On the complexity of approximating the diamond norm

Avraham Ben-Aroya* Amnon Ta-Shma†

Abstract

The *diamond norm* is a norm defined over the space of quantum transformations. This norm has a natural operational interpretation: it measures how well one can distinguish between two transformations by applying them to a state of arbitrarily large dimension. This interpretation makes this norm useful in the study of quantum interactive proof systems.

In this note we exhibit an efficient algorithm for computing this norm, using convex programming. Independently of us, Watrous [Wat09] recently showed a different algorithm to compute this norm.

An immediate corollary of this algorithm is a slight simplification of the argument of Kitaev and Watrous [KW00] that QIP $\subseteq$ EXP.

1 Introduction

How well can one distinguish two quantum transformations? Imagine we have access to some unknown admissible super-operator T and we want to distinguish the case it is T_1 from the case it is T_2 (T_1 and T_2 are known). One possible test to distinguish T_1 from T_2 is preparing an input state $\rho \in D(\mathcal{V})$, applying T on ρ and measuring the result. This corresponds to:

$$\sup \{ \|T_1\rho - T_2\rho\|_{\mathrm{tr}} \ : \ \rho \in D(\mathcal{V}) \}.$$

However, somewhat surprisingly, it turns out that often one can distinguish T_1 and T_2 better, by taking an auxiliary Hilbert space $\mathcal{A}$, preparing an *entangled* input state $\rho \in D(\mathcal{V}\otimes\mathcal{A})$, applying T on the $\mathcal{V}$ register of ρ and then measuring the global result. Therefore, we define:

$$\mathrm{dist}(\rho_1,\rho_2) \ = \ \sup \left\{ \left\| (T_1\otimes I_{L(\mathcal{A})})\rho - (T_2\otimes I_{L(\mathcal{A})})\rho \right\|_{\mathrm{tr}} \ : \ \dim(\mathcal{A}) < \infty, \ \rho \in D(\mathcal{V}\otimes\mathcal{A}) \right\}.$$

Kitaev [Kit97] proved that this phenomena is restricted by dimension and the maximum is attained already with an auxiliary Hilbert space $\mathcal{A}$ of dimension $\dim(\mathcal{A}) \leq \dim(\mathcal{V})$. Define the following functions on general (not necessarily admissible) super-operators $T : L(\mathcal{V}) \to L(\mathcal{W})$:

$$\begin{aligned} \|T\|_{\mathrm{tr}} \ &= \ \sup \{ \|T(X)\|_{\mathrm{tr}} : X \in L(\mathcal{V}), \|X\|_{\mathrm{tr}} = 1 \} \text{ , and,} \\ \|T\|_{\diamond} \ &= \ \left\| T\otimes I_{L(\mathcal{V})} \right\|_{\mathrm{tr}}. \end{aligned}$$

Kitaev showed that both $\|\cdot\|_{\mathrm{tr}}$ and $\|\cdot\|_{\diamond}$ are norms, and furthermore $\mathrm{dist}(T_1,T_2) = \|T_1 - T_2\|_{\diamond}$.

The diamond norm naturally appears when studying the class QIP of languages having a single-prover, multi-round interactive proof protocol between an all-powerful prover and an efficient quantum

*Department of Computer Science, Tel-Aviv University, Tel-Aviv 69978, Israel. Supported by the Adams Fellowship Program of the Israel Academy of Sciences and Humanities, by the European Commission under the Integrated Project QAP funded by the IST directorate as Contract Number 015848 and by USA Israel BSF grant 2004390. Email: abrhambe@post.tau.ac.il.

†Department of Computer Science, Tel-Aviv University, Tel-Aviv 69978, Israel. Supported by the European Commission under the Integrated Project QAP funded by the IST directorate as Contract Number 015848, by Israel Science Foundation grant 217/05 and by USA Israel BSF grant 2004390. Email: amnon@tau.ac.il.

verifier. Kitaev and Watrous [KW00] showed that, without loss of generality, perfect completeness can be achieved and three rounds suffice (starting with the verifier). They also showed that the value of a three round quantum interactive protocol, can be expressed as $\|T\|_\diamond$, for some super-operator T that is naturally defined given the protocol of the verifier. They used this characterization, and the fact that $\|T_1 \otimes T_2\|_\diamond = \|T_1\|_\diamond \cdot \|T_2\|_\diamond$ to show perfect parallel amplification for QIP protocols. Finally, they showed that QIP $\subseteq$ EXP by reducing the problem to an exponential size semi-definite programming problem. Thus QIP is somewhere between PSPACE and EXP (the containment PSPACE = IP $\subseteq$ QIP is immediate).

Another connection between QIP and the diamond norm was given by Rosgen and Watrous [RW05]. They defined the promise problem $\mathsf{QCD}_{a,b}$ (quantum circuit distinguishability) whose input is two admissible super-operators T_1 and T_2, the "yes" instances are pairs (T_1, T_2) for which $\|T_1 - T_2\|_\diamond \geq a$ and the "no" instances are the pairs for which $\|T_1 - T_2\|_\diamond \leq b$. Rosgen and Watrous [RW05] proved that for every $a < b$ the problem $\mathsf{QCD}_{a,b}$ is QIP-complete (see also [Ros08]).

The work of Kitaev and Watrous, as well as the work of Rosgen and Watrous do not imply that approximating the diamond norm itself can be done in P. In this note we prove that the diamond norm can be computed by solving a convex optimization problem, and therefore it is in P. More precisely, if we are given as input a description of $T : L(\mathcal{V}) \to L(\mathcal{V})$ written as a matrix of dimensions $N^2 \times N^2$ (where $N = \dim(\mathcal{V})$), and we are given $\epsilon > 0$, then we can approximate $\|T\|_\diamond$ to within ϵ additive accuracy in time $\mathrm{poly}(N, \log \epsilon^{-1})$. Independently of us, Watrous [Wat09] recently showed a similar result using a semi-definite program.

This claim can also be used to simplify the (somewhat more complicated) proof given in [KW00] that QIP $\subseteq$ EXP. To see this, notice that Kitaev and Watrous already proved that the value of a three round quantum interactive proof game can be captured as the diamond norm of a natural super-operator T. Thus, given such a game, all we need to do is to explicitly write down the description of T (which can be done in PSPACE and therefore in time exponential in $\mathrm{poly}(n)$, where n is the input length of the QIP protocol) and then approximate its diamond norm, in time polynomial in $\exp(\mathrm{poly}(n))$.

Our proof is surprisingly simple. We use an equivalent formulation of the diamond norm, proved by Kitaev, and we notice that it gives a convex program using the joint concavity of the fidelity function. We use a representation for density matrices suggested by Liu [Liu06] in a different context for a similar purpose.

2 Preliminaries

Let $\mathcal{V}, \mathcal{W}$ be two Hilbert spaces. $\mathrm{Hom}(\mathcal{V}, \mathcal{W})$ denotes the set of all linear transformations from $\mathcal{V}$ to $\mathcal{W}$ and is a vector space of dimension $\dim(\mathcal{V}) \cdot \dim(\mathcal{W})$ equipped with the Hilbert-Schmidt inner product $\langle T_1, T_2 \rangle = \mathrm{Tr}(T_1^\dagger T_2)$. $L(\mathcal{V})$ denotes $\mathrm{Hom}(\mathcal{V}, \mathcal{V})$. Let $\{|i\rangle\}$ denote the standard basis for $\mathcal{V}$. The set

$$\{|i\rangle \langle j| \ : \ 1 \leq i, j \leq \dim(\mathcal{V})\}$$

is an orthonormal basis of $L(\mathcal{V})$. When $\dim(\mathcal{V}) = 2^n$, tensor products of Pauli operators form another natural basis for $L(\mathcal{V})$. The Pauli operators are

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \ \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The set $\{\sigma_{i_1} \otimes \ldots \otimes \sigma_{i_n} \ : \ 0 \leq i_1, \ldots, i_n \leq 3\}$ is an orthogonal basis of $L(\mathcal{V})$, and all basis elements have eigenvalues ± 1. For a linear operator L, the spectral norm of L is

$$\|L\| \stackrel{\mathrm{def}}{=} \sup_{x : \|x\| = 1} x^\dagger A^\dagger A x$$

and is equal to the largest singular value of L. For any Pauli operator P, $\|P\| = 1$.

A pure-state is a vector in some Hilbert space. A general quantum system is in a *mixed state*–a probability distribution over pure-states. Let $\{p_i, |\phi_i\rangle\}$ denote the mixed-state in which the pure-state $|\phi_i\rangle$ occurs with probability p_i. The behavior of the mixed-state $\{p_i, |\phi_i\rangle\}$ is completely characterized by its *density matrix* $\rho = \sum_i p_i |\phi_i\rangle\langle\phi_i|$, in the sense that two mixed states with the same density matrix behave the same under any physical operation. Notice that a density matrix over a Hilbert space $\mathcal{V}$ belongs to $L(\mathcal{V})$. Density matrices are positive semi-definite operators and have trace 1. We denote the set of density matrices over $\mathcal{V}$ by $D(\mathcal{V})$.

Trace norm and fidelity. The *trace norm* of a matrix A is defined by

$$\|A\|_{\mathrm{tr}} = \mathrm{Tr}(|A|) = \mathrm{Tr}(\sqrt{A^\dagger A}),$$

which is the sum of the magnitudes of the singular values of A. One way to measure the distance between two density matrices ρ_1 and ρ_2 is by their trace distance $\|\rho_1 - \rho_2\|_{\mathrm{tr}}$. Another useful alternative to the trace metric as a measure of closeness of density matrices is the *fidelity*. For two positive semi-definite operators ρ_1, ρ_2 on the same finite dimensional space $\mathcal{V}$ (not necessarily having trace 1) we define

$$F(\rho_1, \rho_2) \;=\; \left[\mathrm{Tr}\left(\sqrt{{\rho_1}^{1/2}\,\rho_2\,{\rho_1}^{1/2}}\right)\right]^2 \;=\; \|\sqrt{\rho_1}\sqrt{\rho_2}\|_{\mathrm{tr}}^2 .$$

We remark that some authors define $\sqrt{F} = \left\|\sqrt{\rho_1}\sqrt{\rho_2}\right\|_{\mathrm{tr}}$ as the fidelity. Our definition is consistent with [KSV02]. $\sqrt{F}$ is jointly concave, i.e., for every set $\{(\rho_i, \xi_i)\}_{i=1}^k$ of pairs of density matrices and every $0 \le \lambda_1, \ldots, \lambda_k \le 1$ such that $\sum_{i=1}^k \lambda_i = 1$,

$$\sqrt{F}(\sum_{i=1}^k \lambda_i\rho_i, \sum_{i=1}^k \lambda_i\xi_i) \;\ge\; \sum_{i=1}^k \lambda_i\sqrt{F}(\rho_i, \xi_i).$$

A proof appears, e.g., in [NC00, Exercise 9.19].[1] We remark that F is not jointly concave (see [MPH$^+$08, Section 2] for a short survey on what is known about the fidelity function).

The diamond norm. Kitaev gave a different equivalent characterization of the diamond norm as follows. Any $T : L(\mathcal{V}) \to L(\mathcal{V})$ can be written as

$$T(X) = \mathrm{Tr}_{\mathcal{A}}(BXC^\dagger),$$

where $B, C \in \mathrm{Hom}(\mathcal{V}, \mathcal{V}\otimes\mathcal{A})$ and $\dim(\mathcal{A}) \le (\dim(\mathcal{V}))^2$ (see, e.g., [KSV02, page 110] or [Wat04, Lecture 4]). Define two completely positive super-operators $T_1, T_2 : L(\mathcal{V}) \to L(\mathcal{A})$:

$$T_1(X) \;=\; \mathrm{Tr}_{\mathcal{V}}(BXB^\dagger), \tag{1}$$
$$T_2(X) \;=\; \mathrm{Tr}_{\mathcal{V}}(CXC^\dagger). \tag{2}$$

Then the diamond norm of T can be written as

$$\|T\|_\diamond = \max\left\{\sqrt{F}(T_1(\rho), T_2(\xi)) \;:\; \rho, \xi \in D(\mathcal{V})\right\}.$$

The proof of this characterization can be found in [KSV02, Problem 11.10] or in Watrous' lecture notes [Wat04, Lecture 22, Theorem 22.2] (and notice that Watrous defines the fidelity function to be $\sqrt{F}$). Further information on the trace norm and the diamond norm of super-operators can be found in [KSV02].

Convex programming. Maximizing a convex function over a convex domain is, in general, NP-hard (see [FV95] for a survey). In sharp contrast to this, *convex programming*, which is the problem of *minimizing* a convex function over a convex domain, is in P. One of the reasons that convex programming is

[1] Note that in [NC00] the fidelity function is defined to be $\sqrt{F}$. In particular, the joint concavity of the fidelity function proved in [NC00, Exercise 9.19] proves joint concavity of $\sqrt{F}$ according to our notation

easier to solve is due to the fact that in a convex program any *local* optimum equals the *global* optimum. Special cases of convex programming are semi-definite programming and linear programming. Convex programming can be solved in polynomial time using the ellipsoid algorithm [Kha79] or interior-point methods. Often, these algorithms assume a *separation oracle*, i.e., an efficient procedure that given a point tells whether it belongs to the convex set, and if not, gives a half-space that separates the point from the convex set. However, the problem can also be solved using a *membership oracle* [GLS88] (a randomized algorithm is given in [BV04]).

Theorem 2.1. *[GLS88] There exists an algorithm that solves the following problem:*

Input *:*
- *A convex body K, given by a membership oracle,*
- *An integer n, rational numbers $R, r > 0$ and a vector $a_0 \in \mathbb{R}^n$ such that $B(a_0, r) \subseteq K \subseteq B(\overline{0}, R) \subseteq \mathbb{R}^n$.*
- *A convex function $f : K \to \mathbb{R}$ given by an oracle,*
- *A rational number $\epsilon > 0$.*

Output *: A value c such that $|c - \mathsf{opt}| \leq \epsilon$, where $\mathsf{opt} = \min_{x \in K} f(x)$.*

The algorithm runs in time $\mathrm{poly}(n, \log \epsilon^{-1}, \log(R/r))$, and, in particular, makes at most this number of calls to the membership oracle and to the f-oracle.

3 Approximating the diamond norm in P

Theorem 3.1. *Let $\mathcal{V}$ be a Hilbert space of dimension N. Let $T : L(\mathcal{V}) \to L(\mathcal{V})$ be a linear operator given as a matrix of dimension $N^2 \times N^2$ and let $\epsilon > 0$. Then there exists a polynomial time algorithm (in the input length of T and $\log \epsilon^{-1}$) that outputs a value c such that $| c - \|T\|_\diamond | \leq \epsilon$.*

Proof. We assume that N is a power of 2. The input is an operator $T : L(\mathcal{V}) \to L(\mathcal{V})$ given as an $N^2 \times N^2$ matrix. From T we can calculate T_1 and T_2 as in Equations (1) and (2). This computation can be done in time polynomial in the length of T, see, e.g., [Wat04, Lecture 4].

Next, we represent $\rho \in D(\mathcal{V})$ by its Pauli-basis coefficients, but excluding the identity coefficient which is always 1. Thus, we represent $\rho \in D(\mathcal{V})$ as a vector $v(\rho) \in \mathbb{R}^{N^2-1}$, where the ith coordinate of this vector is given by $v_i(\rho) = \mathrm{Tr}(P_{i+1}\rho)$, where P_i is the ith Pauli operator and $P_1 = I$. (Notice that $\mathrm{Tr}(P\rho) \in \mathbb{R}$ for Hermitian P and positive semi-definite ρ.) We let

$$K = \{(v(\rho), v(\xi)) \ : \ \rho, \xi \in D(\mathcal{V})\}.$$

The converse transformation $\Phi : K \to D(\mathcal{V})$ is defined by $\Phi(v) = I + \sum_{i=2}^{N^2} v_i P_i \in D(\mathcal{V})$. Notice that Φ is linear for convex sums, i.e., $\Phi(\sum \lambda_j v_j) = \sum \lambda_j \Phi(v_j)$ for $v_j \in K$ and $0 \leq \lambda_j \leq 1$ whose sum is 1. Clearly,

- K is convex,
- $\overline{0} \in K$,
- $K \subseteq B(\overline{0}, N^2)$, because $|\mathrm{Tr}(\rho P)| \leq \mathrm{Tr}(|\rho P|) \leq \|P\| \,\mathrm{Tr}(\rho) \leq 1$ for any $\rho \in D(\mathcal{V})$ and Pauli operator P,
- Given $v \in \mathbb{R}^{N^2-1}$ we can build the corresponding matrix $\rho = \Phi(v)$ in time polynomial in the length of v, and then check whether $\rho \in D(\mathcal{V})$. Thus we can determine membership in K in time polynomial in the input length.

Claim 3.1. $B(\overline{0}, \frac{1}{2\sqrt{N}}) \subseteq K$.

Proof. Let $v \in \mathbb{R}^{N^2-1}$ be such that $\|v\|_2 \leq \frac{1}{2\sqrt{N}}$ and let $\rho = \Phi(v) = I + \sum_{i=2}^{N^2} v_i P_i$. Clearly ρ is Hermitian and has trace 1. We are left to verify that ρ is positive semi-definite. Fix a unit vector $x \in \mathbb{R}^N$. Then,

$$\begin{aligned} x^\dagger \rho x &= x^\dagger I x + \sum_{i=2}^{N^2} v_i x^\dagger P_i x \geq 1 - \Big| \sum_{i=2}^{N^2} v_i x^\dagger P_i x \Big| \geq 1 - \sum_{i=2}^{N^2} |v_i| \cdot |x^\dagger P_i x| \\ &\geq 1 - \sum_{i=2}^{N^2} |v_i| \cdot \|P_i\| = 1 - \sum_{i=2}^{N^2} |v_i| \geq 1 - \sqrt{N} \|v\|_2 > 0. \end{aligned}$$

□

We define a target function $g : K \to [-1, 0]$ by $g(v, w) = -\sqrt{F}(T_1(\Phi(v)), T_2(\Phi(w)))$. Given $(v, w) \in K$ we can compute $T_1(\Phi(v)), T_2(\Phi(w))$ in time polynomial in N and the representation of the input T, hence g is computable in polynomial time (in its input length).

Claim 3.2. *g is convex over K.*

Proof. For every $0 \leq \lambda_1, \ldots, \lambda_k \leq 1$ such that $\sum_{j=1}^k \lambda_j = 1$,

$$\begin{aligned} g(\sum_{j=1}^k \lambda_j (v_j, w_j)) &= g(\sum_{j=1}^k \lambda_j v_j, \sum_{j=1}^k \lambda_j w_j) = -\sqrt{F}(T_1(\Phi(\sum_{j=1}^k \lambda_j v_j)), T_2(\Phi(\sum_{j=1}^k \lambda_j w_j))) \\ &= -\sqrt{F}(T_1(\sum_{j=1}^k \lambda_j \rho_j), T_2(\sum_{j=1}^k \lambda_j \xi_j)), \end{aligned}$$

where $\rho_j = \Phi(v_j)$, $\xi_j = \rho(w_j)$ and $\rho_j, \xi_j \in D(\mathcal{V})$. Thus, by the joint concavity of $\sqrt{F}$,

$$\begin{aligned} g(\sum_{j=1}^k \lambda_j (v_j, w_j)) &= -\sqrt{F}(\sum_{j=1}^k \lambda_j T_1(\rho_j), \sum_{j=1}^k \lambda_j T_2(\xi_j)) \\ &\leq -\sum_{j=1}^k \lambda_j \sqrt{F}(T_1(\rho_j), T_2(\xi_j)) = \sum_{j=1}^k \lambda_j g(v_j, w_j). \end{aligned}$$

□

Since $\|T\|_\diamond = \min_{(v,w)\in K} g(v, w)$, by Theorem 2.1 we can find a value $c \in \mathbb{R}$ that satisfies $| c - \|T\|_\diamond | \leq \epsilon$, as required. □

References

[BV04] D. Bertsimas and S. Vempala. Solving convex programs by random walks. *Journal of the ACM*, 51(4):540–556, 2004.

[FV95] C.A. Floudas and V. Visweswaran. Quadratic optimization. *Handbook of Global Optimization*, pages 217–269, 1995.

[GLS88] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, New York, 1988.

[Kha79] L. G. Khachiyan. A polynomial algorithm in linear programming. *Soviet Mathematics Doklady*, 20:191–194, 1979.

[Kit97] A.Y. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997.

[KSV02] A.Y. Kitaev, A. Shen, and M.N. Vyalyi. *Classical and Quantum Computation*. American Mathematical Society, 2002.

[KW00] A. Kitaev and J. Watrous. Parallelization, amplification, and exponential time simulation of quantum interactive proof systems. In *ACM Symposium on Theory of Computing (STOC)*, pages 608–617, 2000.

[Liu06] Y.K. Liu. Consistency of local density matrices is QMA-complete. In *APPROX-RANDOM*, pages 438–449, 2006.

[MPH+08] J. A. Miszczak, Z. Puchała, P. Horodecki, A. Uhlmann, and K. Życzkowski. Sub– and super–fidelity as bounds for quantum fidelity. Technical report, quant-ph/0805.2037, 2008.

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000.

[Ros08] B. Rosgen. Distinguishing short quantum computations. In *Proceedings of the 25th International Symposium on Theoretical Aspects of Computer Science*, pages 597–608, 2008.

[RW05] B. Rosgen and J. Watrous. On the hardness of distinguishing mixed-state quantum computations. In *The Annual IEEE Conference on Computational Complexity (COMP)*, pages 344–354, 2005.

[Wat04] J. Watrous. Advanced topics in quantum information processing. Lecture notes, 2004. http://www.cs.uwaterloo.ca/~watrous/lecture-notes/701.

[Wat09] J. Watrous. Semidefinite programs for completely bounded norms, 2009. arXiv:0901.4709.