On the distance between unitary propagators of quantum systems of differing dimensions

Robert L. Kosut
SC Solutions, Sunnyvale, CA
kosut@scsolutions.com

Matthew Grace, Constantin Brif, Herschel Rabitz
Department of Chemistry
Princeton University, Princeton, NJ

Abstract
A distance measure is presented between two unitary propagators of quantum systems of differing dimensions along with a corresponding method of computation. A typical application is to compare the propagator of the actual (real) process with the propagator of the desired (ideal) process; the former being of a higher dimension than the latter. The proposed measure has the advantage of dealing with possibly correlated inputs, but at the expense of working on the whole space and not just the information bearing part as is usually the case, i.e., no partial trace operation is explicitly involved. It is also shown that the distance measure and an average measure of channel fidelity both depend on the size of the same matrix: as the matrix size increases, distance decreases and fidelity increases.

1 Introduction
Much of engineering design is predicated on having a model of the process which captures the essential features (e.g., the design variables and sources of uncertainty), a model of the desired process, and a means of comparing the two. In addition, it is also desirable that the same means of comparison can be determined or validated from experimental data. Distance and fidelity measures are proposed in [GLN05] for comparing what are referred to as real and ideal quantum processes represented as quantum operations [NC00]. Several measures are proposed in [GLN05] along with a set of criteria which aim at making the measure tractable both theoretically as well as experimentally. Here we focus more on the theoretical aspect and in particular using unitary propagators corresponding to the real and ideal systems. The proposed distance measure does, however, have an input output interpretation, and thus, in principal, it may be amenable to an experimental calculation although that is not pursued here.

2 Problem formulation
Consider a finite dimensional, closed, bipartite quantum system consisting of a part, $S$, referred to as the system, which bears the quantum information of interest, coupled to a part $B$, representing the bath or environment. The Hilbert space of the bipartite system is $\mathcal{H}_{SB} = \mathcal{H}_S \otimes \mathcal{H}_B$ where $\mathcal{H}_S$ and $\mathcal{H}_B$ are finite dimensional with $n_S = \dim \mathcal{H}_S$, $n_B = \dim \mathcal{H}_B$, and hence, $n_{SB} = \dim \mathcal{H}_{SB}$. Let $U \in \mathbb{C}^{n_{SB} \times n_{SB}}$ denote the unitary propagator of the (real) bipartite system, and let $G \in \mathbb{C}^{n_S \times n_S}$ denote the unitary propagator of an ideal system acting only on $\mathcal{H}_S$. The question addressed here is this:

How close is $U$ to $G$?

Since it is implicitly assumed that experimental access is confined to the $S$-system, it follows that performance is judged only by how well the reduced state output mimics the desired behavior in $\mathcal{H}_S$ as
represented by $G$. If, for example, $U$ decomposes into a tensor product of unitaries, i.e., $U = G \otimes U_B$, then the $S$ and $B$ systems do not interact and the system is indistinguishable from $G$ acting on $\mathcal{H}_S$. This motivates expressing $U$ as,

$$U = G \otimes \Phi + R$$

(1)

where $\Phi \in \mathbb{C}^{n_B \times n_B}$ is unitary. Since any choice of $\Phi$ determines $R$, it follows that it is always possible to express $U$ in this form. Hence, consider the following distance measure between $U \in \mathbb{C}^{n_S n_B \times n_S n_B}$ and $G \in \mathbb{C}^{n_S \times n_S}$,

$$d(U, G) = \min_{\Phi} \left\{ \| U - G \otimes \Phi \| : \Phi \Phi^\dagger = I_B \right\}$$

(2)

where $\| \cdot \|$ is any matrix norm on $\mathbb{C}^{n_S n_B \times n_S n_B}$. The term inside the norm is $R$ as defined in (1).

The unitary $\Phi \in \mathbb{C}^{n_B \times n_B}$ is thus used solely to find the closest propagator to $U$ in the subspace $\{ G \otimes \Phi : \Phi \Phi^\dagger = I_B \}$ where “closeness” is determined by the choice of norm.

It is clear that using (2) as a distance measure does not consider the partial trace operation which is implicit in other distance (and fidelity) measures via the operator-sum-representation (OSR), e.g., [NC00], [KW04], [GLN05]. As a result (2) is in general a conservative measure. More specifically, (2) is measuring distance over the whole space $\mathcal{H}_S \times \mathcal{H}_B$ rather than just over $\mathcal{H}_S$. For example, suppose the (possibly correlated) pure state $|\psi\rangle$ is acted upon by $U$ and its output is compared to the ideal output of $G \otimes \Phi$ acting on $|\psi\rangle$. The output state error is,

$$e = (U - G \otimes \Phi) |\psi\rangle$$

(3)

It follows that

$$\max_{\langle \psi | \psi \rangle = 1} \langle \psi | e^\dagger e \rangle = \| U - G \otimes \Phi \|_2^2$$

(4)

where $\| \cdot \|_2$ is the induced two-norm, i.e., the maximum singular value of the matrix argument. Suppose the input is random in the sense that $|\psi\rangle = |\psi_i\rangle$ with probability $p_i$ for $i = 1, \ldots, \ell$. Let $E$ denote the expected value operator taken with respect to the input distribution. Then,

$$\max_{\text{Prob}} \left\{ \langle \psi | \psi \rangle = p_i \right\} E \langle \psi | e^\dagger e \rangle = \max_{\rho \geq 0, \text{Tr} \rho = 1} \text{Tr} \left( (U - G \otimes \Phi)^\dagger (U - G \otimes \Phi) \rho \right)$$

$$\max_{\rho \geq 0, \text{Tr} \rho = 1} \text{Tr} \left( (U - G \otimes \Phi)^\dagger (U - G \otimes \Phi) \rho \right)$$

(5)

where $\| \cdot \|_F$ is the Frobenius matrix norm: $\| X \|_F = (\text{Tr} X^\dagger X)^{1/2} = \left( \sum_{i,j} |X_{ij}|^2 \right)^{1/2} = \left( \sum_{i} \sigma_i(X)^2 \right)^{1/2}$ with $\sigma_i(X)$ the $i$-th singular value of $X$. Since the error is defined over states operating on the whole of $\mathcal{H}_S \times \mathcal{H}_B$, (4)-(5) are upper bounds on the error when the actual input is uncorrelated, that is, when either $|\psi\rangle = |\psi_S\rangle \otimes |\psi_B\rangle$ or $\rho = \rho_S \otimes \rho_B$. Despite this, however, there is a benefit: using either the maximum singular value or the Frobenius norm in (2) gives an indication of both the effect of environmental input uncertainty (including unwanted correlations with the information input) and the “distance” of the actual system to an ideal. In the remainder of this note we will show how to compute the distance measure using both of the above matrix norms. We also explore the relation of this distance measure to a typical fidelity measure.

\footnote{If the $S$ and $B$ channels are ordered reversely, then $U = \Phi \otimes G + R$.}
3 Computing the distance measure with the Frobenius norm

Consider the distance measure
\[
d_F(U, G) = \min_{\Phi} \left\{ \frac{1}{\sqrt{2nSn_B}} \|U - G \otimes \Phi\|_F \mid \Phi^\dagger \Phi = I_B \right\}
\]  
(6)

where \(1/\sqrt{2nSn_B}\) is a normalization factor which keeps \(d_F(U, G)\) in the range \([0, 1]\). As shown below,

\[
d_F(U, G) = \left( 1 - \frac{1}{nSn_B} \|\Gamma\|_{\text{Tr}} \right)^{1/2}
\]  
(7)

\[
\Gamma = \sum_{i,j=1}^{n_S} G_{ij}^* U_{(ij)} \in \mathbb{C}^{n_B \times n_B}
\]

where \(\|\cdot\|_{\text{Tr}}\) is the matrix trace-norm,\(^2\) the \(\{U_{(ij)}\}\) are \(n_B \times n_B\) matrix partitions of \(U\),

\[
U = \begin{bmatrix}
U_{(11)} & \cdots & U_{(1n_S)} \\
\vdots & \ddots & \vdots \\
U_{(n_S1)} & \cdots & U_{(n_Sn_S)}
\end{bmatrix} \in \mathbb{C}^{nSn_B \times nSn_B}
\]  
(8)

and \(\{G_{ij}\}\) are the (scalar) elements of \(G\),

\[
G = \begin{bmatrix}
G_{11} & \cdots & G_{1n_S} \\
\vdots & \ddots & \vdots \\
G_{n_S1} & \cdots & G_{n_Sn_S}
\end{bmatrix} \in \mathbb{C}^{n_S \times n_S}
\]  
(9)

Proof

From the definition (6),

\[
\frac{1}{2nSn_B} \|U - G \otimes \Phi\|_F^2 = 1 - \frac{1}{nSn_B} \text{ Re } \text{Tr} \ (G^\dagger \otimes \Phi^\dagger)U
\]

\[
= 1 - \frac{1}{nSn_B} \text{ Re } \text{Tr} \sum_{i,j=1}^{n_S} G_{ij}^* U_{(ij)} \Phi^\dagger
\]  
(10)

\[
= 1 - \frac{1}{nSn_B} \text{ Re } \text{Tr} \Gamma \Phi^\dagger
\]

The first line above follows from the fact that \(U, G, \Phi\) are all unitary. The last two lines follow from the definition of the tensor product together with (7)-(9). To compute \(d_F(U, G)\) is equivalent to finding the maximum value of \(\text{Re } \text{Tr} \Gamma \Phi^\dagger\) over all unitary \(\Phi \in \mathbb{C}^{n_B \times n_B}\). A singular value decomposition of \(\Gamma\) gives \(\Gamma = WSV^\dagger\) with unitary \(W\) and \(V\) and with the singular values in the diagonal matrix \(S = \text{diag}(\sigma_1, \ldots, \sigma_{n_B}) \in \mathbb{R}^{n_B}\), \(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{n_B} \geq 0\). This gives, \(\text{Re } \text{Tr} \Gamma \Phi^\dagger = \text{Re } \text{Tr} WSV^\dagger \Phi^\dagger = \text{Re } \text{Tr} SV^\dagger \Phi^\dagger W = \text{Re } \text{Tr} SB\) with \(B = V^\dagger \Phi^\dagger W\). Clearly \(\Phi\) is unitary if and only if \(B\) is unitary. Hence, \(\text{Re } \text{Tr} SB = \sum_{i=1}^{n_B} s_i \text{ Re } B_{ii}\) will achieve the maximum, \(\text{Re } \text{Tr} S = \text{Tr} S = \|\Gamma\|_{\text{Tr}}\) when \(\text{Re } B = I\), if and only if \(B = I\), which is of course unitary. This completes the proof of (7).

\(^2\)For \(X \in \mathbb{C}^{n \times m}\), \(\|X\|_{\text{Tr}} = \text{Tr} \sqrt{X^\dagger X} = \sum_{i=1}^{\min(n,m)} \sigma_i(X)\) where \(\sigma_i(X)\) is the \(i\)-th singular value of \(X\).
Reversing the channel ordering

By an analogous argument, if the $S$ and $B$ channels are ordered reversely from (1), then the distance measure becomes,

\[
d_F(U, G) = \min_\Phi \left\{ \frac{1}{\sqrt{2n_S n_B}} \|U - \Phi \otimes G\|_\Phi \left| \Phi^\dagger \Phi = I_B \right. \right\}
\]

\[
= \left( 1 - \frac{1}{n_S n_B} \|\Gamma\|_{\text{Tr}} \right)^{1/2}
\]

\[
\Gamma_{ij} = \text{Tr} \ G^\dagger U_{(ij)}, \ i, j = 1, \ldots, n_B
\]

where now $\{U_{(ij)}\}$ are $n_S \times n_S$ matrix partitions of $U$.

\[
U = \begin{bmatrix} U_{(11)} & \cdots & U_{(n_B)} \\ \vdots & \ddots & \vdots \\ U_{(n_B 1)} & \cdots & U_{(n_B n_B)} \end{bmatrix} \in \mathbb{C}^{n_S n_B \times n_S n_B}
\]

Exact tensor product

Suppose $U$ is exactly a tensor product, i.e.,

\[
U = U_S \otimes U_B,
\]

Then,

\[
U_{(ij)} = (U_S)_{ij} U_B
\]

\[
\Gamma = \sum_{i,j=1}^{n_S} G^*_{ij} (U_S)_{ij} U_B = (\text{Tr} \ G^\dagger U_S) U_B
\]

\[
\|\Gamma\|_{\text{Tr}} = \text{Tr} \sqrt{\text{Tr} \ G^\dagger U_S}^2 U_B U_B^\dagger = |\text{Tr} \ G^\dagger U_S| \text{Tr} I_B = n_B |\text{Tr} G^\dagger U_S|
\]

The distance measure becomes,

\[
d_F(U_S \otimes U_B, G) = \left( 1 - \frac{1}{n_S} |\text{Tr} G^\dagger U_S| \right)^{1/2}
\]

If $U$ and $G$ have the same dimensions, that is, $n_B = 1$, then the distance is zero if and only if $U (\equiv U_S)$ and $G$ differ by a scalar phase, i.e., $U_S = e^{i\phi} G$ for some real phase angle $\phi$.

4 Computing the distance measure with the maximum singular value

Consider the distance measure

\[
d_2(U, G) = \min_\Phi \left\{ \frac{1}{\sqrt{2}} \|U - G \otimes \Phi\|_2 \left| \Phi^\dagger \Phi = I_B \right. \right\}
\]

where $1/\sqrt{2}$ is a normalization factor which keeps $d_2(U, G)$ in the range $[0, 1]$. In this case we have not been able to obtain an exact solution, but we can establish the following bounds:

\[
\frac{1}{\sqrt{2}} \|U - G \otimes \Phi\|_2 \leq d_2(U, G) \leq \frac{1}{\sqrt{2}} \|U - G \otimes \hat{\Phi}\|_2
\]
where $\Phi$ is the solution to the convex optimization problem:

$$\begin{align*}
\text{minimize} & \quad \|U - G \otimes \Phi\|_2 \\
\text{subject to} & \quad \Phi^\dagger \Phi \leq I_B
\end{align*}$$

and where $\hat{\Phi}$ is obtained from $\Phi$ via the singular value decomposition of $\Phi$ as follows:

$$\Phi = VSW^\dagger \Rightarrow \hat{\Phi} = VW^\dagger$$

The upper and lower bounds in (17) will be close to each other if the singular values of $\Phi$ are close to unity, that is, if $\Phi$ is close to a unitary.

**Proof**

Problem (18) is a convex optimization because any norm is a convex function, the argument in the norm is affine in the optimization variable $\Phi$, and the constraint set $\Phi^\dagger \Phi \leq I_B$ is convex [BV04]. However, there is no guaranty that the resulting optimizer $\Phi$ is unitary, i.e., on the boundary of the constraint. Since $\Phi$ in (18) is less constrained than in (16), it follows that the lower bound in (17) applies. In consequence, consider $\hat{\Phi}$ from (19) as a unitary approximation to $\Phi$, and since it is not necessarily the optimal solution to (16), the upper bound in (17) follows which completes the proof.

5 Application to optimal control or system design

In the ideal case where $U$ and $G$ are of the same dimension, the term $|\text{Tr } G^\dagger U_S|$ in (15) has often been proposed as an objective function to be maximized for optimal control design (e.g., [PK02]) and also for determining properties of the control landscape [RHR04]. This is equivalent to posing the distance measure (15) as an objective function to be minimized. Where $U$ and $G$ are not of the same dimension, the distance measure (7) has been reported in [GBR+] for optimal control design.

In general, for either control design or system design, $U$ will depend on some parameters. In many cases these parameters are constants, e.g., coefficients of specified time functions which make up the control field, settings of wave-plate angles in a photonic device, geometry variables in a circuit layout, and so on. In these cases, $U \equiv U(\theta)$ where $\theta$ is a (constant) vector to be selected out of a set $\Theta$ to minimize the distance measure (2). Equivalently, consider the following optimization problem:

$$\begin{align*}
\text{minimize} & \quad \|U(\theta) - G \otimes \Phi\| \\
\text{subject to} & \quad \Phi^\dagger \Phi = I_B, \quad \theta \in \Theta
\end{align*}$$

where $\|\cdot\|$ is any matrix norm and the optimization variables are $\Phi$ and $\theta$. Although typically $\Theta$ is a convex set, e.g., $\Theta = \{ \theta \mid \|\theta - \theta_0\| \leq \delta \}$, in general (20) is not a convex optimization. In the first place, it is almost never the case that $U(\theta)$ is a convex function of $\theta$. Suppose for example, that the Hamiltonian which generates $U(\theta)$ depends linearly on $\theta$, i.e., $H(\theta,t) = \sum_k \theta_k H_k(t)$. If $U(\theta)$ is the associated propagator which makes its appearance at some time $\tau$, then it is unlikely that $U(\theta)$ is convex over $\Theta$. Secondly, the constraint $\Phi^\dagger \Phi = I_B$ is not a convex set. This constraint, however, can be relaxed to the convex set $\Phi^\dagger \Phi \leq I_B$ resulting in,

$$\begin{align*}
\text{minimize} & \quad \|U(\theta) - G \otimes \Phi\| \\
\text{subject to} & \quad \Phi^\dagger \Phi \leq I_B, \quad \theta \in \Theta
\end{align*}$$

Since $\|U(\theta) - G \otimes \Phi\|$ for any norm is a convex function of $\Phi$ and $U(\theta)$, and $\Phi^\dagger \Phi \leq I_B$ is a convex set in $\Phi$, it follows that (21) is a convex optimization over $U(\theta)$ and $\Phi$ (see, e.g., [BV04, §3.25]). Again,
this is not a convex optimization over \( \theta \) because \( U(\theta) \) is not a convex function of \( \theta \). Nevertheless, the following iterative scheme will always find a local solution by reducing the distance measure in every step.

**Initialize** \( \hat{\theta} = \theta_0 \)

**Repeat**
1. \( \hat{\Phi} = \arg \min_{\Phi} \left\{ \| U(\hat{\theta}) - G \otimes \Phi \| \mid \Phi^\dagger = I_B \right\} \)
2. \( \hat{\theta} = \arg \min_{\theta} \left\{ \| U(\theta) - G \otimes \hat{\Phi} \| \mid \theta \in \Theta \right\} \)

**Until** \( \| U(\hat{\theta}) - G \otimes \hat{\Phi} \| \) stops decreasing.

In step 1 if the Frobenius norm is used then \( \hat{\Phi} \) can be calculated exactly (7). Under the two-norm we would use the approximation from (19). Step 2 requires using a local solver.

## 6 Distance and fidelity

We will show that the distance measure defined here and a typical measure of fidelity both depend on the size of the matrix \( \Gamma \) as given by either (7) or (11). There are many ways that fidelity has been defined to compare the desired unitary \( G \) with a quantum channel, *e.g.*, [NC00], [KW04], [GLN05]. Specifically, let \( \mathcal{S} \) denote a trace-preserving quantum channel mapping states \( \rho_S \in \mathcal{H}_S \) to \( \tilde{\rho}_S \in \mathcal{H}_S \) with the operator-sum-representation (OSR),

\[
\tilde{\rho}_S = S(\rho_S) = \sum_k S_k \rho_S S_k^\dagger, \quad \sum_k S_k^\dagger S_k = I_S
\]

with (matrix) operation elements \( S_k \in \mathbb{C}^{n_S \times n_S} \). Consider, for example, the worst-case pure state fidelity,

\[
f(S, G) = \min_{|\psi_S\rangle} (G|\psi_S\rangle)^\dagger \tilde{\rho}_S(G|\psi_S\rangle) = \min_{|\psi_S\rangle} \sum_k |\langle\psi_S|G^\dagger S_k|\psi_S\rangle|^2
\]

where \( \tilde{\rho}_S = \sum_k S_k|\psi_S\rangle\langle\psi_S|S_k^\dagger \) is the reduced output state of \( S \) with \( |\psi_S\rangle \) the pure input state. In general calculating \( f(S, G) \) is not easy as it is not a convex optimization; it is, however, bounded as follows:

\[
f(S, G) \leq f(S, G) \leq \overline{f}(S, G)
\]

where

\[
\overline{f}(S, G) = \frac{1}{n_S^2} \sum_k |\text{Tr} G^\dagger S_k|^2
\]

\[
f(S, G) = \min_{\rho_S} \sum_k |\text{Tr} G^\dagger S_k \rho_S|^2
\]

All these fidelities are in \([0,1]\) and equal to one if and only if \( S(\rho_S) = G \rho_S G^\dagger \) for all densities \( \rho_S \). Both bounds are easy to obtain: \( \overline{f}(S, G) \) is direct and \( f(S, G) \) is a convex optimization over all densities (i.e., over all \( \rho_S \in \mathbb{C}^{n_S \times n_S}, \rho_S \geq 0, \text{Tr} \rho_S = 1 \)), and hence, can be numerically obtained. If the worst-case density associated with \( f(S, G) \) is nearly rank one, then \( \overline{f}(S, G) \approx f(S, G) \).

In order to relate \( U \) to \( S \) assume that the input state to \( U \) is the pure uncorrelated state \( |\psi_B\rangle \otimes |\psi_S\rangle \) where \( |\psi_B\rangle \) has elements \( \psi_{Bi}, i = 1, \ldots, n_B \) with \( \sum_i |\psi_{Bi}|^2 = 1 \). It then follows that

\[
S_k = \sum_{i=1}^{n_B} \psi_{Bi} U(\langle ki |), \quad k = 1, \ldots, n_B
\]

\(^3\text{For convenience we use what we referred to previously as the reversed channel ordering.}\)
where the \( U_{(ki)} \) are the \( n_S \times n_S \) matrix partitions of \( U \) as given by (12). The upper bound in (24) is then,
\[
\mathcal{F}(U,G) = \frac{1}{n_S^2} \sum_{i,j=1}^{n_B} \langle \psi_B | \psi_B \rangle \{ \sum_{k=1}^{n_B} \Gamma_{ki} \Gamma_{kj}^* \} = \frac{1}{n_S^2} \langle \psi_B | \Gamma^\dagger \Gamma | \psi_B \rangle
\]  
(27)
with \( \Gamma_{ki} = \text{Tr} \ G^\dagger U_{(ki)} \) from (11). This upper bound on pure state fidelity thus depends on \( \Gamma \) weighted by the state of the environment \( |\psi_B \rangle \). Suppose the environment state is completely random, that is, \( \mathbb{E} \{ \langle \psi_B | \psi_B \rangle \} = I_B/n_B \). Then the average value of the pure state fidelity upper bound becomes,
\[
\mathbb{E} \{ \mathcal{F}(U,G) \} = \frac{1}{n_S^2 n_B} \sum_{k,j=1}^{n_B} \| \Gamma_{ki} \|^2 = \frac{1}{n_S^2 n_B} \| \Gamma \|^2_F
\]  
(28)
The same result holds with \( \Gamma \) defined by (7). Thus the distance measure presented here and a typical channel fidelity both depend on the matrix \( \Gamma \) defined by either (7) or (11), depending on how channel ordering is ascribed. Increasing the size of \( \Gamma \) increases fidelity and decreases distance.

We also mention that the inequality relation of distance and fidelity described in [GLN05] also holds here as well, namely,
\[
(1 - d_F(U,G))^2 \leq \mathbb{E} \{ \mathcal{F}(U,G) \} \leq 1 - d_F(U,G)^2
\]  
(29)

**Proof**

We first show that the bounds follow from \( \| \Gamma \|_2 = \max_i \sigma_i \leq n_S \) where \( \sigma_i \) is the \( i \)-th singular value of \( \Gamma \). Then we will show that \( \| \Gamma \|_2 \leq n_S \) holds.

The upper bound holds if and only if \( \| \Gamma \|_F^2 \leq n_S \| \Gamma \|_T^2 \), or equivalently, \( \sum_i \sigma_i^2 \leq n_S \sum_i \sigma_i \). Assuming that \( \sigma_i \leq n_S \), it immediately follows that \( \sum_i \sigma_i^2 \leq n_S \sum_i \sigma_i \) which establishes the upper bound.

The lower bound is equivalent to \( \| \Gamma \|_T + (1/n_S) \| \Gamma \|_F^2 \leq 2 \sqrt{n_B} \| \Gamma \|_F \). Assuming again that \( \| \Gamma \|_2 \leq n_S \), and hence, \( \| \Gamma \|_F^2 \leq n_S \| \Gamma \|_T^2 \), the lower bound will hold if \( \| \Gamma \|_T \leq \sqrt{n_B} \| \Gamma \|_F \). This is a known inequality for any matrix and thus establishes the lower bound.

To show that \( \sigma_i \leq n_S \), observe that since the elements of \( \Gamma \) (11) are obtained via the trace operation, then by definition they are equivalently obtained from the sum of the diagonal elements of \( G^\dagger U_{(ij)} \). Let \( U' \) denote the matrix whose \( n_S \times n_S \) matrix partitions, \( U'_{(ij)} \), are diagonal, with diagonal elements the diagonals of \( G^\dagger U_{(ij)} \). Thus, \( \Gamma_{ij} = \text{Tr} \ G^\dagger U_{(ij)} = \text{Tr} \ U'_{(ij)} \). Now replace the trace operation by the vector products, that is, \( \Gamma_{ij} = w_i^T U' w_j \) where \( w_i \) is an \( n_S n_B \times 1 \) vector with \( n_S \) ones in elements \( 1 + (i - 1)n_S \) to \( in_S \) and zeros elsewhere. Hence we can write \( \Gamma = W^T U' W \) where \( W = [w_1 \cdots w_{n_B}] \) is \( n_S n_B \times n_B \). Observe that \( \| W \|_2 = \sqrt{n_S} \) and since \( U' \) is constructed from elements of unitary matrices \( G \) and \( U \), then \( \| U' \|_2 \leq 1 \). It therefore follows that \( \| \Gamma \|_2 = \| W^T U' W \|_2 \leq \| W \|_2^2 \| U' \|_2 \leq n_S \). This completes the proof.
Acknowledgements

This work has been funded under the DARPA QuIST Program (Quantum Information Science & Technology). R. Kosut would like to thank Abbas Emami-Naeini (SC Solutions), Paul van Doore (Catholic University of Louvain), and Oliver Kosut (Cornell University) for help with the proof of (29).

References

[BV04] S. Boyd and L. Vandenberghe. Convex Optimization. Cambridge University Press, 2004.

[GBR+10] M. Grace, C. Brif, H. Rabitz, I. Walmsley, R. Kosut, and D. Lidar. Optimal control of logical operations in the presence of decoherence: a two-spin model. Presented at the APS Meeting, Baltimore, MD, March 13-17, 2006.

[GLN05] A. Gilchrist, N. K. Langford, and M. A. Nielsen. Distance measures to compare real and ideal quantum processes. Phys. Rev. A, 71(062310), June 2005. arXiv quant-ph/0408063.

[KW04] D. Kretschmann and R. F. Werner. Tema con varizioni: quantum channel capacity. New Journal of Physics, 6, 26, 2004.

[NC00] M. A. Nielsen and I. L. Chuang. Quantum Computation and Quantum Information. Cambridge, 2000.

[PK02] J. P. Palao and R. Kosloff. Quantum computing by an optimal control algorithm for unitary transformations. Phys. Rev. Lett., 89(188301), 2002.

[RHR04] H. Rabitz, M. Hsieh, and C. Rosenthal. Quantum optimally controlled transition landscapes. Science, 303, 998, 2004.