On Norms and Unitary Design

Ranyiliu Chen

September 27, 2020

This note contains things I know (so far) about unitary design, with some requisite preliminaries about norms.

1 Preliminary: norm for operators and super-operators

To distinguish norms defined on different object (vector, operator, super-operator), in this note I use $|\cdot|$, $\|\cdot\|$, $|||\cdot|||$ to denote norms on a vector, an operator, and a super-operator, respectively.

1.1 The Schatten norm

The matrix norm (also called *operator norm*) I learned before in *Linear Algebra* can be regarded as a norm induced by the vector norm. That is to say, for any normed vector space V with vector norm $|\cdot|_n$, we can always define a norm for the linear operators on V:

$$||A||_n := \sup_{v\in V, |v|\neq 0} \frac{|Av|_n}{|v|_n}$$

If V is an Euclidean space (*i.e.*, $n = 2$) we omit the subscript and denote the corresponding (Euclidean) operator norm as $||\cdot||$.

Although the vector space V can be over the real number field or complex number field, the *operator norms* do not include some very important norms (trace norm, Frobenius norm) in the quantum information theory. Therefore, in most cases we will discard the *operator norm* (don't worry, the Euclidean operator norm will come back soon) and use another useful norm, called the *Schatten norm*, which includes trace norm, Frobenius norm as its special cases.

The *Schatten norm* has two equivalent definitions, one from the perspective of trace and another from the perspective of singular value (in case you forget, any $rank - r$ matrix A has r positive singular values).

Definition 1 (trace version) The *Schatten norm* with parameter $p \in [1, \infty)$ of any linear operator A is

$$\|A\|_p = \left[\mathrm{Tr}\left((A^*A)^{p/2}\right)\right]^{1/p}$$

Definition 2 (singular-value version) The *Schatten norm* with parameter $p \in [1, \infty)$ of any linear operator A is

$$\|A\|_p = |s(A)|_p = |(s_1(A), s_2(A), \ldots, s_r(A))|_p$$

where $s_1(A), s_2(A), \ldots, s_r(A)$ are r positive singular values of A sorted from largest to smallest, and the norm on the RHS is the vector norm.

The *Schatten norm* has many fruitful properties, for example, the *Schatten norms* are non-increasing in p:

$$\|A\|_p \geq \|A\|_q, \text{ for any } p \leq q$$

while the *operator norms* in general are incomparable. This non-increasing property comes directly from **Definition 2**. For more useful properties of the *Schatten norm*, such as the duality relation and its convexity, see [Wat11].

From **Definition 1** we can see that the *Schatten norm* includes trace norm, Frobenius norm as its special cases. Let $p = 1$, the *Schatten norm* reduces to the trace norm (also called the nuclear norm):

$$\|A\|_{\mathrm{Tr}} := \|A\|_1 = \mathrm{Tr}(\sqrt{A^\dagger A})$$

For Hermitian operator H, we have $\|H\|_1 = \sum_i s_i(H) = \sum_i |\lambda_i(H)|$.

Let $p = 2$, the *Schatten norm* reduces to the Frobenius norm (also called the Hilbert–Schmidt norm):

$$\|A\|_{\mathrm{F}} := \|A\|_2 = \sqrt{\mathrm{Tr}(A^\dagger A)}$$

which coincides the norm defined by the inner product. Therefore, the Frobenius norm can be evaluated entry-wisely:

$$\|A\|_{\mathrm{F}} := \|A\|_2 = \sqrt{\sum_{j,k} |A_{jk}|^2}$$

And don't forget our Euclidean operator norm! In fact, the Euclidean operator norm coincides with the limitation $\lim_{p\to\infty} \|A\|_p$. If we extend $p \in [1, \infty)$ to $p \in [1, \infty]$, then the Euclidean operator norm is also included in the *Schatten norm*:

$$\|A\|_\infty := \|A\| = \sup_{v\in V, |v|\neq 0} \frac{|Av|_2}{|v|_2}$$

And also from **Definition 2** we have

$$\|A\|_\infty = \max\{s_i(A)\} = s_1(A)$$

Definition 2 also gives us the following relationship of the above 3 norms:

$$\|A\|_1 \leq \sqrt{d}\|A\|_2 \leq d\|A\|_\infty$$

where d is the dimension of A. Together with the non-increasing property:

$$\|A\|_1 \geq \|A\|_2 \geq \|A\|_\infty$$

these inequalities enable us to bound the gap between definitions of approximate unitary design.

At the last of this section, we introduce the operator-vector correspondence, which is crucial in comparing approximate unitary designs. Define a 'vectorize' mapping *vec* from the linear operators on Hilbert space $\mathcal{H}^d$ to the vectors on $\mathcal{H}^d \otimes \mathcal{H}^d$

$$vec : L(\mathcal{H}^d) \mapsto \mathcal{H}^d \otimes \mathcal{H}^d$$

For standard basis vectors $|a\rangle$ and $|b\rangle$, this mapping simply flips a bra to a ket:

$$vec(|b\rangle\langle a|) = |b\rangle|a\rangle$$

The *vec* mapping is a bijection, and also an isometry, implying that

$$\mathrm{Tr}(B^\dagger A) = \text{inner\_product}(A, B) = \text{inner\_product}(vec(A), vec(B)) = \langle vec(B)|vec(A)\rangle$$

Thus from the definition we have $\|A\|_F = \|A\|_2 = |vec(A)|_2$.

1.2 The induced and Diamond norm for super-operators

Just like that the operator norm can be induced from the vector norm, super-operator norm can also be induced from the Schatten norm. Intuitively, for super-operator $\Psi : L(\mathcal{H}^d) \mapsto L(\mathcal{H}^d)$, this induced norm for super-operators can be defined as:

$$|||\Psi|||_{p\to q} := \sup_{A\in L(\mathcal{H}^{d'}),||A||_q\neq 0} \frac{\|\Psi(A)\|_p}{\|A\|_q}$$

A note on general properties of the induced super-operator norms can be found in [Wat04]. In this note, we constrain our interest in the $p = q = 1$ and $p = q = 2$ case.

When $p = q = 1$, it is unsatisfying that the norm is not stable under tensoring with the identity: for super-operator $T(A) = A^T$, we have $|||T|||_{1\to 1} = 1$ but $|||T \otimes I_2|||_{1\to 1} = 2$. To fix this, the *Diamond norm* is proposed as

$$|||\Psi|||_\diamond := \sup_d |||\Psi \otimes I_d|||_{1\to 1} = \sup_d \sup_{A\in L(\mathcal{H}^d\otimes\mathcal{H}^{d'}),||A||_1\neq 0} \frac{\|(\Psi \otimes I_d)(A)\|_1}{\|A\|_1}$$

A good news is that we can fix $d' = d$ [KSVV02], which saves us one optimization object. In this sense the *Diamond norm* is defined as

$$|||\Psi|||_\diamond := \sup_{A\in L(\mathcal{H}^d\otimes\mathcal{H}^d),||A||_1\neq 0} \frac{\|(\Psi \otimes I_n)(A)\|_1}{\|A\|_1}$$

Apparently $|||\Psi|||_\diamond \geq |||\Psi|||_{1\to 1}$.

When $p = q = 2$, the instability under tensoring with the identity no longer exists. We define the *super-operator 2-norm* as:

$$|||\Psi|||_{2\to 2} := \sup_{A\in L(\mathcal{H}^d),||A||_2\neq 0} \frac{\|\Psi(A)\|_2}{\|A\|_2}$$

There are some bounds for super-operator norms in [vD02], of which I can't find or make a proof. They are listed below:

$$|||\Psi|||_{2\to 2} \leq \sqrt{d}|||\Psi|||_{1\to 1} \leq \sqrt{d}|||\Psi|||_\diamond \text{ and } |||\Psi|||_{1\to 1} \leq \sqrt{d}|||\Psi|||_{2\to 2}$$

$$|||\Psi|||_\diamond = |||\Psi \otimes I_d|||_{1\to 1} \leq d|||\Psi \otimes I_d|||_{2\to 2} = d|||\Psi|||_{2\to 2} \leq d\sqrt{d}|||\Psi|||_{1\to 1}$$

2 Exact unitary design: equivalent definitions

This (and next) section makes a heavy use of averaging (expectation) over a unitary ensemble or a unitary Haar measure. For convenience we denote the expectation of $f(U)$ over a unitary ensemble $\mathcal{E}$ and a unitary Haar measure by $\mathbb{E}_\mathcal{E}[f(U)]$ and $\mathbb{E}_H[f(U)]$, respectively.

The following 4 definitions of unitary t-design are essentially equivalent. We first discuss two of them which require no extra variables.

Definition 3 For arbitrary quantum state ρ, a weighted/unweighted unitary ensemble $\mathcal{E} = \{U_i\}$ forms weighted/unweighted unitary t-design if

$$\mathbb{E}_\mathcal{E}(U^{\otimes t} \otimes U^{\dagger\otimes t}) = \mathbb{E}_H(U^{\otimes t} \otimes U^{\dagger\otimes t}) \tag{1}$$

Definition 4 Denote by $M_{t,t}(U)$ the multiplication (monomials) of t elements of U and t elements of $U^\dagger$. Denote by $P_{t,t}(U)$ the linear combination of $M_{t,t}(U)$ (the homogeneous polynomials). For arbitrary $P_{t,t}(U)$, a weighted/unweighted unitary ensemble $\mathcal{E} = \{U_i\}$ forms weighted/unweighted unitary t-design if

$$\mathbb{E}_{\mathcal{E}}(P_{t,t}(U)) = \mathbb{E}_H(P_{t,t}(U)) \tag{2}$$

Definition 4 implies **Definition 3** apparently because that every matrix element of $U^{\otimes t} \otimes U^{\dagger \otimes t}$ is a monomial of U of degree (t, t). In fact, **Definition 2** implies any definitions of the form of matrix multiplication since every matrix element is a homogeneous polynomial of U. To show that **Definition 3** $\Rightarrow$ **Definition 4**, note that any monomials of U of degree t can be found in the matrix form of $U^{\otimes t} \otimes U^{\dagger \otimes t}$.

When verifying that a set of unitary matrices forms a t-design, **Definition 1** can be more useful since the right hand side of Eq. 1 can be evaluated directly. The evaluation of the RHS Eq. 1 requires the evaluation of the Weingarten function [CS06]. According to following properties of the Weingarten function, we can compute simple cases for $t = 1, 2$ [PM17]:

$$\int_{\mathrm{U}(d)} U_{ij} \bar{U}_{i'j'} dU = \frac{1}{d} \delta_{ii'} \delta_{jj'}$$

$$\begin{aligned}\int_{\mathrm{U}(d)} & U_{i_1 j_1} U_{i_2 j_2} \bar{U}_{i_1' j_1'} \bar{U}_{i_2' j_2'} dU \\ &= \frac{1}{d^2 - 1} (\delta_{i_1 i_1'} \delta_{i_2 i_2'} \delta_{j_1 j_1'} \delta_{j_2 j_2'} + \delta_{i_1 i_2'} \delta_{i_2 i_1'} \delta_{j_1 j_2'} \delta_{j_2 j_1'}) \\ &- \frac{1}{d\,(d^2 - 1)} (\delta_{i_1 i_1'} \delta_{i_2 i_2'} \delta_{j_1 j_2'} \delta_{j_2 j_1'} + \delta_{i_1 i_2'} \delta_{i_2 i_1'} \delta_{j_1 j_1'} \delta_{j_2 j_2'})\end{aligned}$$

Unfortunately, sampling form the ensemble and computation of the LHS seems inefficient since the tensor products expand the dimension exponentially with the number of qubits U acts on. On MacBook Pro it takes 4 minutes to sample 200 times from a PQC acts on 3 qubits and the program crashes for 4 qubits. What's worse, it takes more samples from PQC for the LHS of Eq. 1 to converge when the number of qubits increases (for example, in 5-qubit case it takes about 2000 samples to converge). This inefficiency drives researchers to find other definitions with lower dimension in the expression. Those lower dimension expressions, as a price, will introduce extra variants.

Definition 5 For arbitrary linear **operator** X, a weighted/unweighted unitary ensemble $\mathcal{E} = \{U_i\}$ forms weighted/unweighted unitary t-design if

$$\mathbb{E}_{\mathcal{E}}(U^{\otimes t} X U^{\dagger \otimes t}) = \mathbb{E}_H(U^{\otimes t} X U^{\dagger \otimes t}) \tag{3}$$

Note that in some references the X will be replaced by ρ. This does not indicate that ρ is a density matrix.

Definition 6 (t=2 case) For arbitrary linear **operator** A, B, C, a weighted/unweighted unitary ensemble $\mathcal{E} = \{U_i\}$ forms weighted/unweighted unitary t-design if

$$\mathbb{E}_{\mathcal{E}}(U^\dagger A U B U^\dagger C U) = \mathbb{E}_H(U^\dagger A U B U^\dagger C U) \tag{4}$$

As discussed above, it is obvious that **Definition 4** $\Rightarrow$ **Definition 5** and **Definition 4** $\Rightarrow$ **Definition 6**, since elements of the matrix in **Definition 5 & 6** are homogeneous polynomial of U. To show that **Definition 3** $\Rightarrow$ **Definition 4**, let X be $|j_1, \ldots, j_t\rangle\langle j_1', \ldots, j_t'|$, then elements of $U^{\dagger \otimes t} \rho U^{\otimes t}$ equal to the monomials of degree t of U. To show that **Definition 6** $\Rightarrow$ **Definition 4**, let A, B, C be of the form of $|i\rangle\langle j|$, then elements of $U^\dagger A U B U^\dagger C U$ equal to the monomials of degree t of U. Now we finish the proof of the equivalence of the above 4 definitions of unitary t-design.

3 Approximate unitary design: 'equivalent' definitions

In this section we give 2 definitions of approximate unitary design, and show their equivalency by introducing an auxiliary definition. By equivalency, (unlike the equivalency of exact unitary designs) we mean that, for two equivalent definitions, if an ensemble forms $\epsilon-$approximate unitary design according to one definition, it forms $s\epsilon-$approximate unitary design according to the other.

Definition 7(DIAMOND) Let $G_{\mathcal{E}}(\rho) = \mathbb{E}_{\mathcal{E}}[U^{\otimes t}\rho U^{\dagger\otimes t}]$ and $G_H(\rho) = \mathbb{E}_H[U^{\otimes t}\rho U^{\dagger\otimes t}]$. Note that both $G_{\mathcal{E}}(\cdot)$ and $G_H(\cdot)$ are super-operators. The unitary ensemble $\mathcal{E}$ forms a $\epsilon-$approximate unitary t-design if:

$$|||G_{\mathcal{E}} - G_H|||_{\diamond} \leq \epsilon$$

Definition 8(TRACE) The unitary ensemble $\mathcal{E}$ forms a $\epsilon-$approximate unitary t-design if:

$$\|\mathbb{E}_{\mathcal{E}}[U^{\otimes t} \otimes U^{\dagger\otimes t}] - \mathbb{E}_H[U^{\otimes t} \otimes U^{\dagger\otimes t}]\|_{\mathrm{Tr}} \leq \epsilon$$

where the trace norm is the Schatten 1-norm.

To show that **Definition 7**(DIAMOND) and **Definition 8**(TRACE) are equivalent, we introduce an auxiliary definition.

Definition 9(INDUCED-2) Use the notation $G_{\mathcal{E}}(\cdot)$ and $G_H(\cdot)$ defined in **Definition 3**(DIAMOND) The unitary ensemble $\mathcal{E}$ forms a $\epsilon-$approximate unitary t-design if:

$$|||G_{\mathcal{E}} - G_H|||_{2\to 2} \leq \epsilon$$

The equivalence of **Definition 9**(INDUCED-2) and **Definition 7**(DIAMOND) comes immediately from the bounds of super-operator norms:

$$|||\Psi|||_{2\to 2} \leq \sqrt{d}|||\Psi|||_{\diamond},\ |||\Psi|||_{\diamond} \leq d|||\Psi|||_{2\to 2}$$

To obtain the equivalence of **Definition 9**(INDUCED-2) and **Definition 8**(TRACE), we need the following observation of **Definition 9**(INDUCED-2), using the isometric $vec()$ mapping:

$$
\begin{aligned}
&|||G_{\mathcal{E}} - G_H|||_{2\to 2}\\
=&\sup_{\rho} \frac{1}{\|\rho\|_2}\|\mathbb{E}_{\mathcal{E}}[U^{\otimes t}\rho U^{\dagger\otimes t}] - \mathbb{E}_H[U^{\otimes t}\rho U^{\dagger\otimes t}]\|_2 \text{ (by definition of induced norm)}\\
=&\sup_{\rho_{jk}} \frac{1}{\sum|\rho_{jk}|^2}\|\sum_{j,k}\rho_{jk}(\mathbb{E}_{\mathcal{E}}[U^{\otimes t}|j\rangle\langle k|U^{\dagger\otimes t}] - \mathbb{E}_H[U^{\otimes t}|j\rangle\langle k|U^{\dagger\otimes t}])\|_2 \text{ (expand } \rho \text{ in its entries)}\\
=&\sup_{\rho_{jk}} \frac{1}{\sum|\rho_{jk}|^2}|\sum_{j,k}\rho_{jk}(\mathbb{E}_{\mathcal{E}}[vec(U^{\otimes t}|j\rangle\langle k|U^{\dagger\otimes t})] - \mathbb{E}_H[vec(U^{\otimes t}|j\rangle\langle k|U^{\dagger\otimes t})])|_2\\
&\text{(isometry and linearity of } vec())\\
=&\sup_{\rho_{jk}} \frac{1}{\sum|\rho_{jk}|^2}|\sum_{j,k}\rho_{jk}(\mathbb{E}_{\mathcal{E}}[U^{\otimes t}\otimes U^{\dagger\otimes t}|jk\rangle] - \mathbb{E}_H[U^{\otimes t}\otimes U^{\dagger\otimes t}|jk\rangle])|_2 \text{ (definition of } vec())\\
=&\sup_{|\phi\rangle\in\mathcal{H}^{2n}} \frac{1}{||\phi\rangle|_2}|\mathbb{E}_{\mathcal{E}}[U^{\otimes t}\otimes U^{\dagger\otimes t}|\phi\rangle] - \mathbb{E}_H[U^{\otimes t}\otimes U^{\dagger\otimes t}|\phi\rangle]|_2 \text{ (Let } |\phi\rangle = \sum_{j,k}\rho_{jk}|jk\rangle \in \mathcal{H}^{2n})\\
=&\|\mathbb{E}_{\mathcal{E}}[U^{\otimes t}\otimes U^{\dagger\otimes t}] - \mathbb{E}_H[U^{\otimes t}\otimes U^{\dagger\otimes t}]\|_{\infty}
\end{aligned}
$$

Together with the bounds of the Schatten norm:

$$\|A\|_1 \leq d\|A\|_{\infty},\ \|A\|_1 \geq \|A\|_{\infty}$$

we have the equivalence of **Definition 9**(INDUCED-2) and **Definition 8**(TRACE).

All in all, the following lemma state the equivalence of **Definition 7**(DIAMOND) and **Definition 8**(TRACE).

Lemma 1 For any unitary ensemble $\mathcal{E}$, if it is ϵ-approximate t-design according to **Definition 7**(DIAMOND), then it is $\sqrt{d^t}\epsilon$-approximate t-design according to **Definition 9**(INDUCED-2), therefore $d^{2t}\sqrt{d^t}\epsilon$-approximate t-design according to **Definition 8**(TRACE). If it is ϵ-approximate t-design according to **Definition 8**(TRACE), then it is ϵ-approximate t-design according to **Definition 9**(INDUCED-2), therefore $d^t\epsilon$-approximate t-design according to **Definition 7**(DIAMOND):

$$\epsilon - \text{DIAMOND} \longrightarrow \sqrt{d^t}\epsilon - \text{INDUCED-2} \longrightarrow d^{2t}\sqrt{d^t}\epsilon - \text{TRACE}$$

$$\epsilon - \text{TRACE} \longrightarrow \epsilon - \text{INDUCED-2} \longrightarrow d^t\epsilon - \text{DIAMOND}$$

References

[CS06] Benoit Collins and Piotr Sniady. Integration with respect to the haar measure on unitary, orthogonal and symplectic group. *Communications in Mathematical Physics*, 264(3):773–795, Mar 2006. doi:10.1007/s00220-006-1554-3.

[KSVV02] Alexei Yu Kitaev, Alexander Shen, Mikhail N Vyalyi, and Mikhail N Vyalyi. *Classical and quantum computation.* Number 47. American Mathematical Soc., 2002.

[PM17] Zbigniew Puchala and Jaroslaw Adam Miszczak. Symbolic integration with respect to the haar measure on the unitary groups. *Bulletin of the Polish Academy of Sciences Technical Sciences*, 65(1):21–27, Feb 2017. doi:10.1515/bpasts-2017-0003.

[vD02] Wim van Dam. *On quantum computation theory.* PhD thesis, Inst. for Logic, Language and Computation, University of Amsterdam, 2002.

[Wat04] John Watrous. Notes on super-operator norms induced by schatten norms. *ArXiv e-prints*, 2004, quant-ph/0411077.

[Wat11] John Watrous. 2.3 norms of operators. In *Theory of Quantum Information (lecture notes)*. University of Waterloo, 2011.