Efficient algorithm for multiqudit twirling for ensemble quantum computation

Géza Tóth[1,*] and Juan José García-Ripoll[2]
[1]*Research Institute for Solid State Physics and Optics, Hungarian Academy of Sciences, P. O. Box 49, H-1525 Budapest, Hungary*
[2]*Max Planck Institute for Quantum Optics, Hans-Kopfermann-Straße 1, D-85748 Garching, Germany*
(Received 27 September 2006; published 11 April 2007; corrected 12 April 2007)

We present an efficient algorithm for twirling a multiqudit quantum state. The algorithm can be used for approximating the twirling operation in an ensemble of physical systems in which the systems cannot be individually accessed. It can also be used for computing the twirled density matrix on a classical computer. The method is based on a simple nonunitary operation involving a random unitary. When applying this basic building block iteratively, the mean squared error of the approximation decays exponentially. In contrast, when averaging over random unitary matrices the error decreases only algebraically. We present evidence that the unitaries in our algorithm can come from a very imperfect random source or can even be chosen deterministically from a set of cyclically alternating matrices. Based on these ideas we present a quantum circuit realizing twirling efficiently.

DOI: 10.1103/PhysRevA.75.042311 PACS number(s): 03.67.Lx, 02.70.−c

I. INTRODUCTION

Twirling was first introduced for bipartite systems in Refs. [1,2] in the context of entanglement purification and still appears as part of various quantum-information-processing protocols [3–5]. For example, in the single-party case, twirling makes it possible to obtain the average gate fidelity of a positive map [6]. Later, twirling was generalized to multipartite systems: For a given density matrix ρ the twirled state is defined as [7–9]

$$\mathbf{P}\rho := \int_{U\in \mathrm{U}(d)} U^{\otimes N}\rho(U^{\otimes N})^{\dagger}dU, \tag{1}$$

where $\mathrm{U}(d)$ is the group of d-dimensional unitary matrices, N is the number of qudits, and dU is the normalized Haar measure over $\mathrm{U}(d)$. For the bipartite case one can also consider twirling defined as

$$\mathbf{P}_{\mathrm{iso}}\rho := \int_{U\in \mathrm{U}(d)} U\otimes U^{*}\rho(U\otimes U^{*})^{\dagger}dU, \tag{2}$$

where the asterisk denotes elementwise complex conjugation. States obtained from $\mathbf{P}$ and $\mathbf{P}_{\mathrm{iso}}$ are called Werner states and isotropic states, respectively. Isotropic states are quite useful in quantum-information processing: They are the maximally entangled state mixed with white noise. While in this work we focus on $\mathbf{P}$, our results generalize trivially to the computation of $\mathbf{P}_{\mathrm{iso}}$.

The importance of twirling in the multipartite case is that it transforms a general mixed state into a state that can be characterized with only a few parameters [7,10,11]. Entanglement of formation is known for bipartite isotropic states [12,13] and necessary and sufficient conditions for the entanglement of tripartite Werner states are also known [8]. Therefore, since twirling cannot increase any entanglement monotone, if we can experimentally twirl a state, we can simplify the estimation of its entanglement properties.

Moreover, twirling also appears in various calculations in quantum-information science (e.g., it is used to define a family of quantum states in Ref. [14]). Integrals over $\mathrm{U}(d)$, similar to twirling, appear in many areas of physics [15]. In particular, the computation of integrals of the form

$$\int_{U\in \mathrm{U}(d)} U_{i_1j_1}U_{i_2j_2}\cdots U_{i_mj_m}(U_{k_1l_1}U_{k_2l_2}\cdots U_{k_nl_n})^{*}dU \tag{3}$$

is needed. Such integrals for the $m=n$ case can straightforwardly be obtained from twirling appropriately chosen density matrices. Twirling is also closely related to unitary t-designs which have raised interest recently [16,17].

It seems straightforward to implement twirling: One has to apply a random multilateral unitary rotation to each copy of a state, and then average over the ensemble. However, problems quickly arise when considering practical implementations. Applying different random rotations to different systems of the ensemble requires that we are able to access the systems individually. In practice, very often this also means temporal averaging [18]. We repeat many times the following two steps: (i) Generate the quantum state and (ii) apply a random rotation. The disadvantage of this approach is that the execution time is proportional to the number of systems in the ensemble. Moreover, in many physical realizations of quantum computing, e.g., in a nuclear magnetic resonance (NMR) quantum computer, this approach cannot be used since the systems cannot be individually accessed. From the numerical point of view, the problem is that averaging over the randomly rotated matrices is a very inefficient way for calculating the integral in Eq. (1).

Another approach is using group theory to replace the integral (1) with a sum over a finite number of rotated density matrices [19]. This works for small systems and, for example, for $N=2$ and $d=2$ we need to employ 12 such matrices [2,20]. However, the number of unitaries needed increases rapidly with N and d, making the implementation of twirling for large systems difficult this way [16]. Clearly, this approach does not seem to fit ensemble quantum computing. From the point of view of a realization on a digital

*Electronic address: toth@alumni.nd.edu

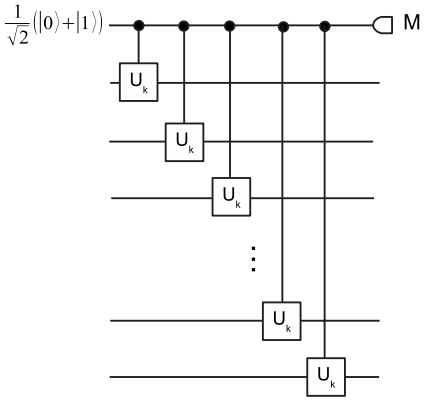

FIG. 1. Twirling can be realized by the repeated application of this basic building block where U_k is a random unitary generated for the kth iteration or a unitary chosen from a cyclically alternating set of unitaries. M represents measurement in the computational basis.

computer, there is the added complexity of computing the required unitaries when compared to averaging over random unitaries.

Numerically, there is another approach for replacing the integration with a discrete sum. The idea is that twirling transforms any quantum state into a $U^{\otimes N}$ invariant state [7]. The density matrix of such a state can be written as $\rho = \Sigma_k \langle R_k \rangle_\rho R_k$ where the R_k basis operators are obtained from orthogonalizing the $N!$ permutation operators [8,9]. Since twirling does not change the expectation values of R_k, we can use these values to reconstruct the final Werner state $\mathbf{P}\rho$ on a classical computer. However, once more, while this approach is feasible for small systems (for $N=2$ and 3 qubits there are two and, respectively, five such matrices [8,9]), for large N the number of permutation matrices increases dramatically.

In this paper we show that a multiqudit state can be approximately twirled by iterating the single nonunitary operation

$$\rho_{k+1} = \frac{1}{2}[\rho_k + U_k^{\otimes N} \rho_k (U_k^{\otimes N})^\dagger], \tag{4}$$

where ρ_k are density matrices and U_k are random unitaries. Using this building block, the error of our approximation decays exponentially with the number of iterations. The exponent of this decay depends neither on the number of qudits nor on their dimension. In contrast, when approximate twirling is realized by averaging density matrices obtained from multilateral random rotations, the convergence is algebraic. We show evidence that the unitary matrices applied can come from a highly imperfect source and also demonstrate through examples that the random unitaries can be replaced by a set of cyclically alternating unitaries, while preserving the exponential convergence. Finally, based on the previous ideas, we present a quantum circuit for twirling by means of controlled unitary gates (see Fig. 1).

Experimental implementation of this operation looks feasible in many physical systems. It is important to stress that, when applied to an ensemble of many systems, our method does not need individual access to the individual systems. We show that for a given set of cyclically alternating unitaries it is possible to obtain general statements for the convergence which are valid for all density matrices. This makes it possible to design algorithms tailored for the operators available in a given physical system. Thus twirling can be one of the quantum algorithms which are especially fitting for realization on a quantum computer. On the other hand, when realizing our algorithm on a classical computer, the programming and computational effort is extremely small.

Our paper is organized as follows. In Sec. II we discuss the usual way twirling is computed on a quantum or a classical computer. In Sec. III we present our proposal, together also with an analysis of the convergence of the approach. In Sec. IV we show that our method is quite robust against the imperfections of the random number generator. In Sec. V we show that, instead of random unitaries, cyclically alternating operators can also be efficiently used for twirling. In Sec. VI we discuss the case of large dimensions. In Sec. VII we explain how to use our ideas for experiments. In Sec. VIII we show how to generalize our method for the numerical integration of useful formulas over the unitary group. Finally, in Sec. IX we discuss connections of our research to existing work.

II. STRAIGHTFORWARD NUMERICAL INTEGRATION

$\mathbf{P}\rho$ can be approximated by an average of a finite number of randomly rotated density matrices

$$\mathbf{P}_M \rho := \frac{1}{M}\left(\rho + \sum_{k=1}^{M-1} U_k^{\otimes N} \rho (U_k^{\otimes N})^\dagger\right). \tag{5}$$

Here M denotes the number of terms and we assume that the unitaries $\{U_k\}$ are distributed uniformly in $\mathrm{U}(d)$ according to the Haar measure. In this section we examine how well $\mathbf{P}_M\rho$ converges to $\mathbf{P}\rho$ for increasing M.

Since we use random matrices, we will obtain a different state for each realization of $\mathbf{P}_M\rho$. To analyze the error, we introduce an expectation value or average over the different choices for U_k as [21]

$$\langle A \rangle := \int A \, dU_1 dU_2 dU_3 \cdots . \tag{6}$$

Using this average, we can analyze how fast $\mathbf{P}_M\rho$ converges to $\mathbf{P}\rho$ for increasing number of unitaries. Simple calculations show that the average error of a particular initial state ρ decreases algebraically as M^{-1},

$$\begin{aligned}\langle \|\mathbf{P}_M\rho - \mathbf{P}\rho\|^2 \rangle &= \langle \|\mathbf{P}_M\rho\|^2 \rangle + \|\mathbf{P}\rho\|^2 - 2\,\mathrm{Tr}(\langle \mathbf{P}_M\rho \rangle \mathbf{P}\rho) \\ &= \langle \|\mathbf{P}_M\rho\|^2 \rangle - \|\mathbf{P}\rho\|^2 = \frac{1}{M}(\|\rho\|^2 - \|\mathbf{P}\rho\|^2),\end{aligned} \tag{7}$$

where $\|A\|^2 := \mathrm{Tr}(A^\dagger A)$ is the Hilbert-Schmidt norm. In the derivation we used that $\langle \mathbf{P}_M\rho \rangle = (1/M)\rho + [(M-1)/M]\mathbf{P}\rho$ and $\mathrm{Tr}(\rho\mathbf{P}\rho) = \mathrm{Tr}[(\mathbf{P}\rho)^2]$.

While computing the error for a given state is illuminating, it is more useful to characterize the convergence of $\mathbf{P}_M$

in a manner that is independent of the initial state. For that, first we will show how to define a matrix describing the action of a linear superoperator and will define a measure of distance between superoperators. Then, we will determine the matrices describing the action of $\mathbf{P}$ and $\mathbf{P}_M$, and will compute the norm of their difference.

Density matrices are vectors in a Hilbert space of complex matrices with the scalar product $\langle \rho, \rho' \rangle := \mathrm{Tr}(\rho\rho')$. Thus it is convenient to switch from matrix notation

$$\rho = \sum_{kl} \rho_{kl} |k\rangle\langle l| \tag{8}$$

and treat the matrices as vectors defined by [22]

$$\vec{\rho} = \sum_{kl} \rho_{kl} |l\rangle \otimes |k\rangle. \tag{9}$$

That is, $\vec{\rho}$ is obtained from ρ by joining its columns consecutively into a column vector. We can use the vector form for any Hermitian operator A, not only for density matrices. Then the expectation value of A can be written as

$$\mathrm{Tr}(A\rho) = (\vec{A})^\dagger \vec{\rho}. \tag{10}$$

Any physically allowed transformation of the density matrix is a linear positive map and it can be written as a matrix acting on $\vec{\rho}$

$$\vec{\rho}' = S\vec{\rho}. \tag{11}$$

Matrix S describes the transformation realized by the superoperator. Both the vectors $\vec{\rho}, \vec{\rho}'$, and the matrix S have to satisfy constraints to ensure the Hermiticity and the positivity of the density matrices. The distance between superoperators can be measured in the form of the Hilbert-Schmidt norm of their difference [23],

$$\|S - \tilde{S}\|^2 := \mathrm{Tr}[(S - \tilde{S})(S - \tilde{S})^\dagger]. \tag{12}$$

In this formalism, the superoperators describing the action of $\mathbf{P}$ and $\mathbf{P}_M$ are, respectively,

$$S_{\mathbf{P}} = \int_{U \in \mathrm{U}(d)} (U^{\otimes N})^* \otimes U^{\otimes N} dU, \tag{13a}$$

$$S_{\mathbf{P}M} = \frac{1}{M}\left(\mathbb{1}_d^{\otimes 2N} + \sum_{k=1}^{M-1} (U_k^{\otimes N})^* \otimes U_k^{\otimes N} \right), \tag{13b}$$

where $\mathbb{1}_d$ denotes a $d \times d$ unit matrix. Based on Eq. (13a), it is easy to see that

$$S_{\mathbf{P}} S_{\mathbf{P}} = S_{\mathbf{P}} = S_{\mathbf{P}}^\dagger. \tag{14}$$

Using these, straightforward calculation shows that

$$\langle \|S_{\mathbf{P}M} - S_{\mathbf{P}}\|^2 \rangle = \langle \|S_{\mathbf{P}M}\|^2 \rangle + \|S_{\mathbf{P}}\|^2 - 2\mathrm{Tr}(S_{\mathbf{P}} \langle S_{\mathbf{P}M} \rangle)$$
$$= \frac{1}{M}(\|\mathbb{1}_d^{\otimes 2N}\|^2 - \|S_{\mathbf{P}}\|^2). \tag{15}$$

Thus the error in the superoperator formalism decays algebraically with increasing number of steps M, irrespective of the initial state (see the Appendix for details).

III. TWIRLING USING A RECURSIVE FORMULA

In order to decrease the error of the result, rather than doubling the number of terms in the summation and computing $\mathbf{P}_{2M}$, we can apply twice the averaging operation with $M-1$ unitaries and calculate $\mathbf{P}_M \mathbf{P}_M \rho$. In this section we show that, even though in both cases $\sim 2M$ random unitaries are needed, the error of the second method is much smaller.

Let us write out the result after two twirlings explicitly:

$$\mathbf{P}_M \mathbf{P}_M \rho = \frac{1}{M^2} \sum_{k=1}^{M-1} \sum_{l=1}^{M-1} \rho + U_k^{\otimes N} \rho [(U_k^{\otimes N}]^\dagger + U_{M-1+l}^{\otimes N} \rho [(U_{M-1+l}^{\otimes N}]^\dagger$$
$$+ (U_{M-1+l} U_k)^{\otimes N} \rho [(U_{M-1+l} U_k)^{\otimes N}]^\dagger, \tag{16}$$

where $\{U_k\}_{k=1}^{M-1}$ and $\{U_k\}_{k=M}^{2M-2}$ are the random unitaries chosen for the first and second twirlings, respectively. Equation (16) is the average of M^2-1 rotated density matrices and the original matrix. We have the same number of terms when computing $\mathbf{P}_{M^2}\rho$. However, the (M^2-1) unitaries are not independent; thus we might expect that the error for $\mathbf{P}_M \mathbf{P}_M \rho$ is larger than that for $\mathbf{P}_{M^2}\rho$.

Let us now consider repeated applications of $\mathbf{P}_m$. For simplicity we will first focus on the $m=2$ case, leaving the $m>2$ case for later. After M iterations, the outcome is

$$\mathbf{Q}_M \rho := \mathbf{P}_2 \mathbf{P}_2 \cdots \mathbf{P}_2 \rho = \left(\prod_{k=1}^{M} \mathbf{P}_2 \right) \rho. \tag{17}$$

Using the definition Eq. (17) we can write the recursive formula

$$\mathbf{Q}_M \rho = \frac{1}{2}[\mathbf{Q}_{M-1}\rho + U_M^{\otimes N} (\mathbf{Q}_{M-1}\rho)(U_M^{\otimes N})^\dagger], \tag{18}$$

where again U_M is a random unitary. As before, we measure the convergence of this operator by the average error in the Hilbert-Schmidt norm,

$$\langle \|\mathbf{Q}_M \rho - \mathbf{P}\rho\|^2 \rangle = \langle \|\mathbf{Q}_M \rho\|^2 \rangle - \|\mathbf{P}\rho\|^2. \tag{19}$$

For computing the error as a function of M, we need the M dependence of the $\langle \|\mathbf{Q}_M \rho\|^2 \rangle$ term on the right-hand side of Eq. (19). For that first we express $\langle \|\mathbf{Q}_M \rho\|^2 \rangle$ with $\langle \|\mathbf{Q}_{M-1}\rho\|^2 \rangle$

$$\langle \|\mathbf{Q}_M \rho\|^2 \rangle = \frac{1}{2}[\langle \|\mathbf{Q}_{M-1}\rho\|^2 \rangle + \langle \|\mathbf{Q}_{M-1}\rho U_M^{\otimes N} (\mathbf{Q}_{M-1}\rho)(U_M^{\otimes N})^\dagger\|^2 \rangle]$$
$$= \frac{1}{2}(\langle \|\mathbf{Q}_{M-1}\rho\|^2 \rangle + \|\mathbf{P}\rho\|^2). \tag{20}$$

Then, from Eq. (20) the M dependence of $\langle \|\mathbf{Q}_M \rho\|^2 \rangle$ can be obtained as

$$\langle \|\mathbf{Q}_M \rho\|^2 \rangle = \|\rho\|^2 + (\|\mathbf{P}\rho\|^2 - \|\rho^2\|)(1 - 2^{-M}). \tag{21}$$

Substituting Eq. (21) into Eq. (19) we obtain

$$\langle \|\mathbf{Q}_M \rho - \mathbf{P}\rho\|^2 \rangle = (\|\rho\|^2 - \|\mathbf{P}\rho\|^2) 2^{-M}. \tag{22}$$

That is, the squared error decays exponentially with M, while according to Eq. (7) the decay was proportional to M^{-1} for the method described in Sec. II. Note that computing $\mathbf{Q}_M \rho$ and $\mathbf{P}_M \rho$ needs the generation of M and $M-1$ random

unitaries, respectively. Thus the computational effort is roughly the same for the two cases.

One can repeat this calculation in the superoperator picture. The definition of $S_{\mathbf{Q}M}$ based on Eq. (18) is

$$S_{\mathbf{Q}M} = \frac{1}{2}\{S_{\mathbf{Q}(M-1)} + [(U_M^{\otimes N})^* \otimes U_M^{\otimes N}]S_{\mathbf{Q}(M-1)}\}. \quad (23)$$

For the error of the approximation we obtain

$$\begin{aligned}\langle \|S_{\mathbf{Q}M} - S_{\mathbf{P}}\|^2 \rangle &= \langle \|S_{\mathbf{Q}M}\|^2 \rangle + \|S_{\mathbf{P}}\|^2 - \langle \mathrm{Tr}(S_{\mathbf{Q}M}^\dagger S_{\mathbf{P}}) \rangle \\ &\quad - \langle \mathrm{Tr}(S_{\mathbf{Q}M} S_{\mathbf{P}}) \rangle \\ &= \langle \|S_{\mathbf{Q}M}\|^2 \rangle - \|S_{\mathbf{P}}\|^2. \end{aligned} \quad (24)$$

For obtaining the M dependence of the error, we need the M dependence of $\|S_{\mathbf{Q}M}\|^2$. This is obtained in two steps. First, we use Eq. (23) to find a recursive relation for $\langle \|S_{\mathbf{Q}M}\|^2 \rangle$,

$$\langle \|S_{\mathbf{Q}M}\|^2 \rangle = \frac{1}{2}(\langle \|S_{\mathbf{Q}(M-1)}\|^2 \rangle + \|S_{\mathbf{P}}\|^2), \quad (25)$$

and then we obtain $\langle \|S_{\mathbf{Q}M}\|^2 \rangle$ without using recursion as

$$\langle \|S_{\mathbf{Q}M}\|^2 \rangle = \|\mathbb{1}_d^{\otimes 2N}\|^2 + (\|S_{\mathbf{P}}\|^2 - \|\mathbb{1}_d^{\otimes 2N}\|^2)(1 - 2^{-M}). \quad (26)$$

Combining Eqs. (26) and (24) we obtain

$$\langle \|S_{\mathbf{Q}M} - S_{\mathbf{P}}\|^2 \rangle = (\|\mathbb{1}_d^{\otimes 2N}\|^2 - \|S_{\mathbf{P}}\|^2) 2^{-M}; \quad (27)$$

thus the error of the superoperator decays exponentially.

Let us now consider combining twirl operations $\mathbf{P}_K$ with more than two unitaries. With that aim we define

$$\mathbf{Q}_{K,M} := (\mathbf{P}_K)^M. \quad (28)$$

On the one hand, when computing the average error for this operator we obtain formulas similar to Eqs. (22) and (27), but with a mean squared error proportional to K^{-M} vs the original, proportional to 2^{-M}. On the other hand, the number of random unitaries required increases, as it is now $K-1$ per iteration step. Based on these we can write the dynamics of the mean squared error as a function of the number of unitaries N_U as

$$\langle \|\mathbf{Q}_{K,M} - \mathbf{P}\|^2 \rangle \propto \exp\left(-\frac{\ln K}{K-1} N_U\right). \quad (29)$$

Hence one can see that, for a given number of unitaries, the smallest error is achieved for $K=2$. For many experiments, this is also a good reasoning since the experimental effort is very often measured in N_U. Thus we will consider the $K=2$ case in the rest of the paper.

We have verified numerically the previous results [24]. For this we focused on the three-qubit case, for which we can compute the twirl operation exactly using the techniques mentioned in the introduction (see the Appendix). In Fig. 2 we plot the error $\langle \|S_{\mathbf{P}M} - S_{\mathbf{P}}\|^2 \rangle$ averaged over 10 000 trajectories. As the figure shows, the simulation results perfectly fit the exponential decay of the error calculated theoretically (27).

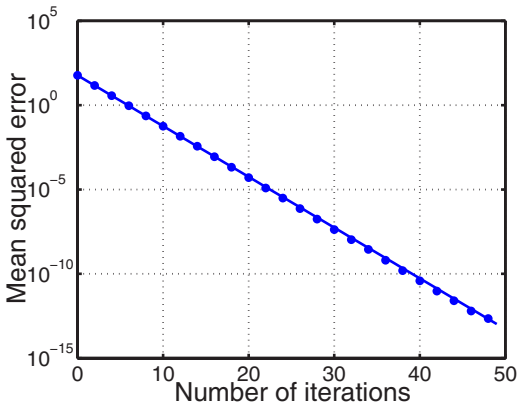

FIG. 2. (Color online) Mean squared error for the recursive method with random matrices when applied on three-qubit states. We plot (dotted) the average over 10 000 realizations and (solid line) the theoretical prediction computed from Eq. (27). For better visibility, the error is shown only for every second iteration.

IV. SENSITIVITY TO THE IMPERFECTIONS OF RANDOM NUMBER GENERATION

In this section we examine what happens if our random number generator does not work perfectly and the random unitaries are not uniformly distributed over U(d). An imperfect random unitary generator can be characterized by the distribution $f(U)$ describing the probability density for getting U. We will show that if $\inf_U f(U) > 0$ then our algorithm still converges to the twirled state and the error decays exponentially.

Let us use a simple model for our faulty distribution in which with probability p_g the unitary is drawn according to the probability distribution $g(U)$ while with probability $(1-p_g)$ it is drawn according to the uniform distribution. The corresponding distribution function $f(U)$ is

$$f(U) := p_g g(U) + (1 - p_g), \quad (30)$$

where we used that $\int_{U\in \mathrm{U}(d)} dU = 1$. Expectation values over this probability distribution are computed as

$$\langle A \rangle_f := \int_{U\in \mathrm{U}(d)} A f(U_1) f(U_2) f(U_3) \cdots dU_1 dU_2 dU_3 \cdots .$$

Let us now examine how the usual method described in Sec. II is affected by such an error of the random number generator. We will define by $\tilde{\mathbf{P}}_M$ the equivalent of Eq. (5) with our biased probability distribution. The mean value of the density matrix obtained from such twirling is

$$\begin{aligned}\langle \tilde{\mathbf{P}}_M \rho \rangle_f = \frac{1}{M}\big(&\rho + (1-p_g)(M-1)\mathbf{P}\rho \;+\; p_g(M-1) \\ &\times \int dW\, g(W) W^{\otimes N} \rho (W^{\otimes N})^\dagger\big).\end{aligned}$$

Here $\int dW$ is an integral over the unitary group U(d). Taking the limit $M \to \infty$ one obtains

$$\langle \tilde{\mathbf{P}}_M \rho \rangle_f \to (1-p_g)\mathbf{P}\rho + p_g \int dW\, g(W) W^{\otimes N} \rho (W^{\otimes N})^\dagger.$$

Thus the expectation value of the operator does not converge to $\mathbf{P}\rho$.

On the contrary, when $\tilde{\mathbf{P}}_2$ is applied M times, the state of the system still converges to the twirled state and the error decays exponentially with M. To show this, let us denote the operation above by $\tilde{\mathbf{Q}}_M$. As before, we measure the convergence of this operator by the average error in the Hilbert-Schmidt norm

$$\langle \|\tilde{\mathbf{Q}}_M \rho - \mathbf{P}\rho\|^2 \rangle_f = \langle \|\tilde{\mathbf{Q}}_M \rho\|^2 \rangle_f - \|\mathbf{P}\rho\|^2. \tag{31}$$

Hence straightforward algebra yields

$$\langle \|\tilde{\mathbf{Q}}_M \rho\|^2 \rangle_f \leq \frac{1+p_g^2}{2} \langle \|\tilde{\mathbf{Q}}_{M-1} \rho\|^2 \rangle_f + \frac{1-p_g^2}{2} \|\mathbf{P}\rho\|^2. \tag{32}$$

For obtaining the upper bound in Eq. (32) we used

$$\mathrm{Tr}[\tilde{\mathbf{Q}}_{M-1} \rho U^{\otimes N} (\tilde{\mathbf{Q}}_{M-1} \rho)(U^{\otimes N})^\dagger] \leq \|\tilde{\mathbf{Q}}_{M-1} \rho\|^2,$$

$$\left\| \int dW\, g(W) W^{\otimes N} (\tilde{\mathbf{Q}}_{M-1} \rho)(W^{\otimes N})^\dagger \right\|^2 \leq \|\tilde{\mathbf{Q}}_{M-1} \rho\|^2, \tag{33}$$

where U is a unitary matrix. From Eq. (32) the M dependence of $\langle \|\tilde{\mathbf{Q}}_M \rho\|^2 \rangle_f$ can be deduced as

$$\langle \|\tilde{\mathbf{Q}}_M \rho\|^2 \rangle_f \leq \|\rho\|^2 + (\|\mathbf{P}\rho\|^2 - \|\rho^2\|) \left[1 - \left(\frac{2}{1+p_g^2} \right)^{-M} \right]. \tag{34}$$

Substituting Eq. (34) into Eq. (31) we obtain

$$\langle \|\tilde{\mathbf{Q}}_M \rho - \mathbf{P}\rho\|^2 \rangle_f \leq (\|\rho\|^2 - \|\mathbf{P}\rho\|^2) \left(\frac{2}{1+p_g^2} \right)^{-M}. \tag{35}$$

The mean square error of the superoperator corresponding to $\tilde{\mathbf{Q}}_M$ also decays proportionally to $[2/(1+p_g^2)]^{-M}$. Thus we have convergence if $p_g < 1$, i.e., if the uniform distribution has a nonzero weight in Eq. (30). For functions $f(U)$ that satisfy

$$\inf_U f(U) > 0, \tag{36}$$

it is always possible to find a decomposition of the type Eq. (30) such that $p_g = 1 - \inf_U f(U)$. Thus for such probability distribution functions our algorithm converges and the error decays exponentially.

Finally, the sufficient condition for the convergence of our method Eq. (36) can also be formulated for the case when $f(U)$ is of the form

$$f(U) = f_r(U) + \sum_k c_k \delta(U - V_k), \tag{37}$$

where $f_r : U(d) \mapsto \mathbb{R}$, $c_k \geq 0$ are constants, δ is the Dirac delta function, and V_k are unitaries. In this case the algorithm converges if

$$\inf_{e(U)} \int_{U \in \mathrm{U}(d)} f(U) e(U) dU > 0, \tag{38}$$

where for the function $e(U)$ we require that $e(U) \geq 0$ and $\int_{U \in \mathrm{U}(d)} e(U) dU = 1$.

V. DETERMINISTIC TWIRLING WITH FEW UNITARIES

The example shown in Sec. IV demonstrated that, even if the random unitaries used in our algorithm come from a very imperfect source, the algorithm may still converge. In this section we will examine what happens if these unitaries are not random but they are chosen deterministically from a small set such that they are cyclically alternating.

Let us consider the two-qubit case. Common sense tells us that we need at least two unitaries since, two unitaries are able to generate the elements of U(2) we need for twirling [25]. We will see that two unitaries are sufficient.

Let us choose the two unitaries as

$$U_x := e^{ic\sigma_x},$$

$$U_z := e^{ic\sigma_z}, \tag{39}$$

where $\sigma_{x/z}$ are Pauli spin matrices and c is a constant. Now we can use the method described at the end of Sec. III to compute the dependence of the superoperator on the number of iterations. We look for the c for which the decay of the error is the fastest. Through numerical optimization we find that the error of the superoperator is the smallest after 50 iterations for $c = 1.0894$. Figure 3(a) shows the results of our numerical calculations for two qubits with this value for c. The dashed line shows the square of the error for the recursive method using random unitaries described in Sec. III. Note that the error for the deterministic method, denoted by disks, decays faster than for the random method.

Let us see a three-qubit example with three cyclically alternating unitaries

$$U_x := e^{i2\pi/3\sigma_x},$$

$$U_y := e^{i2\pi/5\sigma_y},$$

$$U_z := e^{i2\pi/3\sigma_z}. \tag{40}$$

Figure 3(b) shows the results of our numerical calculations. Now the error decays somewhat more slowly than in the case of the random method.

The advantage of our approach is that, by studying the superoperator, we can make general statements, independent of the initial state, about the algorithm. Thus without a thorough group-theoretical study we can show that the error decays exponentially with M and the recursive algorithm with the given alternating unitaries can be used for the efficient twirling of two or three qubits, respectively. Similar calculations can be carried out for several qubits trying out other unitaries or higher dimensions can also be investigated. These calculations can always consider the gates easily available in an experimental implementation.

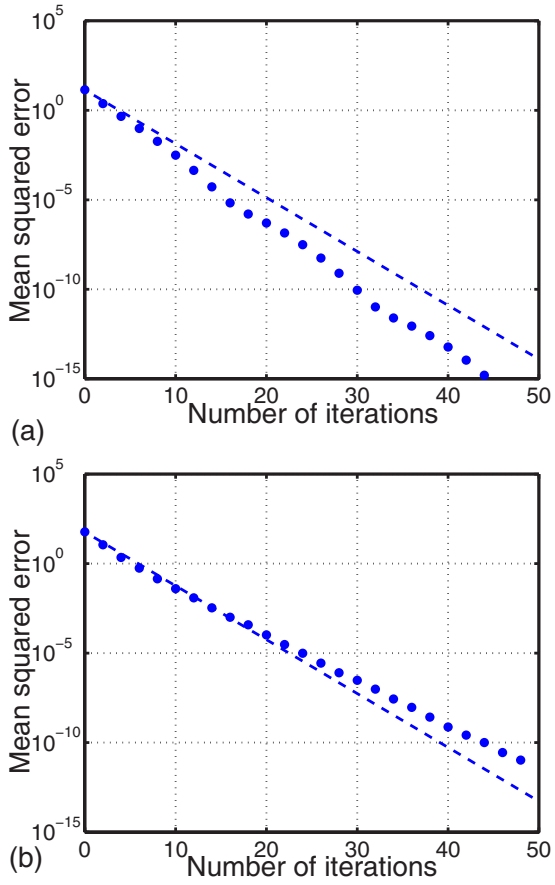

FIG. 3. (Color online) Time dependence of the error for (a) two and (b) three qubits for the deterministic method using two and three unitaries, respectively. For better visibility, the error is shown only for every second iteration. Dashed line indicates the error for the method using random matrices, given in Eq. (27).

VI. TWIRLING FOR QUDITS WITH LARGE DIMENSIONS

In the literature there was a considerable effort to find a method for two-qudit twirling which can be realized with relatively few quantum gates even if the dimension of qudits is large. In this section we show through examples that the number of quantum gates necessary for two-qudit twirling with our algorithm seems to scale better with the dimension of the qudits than for the algorithm generating a d-dimensional random unitary uniformly distributed according to the Haar measure.

A method for generating a random unitary of large dimensions was presented in Ref. [26]. The system was considered to be a multiqubit system which fits well for many physical realizations. The algorithm presented has two steps: (i) Single-qubit random unitaries act on the individual qubits; (ii) a nearest-neighbor Ising interaction acts on the one-dimensional array of qubits. These two steps must be repeated several times. It was found that the number of gates necessary for generating a random unitary this way scales exponentially with the number of qubits n.

In Ref. [3] it was shown that two-qudit twirling over $\mathrm{U}(2^n)$ gives the same result as two-qudit twirling over the Clifford group. This makes efficient twirling possible since the number of gates needed for generating a random Clifford group element scales polynomially with n. The results of Ref. [3] were extended to unitary 2-designs in Ref. [16].

Let us now examine whether it is possible to find an efficient way to realize our algorithm for $d>2$. We also consider the $d=2^n$ case. We look for a simple way for generating an imperfect random unitary such that it still can be used for two-qudit twirling. In particular, we would like that the error does not decay slower than when using unitaries which are uniformly distributed according to the Haar measure.

We use a slight modification of the algorithm presented in Ref. [26]. Our random n-qubit unitary is generated by applying first different random unitaries for each qubit, then making the system evolve under an Ising Hamiltonian with nearest-neighbor interaction realizing

$$U_{\mathrm{Ising}} := \exp\left(i\alpha \sum_k \sigma_z^{(k)} \sigma_z^{(k+1)}\right), \tag{41}$$

where α is a constant and we consider a periodic boundary condition. Thus we use a single iteration of the method presented in Ref. [26]. The gate requirements increase linearly with n for such an algorithm.

Next we show simulations with the density matrix rather than simulations with the superoperator. The reason is that the size of the superoperator is $16^n \times 16^n$ which would make it possible to consider only small systems. We calculate the dynamics obtained from our algorithm for several random density matrices which have a uniform distribution according to the Hilbert-Schmidt measure [27]. In order to compare trajectories corresponding to different density matrices, we compute the normalized error

$$E_{\mathrm{norm}} := \frac{\langle \|\mathbf{Q}_M \rho - \mathbf{P}\rho\|^2 \rangle}{(\|\rho\|^2 - \|\mathbf{P}\rho\|^2)}. \tag{42}$$

It follows from Eq. (22) that for the method using random matrices uniformly distributed over $\mathrm{U}(d)$ we have $E_{\mathrm{norm}} = 2^{-M}$.

Figure 4 shows the results of our calculations for $d=2^3$ and $d=2^4$. We used $\alpha=1.10$ and 1.03, respectively. We find that the error decays almost as in the case of using random unitaries uniformly distributed over $U(d)$ and the difference between the two error curves seems to be subexponential.

VII. EXPERIMENTAL REALIZATION

There are two different situations from the point of view of experimental realizations. In many experiments several copies of a quantum system are available at a time. Very often the systems cannot be individually accessed. However, we would like that these systems undergo *different* multilateral random unitary rotations. In this case, according to our algorithm we have to achieve that at each iteration step half of the systems undergo a unitary rotation $U_k^{\otimes N}$, while the other half does not. While numerically the mixing of the state ρ and the rotated one, $U_k^{\otimes N} \rho (U_k^{\otimes N})^\dagger$ is a matter of adding two matrices, experimentally this mixing can be done with the help of a controlled operation. The corresponding quantum circuit is shown in Fig. 1. For a single step of the

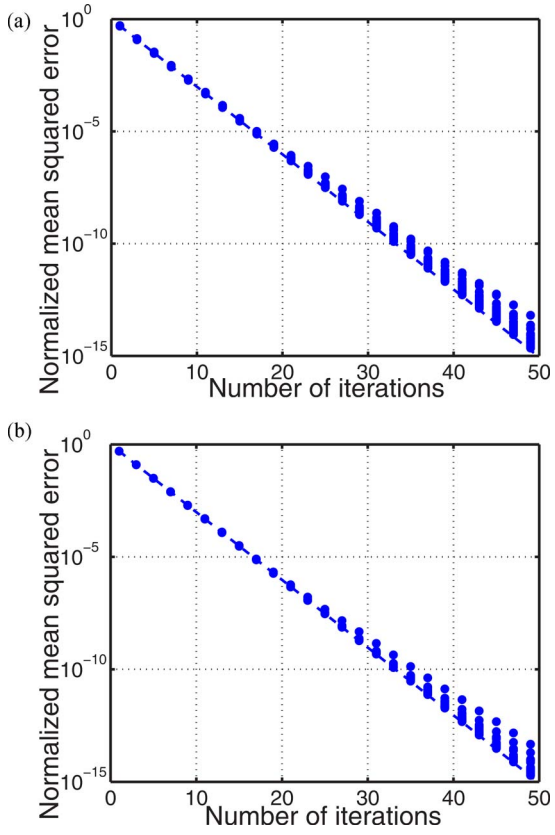

FIG. 4. (Color online) Time dependence of the normalized error of the density matrix for bipartite twirling for $d=$(a) 8 and (b) 16. In both cases 25 realizations are shown. See text for the algorithm used for generating d-dimensional unitaries. For better visibility, the error is shown only for every second iteration. Dashed line indicates the error for the method using random matrices uniformly distributed over U(d).

algorithm we realize $\mathbf{P}_2$. The inputs are the state and an ancilla in a superposition state $(|0\rangle+|1\rangle)/\sqrt{2}$. The same unitary U_k is applied on all qudits of the state, but only when the ancilla qubit is in state 1. Finally, we measure the ancilla and the outcome is $\frac{1}{2}[\rho+U_k^{\otimes N}\rho(U_k^{\otimes N})^\dagger]$. This basic block is applied M times, each time using either a different, random unitary or a unitary from the finite set as in Sec. V. Note that the control qubit can be a qubit which have a short coherence time compared to the other qubits. It rapidly decays to state $|0\rangle$ or $|1\rangle$, and can be used as a sort of classical control for permitting the unitary rotations on the other qudits.

In other experiments only a single copy of the state is produced. For having an ensemble average of some quantity, the experiment must be repeated many times. In this case our method can be used the following way. Each time after the single copy of the state is created, with 50% probability we apply $U_1^{\otimes N}$, then with 50% probability we apply $U_2^{\otimes N}$, etc. The number of gates needed in average is the half of the number of iterations. If we use the deterministic version of our method described in Sec. V then it makes it possible to twirl with a few single-qubit gates. This is an advantage in some systems. For example, when using photons created with parametric down-conversion and postselection, the single-qubit gates can be realized with wave plates.

VIII. NUMERICAL INTEGRATION OVER U(d)

As we have already mentioned, our method can be used for integrating numerically expressions of the type Eq. (3). In this section we discuss how to generalize our approach for integrating expressions of the type

$$I := \int_{U\in \mathrm{U}(d)} \mathrm{Tr}(A_1U)\mathrm{Tr}(A_2U)\cdots\mathrm{Tr}(A_mU) \times \mathrm{Tr}(B_1U^\dagger)\mathrm{Tr}(B_2U^\dagger)\cdots\mathrm{Tr}(B_nU^\dagger)dU, \quad (43)$$

where A_k and B_k are $d\times d$ matrices.

Based on the main ideas of the paper, Eq. (43) can be computed in two steps. (i) First we need to obtain

$$M := \int_{U\in \mathrm{U}(d)} U^{\otimes m}\otimes(\mathrm{U}^\dagger)^{\otimes n}dU. \quad (44)$$

This can be done by iterating the formula

$$M_{k+1}=\frac{1}{2}[\mathbb{1}_d^{\otimes(m+n)}+U_k^{\otimes m}\otimes(U_k^\dagger)^{\otimes n}]M_k, \quad (45)$$

where $M_0=\mathbb{1}$ and U_k are random unitaries. The series M_k will converge very fast to M. (ii) The second step in computing Eq. (43) is

$$I=\mathrm{Tr}(MA_1\otimes A_2\otimes\cdots\otimes A_m\otimes B_1\otimes B_2\otimes\cdots\otimes B_n). \quad (46)$$

Note that M does not depend on A_k and B_k. Thus, when we compute Eq. (43) for several $\{A_k\}$ and $\{B_k\}$, we have to compute M only once.

These ideas seem to work also when integrating over a subgroup of U(d), in particular, over the special unitary group SU(d). Such integrals appear, for example, in quantum chromodynamics [28,29].

IX. DISCUSSION

First let us discuss the importance of the fact that our algorithm does not require an individual access to the systems of the ensemble. This characteristic is important since we are presenting the realization of a superoperator mapping a density matrix to another density matrix. Ideally, we want that this mapping works even if the density matrix describes an ensemble of very many systems. A method which requires an individual access to the systems of the ensemble cannot handle this situation. When realizing a superoperator in a physical system, it is also advantageous that if a pure state is mapped to a mixed one then this mixed state is realized as the reduced state of a pure state of a larger system [30]. The usual method is not able to create such a purification of the output density matrix. In contrast, our method can handle a very large ensemble. Also, when we apply the quantum circuit proposed in this paper for twirling, and we omit the measurements then we get a pure state. The twirled state is the reduced state of this pure state.

The algorithm presented in this paper is intimately related to other works on random matrices. For instance, Ref. [31] studies the statistical properties of unitary matrices com-

posed as the product of random unitaries. In Ref. [32] it is proved that the product of a series of random unitaries with nonuniform distribution converges exponentially fast to the uniform distribution in many cases.

The relation of our paper to these works is the following. We also used composed ensembles of unitary matrices. However, when looking at Eq. (16), we can see that our composed unitaries are not independent and they are composed from a small set of random matrices. Thus, especially in the first part of the paper, the main goal was to realize twirling on an ensemble of *many* systems using only a *few* random unitaries, rather than realizing twirling using unitaries from an imperfect random source. Note that we assumed that our unitaries were drawn from a perfect random source providing unitaries distributed uniformly in U(d). In the second half of our paper we found that our results can also be applied to the case of an imperfect source or for a deterministic algorithm.

The key point of our algorithm is mixing of a subensemble in the original state and the other subensemble in which all the systems undergo the same multilateral rotation. This mixing can be realized efficiently both in a classical computer and in a quantum computer. Let us analyze the role of mixing pointing out something seemingly paradoxical. Let us consider redefining $\mathbf{P}_2$ as the application of a unitary U which with 50% probability it is the identity and with 50% probability it is uniformly distributed. That would amount to applying a unitary with a distribution

$$h(U) := \frac{1}{2}\delta(U-\mathbb{I}) + \frac{1}{2}. \tag{47}$$

However, if we compute the error of a single application of this "new" $\mathbf{P}_2$ we find

$$\begin{aligned}\langle \|\hat{\mathbf{P}}_2\rho - \mathbf{P}\rho\|^2\rangle_h &= \int \|U^{\otimes N}\rho(U^{\otimes N})^\dagger - \mathbf{P}\rho\|^2 h(U)dU \\ &= \int (\|\rho\|^2 - \|\mathbf{P}\rho\|^2)h(U)dU \\ &= \|\rho\|^2 - \|\mathbf{P}\rho\|^2,\end{aligned} \tag{48}$$

which, unlike Eq. (22), does not give us an exponential convergence. Of course, this is because in the algorithm described in this paragraph there is only one component present: The application of a random unitary. The other component, namely, mixing of two subensembles, is missing.

Moreover, while Ref. [32] studied the convergence of the distribution of the composed unitaries to the uniform distribution, we studied the convergence of a certain quantum operation, namely, twirling. In particular, we computed the exponent of this convergence and found that it does not depend on N or d. When studying unitaries composed from random ones with a nonuniform distribution, clearly the requirements for the convergence of the operator built from these unitaries are much weaker than the requirements for the convergence of the distribution of the unitaries. It is also easier to get general statements on the convergence of the operator for a quite wide class of faulty random unitary generators. Indeed, we proved that convergence is reached even in the case of a very poor random source.

Note that recursive algorithms can also be applied to summing over discrete groups. For example, it has already been discussed in Ref. [3] that a random element of the Clifford group can be generated by a sequence of $O(n^8)$ operations. At each step, with 1/2 probability nothing happens, and with 1/2 probability a random element of the generating set is executed. Another example is discussed in Refs. [33,34]. An N-qubit state can be depolarized by summing over the stabilizer group [35,36] of a Greenberger-Horne-Zeilinger [37] state or a graph state [38,39] as

$$\rho = \sum_{k=1}^{2^N} S_k \rho_0 S_k^\dagger, \tag{49}$$

where $\{S_k\}$ are the group elements. Exploiting that the stabilizer group is commutative and that $S_k^2=1$, the operation Eq. (49) can be realized in N steps. At step k with 1/2 probability nothing happens, and with 1/2 probability g_k is executed. Here g_k are the N generators of the stabilizer group.

Finally, twirling a completely positive map [16], rather than a quantum state, is also a useful procedure. Twirling a map makes it possible, for example, to estimate the average fidelity of a physical implementation of the map [6]. It can be done with a slight modification of our algorithm in three steps: (i) Applying the circuit Fig. 1 several times with unitaries $U_1, U_2, \ldots, U_M$, (ii) applying the map, and (iii) applying again the circuit Fig. 1 with unitaries $U_M^\dagger, U_{M-1}^\dagger, \ldots, U_1^\dagger$. The control qubit for rotations U_k and $U_k^\dagger$ must be the same.

X. CONCLUSIONS

Summing up, we have presented a very efficient approach for the realization of twirling. Although it is based on random matrices, it converges very fast, that is, the error decays exponentially with increasing number of steps during the iteration. Together with the simplicity of the method, this means that our algorithm requires very little experimental or computational effort, for the implementation in either a quantum or a classical computer. We have demonstrated the robustness of the algorithm, which converges both in the case of an imperfect random number generator and if the unitaries are chosen deterministically from a small set. In the future, it would be interesting to extend our approach to operations which realize twirling with a subgroup of U(d). In particular, we would like to apply the method described in Sec. VIII for integrating numerically over the SU(d) group and look for possible applications.

ACKNOWLEDGMENTS

We thank M. M. Wolf for many useful discussions. We also thank P. Hayden and U. Wenger for helpful comments. We acknowledge the support of the EU projects RESQ and QUPRODIS and the Kompetenznetzwerk Quanteninformationsverarbeitung der Bayerischen Staatsregierung. J.J.G.R. was supported by the Programa Ramon y Cajal of the Spanish Ministry of Education. G.T. acknowledges the support of the European Union (Grant No. MERG-CT-2005-029146), the National Research Fund of Hungary OTKA under Con-

tract No. T049234, and the Hungarian Academy of Sciences (Janos Bolyai Programme).

APPENDIX: COMPUTING THE SUPEROPERATORS CORRESPONDING TO TWIRLING

For evaluating the right-hand side of Eq. (15), we need to know $\langle \|S_{\mathbf{P}}\|^2 \rangle$. For studying the convergence through simulations of individual trajectories we need even the matrix $S_{\mathbf{P}}$. In this appendix, we will study the general properties of $S_{\mathbf{P}}$ and determine it explicitly for small systems.

Based on Eq. (14) we can write

$$\langle \|S_{\mathbf{P}}\|^2 \rangle = \mathrm{Tr}(S_{\mathbf{P}} S_{\mathbf{P}}^{\dagger}) = \mathrm{Tr}(S_{\mathbf{P}}). \tag{A1}$$

Based on Eq. (14) we also know that $S_{\mathbf{P}}$ is a projector matrix with eigenvalues 0 and 1; thus $\langle \|S_{\mathbf{P}}\|^2 \rangle$ must be an integer. We can get to know more about $S_{\mathbf{P}}$ by recalling that twirling produces a Werner state and such a state is a linear combination of permutation operators. Since these permutation matrices are not linearly independent, one has to first orthogonalize them. Let us assume that $\{R_k\}_{k=1}^{N_R}$ are the matrices obtained this way, satisfying $\mathrm{Tr}(R_k R_l) = \delta_{kl}$ where δ is the Kronecker symbol. Now it is easy to see that we can write the twirled matrix as

$$\mathbf{P}\rho = \sum_k \mathrm{Tr}(\rho R_k) R_k. \tag{A2}$$

Hence using Eq. (10), we obtain

$$S_{\mathbf{P}} = \sum_k \vec{R}_k (\vec{R}_k)^{\dagger}. \tag{A3}$$

Thus

$$\langle \|S_{\mathbf{P}}\|^2 \rangle = \mathrm{Tr}(S_{\mathbf{P}}) = N_R. \tag{A4}$$

Now, let us determine $S_{\mathbf{P}}$ explicitly for two and three qudits. For $N=2$ we have $N_R=2$ and a possible choice of the basis matrices is [7]

$$R_1 := \frac{\mathbb{1}_d \otimes \mathbb{1}_d + V_{12}}{\sqrt{d(d+1)}},$$

$$R_2 := \frac{\mathbb{1}_d \otimes \mathbb{1}_d - V_{12}}{\sqrt{d(d-1)}}. \tag{A5}$$

Here V_{12} is the permutation matrix exchanging the two qudits and d is the dimension of the qudits. Hence $S_{\mathbf{P}}$ can be reconstructed based on Eq. (A3). For $N=3$ and $d=2$ we have $N_R=5$, while for $d>2$ we have $N_R=6$. The basis matrices can be found in Refs. [8,9].

[1] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996).
[2] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
[3] D. DiVincenzo, D. Leung, and B. Terhal, IEEE Trans. Inf. Theory **48**(3), 580 (2002).
[4] H. Bombin and M. A. Martin-Delgado, e-print quant-ph/0503013.
[5] W. Dür, M. Hein, J. I. Cirac, and H.-J. Briegel, Phys. Rev. A **72**, 052326 (2005).
[6] J. Emerson, R. Alicki, and K. Życzkowski, J. Opt. B: Quantum Semiclassical Opt. **7**, S347 (2005).
[7] R. F. Werner, Phys. Rev. A **40**, 4277 (1989).
[8] T. Eggeling and R. F. Werner, Phys. Rev. A **63**, 042111 (2001).
[9] T. Eggeling, Ph.D. thesis, Technical University of Braunschweig, Germany, 2003 (unpublished).
[10] One can consider a subset of such states with even fewer parameters. See, for example, D. Chruściński and A. Kossakowski, Phys. Rev. A **73**, 062314 (2006); **73**, 062315 (2006).
[11] It is also possible to depolarize a state into other forms which can be described with few parameters, and their entanglement properties are also known. See W. Dür, J. I. Cirac, and R. Tarrach, Phys. Rev. Lett. **83**, 3562 (1999); W. Dür and J. I. Cirac, Phys. Rev. A **61**, 042314 (2000).
[12] B. M. Terhal and Karl Gerd H. Vollbrecht, Phys. Rev. Lett. **85**, 2625 (2000).
[13] S.-M. Fei and X. Li-Jost, Phys. Rev. A **73**, 024302 (2006).
[14] G. Tóth and A. Acín, Phys. Rev. A **74**, 030306(R) (2006).
[15] P. W. Brouwer and C. W. J. Beenakker, J. Math. Phys. **37**, 4904 (1996); S. Aubert and C. S. Lam, *ibid.* **44**, 6112 (2003).
[16] C. Dankert, R. Cleve, J. Emerson, and E. Livine, e-print quant-ph/0606161.
[17] D. Gross, K. Audenaert, and J. Eisert, e-print quant-ph/0611002.
[18] M. S. Anwar, L. Xiao, A. J. Short, J. A. Jones, D. Blazina, S. B. Duckett, and H. A. Carteret, Phys. Rev. A **71**, 032327 (2005).
[19] We note that replacing the integration by a sum is not the only way to realize twirling. Twirling has been implemented in an ensemble quantum computer (i.e., in an NMR quantum computer) using a continuous set of rotations generated by a field gradient over the sample [18].
[20] P. K. Aravind, Phys. Lett. A **233**, 7 (1997).
[21] Note that $\langle A \rangle$ is *not* the quantum-mechanical expectation value computed as $\mathrm{Tr}(\rho A)$.
[22] We use the following definition for the tensor product

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & a_{13}B & \cdots \\ a_{21}B & a_{22}B & a_{23}B & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}.$$

[23] In this paper we will present bounds for the Hilbert-Schmidt norm of the error, i.e., the difference between twirling and our approximation. Other norms can also be used for measuring the difference between the S matrices representing linear superoperators. The bounds we present can be used to construct bounds also for these norms. In particular, the trace norm is defined as $\|A\|_{\mathrm{tr}} := \mathrm{Tr}\sqrt{A^{\dagger}A}$. Since for $n \times n$ matrices we have $\|A\|_{\mathrm{tr}}^2 \leq n\|A\|^2$, an upper bound for the Hilbert-Schmidt norm of the error is also a valid bound for the trace norm of the error. Instead of asking how large is the difference between the ma-

trices representing two linear superoperators, one can look for the superoperator norm of the difference between the two superoperators. Let $\mathcal{T}$ denote a mapping from a density matrix to another density matrix. Then its superoperator norm is defined as $\|\mathcal{T}\|_{\mathrm{so}} := \sup_{X\neq 0}\|\mathcal{T}X\|_{\mathrm{tr}}/\|X\|_{\mathrm{tr}}$. For the norm of the difference between two superoperators it can be proven that $\|\mathcal{T}_1-\mathcal{T}_2\|_{\mathrm{so}}^2 \leq d^N\|S_{\mathcal{T}_1}-S_{\mathcal{T}_2}\|^2$, where $\mathcal{T}$ acts on N qudits of dimension d. Here we used that for an $n\times n$ matrix A we have $\|A\|^2 \leq \|A\|_{\mathrm{tr}}^2 \leq n\|A\|^2$. Finally, another norm that can be used for measuring the difference between superopertors is the diamond norm described in D. Aharonov, A. Kitaev, and N. Nisan, e-print quant-ph/9806029. The diamond norm is defined as $\|\mathcal{T}\|_{\diamond} := \|\mathcal{T}\otimes(1_d^{\otimes N})\|_{\mathrm{so}}$. Here the tensor product by 1 represents the extension of the superoperator $\mathcal{T}$ to a larger space. For our norm $\|S_{\mathcal{T}\otimes(1_d^{\otimes N})}\|^2 = d^{2N}\|S_{\mathcal{T}}\|^2$. Hence it also follows that $\|\mathcal{T}_1-\mathcal{T}_2\|_{\diamond}^2 \leq d^{4N}\|S_{\mathcal{T}_1}-S_{\mathcal{T}_1}\|^2$.

[24] The simulations were done with the help of the computer code QUBIT4MATLAB V2.0, http://www.mathworks.com/matlabcentral/fileexchange. The command TWIRL.M realizes twirling using the recursive method explained in this paper.

[25] Clearly, it is enough if we are able to generate the elements of the special unitary group SU(2).

[26] J. Emerson, Y. Weinstein, M. Saraceno, S. Lloyd, and D. Cory, Science **302**, 2098 (2003).

[27] K. Życzkowski and H.-J. Sommers, J. Phys. A **34**, 7111 (2001).

[28] Urs Wenger (private communication).

[29] G. Parisi, R. Petronzio, and F. Rapuano, Phys. Lett. **128B**, 418 (1983); Ph. de Forcrand and C. Roiesnel, *ibid.* **151B**, 79 (1985).

[30] For an example in which a purification of a mixed state is generated by a quantum circuit, see P. Hyllus, C. M. Alves, D. Bruss, and C. Macchiavello, Phys. Rev. A **70**, 032316 (2004).

[31] M. Poźniak, K. Życzkowski, and M. Kuś, J. Phys. A **31**, 1059 (1998).

[32] J. Emerson, E. Livine, and S. Lloyd, Phys. Rev. A **72**, 060302(R) (2005).

[33] H. Aschauer, W. Dür, and H.-J. Briegel, Phys. Rev. A **71**, 012319 (2005).

[34] B. Kraus, N. Gisin, and R. Renner, Phys. Rev. Lett. **95**, 080501 (2005).

[35] D. Gottesman, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1997 (unpublished).

[36] D. Gottesman, Phys. Rev. A **57**, 127 (1998).

[37] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Am. J. Phys. **58**, 1131 (1990).

[38] M. Hein, J. Eisert, and H. J. Briegel, Phys. Rev. A **69**, 062311 (2004).

[39] W. Dür, H. Aschauer, and H.-J. Briegel, Phys. Rev. Lett. **91**, 107903 (2003).