of manipulating the numerical range of a unitary matrix. This task has far-reaching impact on the study of discrimination of quantum measurements. We achieve the aforementioned manipulation by introducing a method which allows us to find a unitary matrix whose numerical range contains the origin where at the same time the distance between unitary matrix and its perturbation is relative small in given metric.

Keywords: numerical range, perturbation of the numerical range, unitary matrices, quantum unitary channels

1 Introduction

One of the most important tasks in quantum information theory is a problem of distinguishability of quantum channels [1,2]. Imagine we have an unknown device, a black-box. The only information we have is that it performs one of two channels, say Φ and Ψ. We want to tell whether it is possible to discriminate Φ and Ψ perfectly, i.e. with probability equal to one. Helstrom's result [3] gives the analytical formula for upper bound of probability of discrimination quantum channels using the special operators norm called diamond norm or sometimes referred to as completely bounded trace norm [4]. The Holevo-Helstrom theorem says that the quantum channels Ψ and Φ are perfectly distinguishable if and only if the distance between them is equal two by using diamond norm. In general, numerical computing of diamond norm is a complex task. Therefore, researchers were limited to smaller classes of quantum channels. One of the first results was the study of discrimination of unitary channels $\Phi_U : \rho \to U\rho U^\dagger$ where ρ is a quantum state. The sufficient condition for perfect discrimination of unitary channels Φ_U and Φ_1 is that zero belongs to the numerical range of unitary matrix U [5].

The situation in which zero belongs to numerical range of unitary matrix U paves the way toward simple calculating of probability of discrimination unitary channels without the necessity of computing the diamond norm. Now consider the following scenario. We have two quantum channels Φ_U and Φ_1 such that zero does not belong to the numerical range of U. Hence, we know that we

cannot distinguish between Φ_U and $\Phi_\mathbf{1}$ perfectly. Therefore, we can assume some kind of noise and consider the unitary channel Φ_V instead of Φ_U such that the distance between unitary matrix V and U is relative small where at the same time zero belongs in numerical range of V. Such a unitary matrix V will be called perturbation of U.

In this work we are interested in determining the perturbation form of V. Our motivation is two-fold. On the one hand considering the unitary channels Φ_V and $\Phi_\mathbf{1}$ we know that they will be perfectly distinguishable. On the other hand our method of computing V does not change the measurement result in standard basis.

Our work is naturally divided into three parts. In the first part we show the mathematical preliminaries needed to present our main result. The second part presents the theorem which gives us the method of manipulation of numerical range of unitary matrices. In third part we show the example illustrative our theorem. Concluding remarks and some future work are presented in the end of our work.

2   Mathematical preliminaries

Let us introduce the following notation. Let $\mathbb{C}^d$ be complex d-dimensional vector space. We denote the set of all matrix operators by $\mathrm{L}(\mathbb{C}^{d_1}, \mathbb{C}^{d_2})$ while the set of isometries by $\mathrm{U}(\mathbb{C}^{d_1}, \mathbb{C}^{d_2})$. It easy to see that every square isometry is a unitary matrix. The set of all unitary matrices we will be denoted by $U \in \mathrm{U}(\mathbb{C}^d)$. We will be also interested in diagonal matrices and diagonal unitary matrices denoted by $\mathrm{Diag}(\mathbb{C}^d)$ and $\mathrm{DU}(\mathbb{C}^d)$ respectively. Next classes of matrices that will be used in this work are Hermitian matrices denoted by $\mathrm{Herm}(\mathbb{C}^d)$. All of the above-mentioned matrices are normal matrices i.e. $AA^\dagger = A^\dagger A$. Every normal matrix A can be expressed as a linear combination of projections onto pairwise orthogonal subspaces

$$A = \sum_{i=1}^{k} \lambda_i \, |x_i\rangle\!\langle x_i| \,, \tag{1}$$

where scalar $\lambda_i \in \mathbb{C}$ is an eigenvalue of A and $|x_i\rangle \subset \mathbb{C}^d$ is an eigenvector corresponding to the eigenvalue λ_i. This expression of a normal matrix A is called a spectral decomposition of A [6]. Many interesting and useful norms, not only for normal matrices, can be defined on spaces of matrix operators. In this work we will mostly be concerned with a family of norms called Schatten [7] p-norms defined as

$$||A||_p = \left(\mathrm{tr}\left(\left(A^\dagger A \right)^{\frac{p}{2}} \right) \right)^{\frac{1}{p}} \tag{2}$$

for any $A \in \mathrm{L}(\mathbb{C}^{d_1}, \mathbb{C}^{d_2})$. The Schatten ∞-norm is defined as

$$||A||_\infty = \max \left\{ ||A\,|u\rangle\,|| : |u\rangle \in \mathbb{C}^{d_1}, ||\,|u\rangle\,|| \leq 1 \right\}. \tag{3}$$

For a given square matrix A the set of all eigenvalues of A will be denoted by $\lambda(A)$ and $r(\lambda_i)$ will denote the multiplicity of each eigenvalue $\lambda_i \in \lambda(A)$. For any

square matrix A, one defines its numerical range [8,9] as a subset of the complex plane

$$W(A) = \{z \in \mathbb{C} : z = \langle \psi | A | \psi \rangle, | \psi \rangle \in \mathbb{C}^d, \langle \psi | \psi \rangle = 1\}. \tag{4}$$

It is easy to see that $\lambda(A) \subseteq W(A)$. One of the most important properties of $W(A)$ is its convexity which was shown by Hausdorff and Toeplitz [10,11]. For any normal matrix A the set $W(A)$ is a convex hull of spectrum of A which will be denoted by $\text{conv}(\lambda(A))$. Another well-known property of $W(U)$ for any unitary matrix $U \in \mathrm{U}(\mathbb{C}^d)$ is the fact that its numerical range forms a polygon whose vertices are eigenvalues of U lying in unit circle on complex plane. In our work we introduce the counterclockwise order of eigenvalues of unitary matrix U [12] such that we choose any eigenvalue named $\lambda_1 \in \lambda(U)$ on the unit circle and next eigenvalues are labeled counterclockwise.

In our setup we consider the space $\mathrm{L}(\mathbb{C}^d)$. Imagine that the matrices are points in space $\mathrm{L}(\mathbb{C}^d)$ and the distance between them is bounded by small constant $0 < c \ll 1$. We will take two unitary matrices - matrix $U \in \mathrm{U}(\mathbb{C}^d)$ and its perturbation $V \in \mathrm{U}(\mathbb{C}^d)$ i.e. $||U - V||_\infty \leq c$ by using ∞-Schatten norm. We want to determine the path connecting these points given by smooth curve. To do so, we fix continuous parametric (by parameter t) curve $U(t) \in \mathrm{U}(\mathbb{C}^d)$ for any $t \in [0,1]$ with boundary conditions $U(0) := U$ and $U(1) := V$. The most natural and also the shortest curve connecting U and V is geodesic [13] given by

$$t \to U \exp\left(t \operatorname{Log}\left(U^\dagger V\right)\right), \tag{5}$$

where Log is the matrix function such that it changes eigenvalues $\lambda \in \lambda(U)$ into $\log(\lambda(U))$, where $-i \log(\lambda(U)) \subset (-\pi, \pi]$.

We will study how the numerical range $W(U(t))$ will be changed depending on parameter t. Let $H := -i \operatorname{Log}\left(U^\dagger V\right)$. Let us see that $H \in \operatorname{Herm}(\mathbb{C}^d)$ and $W(H) \subset (-\pi, \pi]$ for any $U, V \in \mathrm{U}(\mathbb{C}^d)$. Taking the spectral decomposition $H = YDY^\dagger, Y \in \mathrm{U}(\mathbb{C}^d), D \in \operatorname{Diag}(\mathbb{C}^d)$ we get

$$\begin{aligned} W\left(U \exp\left(itH\right)\right) &= W\left(U \exp\left(itYDY^\dagger\right)\right) = W\left(UY \exp\left(itD\right)Y^\dagger\right) \\ &= W\left(Y^\dagger UY \exp\left(itD\right)\right) = W\left(\widetilde{U} \exp\left(itD\right)\right), \end{aligned} \tag{6}$$

where $\widetilde{U} := Y^\dagger UY \in \mathrm{U}(\mathbb{C}^d)$. Hence, without loss of generality in our analysis we will assume that H is a diagonal matrix. Moreover, we can assume that $D \geq 0$ which follows from simple calculations

$$\begin{aligned} W\left(U \exp\left(itD\right)\right) &= W\left(U \exp\left(itD_+\right)\left(\exp\left(it\alpha \mathbb{1}\right)\right)\right) = W\left(e^{it\alpha}U \exp\left(itD_+\right)\right) \\ &\equiv W\left(U \exp\left(itD_+\right)\right). \end{aligned} \tag{7}$$

Let us see that the numerical range of $U(t)$ for any $t \in [0,1]$ is invariant to above calculations although the trajectory of $U(t)$ is changed. Therefore, we will consider the curve

$$t \to U \exp\left(itD_+\right), \tag{8}$$

where $t \in [0,1]$ and $U \in \mathrm{U}(\mathbb{C}^d)$, $D_+ \in \operatorname{Diag}(\mathbb{C}^d)$ such that $D_+ \geq 0$.

3   Main result

In this section we will focus on the behavior of the spectrum of the unitary matrices $U(t)$, which will reveal the behavior of $W(U(t))$ for relatively small parameter t. Without loss of generality we can assume that $\mathrm{tr}\,(D_+) = 1$. Together with the fact that $D_+ \in \mathrm{Diag}(\mathbb{C}^d)$ and $D_+ \geq 0$ we can note that $D_+ = \sum_{i=1}^d p_i\,|i\rangle\langle i|$, where $p \in \mathbb{C}^d$ is a probability vector. Let us also define the set

$$S_\lambda^M = \left\{|x\rangle \in \mathbb{C}^d : (\lambda \mathbb{1}_d - M)\,|x\rangle = 0, \|\,|x\rangle\,\|_2 = 1\right\} \tag{9}$$

for some matrix $M \in \mathrm{L}(\mathbb{C}^d)$ which consists of unit eigenvectors corresponding to the eigenvalue λ of the matrix M. We denote by $k = r(\lambda)$ the multiplicity of eigenvalue λ whereas by $I_{M,\lambda} \in U(\mathbb{C}^k, \mathbb{C}^d)$ we denote the isometry which columns are formed by eigenvectors corresponding to eigenvalue λ of a such matrix M. Let $\lambda(t), \beta(t) \in \mathbb{C}$ for $t \geq 0$. We will write $\lambda(t) \approx \beta(t)$ for relatively small $t \geq 0$, whenever $\lambda(0) = \beta(0)$ and $\frac{\partial}{\partial t}\lambda(0) = \frac{\partial}{\partial t}\beta(0)$.

Theorem 1. *Let $U \in \mathrm{U}(\mathbb{C}^d)$ be a unitary matrix with spectral decomposition*

$$U = \sum_{j=1}^d \lambda_j\,|x_j\rangle\langle x_j|. \tag{10}$$

Assume that the eigenvalue $\lambda \in \lambda(U)$ is such that $r(\lambda) = k$. Let us define a matrix $E(t)$ given by

$$E(t) = \exp(itD_+) = \sum_{i=1}^d e^{ip_i t}\,|i\rangle\langle i| \in \mathrm{DU}(\mathbb{C}^d), \quad t \geq 0. \tag{11}$$

Let $\lambda(t) := \lambda(UE(t))$ and let every $\lambda_j(t) \in \lambda(t)$ corresponds to eigenvector $|x_j(t)\rangle$. Assume that $\lambda_1(t), \ldots, \lambda_k(t)$ are such eigenvalues that $\lambda_j(t) \to \lambda$, as $t \to 0$. Then:

(a) If $\min\limits_{|x\rangle \in S_\lambda^U} \sum\limits_{i=1}^d p_i|\langle i|x\rangle|^2 = 0$, then λ is an eigenvalue of $UE(t)$.
(b) If $|\{p_i : p_i > 0\}| = l < k$, then λ is an eigenvalue of $UE(t)$ and $r(\lambda) \geq k - l$.
(c) Each eigenvalue of product $UE(t)$ moves counterclockwise or stays in the initial position as parameter t increases.
(d) If $k = 1$, then

$$\lambda_1(t) \approx \lambda \exp\left(it \sum_{i=1}^d p_i|\langle i|x_1\rangle|^2\right)$$

for small $t \geq 0$.
(e) Let $Q := I_{U,\lambda}^\dagger D_+ I_{U,\lambda}$ and $\lambda_1(Q) \leq \lambda_2(Q) \leq \ldots \leq \lambda_k(Q)$. Then we have

$$\lambda_j(t) \approx \lambda \exp\left(i\lambda_j(Q)t\right)$$

for small $t \geq 0$ and eigenvector $|x_j\rangle$ corresponding to $\lambda_j \in \lambda(U)$ is given by

$$|x_j\rangle = I_{U,\lambda}\,|v_j\rangle,$$

where $|v_j\rangle \in S_{\lambda_j(Q)}^Q$.

(f) For each $j = 1, \ldots, d$ we have

$$\frac{\partial}{\partial t}\lambda_j(t) = i\lambda_j(t)\sum_{i=1}^{d} p_i |\langle i|x_j(t)\rangle|^2.$$

Moreover,

$$\sum_{j=1}^{d}\left|\frac{\partial}{\partial t}\lambda_j(t)\right| = 1.$$

This theorem gives us equations which one can use to predict behavior of $W(UE(t))$. Observe the postulate (f) fully determines the movement of the spectrum. However, this is a theoretical statement and in practice determining the function $t \mapsto |x_j(t)\rangle$ is a numerically complex task. The postulates $(a) - (e)$ play a key role in numerical calculations of $W(UE(t))$. The most important fact comes from (c) which says that all eigenvalues move in the same direction or stay in the initial position. The instantaneous velocity of a given eigenvalue in general case is given in (e), while in the case of eigenvalue with multiplicity equal one, the instantaneous velocity is determined by (d). We can see that whenever the spectrum of the matrix U is not degenerated, calculating these velocities is easy. What is more, when some eigenvalue is degenerated, the postulate (e) not only gives us method to calculate the trajectory of this eigenvalue, but also determines the form of corresponding eigenvector. It is worth noting that the postulates $(d), (e)$ give us only an approximation of the velocities, so despite being useful in numerical calculations, these expressions are valid only in the neighborhood of $t = 0$. Moreover, sometimes we are able to precisely specify this velocities. This happens in the cases presented in $(a), (b)$. Whenever the calculated velocity is zero we know for sure that this eigenvalue will stay in the initial position. According to the postulate (b) the same happens when the multiplicity of the eigenvalue is greater than the number of positive elements of vector p.

4   Example

We start with sampling some random unitary matrix $U \in \mathrm{U}(\mathbb{C}^3)$ such that $0 \notin W(U)$

$$U = \begin{bmatrix} 0.267868 + 0.026891i & 0.752935 - 0.510663i & -0.314404 - 0.0313982i \\ -0.83413 - 0.0693252i & 0.245915 - 0.275811i & 0.34174 - 0.214685i \\ 0.472125 + 0.0635826i & 0.0211772 - 0.18793i & 0.795835 - 0.322391i \end{bmatrix}.$$

(12)

We would like to construct matrix $E \in \mathcal{DU}(\mathbb{C}^d)$ to obtain $0 \in W(UE)$. The first feature we need to analyze is the numerical range of U given as $\mathrm{conv}(\{\lambda_1, \lambda_2, \lambda_3\})$ for $\lambda_1, \lambda_2, \lambda_3 \in \lambda(U)$. The Figure 1 presents numerical range $W(U)$, which boundary is determined by dashed triangle. The most distant pair of eigenvalues, which determine the distance between $W(U)$ and the origin is the pair
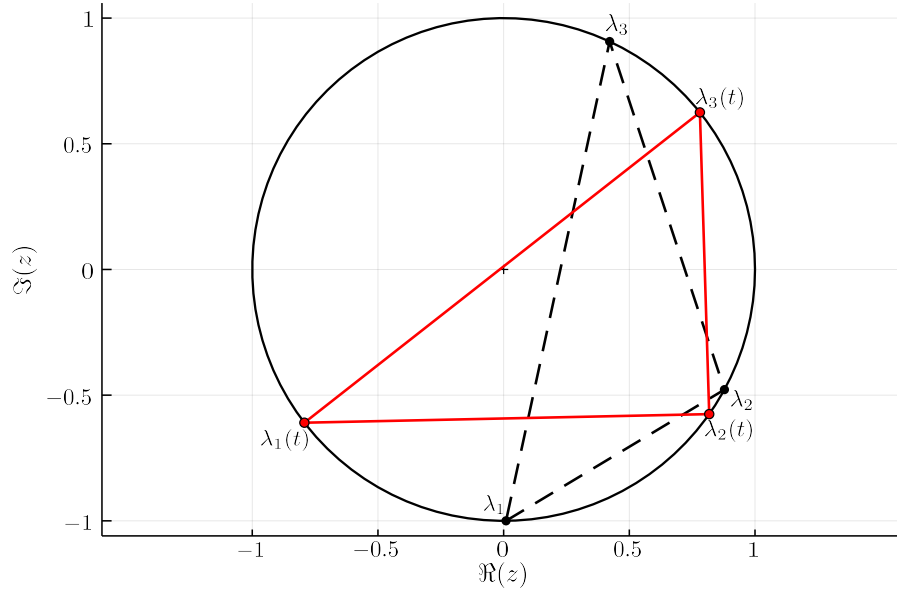
Fig. 1: The figure consists two triangles representing boundaries of the numerical ranges: $W(U)$ (dashed black) and $W(UE(t))$ for $t = 1.5$ (solid red).

(λ_1, λ_3). Therefore, our analysis will be focused on eigenvectors corresponding to λ_1 and λ_3. Let us calculate eigenvectors of matrix U and then according to the postulate (d) choose appropriate probability vector p. The squared modules of eigenvectors entries form the matrix

$$Q = \begin{bmatrix} 0.426542 & 0.543517 & 0.0299407 \\ 0.0480551 & 0.105588 & 0.846357 \\ 0.525403 & 0.350895 & 0.123702 \end{bmatrix} \tag{13}$$

Rows of the matrix Q correspond to the considered eigenvectors, i.e. first row corresponds to λ_1, second to λ_2 and third to λ_3. Analyzing the first and the third row, we can see the greatest difference in values is in the second column, namely between values $Q_{1,2} = 0.543517$ and $Q_{3,2} = 0.350895$. If we consider probability vector $p = (0, 1, 0)$ with associated matrix $E(t) = \sum_{i=1}^{3} e^{ip_i t} |i\rangle\langle i|$, then according to Theorem 1, the values $Q_{1,2}$ and $Q_{3,2}$ will indicate the instantaneous velocities of eigenvalues λ_1 and λ_3, respectively. Therefore, the first eigenvalue will move faster than the third, and hence, in order to minimize the distance between the numerical range and the origin, we have to rotate spectrum clockwise. This can be done by substituting $E(t) \leftarrow E(t)^\dagger$, namely $E(t) = \sum_{i=1}^{3} e^{-ip_i t} |i\rangle\langle i|$. The postulate (d) of Theorem 1 provides our calculations accurate only in the neighborhood of $t = 0$, but this may be not sufficient to satisfy $0 \in W(UE(t))$. Therefore, we utilize the postulate (f) and calculate instantaneous velocities

$|\langle 2|x_1(t)\rangle|^2$ and $|\langle 2|x_3(t)\rangle|^2$ of eigenvalues $\lambda_1(t)$ and $\lambda_3(t)$ accurately for each $t > 0$. As long as $|\langle 2|x_1(t)\rangle|^2 > |\langle 2|x_3(t)\rangle|^2$ the eigenvalue $\lambda_1(t)$ will be faster than $\lambda_3(t)$, which will progressively minimize the value $\min|W(UE(t))|$. The Figure 2 shows velocities $|\langle 2|x_1(t)\rangle|^2$ and $|\langle 2|x_3(t)\rangle|^2$ calculated for $t \in [0, 1.5]$. As we can see, for $t \in [0, 1.5]$ the relation $|\langle 2|x_1(t)\rangle|^2 > |\langle 2|x_3(t)\rangle|^2$ is satisfied,
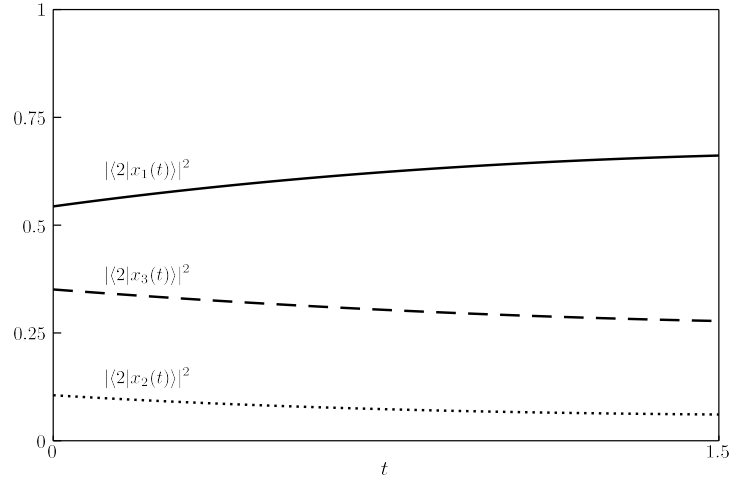

Fig. 2: The instantaneous velocities of eigenvalues $\lambda_1(t)$ (solid line), $\lambda_2(t)$ (dotted line) and $\lambda_3(t)$ (dashed line) of matrices $U(t)$. The eigenvectors $|x_1(t)\rangle, |x_2(t)\rangle, |x_3(t)\rangle$ correspond to eigenvalues $\lambda_1, \lambda_2, \lambda_3$, respectively.

so it remains to check the motion of $W(U(t))$, where we define $U(t) := UE(t)$. First of all, in the Figure 1 we showed the numerical range $W(UE(1.5))$ and we can see that $0 \in W(UE(1.5))$. The behavior of function $t \mapsto W(U(t))$ is presented in Figure 3. At last, numerical calculations reveal that the time t when zero enters numerical range $W(U(t))$ is approximately $t \approx 1.45$ (see Figure 4 for more details).

5   Conclusion and discussion

In this work we considered an approach to manipulation of the numerical range of unitary matrices. That was done by multiplying given unitary matrix U by some unitary matrix E which is diagonal in the fixed computational basis (we took the standard basis) and is relatively close to identity matrix. We established differential equations describing behavior of eigenvalues and presented their approximated solutions, which we find useful in numerical calculations. Our motivation was to find for given unitary matrix U the closest unitary matrix of the form $V = UE$ such that channels Φ_V and $\Phi_\mathbf{1}$ are perfectly distinguishable. It is
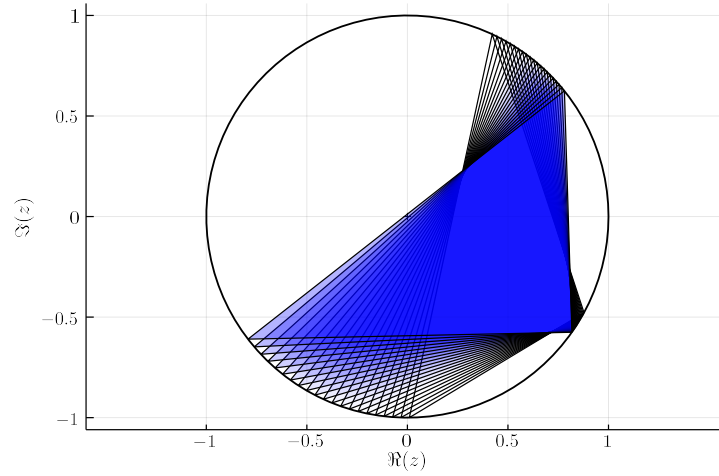
Fig. 3: Time dependence of $W(U(t))$ for $t \in [0, 1.5]$.
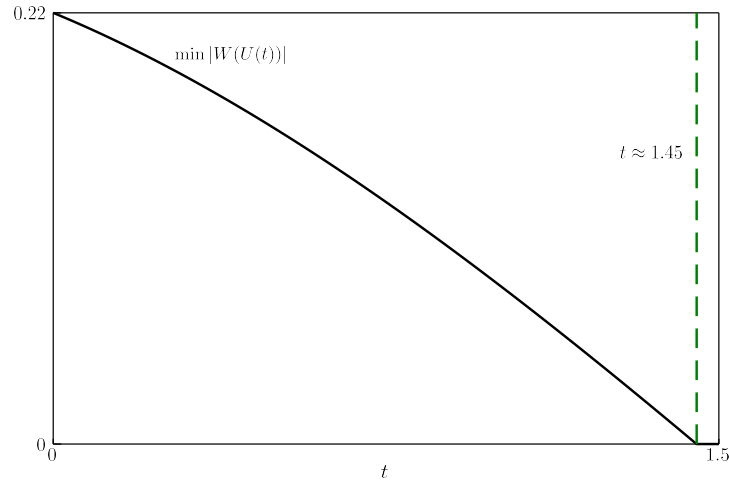

Fig. 4: The distance between numerical range $W(U(t))$ and the origin, calculated for $t \in [0, 1.5]$. Zero enters $W(U(t))$ around value ≈ 1.45.

important to stress that applying channel Φ_V to the quantum states leaves their classical distribution unchanged.

The results in Theorem 1 are suitable to solve various tasks. For example, one would like to maximize the distance between the numerical range and the point zero. Such task was introduced in [14] and plays crucial role in calculating diamond norm of difference of two von Neumann measurements.

In the end, we would like to note that solving numerical tasks with the use of postulates (d) and (e) of Theorem 1 requires one to find the bounds on maximal value of the parameter t still allowing to apply these results. Such bounds should depend on variability of functions $t \mapsto |\langle i|x_j(t)\rangle|^2$. Numerical analysis (e.g. Figure 2) shows smoothness of functions of this type, giving an argument for existence of practical bounds and hence providing direction of future research.

Acknowledgements

References

1. C. W. Helstrom, "Quantum detection and estimation theory," *Academic press*, 1976.
2. A. Jenčová, "Base norms and discrimination of generalized quantum channels," *Journal of Mathematical Physics*, vol. 55, no. 2, p. 022201, 2014.
3. C. W. Helstrom, "Quantum detection and estimation theory," *Journal of Statistical Physics*, vol. 1, no. 2, pp. 231–252, 1969.
4. D. Aharonov, A. Kitaev, and N. Nisan, "Quantum circuits with mixed states," *Proceedings of the thirtieth annual ACM symposium on Theory of computing*, pp. 20–30, 1998.
5. J. Watrous, "The theory of quantum information," *Cambridge University Press*, 2018.
6. B. C. Hall, "Quantum theory for mathematicians," *Springer*, vol. 267, 2013.
7. J. N. Franklin, "Matrix theory," *Courier Corporation*, 2012.
8. F. D. Murnaghan, "On the field of values of a square matrix," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 18, no. 3, p. 246, 1932.
9. K. Gustafson and D. Rao, "The field of values of linear operators and matrices, universitext," 1997.
10. O. Toeplitz, "Das algebraische analogon zu einem satze von fejér," *Mathematische Zeitschrift*, vol. 2, no. 1-2, pp. 187–197, 1918.
11. F. Hausdorff, "Der wertvorrat einer bilinearform," *Mathematische Zeitschrift*, vol. 3, no. 1, pp. 314–316, 1919.
12. R. Bhatia, "Matrix analysis," *Springer Science & Business Media*, vol. 169, 2013.
13. J. Antezana, G. Larotonda, and A. Varela, "Optimal paths for symmetric actions in the unitary group," *Communications in Mathematical Physics*, vol. 328, no. 2, pp. 481–497, 2014.
14. Z. Puchała, L. Pawela, A. Krawiec, and R. Kukulski, "Strategies for optimal single-shot discrimination of quantum measurements," *Physical Review A*, vol. 98, no. 4, p. 042103, 2018.

A   Proof of Theorem 1

Proof. (a) This fact is implicated by equation $\sum\limits_{i=1}^{d} p_i |\langle i|x\rangle|^2 = 0$ for some eigenvector $|x\rangle$ of eigenvalue λ. Eventually, we obtain $UE(t)|x\rangle = U|x\rangle = \lambda|x\rangle$.

(b) We will show that there are at least $k-l$ orthogonal eigenvectors $|x\rangle$ of eigenvalue λ, for which $\sum\limits_{i=1}^{d} p_i |\langle i|x\rangle|^2 = 0$, so (a) will imply (b). W.l.o.g. assume $p_i > 0$ for $i = 1,\ldots,l$. For each matrix $W \in U(\mathbb{C}^k)$ the columns of isometry $I_{U,\lambda}W$ consist of eigenvectors of eigenvalue λ. We can choose such a matrix W for which first $k-l$ columns are orthogonal to each of vectors $I_{U,\lambda}^{\top}|1\rangle,\ldots,I_{U,\lambda}^{\top}|l\rangle$. One can note that for $|x\rangle \in \{I_{U,\lambda}W|1\rangle,\ldots,I_{U,\lambda}W|k-l\rangle\}$ we obtain $\sum\limits_{i=1}^{d} p_i |\langle i|x\rangle|^2 = 0$.

(c) Fix some eigenvalue λ with $r(\lambda) = k$. We introduce the notation of $\Pi = I_{U,\lambda}I_{U,\lambda}^{\dagger}$. We consider the subspace $\{|x\rangle : D_+|x\rangle = 0, \Pi|x\rangle = |x\rangle\}$ along with projection Π_0 on this subspace. Denote $\Pi_+ = \Pi - \Pi_0$. Let

$$c := \min_{|x\rangle : \Pi_+|x\rangle = |x\rangle, \||x\rangle\|_2 = 1} \langle x| D_+ |x\rangle. \tag{14}$$

One can note that $c > 0$. Take unit vector $|x\rangle$ and define $|v\rangle = (\mathbb{1}_d - \Pi)|x\rangle$. First of all, we will show that $\operatorname{tr}(\mathbb{1}_d - \Pi)|x\rangle\langle x| \in O(\epsilon)$ if $|\lambda - \langle x|U|x\rangle| = \epsilon$. Direct calculations reveal that

$$
\begin{aligned}
\epsilon = |\lambda - \langle x|U|x\rangle| &= |\lambda \operatorname{tr}|x\rangle\langle x| - \operatorname{tr} U\Pi|x\rangle\langle x| - \operatorname{tr} U(\mathbb{1}_d - \Pi)|x\rangle\langle x|\,| \\
&= |\lambda \operatorname{tr}(\mathbb{1}_d - \Pi)|x\rangle\langle x| - \operatorname{tr}(\mathbb{1}_d - \Pi)U(\mathbb{1}_d - \Pi)|x\rangle\langle x|\,|.
\end{aligned}
\tag{15}
$$

If $\|v\| = 0$ the statement is true, so assume $\|v\| \neq 0$. We obtain

$$\epsilon = |\lambda \langle v|v\rangle - \langle v|U|v\rangle| \geq \langle v|v\rangle \operatorname{dist}(\lambda, \operatorname{conv}(\lambda_{k+1},\ldots,\lambda_d)) > 0, \tag{16}$$

which finishes this part of proof.

In the second part we will check the behavior of points $\langle x|U|x\rangle$ in the neighborhood of point λ. We assume that $|\lambda - \langle x|U|x\rangle| = \epsilon$ for relatively small $\epsilon \geq 0$, so $\operatorname{tr}(\mathbb{1}_d - \Pi)|x\rangle\langle x| \in O(\epsilon)$. The derivative of trajectory of such a point is

$$\frac{\partial}{\partial t}(\langle x|UE(t)|x\rangle)(0) = i\langle x|UD_+|x\rangle. \tag{17}$$

We can rewrite the above as

$$
\begin{aligned}
i\langle x|UD_+|x\rangle = i\langle x|(\Pi_+ + \Pi_0 + \mathbb{1}_d - \Pi)UD_+(\Pi_+ + \Pi_0 + \mathbb{1}_d - \Pi)|x\rangle &= \\
= i\lambda\langle x|\Pi_+D_+\Pi_+|x\rangle + i\langle x|\Pi_+UD_+(\mathbb{1}_d - \Pi)|x\rangle& \\
+ i\langle x|(\mathbb{1}_d - \Pi)UD_+\Pi_+|x\rangle + i\langle x|(\mathbb{1}_d - \Pi)UD_+(\mathbb{1}_d - \Pi)|x\rangle).&
\end{aligned}
\tag{18}
$$

The above equation means that the instantaneous velocity of point $\langle x|U|x\rangle$ is the sum of the velocity

$$i\lambda\langle x|\Pi_+D_+\Pi_+|x\rangle \tag{19}$$

which is responsible for counterclockwise movement and which speed is $\langle x|\, \varPi_+ D_+ \varPi_+\, |x\rangle$ and the "noise" velocity which for the most pessimistic scenario can be rotated in any direction and which speed is at most

$$|\langle x|\, \varPi_+ U D_+ (\mathbb{1}_d - \varPi)\, |x\rangle\,|+|\langle x|\, (\mathbb{1}_d - \varPi) U D_+ \varPi_+\, |x\rangle\,|+|\langle x|\, (\mathbb{1}_d - \varPi) U D_+ (\mathbb{1}_d - \varPi)\, |x\rangle\,|. \tag{20}$$

The speed of velocity with direction $i\lambda$ can be lower bounded by

$$c\,\langle x|\, \varPi_+\, |x\rangle\,, \tag{21}$$

while the speed of the second velocity can be upper bounded by

$$2\sqrt{\langle x|\, \varPi_+\, |x\rangle}\sqrt{\langle x|\, (\mathbb{1}_d - \varPi)\, |x\rangle} + \langle x|\, (\mathbb{1}_d - \varPi)\, |x\rangle \le 2\sqrt{\langle x|\, \varPi_+\, |x\rangle}\sqrt{\epsilon} + \epsilon. \tag{22}$$

There exists constant $d > 0$ depending on the geometry of the numerical range of the matrix U, such that if

$$c\,\langle x|\, \varPi_+\, |x\rangle \ge d(2\sqrt{\langle x|\, \varPi_+\, |x\rangle}\sqrt{\epsilon} + \epsilon), \tag{23}$$

then the point $\langle x|\, U\, |x\rangle$ moves counterclockwise. This is true if

$$\langle x|\, \varPi_+\, |x\rangle \ge \frac{cd + 2d^2 + 2d\sqrt{d^2 + cd}}{c^2}\epsilon. \tag{24}$$

In the case when the above inequality does not hold, the speed of the second velocity is upper bounded by a some linear function of variable ϵ. That means there exists t_0 such that for $t \le t_0$ there can not exists eigenvalue $\widetilde{\lambda} \in UE(t)$ which $\widetilde{\lambda} \to \lambda$ as $t \to 0$ and $\widetilde{\lambda}$ is before λ in the counterclockwise order. To finish the proof we can see the above holds for any $t \ge 0$ due to fact that $E(t_0 + t) = E(t_0)E(t)$.

(d) To see this we first need to describe local dynamics of point $\beta(t) = \langle x_1|\, UE(t)\, |x_1\rangle = \lambda\,\langle x_1|\, E(t)\, |x_1\rangle$. One can note that $\beta(0) = \lambda$ and $\frac{\partial}{\partial t}\beta(0) = i\lambda \sum_{i=1}^{d} p_i|\,\langle i|x_1\rangle\,|^2$. That means $\beta(t) \approx \lambda \exp\!\left(it \sum_{i=1}^{d} p_i|\,\langle i|x_1\rangle\,|^2\right)$. To see that $\beta(t) \approx \lambda_1(t)$ we need to utilize the following facts:

- Eigenvalues of $UE(t)$ are continuous functions.
- $\beta(t) \in W(UE(t))$.
- Trajectory of $\beta(t)$ is curved in such a way that holds $\frac{1 - |\beta(t)|}{|\beta(t) - \lambda|} \to 0$.

The above means that if $R(t) \subset \{|z| = 1\}$ is an arch in which we can potentially find eigenvalue $\lambda_1(t)$ according to the fact that $\beta(t) \in W(UE(t))$, then it is true that $\frac{|R(t)|}{|\beta(t) - \lambda|} \to 0$ and consequently $\lambda_1(t) \approx \beta(t)$ for small $t \ge 0$.

(e) The see this point we need to utilize the postulate (c) along with the proof of the postulate (d) for eigenvectors $|x\rangle \in \{I_{U,\lambda}\,|v_1\rangle, \ldots, I_{U,\lambda}\,|v_k\rangle\}$, where $Q := I_{U,\lambda}^{\dagger} D_+ I_{U,\lambda}$ and $|v_j\rangle \in S_{\lambda_j(Q)}^{Q}$ for $j = 1, \ldots, k$.

(f) This relation follows from (e).