



Optimal approximation to unitary quantum operators with linear optics

Juan Carlos Garcia-Escartin¹ · Vicent Gimeno² ·
Julio José Moyano-Fernández³

Received: 6 January 2021 / Accepted: 10 September 2021 / Published online: 21 September 2021
© The Author(s) 2021

Abstract

Linear optical systems acting on photon number states produce many interesting evolutions, but cannot give all the allowed quantum operations on the input state. Using Toponogov's theorem from differential geometry, we propose an iterative method that, for any arbitrary quantum operator U acting on n photons in m modes, returns an operator \tilde{U} which can be implemented with linear optics. The approximation method is locally optimal and converges. The resulting operator \tilde{U} can be translated into an experimental optical setup using previous results.

Keywords Optical implementations of quantum evolutions · Quantum Optics · Toponogov's Theorem

1 Introduction

Linear optical devices under quantum light show a rich behaviour and have different applications in experiments on the foundations of quantum optics and quantum infor-

✉ Juan Carlos Garcia-Escartin
juagar@tel.uva.es

Vicent Gimeno
gimenov@uji.es

Julio José Moyano-Fernández
moyano@uji.es

¹ Departamento de Teoría de la Señal y Comunicaciones e Ingeniería Telemática. ETSI de Telecomunicación, Universidad de Valladolid Campus Miguel Delibes, Paseo Belén 15. 47011, Valladolid, Spain

² Universitat Jaume I, Campus de Riu Sec, Departament de Matemàtiques & Institut Universitari de Matemàtiques i Aplicacions de Castelló-IMAC, 12071 Castellón de la Plana, Spain

³ Universitat Jaume I, Campus de Riu Sec, Departamento de Matemáticas & Institut Universitari de Matemàtiques i Aplicacions de Castelló-IMAC, 12071 Castellón de la Plana, Spain

mation [1–3]. While they can be built with relatively simple optical elements like beam splitters and phase shifters [4–8], their behaviour for photon number states cannot be accurately reproduced by any classical system. One clear example is the boson sampling problem, for which quantum systems can give efficient solutions which cannot be produced by any classical method [9].

Passive lossless linear optical systems where the number of photons is preserved are particularly attractive: they are simple, well understood, and they can be translated to standard experimental setups.

There are many results on the synthesis of linear systems from their classical description [4,7,8] and that analyze the evolution of multiple photons in those devices [10–18]. However, there are relatively few methods for the design of tailored quantum evolutions for multiple photons. We have previously presented an inverse method which can tell if any desired quantum evolution on n photons can be achieved with a linear optical system or not, giving the corresponding system when it is possible [19].

In this paper, we complete this design method with a procedure that gives the best possible approximation to any quantum unitary that cannot be achieved using only photon preserving linear optical systems. The approximation starts from an initial guess and returns a matrix in its neighbourhood which minimizes the matrix distance to the desired unitary but can be achieved using linear optics alone. The result is based on Toponogov's comparison theorem [20] from differential geometry.

Section 2 introduces the mathematical description of linear optical systems and previously known results we will use. Section 3 describes the structure of the image algebra and its complement and states two theorems that will become useful later. Section 4 gives the basic concepts from differential geometry used in the proof and the notation for the rest of the paper. Section 5 defines the bi-invariant Riemannian metric in which the results are given. Section 6 shows how to apply Toponogov's theorem to reduce the problem of approximating a unitary to finding a geodesic in the correct manifold. Section 7 discusses some tricks related to the generation of random unitaries which are needed to explore the image group and to be able to compute a valid matrix logarithm for any desired U . Section 8 describes the iterative method that produces the desired approximation. Section 9 gives an application example. Finally, Sect. 10 gives a general overview of the method and comments on some practical problems and possible improvements for the approximation algorithm.

2 Mathematical description of linear optics interferometers

All through this paper we restrict to lossless linear optical systems where the total number of photons is preserved. Classically, the evolution of the electrical field in m orthogonal modes going through such a linear optical system is perfectly described by a unitary $m \times m$ matrix, S , called the *scattering matrix* of the system [21]. Linear optical systems including loss and amplification can be described using quasiunitary matrices [22].

The evolution of n photons distributed through these m possible modes is given by an $M \times M$ unitary *evolution matrix* U acting on the $M = \binom{m+n-1}{n}$ states of the resulting Hilbert space.

The *photonic homomorphism* $\varphi : U(m) \rightarrow U(M)$ gives the evolution matrix U which corresponds to a scattering matrix S describing the photon preserving linear optical system. $U = \varphi(S)$ can be computed from different equivalent methods [10–12].

Any unitary matrix can be written as an exponential $U = e^{iH}$ for a Hermitian matrix H . In linear optical devices, we will call this matrix the *effective Hamiltonian* H_U of the linear system. Similarly, $S = e^{iH_S}$.

The image of the photonic homomorphism, $\text{im}(\varphi)$, is a subgroup of $U(M)$ which contains all the quantum evolutions that are allowed for n photons a linear optical system with m modes. The image subgroup is a representation of $U(m)$ in $U(M)$ and maps each possible classical scattering matrix S describing a linear system into the quantum evolution $U = \varphi(S)$ it induces for n photons.

The evolution in the corresponding unitary algebras from iH_S to iH_U is given by the differential of φ , $d\varphi : \mathfrak{u}(m) \rightarrow \mathfrak{u}(M)$, for which there are also explicit expressions [13–18].

From the point of view of system design, a natural question is whether any given quantum evolution $U \in U(M)$ can be realized using only linear optics. From a simple dimensional argument, it is clear that, except when $m = 1$ or $n = 1$, there must be some impossible operations [23].

In a previous work, we have given an explicit inverse method to find the S corresponding to any $U \in \text{im}(\varphi)$ which can be implemented using linear optics [19].

Here, we address the problem of approximating $U \notin \text{im}(\varphi)$. We give a method to find the linear optics system with an evolution matrix $\tilde{U} \in \text{im}(\varphi)$ which minimizes the distance to U locally.

3 The image algebra and its orthogonal complement

If we study the induced map $d\varphi : \mathfrak{u}(m) \rightarrow \mathfrak{u}(M)$, we can decompose the Lie algebra $\mathfrak{u}(M)$ orthogonally so that

$$\mathfrak{u}(M) = \text{im } d\varphi \oplus (\text{im } d\varphi)^\perp, \tag{3.1}$$

where $(\text{im } d\varphi)^\perp$ is the orthogonal complement of $\text{im } d\varphi$ with respect to the metric

$$\langle u, v \rangle = \frac{1}{2} \text{tr}(u^\dagger v + v^\dagger u). \tag{3.2}$$

For this metric, we can prove a couple of useful facts.

Theorem 1 *For $U \in U(M)$ such that $U \notin \text{im } \varphi$, let $v \in \mathfrak{u}(M)$ be the principal logarithm of U . Let*

$$v = v_T + v_N \tag{3.3}$$

be the orthogonal decomposition of v , with a tangent component $v_T \in \text{im } d\varphi$ and a normal component $v_N \in (\text{im } d\varphi)^\perp$. Then,

1. $U_a = \exp(v_T) \in \text{im } \varphi$.
2. $\|U - U_a\| \leq \|v_N\|$.

Therefore, for any normalized $|\psi\rangle$ with $\langle\psi|\psi\rangle = 1$, we have

$$1 \geq |\langle U\psi | U_a\psi \rangle| \geq 1 - \frac{\|v_N\|^2}{2}. \quad (3.4)$$

The proof is given by introducing a bi-invariant metric and reducing the issue to a problem in plane geometry thanks to Toponogov's comparison theorem [20]. Later, with this theorem, we can give a recursive method to find a locally optimal approximation and show it converges.

4 Prerequisites and Notation

For the very basic notions in differential geometry, such as manifold, curve, tangent space, the reader is referred to the books of Do Carmo [24] or Sakai [25].

A Riemannian metric on a differentiable manifold M is a correspondence which associates to each point p on M an inner product $\langle \cdot, \cdot \rangle_p$ on the tangent space T_pM which varies differentiably in the sense that, for any pair of vector fields X and Y which are differentiable in a neighbourhood V of M , the function $\langle X, Y \rangle$ is differentiable on V . The metric with which a Riemannian manifold M is endowed may come from a distance. Given two points $p, q \in M$, the distance $d(p, q)$ between them is defined to be the infimum of the lengths of all curves joining p and q which are piecewise differentiable.

Two fundamental concepts of Riemannian geometry are those of geodesic and curvature. Roughly speaking, a geodesic is a curve minimizing the distance between two nearby points. More precisely, let I be a closed interval in \mathbb{R} ; a parametrized curve $\gamma : I \rightarrow M$ is called a geodesic at t_0 if the covariant derivative $\frac{D}{dt} \left(\frac{d\gamma}{dt} \right)$ vanishes at the point t_0 (see, i.e. [24], Definition 2.1); if γ is a geodesic at t for all $t \in I$, then γ is called a geodesic. If $[a, b] \subseteq I$ and $\gamma : I \rightarrow M$ is a geodesic, the restriction of γ to $[a, b]$ is called a *geodesic segment joining* $\gamma(a)$ to $\gamma(b)$. By abuse of language it is often referred to the image $\gamma(I)$ of a geodesic γ as a geodesic.

A *minimal geodesic* between p and q is the shortest one joining p and q . It is easily seen that if there exists a minimal geodesic γ joining p to q , then $d(p, q)$ equals the length $\ell(\gamma)$ of γ . This conditional *if* holds under the hypothesis of completeness: a Riemannian manifold M is said to be (geodesically) complete if for every $p \in M$ the exponential map \exp_p is defined for the whole tangent space T_pM , i.e. if any geodesic $\gamma(t)$ starting from p is defined for all $t \in \mathbb{R}$, and the statement is:

Theorem 2 (Hopf-Rinow) *Let M be a Riemannian manifold and let $p \in M$. Then M is geodesically complete if and only if it is complete as a metric space. Moreover, this implies that for any $q \in M$ there exists a minimizing geodesic γ joining p to q .*

An important consequence of Theorem 2 is the following, see [25], Corollary 1.4 and Problem 1 of Chapter III:

Corollary 1 *A C^∞ manifold M is compact if and only if any Riemannian metric on M is complete.*

On the other hand, the concept of curvature we will refer to is that of sectional curvature. According to Milnor [26] (p. 295), the sectional curvature of the tangential 2-plane spanned by some orthogonal unit vectors u and v can be described geometrically as the Gaussian curvature, at the point, of the surface swept out by all geodesics having a linear combination of u and v as tangent vector.

We are interested in Riemannian manifolds with additional algebraic structure: Lie groups. A Lie group is a group G with a differentiable structure such that the mapping $G \times G \rightarrow G$ given by $(x, y) \mapsto xy^{-1}$, $x, y \in G$, is differentiable. It follows that *translations from the left* L_x resp. *translations from the right* R_x given by $L_x : G \rightarrow G, L_x(y) = xy$ resp. $R_x : G \rightarrow G, R_x(y) = yx$ are diffeomorphisms.

A Riemannian metric on G is said to be *left invariant* resp. *right invariant* if for all $p, g \in G$ and for all $u, v \in T_pG$ it holds that

$$\langle u, v \rangle_p = \langle d(L_g)(u), d(L_g)(v) \rangle_{L_g(p)} \quad \text{resp.} \quad \langle u, v \rangle_p = \langle d(R_g)(u), d(R_g)(v) \rangle_{R_g(p)}.$$

A Riemannian metric is called *bi-invariant* if it is both left and right invariant. Any compact Lie group can be endowed with a bi-invariant metric [24, Exercise 7].

We also consider the Lie algebra \mathcal{G} of G , which consists of the vectors in T_eG with e the neutral element of G and with a well-known additional structure provided by a commutator (or Lie bracket) in the usual way.

We will focus on the Lie group $U(M)$ as a differentiable manifold, for a positive integer M ; its Lie algebra will be denoted by $\mathfrak{u}(M)$. Identity matrices of any size will be denoted by Id .

5 A bi-invariant Riemannian metric

In this section, we endow $U(M)$ with a Riemannian structure which will be useful later on.

For $u, v \in \mathfrak{u}(M)$, we define an inner product

$$\langle u, v \rangle := \frac{1}{2} \text{tr}(u^\dagger v + v^\dagger u). \tag{5.1}$$

This definition does actually correspond to a positive definite symmetric bilinear form: The bilinearity is an easy exercise, and moreover:

1. Since $u, v \in \mathfrak{u}(M)$, then $u^\dagger = -u, v^\dagger = -v$, and therefore

$$\langle u, v \rangle := \frac{1}{2} \text{tr}(-uv - vu) = -\text{tr}(uv).$$

2. The symmetry of (5.1) is clear, since $\langle u, v \rangle = -\text{tr}(uv) = -\text{tr}(vu) = \langle v, u \rangle$.

3. The positive definiteness follows from the fact that

$$\langle u, u \rangle = \text{tr}(u^\dagger u) = \|u\|^2 \geq 0,$$

where $\|u\| = \sqrt{\sum_{i,j} |u_{ij}|^2}$ is the Frobenius norm of u , which is nonnegative and has all the required norm properties [27].

The metric defined above is Riemannian and bi-invariant. Bi-invariant metrics are useful for us because of the following.

Theorem 3 (Milnor [26]) *Every compact Lie group admits a bi-invariant metric, which has nonnegative sectional curvature.*

In the case of a bi-invariant metric, the sectional curvature admits an easier formula, see [26, p. 323, Eqn. (7.3)]:

$$\kappa(u, v) = \frac{1}{4} \langle [u, v], [u, v] \rangle.$$

Furthermore, in a Lie group admitting a bi-invariant metric, geodesic curves have an easy description: they coincide with the exponential map. More precisely, for $p \in U(M)$, a geodesic curve $\gamma : I \rightarrow U(M)$ such that $\gamma(0) = p$ and $\dot{\gamma}(0) = u$ is of the form

$$\gamma(t) = \exp(up^{-1}t) \cdot p \tag{5.2}$$

In fact, for $p, q \in U(M)$, there exists a geodesic γ joining p and q with $\gamma(0) = p$ such that

$$\begin{aligned} \ell(\gamma([0, t])) &= \int_0^t \sqrt{\langle \dot{\gamma}(t), \dot{\gamma}(t) \rangle} dt \\ &= \int_0^t \sqrt{\langle u, u \rangle} dt = \int_0^t \|u\| dt \\ &= \|u\| t. \end{aligned} \tag{5.3}$$

A geodesic segment $\gamma : [0, 1] \rightarrow U(M)$ is called *minimal* if it realizes a distance for any $t \in [0, 1]$, i.e.,

$$d(\gamma(0), \gamma(t)) = \ell(\gamma([0, t])). \tag{5.4}$$

By equation (5.2) a geodesic segment $\gamma : [0, 1] \rightarrow U(M)$ joining $\gamma(0) = p$ and $\gamma(1) = q$ can be always obtained as

$$\gamma(t) = \exp(vt)p, \quad \text{with } v \text{ such that } \exp(v) = qp^{-1}.$$

The concept of a minimal geodesic is related to the concept of principal logarithm in the following way

Lemma 1 *Let $p, q \in U(M)$, let w be a principal logarithm of qp^{-1} . Then, the geodesic segment*

$$\gamma : [0, 1] \rightarrow U(M), \quad t \mapsto \gamma(t) = \exp(wt)p$$

is a minimal geodesic segment joining p and q .

Recall iK is called a *principal logarithm* of a unitary matrix $M \in U(M)$ if

$$K^\dagger = K, \quad \exp(iK) = M, \quad \text{and the eigenvalues of } K \text{ are in } (-\pi, \pi].$$

There are efficient algorithms that can compute the principal logarithm of a unitary matrix [28]. We choose this definition of principal logarithm over the one for the interval $(-\pi, \pi)$ so that there is always a principal matrix logarithm, even for matrices with real negative eigenvalues (-1). Some properties, like infinite differentiability, are lost with this definition, but they are not used in our result.

Proof of Lemma 1 The length of a geodesic segment $\gamma(t) = \exp(vt)p$ with $\exp(v) = qp^{-1}$ is (by equation (5.3))

$$\ell(\gamma([0, 1])) = \|vp\| = \|v\|.$$

The distance $d(p, q)$ between p and q is the shortest length of the curves joining p and q . In the case of a complete metric this shortest length is attained by a geodesic segment joining p and q . Hence, the geodesic segment $\gamma(t) = \exp(vt)p$ is minimal if and only if

$$\|v\| = \min\{\|w\| : \exp(w) = qp^{-1}\}.$$

But the equation $\exp(w) = qp^{-1}$ has the following family of solutions

$$w = U[\log(\lambda_i)\delta_{ij}]U^\dagger$$

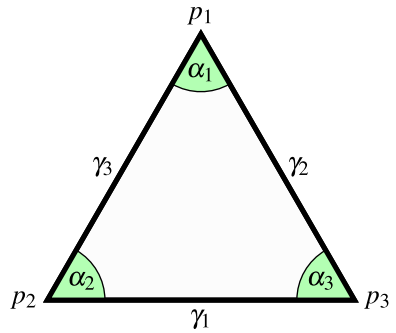
with $qp^{-1} = U\Lambda U^\dagger$, U being a unitary matrix, $[\Lambda]_{ij} = (\lambda_i)\delta_{ij}$ being the diagonal matrix of eigenvalues of qp^{-1} , and $\log(\lambda_i)$ being any logarithm of λ_i . Observe that

$$\|w\| = \sqrt{\sum_{i=1}^M |\log(\lambda_i)|^2}.$$

This expression only depends on the list of eigenvalues of qp^{-1} and their logarithms. Since qp^{-1} is unitary

$$\log(\lambda_i) = i(k_i + 2\pi l_i), \quad \text{with } k_i \in (-\pi, \pi], \text{ and } l_i \in \mathbb{Z}.$$

Fig. 1 Geodesic triangle $\Delta(p_1 p_2 p_3)$



Then

$$\{\|w\| : \exp(w) = qp^{-1}\} = \left\{ \sqrt{\sum_{i=1}^M |(k_i + 2\pi l_i)|^2} : l_i \in \mathbb{Z} \right\},$$

with a minimum when $l_i = 0$ for $i = 1, \dots, M$, which corresponds to a principal logarithm. □

6 An application of Toponogov’s comparison theorem

Riemannian manifolds whose curvature is bounded below may be investigated by applying Toponogov’s comparison theorem. We first need to define triangles on the Riemannian manifold.

Definition 1 A geodesic triangle $T = \Delta(p_1 p_2 p_3)$ of a Riemannian manifold M is a set consisting of three segments of minimal geodesics, which are called the sides of T , say

$$\gamma_1 : [0, 1] \rightarrow M, \gamma_2 : [0, 1] \rightarrow M, \text{ and } \gamma_3 : [0, 1] \rightarrow M,$$

such that $\gamma_i(1) = \gamma_{i+1}(0)$ for $i = 1, 2$, and $\gamma_3(1) = \gamma_1(0)$. The endpoints p_1, p_2 and p_3 are called the vertices of the triangle. The angle between the tangent vectors to γ_{i-1} and γ_{i+1}^{-1} at p_i is called the angle of T at p_i and denoted by $\alpha_i = \angle(p_{i-1} p_i p_{i+1})$ or $\angle p_i$. The perimeter ℓ is defined as the sum $\ell(\gamma_1) + \ell(\gamma_2) + \ell(\gamma_3)$; if we consider, in addition, that the two sides γ_2, γ_3 are minimal geodesics, and the side γ_1 is a geodesic segment, not necessarily minimal, with $\ell(\gamma_1) \leq \ell(\gamma_2) + \ell(\gamma_3) = d(p_1, p_3) + d(p_1, p_2)$, then the set is said to be a *generalized geodesic triangle* (Fig. 1).

Let us set the part of Toponogov’s comparison theorem we are interested in, cf. [25, Theorem 4.2 in Chapter IV]:

Theorem 4 (Toponogov) *Let M be a complete Riemannian manifold whose sectional curvatures satisfy $\kappa \geq \delta$ everywhere for some constant δ . Denote by M_δ^2 the 2-dimensional complete simply connected Riemannian manifold of constant curvature*

δ . Consider a generalized geodesic triangle $\Delta(p_1 p_2 p_3)$ such that γ_2, γ_3 are minimal and $\ell(\gamma_1) \leq \pi/\sqrt{\delta}$. Then the perimeter $\ell \leq 2\pi/\sqrt{\delta}$ and there exists a geodesic triangle $\Delta(\tilde{p}_1 \tilde{p}_2 \tilde{p}_3)$ in M_δ^2 with the same side lengths $\ell(\tilde{\gamma}_i) = \ell(\gamma_i)$, for $i = 1, 2, 3$ and satisfying $\alpha_2 \geq \tilde{\alpha}_2$ and $\alpha_3 \geq \tilde{\alpha}_3$.

Remark It is to assume that $\pi/\sqrt{\delta} = +\infty$ when $\delta \leq 0$ in the theorem above.

Theorem 4 allows us to compare triangles of $U(M)$ with triangles in \mathbb{R}^2 by setting $\delta = 0$, since in this case $M_0^2 = \mathbb{R}^2$. In our situation it is $p_1 = U, p_3 = Id$ and $p_2 = U_a = \exp(v_T)$ is our approximation matrix in $\text{im}(\varphi)$ (recall that v_T is the tangential component of the principal logarithm v of U). Set $\gamma_1(t) = U_a \exp(-v_T t), \gamma_2(t) = \exp(vt)$ and $\gamma_3(t)$ a minimal geodesic segment joining U with U_a . Then γ_2 and γ_3 are minimal geodesic segments and γ_1 is a geodesic segment (not necessarily minimal). Set $\ell_1 = \ell(\gamma_1([0, 1])) = \|v_T\|, \ell_2 = \ell(\gamma_2([0, 1])) = d(U, Id)$ and $\ell_3 = \ell(\gamma_3([0, 1])) = d(U, U_a)$. Observe that

$$\ell_1 = \|v_T\| \leq \|v\| + d(U, U_a) = \ell_2 + \ell_3.$$

Hence, by Theorem 4, there exists a geodesic triangle in \mathbb{R}^2 with sides ℓ_1, ℓ_2 and ℓ_3 of lengths ℓ_1, ℓ_2 resp. ℓ_3 , such that $\alpha_2 \geq \tilde{\alpha}_2$ and $\alpha_3 \geq \tilde{\alpha}_3$. We want to estimate the distance ℓ_3 . First of all notice that $\ell_3 = \ell_2 - \ell_1$; hence, the law of cosinus implies that

$$\|\ell_3\|^2 = \|\ell_2 - \ell_1\|^2 = \|\ell_1\|^2 + \|\ell_2\|^2 - 2\|\ell_1\| \|\ell_2\| \cos \angle(\ell_1, \ell_2)$$

and

$$\ell_3^2 = \ell_1^2 + \ell_2^2 - 2\ell_1 \ell_2 \cos \tilde{\alpha}_3. \tag{6.1}$$

Since $\cos \alpha_3 = \frac{\langle v, v_T \rangle}{\|v\| \|v_T\|} = \frac{\|v_T\|}{\|v\|} \geq 0$, it follows that $\tilde{\alpha}_3 \leq \alpha_3 \leq \frac{\pi}{2}$ which implies $-\cos \tilde{\alpha}_3 \leq -\cos \alpha_3$ and, so,

$$\begin{aligned} \ell_3^2 &\leq \ell_1^2 + \ell_2^2 - 2\ell_1 \ell_2 \cos \alpha_3 = \|v_T\|^2 + \|v\|^2 - 2\|v_T\| \|v\| \cos \alpha_3 \\ &= \|v\|^2 + \|v_T\|^2 - 2\langle v, v_T \rangle = \|v - v_T\|^2 = \|v_N\|^2. \end{aligned} \tag{6.2}$$

Figure 2 shows a graphical representation of our scenario.

Now we want to see the bi-invariant metric in $u(M)$ defined in Sect. 5 as a metric in $U(M)$. First observe that the Riemannian manifold $(U(M), \langle, \rangle)$ is a Riemannian subvariety of the manifold $\mathcal{M}_n(\mathbb{C})$ of the complex $n \times n$ -matrices endowed with the Euclidean inner product. If we write $d_{U(M)}$ for the distance we use on $U(M)$, and $d_{\mathcal{M}_n}$ for the one on $\mathcal{M}_n(\mathbb{C})$, then

$$\ell_3 \geq d_{U(M)}(U_a, U) \geq d_{\mathcal{M}_n}(U_a, U).$$

The inequality holds since “distance” between two points is the infimum of the length of any two curves joining the points. Now it is easy to find a minimal geodesic for

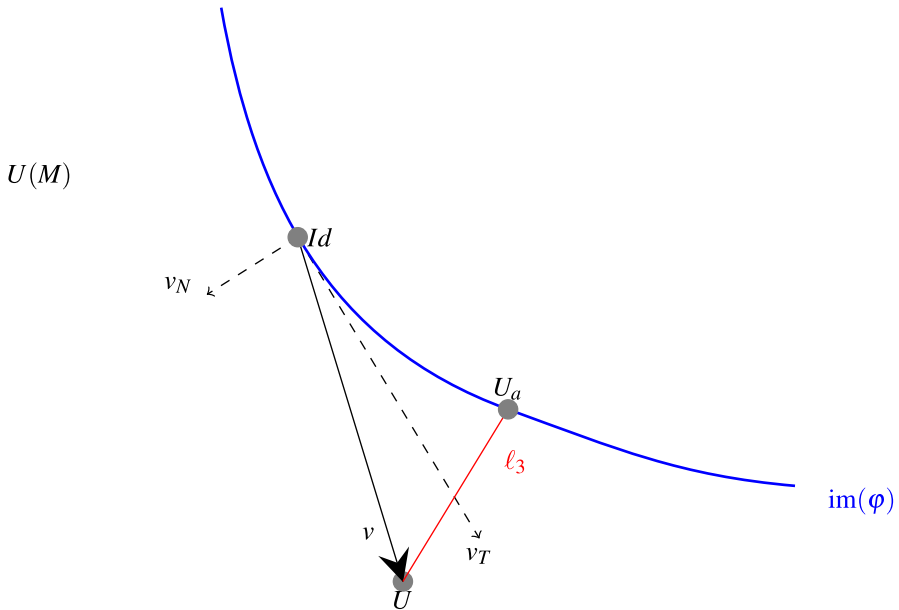


Fig. 2 Submanifold $\text{im}(\varphi)$ in $U(M)$, matrix U and approximation U_a

$\mathcal{M}_n(\mathbb{C})$ joining U_a and U , namely the segment $\gamma_4(t) = U_a + (U - U_a)t$ for $t \in [0, 1]$. Since $\dot{\gamma}_4(t) = U - U_a$, we find that $d_{\mathcal{M}_n}(U_a, U) = \|\dot{\gamma}_4(t)\| \cdot 1 = \|U - U_a\|$ and, therefore,

$$\ell_3 \geq \|U - U_a\| .$$

This inequality together with (6.2) yields

$$\|U - U_a\|^2 \leq \ell_3^2 \leq \|v_N\|^2 ,$$

and we obtain, finally,

$$\|U - U_a\| \leq \|v_N\| .$$

7 Random unitary matrices and a basis for the image

Before looking into the iterative process that gives the locally optimal approximation to any desired $U \notin \text{im}\varphi$, we need to consider a few useful tricks.

While in the previous section we have considered the identity matrix as our starting point in the image group, the results hold for any $U_0 \in \text{im}\varphi$. The identity has some advantages: it is always in the image, for any values of n and m , and, in terms of computation, it is trivial to generate the $M \times M$ identity matrix.

However, in general, the landscape of the image group is unknown and numerical experiments show there are multiple local minima. In a 2D space, we can picture the image as an irregular profile with mountains and valleys. For any given starting point, the iterative process ends in a local minimum, but, as we do not know how many minima exist, we need a series of random starting points to sample multiple approximations, each the closest to the intended U in its local neighbourhood.

In order to generate a random unitary $U_r \in \text{im}\varphi$, we choose a random $S_r \in U(m)$ uniformly from all the possible matrices and compute $U_r = \varphi(S_r)$. The random unitaries in $U(m)$ can be generated from samples of a normal distribution [29,30], and the photonic homomorphism is known [11,12].

Finally, we need a way to project the logarithm of any $U \in U(M)$ to the image algebra $\text{im}d\varphi$. The method is similar to the generation of random unitaries. We start by working on the algebra corresponding to the scattering matrices, $\mathfrak{u}(m)$, where we can write down a known basis and, using the differential map $d\varphi$, explicitly given in [18], we can obtain a basis for $\text{im}d\varphi$, which is a subalgebra of $\mathfrak{u}(M)$. This basis is not necessarily orthonormal, but it can be orthogonalized and normalized. The details of the whole procedure can be found in [19]. This basis, together with the inner product of Eq. (5.1), is enough to obtain the desired projection on the image algebra.

8 An iterative process for the approximation

In Sect. 6 we have constructed an approximation matrix $U_1 := U_a \in \text{im}(\varphi)$, starting from $U = \exp(v)$ as

$$U_1 = \exp((\log U)_T), \quad \text{with} \quad \|U - U_1\| \leq \|(\log U)_N\|.$$

Now, we can repeat this by taking a new approximation $U_2 \in \text{im}(\varphi)$ by considering a geodesic triangle of vertices Id , $U_1^{-1}U$ and U_2 , where $U_1^{-1}U = \exp(-v_T) \exp v$ and

$$U_2 = U_1 \exp((\log U_1^{-1}U)_T) \quad \text{with} \quad \|U - U_2\| \leq \|(\log U_1^{-1}U)_N\|.$$

Iteratively,

$$\begin{cases} U_0 = Id \\ U_n = U_{n-1} \exp((\log(U_{n-1}^{-1}U)_T)) \end{cases}$$

with

$$\|U - U_n\| \leq \|(\log(U_{n-1}^{-1}U)_N)\|.$$

8.1 Convergence

The method described above converges. Consider the sequence $\{U_i\}$ with $i \geq 0$ and $U_0 = Id$. For the first step we know that $d(U_1, U) = \|v^1\|$, $d(U_2, U) \leq \|v_N^1\|$, and $d(U_1, U_2) \leq \|v_T^1\|$. In general, we have:

1. $d(U_i, U) = \|v^i\|$,
2. $d(U_{i+1}, U) \leq \|v_N^i\|$,
3. $d(U_i, U_{i+1}) \leq \|v_T^i\|$.

Proposition 1 *We have that*

$$d(U_{i+1}, U) \leq d(U_i, U). \tag{8.1}$$

Furthermore, the equality holds if and only if $U_i = U_{i+1}$.

Proof By a successive application of inequalities (2) and (1) above we obtain

$$d(U_{i+1}, U) \stackrel{(2)}{\leq} \|v_N^i\| \leq \|v^i\| \stackrel{(1)}{=} d(U_i, U).$$

Now, if equality (8.1) holds, then $\|v_N^i\| = \|v^i\|$ and so $\|v_T^i\| = 0$; inequality (3) above allows us to conclude that $U_i = U_{i+1}$. The converse is trivial. \square

Let us define

$$d : \mathbb{N} \rightarrow \mathbb{R}, \quad i \mapsto d_i := d(U_i, U).$$

Proposition 2 *The sequence $\{d_i\}$ is convergent.*

Proof Assume that $U_i \neq U_{i+1}$ for every i (otherwise, there would exist $n \in \mathbb{N}$ such that $U_n = U_{n+1}$, and therefore $d_n = d_{n+1} = \dots$ and the sequence converges to d_n). By Proposition 1 the sequence $\{d_i\}$ is decreasing; in particular $d_i < d_1$ for all i . This together with the fact that $d_i \geq 0$ for all i (i.e. the sequence $\{d_i\}$ is bounded) implies the convergence. \square

In fact, the approximation given by the described method is the best possible one in the following sense. Since $\{d_i\}$ is a convergent sequence, then it is a Cauchy sequence. This means that for every $\epsilon > 0$ there exists $n_\epsilon \in \mathbb{N}$ with

$$0 < d_i - d_{i+1} < \epsilon, \quad \text{for all } i > n_\epsilon.$$

Since, by (1) and (2) we have that $d_i = \|v^i\|$ and $d_{i+1} \leq \|v_N^i\|$, it is easily deduced that

$$0 \leq \|v^i\| - \|v_N^i\| \leq d_i - d_{i+1} < \epsilon. \tag{8.2}$$

We first observe the following inequality:

Lemma 2

$$\|v^i - v_N^i\|^2 \leq (\|v^i\| - \|v_N^i\|) \cdot 2 \cdot d_1.$$

Proof

$$\begin{aligned} \|v^i - v_N^i\|^2 &= \|v^i\|^2 + \|v_N^i\|^2 - 2\langle v^i, v_N^i \rangle \\ &= \|v^i\|^2 + \|v_N^i\|^2 - 2\langle v_T^i + v_N^i, v_N^i \rangle \\ &= \|v^i\|^2 + \|v_N^i\|^2 - 2\|v_N^i\|^2 = \|v^i\|^2 - \|v_N^i\|^2 \\ &= (\|v^i\| - \|v_N^i\|)(\|v^i\| + \|v_N^i\|). \end{aligned}$$

Since $\|v_N^i\| \leq \|v^i\|$, it holds that

$$\begin{aligned} (\|v^i\| - \|v_N^i\|)(\|v^i\| + \|v_N^i\|) &\leq (\|v^i\| - \|v_N^i\|)2\|v^i\| \\ &\stackrel{(1)}{=} (\|v^i\| - \|v_N^i\|)2d_1. \end{aligned}$$

Proposition 1 implies that $d_i \leq d_1$; hence, we get

$$\|v^i - v_N^i\|^2 \leq (\|v^i\| - \|v_N^i\|)2d_1. \quad \square$$

Now, inequality (8.2) together with Lemma 2 implies that

$$0 \leq \|v^i - v_N^i\|^2 \leq 2\epsilon d_1.$$

On the other hand, $\|v^i - v_N^i\|^2 = \|v_T^i\|^2$ and, together with inequality (3), this yields

$$d(U_i, U_{i+1}) \leq \|v^i - v_N^i\|^2 \leq 2\epsilon d_1.$$

Therefore the sequence $\{U_i\}$ is a Cauchy sequence itself. This allows us to apply Theorem 2 (Hopf-Rinow) in the complete metric space (M, g) , and so the sequence $\{U_i\}$ converges to a matrix $\tilde{U} \in M$, and $\|v_\infty^T\| = 0$. Hence the geodesic joining \tilde{U} with U is normal.

9 Application example

We can see an example of the procedure with the quantum Fourier transform

$$QFT |x\rangle = \frac{1}{\sqrt{M}} \sum_{y=0}^{M-1} e^{\frac{i2\pi xy}{M}} |y\rangle. \tag{9.1}$$

We choose the QFT not only for being a useful transformation in quantum information and quantum optics, but also because it can be considered one of the most difficult transformations for linear optics.

We start from the QFT matrix for $M = 3$ ($n = m = 2$)

$$\begin{aligned}
 U &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \\ 1 & e^{-i\frac{2\pi}{3}} & e^{-i\frac{4\pi}{3}} \\ 1 & e^{-i\frac{4\pi}{3}} & e^{-i\frac{8\pi}{3}} \end{pmatrix} \\
 &\approx \begin{pmatrix} 0.57735 & 0.57735 & 0.57735 \\ 0.57735 & -0.28868 - 0.5i & -0.28868 + 0.5i \\ 0.57735 & -0.28868 + 0.5i & -0.28868 - 0.5i \end{pmatrix}.
 \end{aligned}
 \tag{9.2}$$

This evolution is impossible to achieve with lossless linear optics alone, which can be checked with the method in [19].

For the implementation of U , we consider the states in the ordered basis $\{|20\rangle, |02\rangle, |11\rangle\}$. All the results are given with 5 significant digits for matrix entries and 10 significant digits for distances, rounding the imaginary or real parts to zero when they are much smaller than the surrounding terms.

The starting point is the identity matrix, for which $\|U - \text{Id}\| = 2.449489743$. In the first step of our procedure, we take the projection of the principal logarithm of U into the image algebra:

$$\log(U) = \begin{pmatrix} -0.6639i & 0.9069i & 0.9069i \\ 0.9069i & -2.0242i & -0.45345i \\ 0.9069i & -0.45345i & -2.0242i \end{pmatrix}$$

and

$$\log(U)_T = \begin{pmatrix} -0.89062i & 0 & 0.22672i \\ 0 & -2.251i & 0.22672i \\ 0.22672i & 0.22672i & -1.5708i \end{pmatrix}$$

so that

$$U_1 = \begin{pmatrix} 0.61786 - 0.75486i & 0.024514i & 0.20595 + 0.073541i \\ 0.024514i & -0.61786 - 0.75486i & 0.20595 - 0.073541i \\ 0.20595 + 0.073541i & 0.20595 - 0.073541i & -0.95097i \end{pmatrix}$$

with $\|U - U_1\| = 1.770101749$.

After 10 steps, we have:

$$U_{10} = \begin{pmatrix} 0.86432 - 0.50294i & 2.889 \cdot 10^{-6}i & 0.0020837 + 0.0011983i \\ 2.889 \cdot 10^{-6}i & -0.86432 - 0.50294i & 0.0020837 - 0.0011983i \\ 0.0020837 + 0.0011983i & 0.0020837 - 0.0011983i & -0.99999i \end{pmatrix}$$

with $\|U - U_{10}\| = 1.732054756$. Further iterations only produce marginal improvements in the distance to the target matrix beyond the fifth decimal place.

We stop after 10 additional steps and use as an approximation

$$U_{20} = \begin{pmatrix} 0.86601 - 0.50003i & 0 & 0 \\ 0 & -0.86601 - 0.50003i & 0 \\ 0i & 0 & -1.0i \end{pmatrix}$$

with $\|U - U_{20}\| = 1.732050808$.

9.1 Random initial matrices

In fact, the results hold for any arbitrary $U_0 \in \text{im}\varphi$. Now the procedure becomes computationally more involved, as we need an explicit evaluation of $U_r = \varphi(S_r)$ for random S_r matrices, with a complexity which grows combinatorially in n and m . However, as we see in this second example, we can explore different local optima.

For the QFT matrix in Eq. (9.2), if we start at random points, the approximation gravitates towards three solutions. Two of them are at the same distance from the QFT matrix than the approximation we found starting with the identity matrix, 1.7320, with

$$U_a^1 = \begin{pmatrix} 0.86602 - 0.5i & 0 & 0 \\ 0 & -0.86602 - 0.5i & 0 \\ 0 & 0 & -1.0i \end{pmatrix}$$

and

$$U_a^2 = \begin{pmatrix} 0 & 0.86602 + 0.5i & 0 \\ 0.86602 + 0.5i & 0 & 0 \\ 0 & 0 & -0.86602 - 0.5i \end{pmatrix}.$$

There is a third solution with a distance 0.85675 to the QFT for the matrix:

$$U_a^3 = \begin{pmatrix} 0.43301 + 0.25i & 0.43301 - 0.25i & 0.70711 \\ 0.43301 - 0.25i & -0.5i & -0.35355 + 0.61237i \\ 0.70711 & -0.35355 + 0.61237i & 0 \end{pmatrix}.$$

These solutions seem to be found with equal probability. For a run of 1000 experiments we found U_a^1 311 times, U_a^2 374 times and U_a^3 315 times. Further experiments showed a similar behaviour.

The best approximation to the 3×3 QFT matrix U_a^3 comes from a scattering matrix:

$$S_a^3 = \begin{pmatrix} 0.68301 + 0.18301i & 0.68301 - 0.18301i \\ 0.68301 - 0.18301i & -0.5 + 0.5i \end{pmatrix}.$$

In general, depending on the starting local point, there are different local optima. Sometimes, two different approximation matrices in the image group give the same distance to the target unitary, as it happens in this example for U_a^1 and U_a^2 .

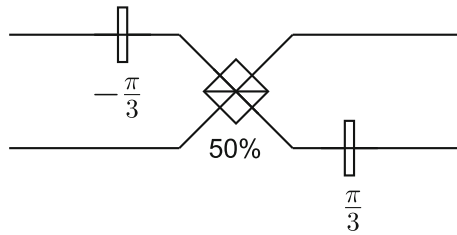


Fig. 3 Linear optics interferometer approximating the QFT for a system with two inputs and two photons. The system uses two phase shifters (with shifts $-\frac{\pi}{3}$ and $\frac{\pi}{3}$ in the upper and lower mode, respectively) and one balanced beam splitter (centre). This configuration gives one possible approximation, U_a^3 , which has matrix distance 0.85675 to the QFT

9.2 Translation to an optical setup

The best approximation to the QFT comes from a scattering matrix S_a^3 which corresponds, up to a $\frac{5\pi}{12}$ global phase, to a balanced two input beam splitter with a scattering matrix

$$S_{BS} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$$

preceded by a $\frac{-\pi}{3}$ phase shifter in the first port and followed by a $\frac{\pi}{3}$ in the second port. The phase shifters have scattering matrices

$$S_{-\frac{\pi}{3}} = \begin{pmatrix} e^{-i\frac{\pi}{3}} & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad S_{\frac{\pi}{3}} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{3}} \end{pmatrix},$$

respectively.

The experimental configuration corresponding to this approximation to the QFT is shown in Fig. 3.

10 Summary, recommendations and future improvements

We have given an iterative method that finds a linear optical setup that approximates any arbitrary quantum evolution for n photons in m modes. The approximation is optimal in the local neighbourhood of the initial guess. Once we have the closest unitary that can be implemented, \tilde{U} , we can use a previous method [19] to obtain a scattering matrix \tilde{S} that gives the approximated evolution and there are multiple algorithms that give the physical setup corresponding to \tilde{S} using only beam splitters and phase shifters [4,7].

This gives a design method which starts from the desired evolution instead of the usual analysis methods which, given a classical description of the linear optical system, find out the evolution it induces on quantum states.

The proposed algorithm is based on results from differential geometry, in particular, Toponogov’s theorem. They show the method will converge and find local optima.

There are a few practical details worth mentioning. First, numerically, we find that if we try the method on an evolution which is possible to obtain from a linear optics system, $U \in \text{im}\varphi$, sometimes it will converge to a matrix close the actual solution, but, depending on the local landscape, it might fall into a variety of different matrices all at a similar large distance. In practical applications, the recommendation would be, first, check whether an exact implementation exists (with our previous algorithm [19]) and, if there is none, look for an approximation.

There are also some open problems. The presented results are only valid for lossless linear systems. Linear optical systems including losses and squeezing can be described using quasiunitary matrices [13,22], which can also be studied using group theory. However, there appear some technical complications and the theory of these linear system is not as developed as in the lossless case. We leave for future research the adaptation of the presented design methods to these scenarios.

Additionally, from the structure of the involved groups and algebras, it is not clear how many local optima exist for any given transformation φ . We have proposed a randomized way of exploring the state space of the potential approximation matrices, and, for the limited dimensions that can be numerically explored, it seems to work well, but there is no guarantee the method finds a global minimum. Any further knowledge of the structure of the image group would help in the search for a global minimum or, at least, in finding a probabilistic bound on the optimal approximation. Additionally, a lower bound on $\|U - U_a\|$, as opposed to our upper norm, would help to determine the global optimum.

Even in its present form, the proposed algorithm, when combined with previous results, can assist in the design of quantum optical operations and has applications to quantum information and quantum optics experiment design.

Acknowledgements The first author has been funded by the Spanish Government and FEDER grant PID2020-119418GB-I00 (MICINN) and Junta de Castilla y León (project VA296P18). The second author has been partially supported by the Research Program of the University Jaume I—Project UJI-B2018-35, as well as by the Spanish Government and FEDER grants PID2020-115930GA-I00 (MICINN) and MTM2017-84851-C2-2 (MINECO). The third author was partially supported by the Spanish Government, Ministerios de Ciencia e Innovación y de Universidades, grant PGC2018-096446-B-C22, as well as by Universitat Jaume I, grant UJI-B2018-10.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Loudon, R.: The Quantum Theory of Light, 3rd edn. Oxford University Press, Great Clarendon Street, Oxford, UK (2000)

2. Gerry, C., Knight, P.: *Introductory quantum optics*. Cambridge University Press, Cambridge (2004)
3. Leonhardt, U.: *Essential quantum optics: from quantum measurements to black holes*. Cambridge University Press, Cambridge (2010)
4. Reck, M., Zeilinger, A., Bernstein, H.J., Bertani, P.: Experimental realization of any discrete unitary operator. *Phys. Rev. Lett.* **73**(1), 58 (1994)
5. Bouland, A., Aaronson, S.: Generation of universal linear optics by any beam splitter. *Phys. Rev. A* **89**, 062316 (2014)
6. Sawicki, A.: Universality of beamsplitters. *Quantum Inf. Comput.* **16**(3 & 4), 0291–0312 (2016)
7. Clements, W.R., Humphreys, P.C., Metcalf, B.J., Kolthammer, W.S., Walmsley, I.A.: Optimal design for universal multiport interferometers. *Optica* **3**(12), 1460–1465 (2016)
8. de Guise, H., Di Matteo, O., Sánchez-Soto, L.L.: Simple factorization of unitary transformations. *Phys. Rev. A* **97**, 022328 (2018)
9. Aaronson, S., Arkhipov, A.: The computational complexity of linear optics. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, STOC '11, pp. 333–342, New York, NY, USA, (2011). ACM
10. Caianiello, E.R.: On quantum field theory – I: Explicit solution of Dysons equation in electrodynamics without use of Feynman graphs. *Il Nuovo Cimento* (1943-1954) **10**(12), 1634–1652 (1953)
11. Scheel, S.: Permanents in linear optical networks. *quant-ph/0406127*, (2004)
12. Skaar, J., García Escartin, J.C., Landro, H.: Quantum mechanical description of linear optics. *Am. J. Phys.* **72**(11), 1385–1391 (2004)
13. Leonhardt, U., Neumaier, A.: Explicit effective hamiltonians for general linear quantum-optical networks. *J. Opt. B: Quantum Semiclass. Opt.* **6**(1), L1 (2004)
14. Fan, H.Y., Xiao, M.: Construction of optical networks by virtue of the IWOP technique. *Quantum Semiclass. Opt.: J. Eur. Opt. Soc. Part B* **9**(1), 53–58 (1997)
15. Aniello, P., Cagli, R.C.: An algebraic approach to linear-optical schemes for deterministic quantum computing. *J. Opt. B: Quantum Semiclass. Opt.* **7**(12), S711–S720 (2005)
16. Biedenharn, L.C., Louck, J.D.: *Angular momentum in quantum physics. Theory and application*, volume 8 of *Encycl. Math. Appl.* Addison-Wesley, (1981)
17. Aniello, P., Lupo, C., Napolitano, M.: Exploring representation theory of unitary groups via linear optical passive devices. *Open Syst. Inf. Dyn.* **13**(4), 415–426 (2006)
18. Garcia-Escartin, J.C., Gimeno, V., Moyano-Fernández, J.J.: Multiple photon effective Hamiltonians in linear quantum optical networks. *Opt. Commun.* **430**, 434–439 (2019)
19. Garcia-Escartin, J.C., Gimeno, V., Moyano-Fernández, J.J.: Method to determine which quantum operations can be realized with linear optics with a constructive implementation recipe. *Phys. Rev. A* **100**, 022301 (2019)
20. Cheeger, J., Ebin, D. G.: *Comparison theorems in Riemannian geometry*, volume 365. American Mathematical Soc., (2008)
21. Pozar, D.M.: *Microwave Engineering*, 4th edn. Wiley, New York (2004)
22. Tischler, N., Rockstuhl, C., Slowik, K.: Quantum optical realization of arbitrary linear transformations allowing for loss and gain. *Phys. Rev. X* **8**, 021017 (2018)
23. Moyano-Fernández, J.J., Garcia-Escartin, J.C.: Linear optics only allows every possible quantum operation for one photon or one port. *Opt. Commun.* **382**, 237–240 (2017)
24. Carmo, M.P. do.: *Riemannian geometry*. Birkhäuser, (1992)
25. Sakai, T.: *Riemannian geometry*, volume 149. American Mathematical Society, (1996)
26. Milnor, J.: Curvatures of left invariant metrics on Lie groups. *Adv. Math.* **21**(3), 293–329 (1976)
27. Horn, R.A., Johnson, C.R.: *Matrix analysis*. Cambridge University Press, Cambridge (1985)
28. Loring, T.A.: Computing a logarithm of a unitary matrix with general spectrum. *Numer. Linear Algebra Appl.* **21**(6), 744–760 (2014)
29. Mezzadri, F.: How to generate random matrices from the classical compact groups. *Not. Am. Math. Soc.* **54**(5), 592–604 (2007)
30. Tóth, G.: Qubit4Matlab v3.0: A program package for quantum information science and quantum optics for Matlab. *Comput. Phys. Commun.* **179**(6), 430–437 (2008)