Weak approximate unitary designs and applications to quantum encryption

Cécilia Lancien[1] and Christian Majenz[2]

[1]Institut de Mathématiques de Toulouse & CNRS, Université Paul Sabatier, 118 route de Narbonne, F-31062 Toulouse Cedex 9, France.

[2]QuSoft and Centrum Wiskunde & Informatica, Science Park 123, 1098 XG Amsterdam, the Netherlands.

November 8th 2019

Unitary t-designs are the bread and butter of quantum information theory and beyond. An important issue in practice is that of efficiently constructing good approximations of such unitary t-designs. Building on results by Aubrun (Comm. Math. Phys. 2009), we prove that sampling $d^t \mathrm{poly}(t, \log d, 1/\epsilon)$ unitaries from an exact t-design provides with positive probability an ϵ-approximate t-design, if the error is measured in one-to-one norm distance of the corresponding t-twirling channels. As an application, we give a partially derandomized construction of a quantum encryption scheme that has roughly the same key size and security as the quantum one-time pad, but possesses the additional property of being non-malleable against adversaries without quantum side information.

1 Introduction

Random unitaries, drawn from the Haar measure on the unitary group, play an important role in many aspects of theoretical quantum information science. For instance, most results on quantum source and channel coding are obtained with Haar-random coding strategies [HHWY08, ADHW09, BCR11] using the decoupling technique [HOW07, SDTR13, DBWR14, MBD+17]. The columns and rows of Haar random unitaries are Haar random unit vectors and have also found many applications in quantum information theory, e.g. for constructing quantum money schemes [JLS18, AMR19]. However, it is infeasible in practice to even just approximate Haar random unitaries, the randomness and number of gates necessary to sample and implement them being exponential in the number of qubits they act on.

In most situations, unitary t-designs, the quantum analogues of t-wise independent functions, come to the rescue [DCEL09]. A unitary t-design is a measure on the unitary group that reproduces the Haar measure up to the t-th moment. This means that a random unitary sampled from a t-design can replace a Haar-random unitary in any situation where it is only applied t times. For practical purposes, one would like this measure to be more economical than the Haar measure (for instance to have finite, as small as possible, support). Often even approximate versions of unitary t-designs (in the right metrics)

Cécilia Lancien: clancien@math.univ-toulouse.fr
Christian Majenz: christian.majenz@cwi.nl

are already sufficient. In quantum information theory and related fields the most common metric between measures on the unitary group is the completely bounded one-to-one norm, or diamond norm, on the induced t-twirling channels. The t-twirling channel associated to a measure is the channel that can be implemented by sampling a unitary according to the measure, and then applying it to each sub-system of a t-partite input system.

In [HLSW04], approximate 1-designs have been studied using a metric based on the (not completely bounded) one-to-one norm. There, it is shown that approximate 1-designs in this weaker sense can be made of much fewer unitaries, and that they still have interesting applications, such as unconditionally secure encryption of quantum data when confidentiality is only desired against adversaries without quantum side information. The former result is shown by proving that sampling a small number of independent Haar-random unitaries provides with high probability an approximate 1-design. This construction was subsequently partially derandomized in [Aub09].

Let us mention one last result which was known prior to this work. It was shown in [LW17] that, in fact, any channel can be approximated in one-to-one norm by a channel having few Kraus operators. However, this does not tell us whether it can be further imposed that the Kraus operators of this approximating channel are of a specific form (such as e.g. being tensor powers of unitaries sampled from a simple enough distribution, which is what we are interested in here).

Our contribution

In this work, we generalize the approach of [Aub09] to construct small approximate t-designs, for any given t, in one-to-one norm distance. In addition, for $t = 2$, we show that the approach extends to designs where the goal is to approximate the channel twirl, i.e. the transformation of quantum channels obtained by sampling a unitary, applying it to the input state before the channel acts on it, and undoing this action afterwards. Here, the appropriate distance is the one stemming from the operator norm induced by the diamond norm, which we call diamond-to-diamond norm. To prove the approximation result on the so-called $U^{\otimes t}$-twirl, we use basic representation theory of the unitary group, including the Weyl dimension formula, to show that this channel has small one-to-operator norm. This allows us to apply the powerful probabilistic and functional analytic tools developed in [Aub09]. For the channel twirl, the invariant space spanned by the identity, as well as the off-diagonal terms involving this invariant space, require a careful analysis. Along the way, we also construct a design that approximates the so-called $U \otimes \bar{U}$-twirl, the image of the channel twirl under the Choi-Jamiołkowski isomorphism.

An application

Subsequently, we apply our results in a cryptographic context. We show, that an approximate channel-twirl design in the diamond-to-diamond norm metric can be used to construct a quantum encryption scheme that is as secure as the quantum one-time pad and has (essentially) the same key length, but also is non-malleable against adversaries without quantum side information.

Related work

Unitary t-designs exist for all t and all dimensions [Kan15][1]. For $t > 3$, time-efficient constructions are, however, only known for approximate unitary t-designs [BHH16]. The sub-sampling technique that we use, following [Aub09], i.e. the strategy of sampling (a small number of) random unitaries from an exact design, was first introduced in [ABW09] to show the existence of small approximate 2-designs.

Non-malleability for quantum encryption was first introduced and characterized in [ABW09]. In this work it was also shown that the notion of quantum non-malleability is equivalent to the notion of approximate unitary 2-designs, under the condition that the encryption algorithm be unitary. Subsequently, non-malleability for quantum encryption has been further studied in [AM17, MSvW19].

Notation and standard definitions

Let us gather here notation that we will be using throughout the whole paper. Given $d \in \mathbf{N}$, we denote by $L(d)$ the set of linear operators on $\mathbf{C}^d$, by $D(d)$ the set of quantum states (i.e. positive semidefinite and trace 1 operators) on $\mathbf{C}^d$, and by $U(d)$ the set of unitary operators on $\mathbf{C}^d$. We additionally denote by $\mathcal{L}(d)$ the set of linear operators on $L(d)$, and by $\mathcal{C}(d)$ the set of quantum channels (i.e. completely positive and trace-preserving operators) on $L(d)$. Let us conclude with some standard notation/definitions from probability theory. Given a random variable X, we denote by $\mathbf{E}\, X$ its average and by $\mathbf{P}(X \in E)$ the probability that X satisfies event E. We say that ε is a Bernoulli random variable if $\mathbf{P}(\varepsilon = +1) = \mathbf{P}(\varepsilon = -1) = 1/2$.

2 Representation theoretic preliminaries

Given $t \in \mathbf{N}$ let S_t be the permutation group of $\{1, \ldots, t\}$. The irreducible representations $[\lambda]$ of S_t are called *Specht modules* and are indexed by integer partitions of t, denoted as $\lambda \vdash t$. Such a partition is represented as a tuple $\lambda = (\lambda_1, ..., \lambda_r) \in \mathbf{N}^r$, for some $r \in \mathbf{N}$, with $\lambda_1 \geqslant \cdots \geqslant \lambda_r$ and $\sum_{i=1}^r \lambda_i = t$.

Given $d \in \mathbf{N}$ let $U(d)$ be the unitary group of $\mathbf{C}^d$. The polynomial irreducible representations V_λ of $U(d)$ are called *Weyl modules* and are indexed by integer partitions of any number $t \in \mathbf{N}$ into exactly d parts (some of which might be 0), denoted as $\lambda \vdash (t, d)$. The dimension of the Weyl module V_λ is given by the Weyl dimension formula

$$m_\lambda = \prod_{\substack{i,j \in \{1,\ldots,d\} \\ i<j}} \frac{\lambda_i - \lambda_j + j - i}{j - i}. \tag{1}$$

A particular vector space that carries representations of both S_t and $U(d)$ is $\left(\mathbf{C}^d\right)^{\otimes t}$. The corresponding actions are defined as

$$\forall\ \sigma \in S_t,\ \sigma.|\phi_1\rangle \otimes \cdots \otimes |\phi_t\rangle = |\phi_{\sigma^{-1}(1)}\rangle \otimes \cdots \otimes |\phi_{\sigma^{-1}(t)}\rangle,$$
$$\forall\ U \in U(d),\ U.|\phi\rangle = U^{\otimes t}|\phi\rangle.$$

The two actions commute, i.e. $\left(\mathbf{C}^d\right)^{\otimes t}$ decomposes into a direct sum of irreducible representations (irreps) of the product group $S_t \times U(d)$. These irreps are just tensor products of

[1] In [Kan15], the existence of exact designs is proven in a much more general context, see [AMR19, Corollary 2] for a straightforward application to the unitary case.

an irrep of S_t with an irrep of $U(d)$. What is more, the corresponding representations of the group algebras of S_t and $U(d)$ are double commutants, implying that the decomposition is multiplicity free.

Theorem 2.1 (Schur-Weyl duality)**.** *Let S_t and $U(d)$ act on $\left(\mathbf{C}^d\right)^{\otimes t}$ as described above. The direct sum decomposition into irreducible representations of $S_t \times U(d)$ is multiplicity free, and is given by*

$$\left(\mathbf{C}^d\right)^{\otimes t} \cong \bigoplus_{\lambda \vdash (t,d)} V_\lambda \otimes [\lambda]. \tag{2}$$

Define the quantum channel $T^{(t)}$ on $(\mathbf{C}^d)^{\otimes t}$ as

$$T^{(t)} : X \in L(d^t) \mapsto \int_{U \in U(d)} U^{\otimes t} X U^{*\otimes t} dU \in L(d^t), \tag{3}$$

where dU stands for the Haar measure on $U(d)$. The channel $T^{(t)}$ is often referred to as a *twirling channel*. It is obviously covariant with respect to the action of $U(d)$. Hence, denoting by W the isomorphism between the right and left hand sides of equation (2) above, Schur's Lemma implies that

$$W T^{(t)}(W^*(\cdot)W)W^* = \sum_{\lambda \vdash (t,d)} \tau_{V_\lambda} \otimes \mathrm{Tr}_{V_\lambda} \left[P_\lambda(\cdot)P_\lambda\right], \tag{4}$$

where P_λ is the projector onto $V_\lambda \otimes [\lambda]$ in $\bigoplus_{\lambda \vdash (t,d)} V_\lambda \otimes [\lambda]$ and $\tau_{V_\lambda} = \mathbf{1}_{V_\lambda}/m_\lambda$ is the maximally mixed state on V_λ.

Let us make things slightly more explicit in the case $t = 2$. We have

$$\left(\mathbf{C}^d\right)^{\otimes 2} \cong \wedge^2(d) \oplus \vee^2(d),$$

where $\wedge^2(d)$ and $\vee^2(d)$ are, respectively, the symmetric and anti-symmetric subspaces of $\left(\mathbf{C}^d\right)^{\otimes 2}$. The corresponding projectors are $P_{\wedge^2(d)} = (\mathbf{1} + F)/2$ and $P_{\vee^2(d)} = (\mathbf{1} - F)/2$, where F denotes the so-called *flip operator*. And the action of $T^{(2)}$ can be explicitly written as, for any $X \in L(d^2)$,

$$T^{(2)}(X) = \frac{2}{d(d+1)} \mathrm{Tr}\left(P_{\wedge^2(d)} X P_{\wedge^2(d)}\right) P_{\wedge^2(d)} + \frac{2}{d(d-1)} \mathrm{Tr}\left(P_{\vee^2(d)} X P_{\vee^2(d)}\right) P_{\vee^2(d)}.$$

Fix a basis $B = \{|i\rangle\}_{i=0}^{d-1}$ for $\mathbf{C}^d$ (which we refer to as the computational basis). Let T be the transposition in this basis and denote by X^Γ the partial transposition of X (i.e. $X^\Gamma = \mathrm{id} \otimes \mathrm{T}(X)$). It is easy to check that, for any $X \in L(d^2)$,

$$T^{(2)}(X)^\Gamma = \left(\int_{U \in U(d)} U \otimes U \, X \, U^* \otimes U^* dU\right)^\Gamma = \int_{U \in U(d)} U \otimes \bar{U} \, X^\Gamma \, U^* \otimes \bar{U}^* dU.$$

Let us define the quantum channel $T^{(1,1)}$ on $(\mathbf{C}^d)^{\otimes 2}$ as

$$\forall \; X \in L(d^2), \; T^{(1,1)}(X) = \int_{U \in U(d)} U \otimes \bar{U} \, X \, U^* \otimes \bar{U}^* dU. \tag{5}$$

By the preceding discussion, we know that $T^{(1,1)}(X)$ can be written as a linear combination of $P^\Gamma_{\wedge^2(d)}$ and $P^\Gamma_{\vee^2(d)}$. Now, $\mathbf{1}^\Gamma = \mathbf{1}$ and $F^\Gamma = d|\psi\rangle\langle\psi|$, where

$$|\psi\rangle = \frac{1}{\sqrt{d}} \sum_{i=0}^{d-1} |ii\rangle$$

is the *standard maximally entangled state* with respect to B. So equivalently, $T^{(1,1)}(X)$ can be written as a linear combination of $|\psi\rangle\langle\psi|$ and $Q = \mathbf{1} - |\psi\rangle\langle\psi|$, which are orthogonal to one another. More specifically

$$\forall\ X \in L(d^2),\ T^{(1,1)}(X) = \langle\psi|X|\psi\rangle|\psi\rangle\langle\psi| + \frac{1}{d^2-1}\operatorname{Tr}(QX)Q. \tag{6}$$

3 Several channel approximation results

3.1 Approximating the twirling channel $T^{(t)}$

Let $t \in \mathbf{N}$ be such that $t \leqslant d/2$. The goal here is to show that the twirling channel $T^{(t)}$, as defined by equation (3), can be approximated with 'few' Kraus operators sampled from a 'simple' probability measure. We will be able to prove such approximation in a strong sense, namely in one-to-infinity norm.

A probability measure μ on $U(d)$ is called a t-design if

$$\forall\ X \in L(d^t),\ \int_{U\in U(d)} U^{\otimes t} X U^{*\otimes t} d\mu(U) = T^{(t)}(X).$$

We will show the following result:

Theorem 3.1. *Let $0 < \epsilon < 1$. Assume that the probability measure μ on $U(d)$ is a t-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. There exists a universal constant $C > 0$ such that, if $n \geqslant C(td)^t(t\log d)^6/\epsilon^2$, then with probability at least $1/2$, we have*

$$\forall\ \rho \in D(d^t),\ \left\| \frac{1}{n}\sum_{i=1}^{n} U_i^{\otimes t}\rho U_i^{*\otimes t} - T^{(t)}(\rho)\right\|_\infty \leqslant \frac{\epsilon}{d^t}.$$

Theorem 3.1 generalizes [Aub09, Theorem 2] to t-designs for any $t \in \mathbf{N}$ rather than only for 1-designs. We actually follow the exact same proof strategy as that of [Aub09, Theorem 2]. The only additional technical lemma that we need in the case $t > 1$ is one that tells us that $T^{(t)}$ has a small $(1 \to \infty)$-norm (a fact which is obvious for $t = 1$).

Lemma 3.2. *The quantum channel $T^{(t)}$ is such that*

$$\sup_{\rho\in D(d^t)} \left\|T^{(t)}(\rho)\right\|_\infty \leqslant \left(\frac{2t}{d}\right)^t.$$

Proof. By equation (4), the operator norm in question is just given by the inverse of the minimal dimension of an irrep V_λ,

$$\sup_{\rho\in D(d^t)} \left\|T^{(t)}(\rho)\right\|_\infty = \frac{1}{\min_{\lambda\vdash(t,d)} m_\lambda}.$$

Indeed, let us denote by λ^* the partition minimizing m_λ. It is clear that if $|\phi_{\lambda^*}\rangle \in V_{\lambda^*}$ and $|\varphi_{\lambda^*}\rangle \in [\lambda^*]$, then $\left\|T^{(t)}(|\phi_{\lambda^*}\rangle\langle\phi_{\lambda^*}| \otimes |\varphi_{\lambda^*}\rangle\langle\varphi_{\lambda^*}|\right\|_\infty = 1/m_{\lambda^*}$. And this is obviously maximizing $\left\|T^{(t)}(\rho)\right\|_\infty$ as $T^{(t)}$ begins with a pinching with respect to the direct sum decomposition (2). We go on to find a lower bound on m_{λ^*} using the formula (1). To this end we first note that λ^* is a partition of t into d parts, so $\lambda_i^* = 0$ for all $i > t$. Noting that

all the factors in the product in equation (1) are lower bounded by 1, and only keeping factors such that $i \leq t < j$ we get

$$\begin{aligned} m_{\lambda^*} &\geq \prod_{\substack{i,j\in\{1,\dots,d\}\\ i\leq t<j}} \frac{\lambda_i^* + j - i}{j-i} \\ &= \prod_{i=1}^{t}\prod_{j=t+1}^{d} \frac{\lambda_i^* + j - i}{j-i} \\ &= \prod_{i=1}^{t} \frac{(\lambda_i^* + d - i)!(t-i)!}{(\lambda_i^* + t - i)!(d-i)!} \\ &= \prod_{i=1}^{t}\prod_{\alpha=1}^{\lambda_i^*} \frac{d-i+\alpha}{t-i+\alpha}. \end{aligned}$$

As a final step we use that $(a-x)/(b-x) \geq a/b$ for all $a, b, x \in \mathbf{R}$ such that $a \geq b > x \geq 0$, and that $\alpha \leq t$. We thus conclude that

$$\prod_{i=1}^{t}\prod_{\alpha=1}^{\lambda_i^*} \frac{d-i+\alpha}{t-i+\alpha} \geq \prod_{i=1}^{t}\prod_{\alpha=1}^{\lambda_i^*} \frac{d+\alpha}{t+\alpha} \geq \prod_{i=1}^{t}\prod_{\alpha=1}^{\lambda_i^*} \frac{d}{2t} = \prod_{i=1}^{t} \left(\frac{d}{2t}\right)^{\lambda_i^*} \geqslant \left(\frac{d}{2t}\right)^{t},$$

where the last inequality is because $d \geqslant 2t$ by assumption and $\lambda_i^* \geqslant 1$. □

We then need the technical result below, which is an immediate corollary of [Aub09, Lemma 5] (which itself makes crucial use of Dudley's inequality and a duality argument for entropy numbers).

Lemma 3.3. *Let $U_1, \dots, U_n \in U(d)$. For $\varepsilon_1, \dots, \varepsilon_n$ independent Bernoulli random variables, we have*

$$\mathbf{E}\left(\sup_{\rho\in D(d^t)} \left\|\sum_{i=1}^{n} \varepsilon_i U_i^{\otimes t} \rho U_i^{*\otimes t}\right\|_\infty\right) \leqslant C(t\log d)^{5/2}(\log n)^{1/2} \sup_{\rho\in D(d^t)} \left\|\sum_{i=1}^{n} U_i^{\otimes t} \rho U_i^{*\otimes t}\right\|_\infty^{1/2},$$

where $C > 0$ is a universal constant.

Proof. This follows directly from [Aub09, Lemma 5], applied with d^t playing the role of d and $U_i^{\otimes t}$ playing the role of U_i, $1 \leqslant i \leqslant n$. □

With these two preliminary lemmas at hand, we are now in position to prove Theorem 3.1.

Proof of Theorem 3.1. Let $V_1, \dots, V_n$ be independent copies of $U_1, \dots, U_n$ and let $\varepsilon_1, \dots, \varepsilon_n$ be independent Bernoulli random variables. Setting

$$M = \sup_{\rho\in D(d^t)} \left\|\frac{1}{n}\sum_{i=1}^{n} U_i^{\otimes t} \rho U_i^{*\otimes t} - T^{(t)}(\rho)\right\|_\infty,$$

we then have

$$\begin{aligned}
\mathbf{E}\, M &= \mathbf{E}_U \left(\sup_{\rho \in D(d^t)} \left\| \frac{1}{n} \sum_{i=1}^n U_i^{\otimes t} \rho U_i^{*\otimes t} - \mathbf{E}_V \left(\frac{1}{n} \sum_{i=1}^n V_i^{\otimes t} \rho V_i^{*\otimes t} \right) \right\|_\infty \right) \\
&\leqslant \mathbf{E}_{U,V} \left(\sup_{\rho \in D(d^t)} \left\| \frac{1}{n} \sum_{i=1}^n \left(U_i^{\otimes t} \rho U_i^{*\otimes t} - V_i^{\otimes t} \rho V_i^{*\otimes t} \right) \right\|_\infty \right) \\
&= \mathbf{E}_{U,V,\varepsilon} \left(\sup_{\rho \in D(d^t)} \left\| \frac{1}{n} \sum_{i=1}^n \varepsilon_i \left(U_i^{\otimes t} \rho U_i^{*\otimes t} - V_i^{\otimes t} \rho V_i^{*\otimes t} \right) \right\|_\infty \right) \\
&\leqslant 2\ \mathbf{E}_{U,\varepsilon} \left(\sup_{\rho \in D(d^t)} \left\| \frac{1}{n} \sum_{i=1}^n \varepsilon_i U_i^{\otimes t} \rho U_i^{*\otimes t} \right\|_\infty \right),
\end{aligned}$$

where the first inequality is by Jensen's inequality, the second equality is by symmetry, and the third inequality is by the triangle inequality.

Hence, by Lemma 3.3, we get

$$\begin{aligned}
\mathbf{E}\, M &\leqslant \frac{2C}{\sqrt{n}} (t \log d)^{5/2} (\log n)^{1/2}\, \mathbf{E} \left(\sup_{\rho \in D(d^t)} \left\| \frac{1}{n} \sum_{i=1}^n U_i^{\otimes t} \rho U_i^{*\otimes t} \right\|_\infty^{1/2} \right) \\
&\leqslant \frac{2C}{\sqrt{n}} (t \log d)^{5/2} (\log n)^{1/2}\, \mathbf{E} \left(M + \left(\frac{2t}{d} \right)^t \right)^{1/2} \\
&\leqslant \frac{2C}{\sqrt{n}} (t \log d)^{5/2} (\log n)^{1/2} \left(\mathbf{E} \left(M + \left(\frac{2t}{d} \right)^t \right) \right)^{1/2},
\end{aligned}$$

where the second inequality is by Lemma 3.2 while the third inequality is by Jensen's inequality.

Now, it is easy to check that, given $X, \alpha, \beta \geqslant 0$, if $X \leqslant \alpha \sqrt{X + \beta}$, then $X \leqslant \alpha^2 + \alpha \sqrt{\beta}$. Therefore, we eventually obtain

$$\mathbf{E}\, M \leqslant \frac{4C^2}{n} (t \log d)^5 \log n + \frac{2C}{\sqrt{n}} (t \log d)^{5/2} (\log n)^{1/2} \left(\frac{2t}{d} \right)^{t/2}.$$

And the latter quantity is smaller than ϵ / d^t as soon as n is larger than $C'(td)^t (t \log d)^6 / \epsilon^2$.

To conclude, we just have to use Markov's inequality, which guarantees that, if $\mathbf{E}\, M \leqslant \epsilon / d^t$, then

$$\mathbf{P} \left(M \leqslant \frac{2\epsilon}{d^t} \right) \geqslant 1 - \frac{\mathbf{E}\, M}{2\epsilon / d^t} \geqslant \frac{1}{2}.$$

This is exactly what we wanted to show (after relabelling 2ϵ in ϵ and $4C'$ in C). □

Remark 3.4. *Note that, up to a* $\mathrm{poly}(t, \log d)$ *factor, the result of Theorem 3.1 is optimal, in the sense that it is impossible to approximate the twirling channel* $T^{(t)}$ *with less than order* d^t *operators. This is true even if we only require* ϵ*-approximation in* $(1 \to 1)$*-norm rather than* ϵ / d^t*-approximation in* $(1 \to \infty)$*-norm. Indeed, the following general result was shown in [LW17, Section 5.1]: If* $T, \hat{T}$ *are channels on* $L(d')$ *which are* ϵ*-close in* $(1 \to 1)$*-norm, then the Kraus rank* $r(\hat{T})$ *of* $\hat{T}$ *(i.e. the minimal number of Kraus operators for* $\hat{T}$*) satisfies*

$$\log r(\hat{T}) \geqslant (1 - \epsilon) \max_{\rho \in D(d')} |S(T(\rho)) - S(\rho)|,$$

where $S(\cdot)$ is the von Neumann entropy. In particular, if T is such that, for all $\rho \in D(d')$, $\|T(\rho)\|_\infty \leqslant c/d'$, then

$$\max_{\rho \in D(d')} |S(T(\rho)) - S(\rho)| \geqslant \log\left(\frac{d'}{c}\right),$$

and hence necessarily

$$r(\hat{T}) \geqslant \left(\frac{d'}{c}\right)^{1-\epsilon}.$$

In the case of the channel $T^{(t)}$ on $L(d^t)$, we know by Lemma 3.2 that, for all $\rho \in D(d^t)$, $\|T^{(t)}(\rho)\|_\infty \leqslant (2t/d)^t$. So if a channel $\hat{T}^{(t)}$ is ϵ-close to $T^{(t)}$ in $(1 \to 1)$-norm, then it has to satisfy $r(\hat{T}^{(t)}) \geqslant (d/2t)^{(1-\epsilon)t}$.

3.2 Approximating the twirling channel $T^{(1,1)}$

The goal here is to show that the twirling channel $T^{(1,1)}$, as defined by equation (5), can be approximated with 'few' Kraus operators sampled from a 'simple' probability measure. We will only be able to prove such approximation in a weaker sense than in the case of $T^{(t)}$ treated before, namely in one-to-one norm.

If μ is a 2-design on $U(d)$, then, by equation (5), we have that

$$\forall\ X \in L(d^2),\ \int_{U \in U(d)} U \otimes \bar{U}\, X\, U^* \otimes \bar{U}^* d\mu(U) = T^{(1,1)}(X).$$

We will show the following result:

Theorem 3.5. *Let $0 < \epsilon < 1$. Assume that the probability measure μ on $U(d)$ is a 2-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. There exists a universal constant $C > 0$ such that, if $n \geqslant Cd^2(\log d)^6/\epsilon^2$, then with probability at least $1/2$, we have*

$$\forall\ \rho \in D(d^2),\ \left\| \frac{1}{n} \sum_{i=1}^{n} U_i \otimes \bar{U}_i\, \rho\, U_i^* \otimes \bar{U}_i^* - T^{(1,1)}(\rho) \right\|_1 \leqslant \epsilon.$$

The way we prove Theorem 3.5 is by first analysing separately the cases where the input state is the maximally entangled state or a state orthogonal to it. This is the content of Propositions 3.6 and 3.7 below.

Proposition 3.6. *Assume that the probability measure μ on $U(d)$ is a 2-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. Then,*

$$\frac{1}{n} \sum_{i=1}^{n} U_i \otimes \bar{U}_i\, |\psi\rangle\langle\psi|\, U_i^* \otimes \bar{U}_i^* = T^{(1,1)}(|\psi\rangle\langle\psi|).$$

Proof. We just have to notice that, for any $U \in U(d)$, $U \otimes \bar{U}|\psi\rangle = |\psi\rangle$. And thus,

$$\frac{1}{n} \sum_{i=1}^{n} U_i \otimes \bar{U}_i\, |\psi\rangle\langle\psi|\, U_i^* \otimes \bar{U}_i^* = |\psi\rangle\langle\psi| = T^{(1,1)}(|\psi\rangle\langle\psi|),$$

as announced. □

Proposition 3.7. *Let $0 < \epsilon < 1$. Assume that the probability measure μ on $U(d)$ is a 2-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. There exists a universal constant $C > 0$ such that, if $n \geqslant Cd^2(\log d)^6/\epsilon^2$, then with probability at least $1/2$, we have*

$$\forall\ \rho \in D(d^2),\ \rho \perp \psi,\ \left\| \frac{1}{n} \sum_{i=1}^{n} U_i \otimes \bar{U}_i\, \rho\, U_i^* \otimes \bar{U}_i^* - T^{(1,1)}(\rho) \right\|_\infty \leqslant \frac{\epsilon}{d^2}.$$

In order to prove Proposition 3.7 we follow the same route as to prove Theorem 3.1. We thus begin by observing that $T^{(1,1)}$ has a small $(1 \to \infty)$-norm on the orthogonal of the maximally entangled state, which is the analogue of Lemma 3.2 in the study of $T^{(t)}$.

Lemma 3.8. *The quantum channel $T^{(1,1)}$ is such that*

$$\sup_{\rho \in D(d^2),\, \rho \perp \psi} \left\| T^{(1,1)}(\rho) \right\|_\infty = \frac{1}{d^2-1}.$$

Proof. By equation (6), we see that, for any state ρ orthogonal to $|\psi\rangle\langle\psi|$, $T^{(1,1)}(\rho) = Q/(d^2-1)$, so that $\|T^{(1,1)}(\rho)\|_\infty = 1/(d^2-1)$. □

We then need the technical result below, which is the analogue of Lemma 3.3 in the study of $T^{(t)}$.

Lemma 3.9. *Let $U_1, \ldots, U_n \in U(d)$. For $\varepsilon_1, \ldots, \varepsilon_n$ independent Bernoulli random variables, we have*

$$\begin{aligned}
&\mathbf{E}\left(\sup_{\rho \in D(d^2),\, \rho \perp \psi} \left\| \sum_{i=1}^n \varepsilon_i U_i \otimes \bar{U}_i \, \rho \, U_i^* \otimes \bar{U}_i^* \right\|_\infty \right) \\
&\leqslant C (\log d)^{5/2} (\log n)^{1/2} \sup_{\rho \in D(d^2),\, \rho \perp \psi} \left\| \sum_{i=1}^n U_i \otimes \bar{U}_i \, \rho \, U_i^* \otimes \bar{U}_i^* \right\|_\infty^{1/2},
\end{aligned}$$

where $C > 0$ is a universal constant.

Proof. This follows directly from [Aub09, Lemma 5], applied with $d^2 - 1$ playing the role of d and $U_i \otimes \bar{U}_i$ playing the role of U_i, $1 \leqslant i \leqslant n$. □

With Lemmas 3.8 and 3.9 at hand it is straightforward to prove Proposition 3.7, starting from the same symmetrization trick than the one which allows to prove Theorem 3.1 from Lemmas 3.2 and 3.3. We therefore do not repeat the proof here.

So we can now combine Propositions 3.6 and 3.7 to get Theorem 3.5.

Proof of Theorem 3.5. By convexity of $\|\cdot\|_1$ and extremality of pure states amongst all states, it is enough to prove that the result is true for all pure input states. Given $|\varphi\rangle$ a unit vector, we can write it as $|\varphi\rangle = \alpha|\psi\rangle + \beta|\psi'\rangle$, where $\alpha = \langle\psi|\varphi\rangle$, $|\alpha|^2 + |\beta|^2 = 1$ and $|\psi'\rangle$ is a unit vector orthogonal to $|\psi\rangle$. Defining

$$\Delta : X \in L(d^2) \mapsto \frac{1}{n} \sum_{i=1}^n U_i \otimes \bar{U}_i \, X \, U_i^* \otimes \bar{U}_i^* - T^{(1,1)}(X) \in L(d^2), \tag{7}$$

we then have

$$\begin{aligned}
\|\Delta(|\varphi\rangle\langle\varphi|)\|_1 &= \| |\alpha|^2 \Delta(|\psi\rangle\langle\psi|) + |\beta|^2 \Delta(|\psi'\rangle\langle\psi'|) + \alpha\bar{\beta}\Delta(|\psi\rangle\langle\psi'|) + \bar{\alpha}\beta\Delta(|\psi'\rangle\langle\psi|) \|_1 \\
&\leqslant |\alpha|^2 \|\Delta(|\psi\rangle\langle\psi|)\|_1 + |\beta|^2 \|\Delta(|\psi'\rangle\langle\psi'|)\|_1 + 2|\alpha||\beta| \|\Delta(|\psi\rangle\langle\psi'|)\|_1.
\end{aligned}$$

First, we know from Proposition 3.6 that $\|\Delta(|\psi\rangle\langle\psi|)\|_1 = 0$, while we know from Proposition 3.7 that, with probability at least $3/4$, for any $|\psi'\rangle$ orthogonal to $|\psi\rangle$, $\|\Delta(|\psi'\rangle\langle\psi'|)\|_1 \leqslant d^2 \|\Delta(|\psi'\rangle\langle\psi'|)\|_\infty \leqslant \epsilon$. Second, we know that we can write $|\psi'\rangle = X \otimes \mathbf{1}|\psi\rangle$ for some X

such that $\mathrm{Tr}(X) = 0$ and $\|X\|_2 = \sqrt{d}$. That way, since for any $U \in U(d)$, $U \otimes \bar{U}|\psi\rangle = |\psi\rangle$ and $UX \otimes \bar{U}|\psi\rangle = UXU^* \otimes \mathbf{1}|\psi\rangle$, we get

$$\begin{aligned}\|\Delta(|\psi\rangle\langle\psi'|)\|_1 &= \left\| |\psi\rangle\langle\psi| \left(\frac{1}{n}\sum_{i=1}^n U_i X U_i^* - T^{(1)}(X) \right) \otimes \mathbf{1} \right\|_1 \\ &\leqslant \left\| \left(\frac{1}{n}\sum_{i=1}^n U_i X U_i^* - T^{(1)}(X) \right) \otimes \mathbf{1} \right\|_\infty \\ &= \left\| \frac{1}{n}\sum_{i=1}^n U_i X U_i^* - T^{(1)}(X) \right\|_\infty .\end{aligned}$$

Now, we know from Theorem 3.1 (for $t = 1$) that, with probability at least $3/4$, for any X such that $\|X\|_2 = \sqrt{d}$,

$$\left\| \frac{1}{n}\sum_{i=1}^n U_i X U_i^* - T^{(1)}(X) \right\|_\infty \leqslant \frac{\epsilon}{d}\|X\|_1 \leqslant \frac{\epsilon}{\sqrt{d}}\|X\|_2 = \epsilon$$

(actually as soon as $n \geqslant Cd(\log d)^6/\epsilon^2$, hence a fortiori for $n \geqslant Cd^2(\log d)^6/\epsilon^2$).

Putting everything together we eventually obtain that, with probability at least $1/2$, for any $|\varphi\rangle$,

$$\|\Delta(|\varphi\rangle\langle\varphi|)\|_1 \leqslant \frac{3\epsilon}{2},$$

which, up to re-labelling $3\epsilon/2$ in ϵ, is exactly what we wanted to prove. □

Remark 3.10. *It can be shown that the result of Theorem 3.5 is optimal, up to a* $\mathrm{poly}(\log d)$ *factor, just as the one of Theorem 3.1. Indeed, using again the result of [LW17, Section 5.1], together with Lemma 3.8, we see that, if a channel* $\hat{T}^{(1,1)}$ *is* ϵ*-close to* $T^{(1,1)}$ *in* $(1 \to 1)$*-norm, then it has to satisfy* $r(\hat{T}^{(1,1)}) \geqslant (d^2 - 1)^{1-\epsilon}$.

3.3 Approximating the twirling super-channel Θ

We are now interested in a slightly different kind of twirling, namely one that acts on channels rather than states. We thus define the quantum super-channel Θ on $\mathbf{C}^d$ as

$$\Theta : \mathcal{M} \in \mathcal{L}(d) \mapsto \left(\Theta(\mathcal{M}) : X \in L(d) \mapsto \int_{U \in U(d)} U\mathcal{M}(U^*XU)U^* dU \right). \tag{8}$$

Similarly as before, we here want to show that Θ can be approximated by sampling 'few' unitaries from a 'simple' probability measure. We will be able to prove approximation in completely bounded one-to-one norm (also known as diamond norm) for all input channel.

More precisely, denoting by $\mathrm{id} : L(d) \to L(d)$ the identity map on $L(d)$, we will show the following result:

Theorem 3.11. *Let* $0 < \epsilon < 1$. *Assume that the probability measure* μ *on* $U(d)$ *is a* 2*-design, and let* $U_1, \ldots, U_n$ *be sampled independently from* μ. *There exists a universal constant* $C > 0$ *such that, if* $n \geqslant Cd^2(\log d)^6/\epsilon^2$, *then with probability at least* $1/2$, *we have, for all* $\mathcal{N} \in \mathcal{C}(d)$ *and all* $\rho \in D(d^2)$,

$$\left\| \frac{1}{n}\sum_{i=1}^n \mathbf{1} \otimes U_i \,\mathrm{id} \otimes \mathcal{N}(\mathbf{1} \otimes U_i^* \,\rho\, \mathbf{1} \otimes U_i)\, \mathbf{1} \otimes U_i^* - \mathrm{id} \otimes \Theta(\mathcal{N})(\rho) \right\|_1 \leqslant \epsilon.$$

Proof. By convexity of $\|\cdot\|_1$ and extremality of pure states amongst all states, it is enough to prove that the result is true for all pure input states (and all input channels). Let $\mathcal{N}$ be a channel and $|\varphi\rangle$ be a pure state, which we can write as $|\varphi\rangle = X \otimes \mathbf{1}|\psi\rangle$ for some X such that $\|X\|_2 = \sqrt{d}$. Now, for any $U \in U(d)$, $X \otimes U^*|\psi\rangle = X\bar{U} \otimes \mathbf{1}|\psi\rangle$, so that

$$\begin{aligned}
&\mathbf{1} \otimes U \,\mathrm{id} \otimes \mathcal{N}(\mathbf{1} \otimes U^* \,|\varphi\rangle\langle\varphi|\, \mathbf{1} \otimes U)\, \mathbf{1} \otimes U^* \\
&= \mathbf{1} \otimes U \,\mathrm{id} \otimes \mathcal{N}(X \otimes U^* \,|\psi\rangle\langle\psi|\, X^* \otimes U)\, \mathbf{1} \otimes U^* \\
&= \mathbf{1} \otimes U \,\mathrm{id} \otimes \mathcal{N}(X\bar{U} \otimes \mathbf{1} \,|\psi\rangle\langle\psi|\, \bar{U}^* X^* \otimes \mathbf{1})\, \mathbf{1} \otimes U^* \\
&= X \otimes \mathbf{1}\, U \otimes \bar{U} \mathrm{id} \otimes \mathcal{N}(|\psi\rangle\langle\psi|) U^* \otimes \bar{U}^*\, X^* \otimes \mathbf{1}.
\end{aligned}$$

Therefore, defining Δ as in equation (7), we have

$$\begin{aligned}
&\left\| \frac{1}{n} \sum_{i=1}^{n} \mathbf{1} \otimes U_i \,\mathrm{id} \otimes \mathcal{N}(\mathbf{1} \otimes U_i^* \,\rho\, \mathbf{1} \otimes U_i)\, \mathbf{1} \otimes U_i^* - \mathrm{id} \otimes \Theta(\mathcal{N})(\rho) \right\|_1 \\
&= \|X \otimes \mathbf{1}\, \Delta(\mathrm{id} \otimes \mathcal{N}(|\psi\rangle\langle\psi|)\, X^* \otimes \mathbf{1}\|_1. \qquad (9)
\end{aligned}$$

We now proceed exactly as in the proof of Theorem 3.5. First, by Proposition 3.6, $\Delta(|\psi\rangle\langle\psi|) = 0$, so that

$$\|X \otimes \mathbf{1}\, \Delta(|\psi\rangle\langle\psi|)\, X^* \otimes \mathbf{1}\|_1 = 0.$$

Second, by Proposition 3.7, with probability at least 3/4, for any $|\psi'\rangle$ orthogonal to $|\psi\rangle$, $\|\Delta(|\psi'\rangle\langle\psi'|)\|_\infty \leqslant \epsilon/d^2$, so that

$$\begin{aligned}
\|X \otimes \mathbf{1}\, \Delta(|\psi'\rangle\langle\psi'|)\, X^* \otimes \mathbf{1}\|_1 &\leqslant \|X \otimes \mathbf{1}\|_2 \|X^* \otimes \mathbf{1}\|_2 \|\Delta(|\psi'\rangle\langle\psi'|)\|_\infty \\
&= \|X\|_2^2 \|\mathbf{1}\|_2^2 \|\Delta(|\psi'\rangle\langle\psi'|)\|_\infty \\
&\leqslant \epsilon,
\end{aligned}$$

where the first inequality is by Hölder inequality while the last inequality is simply recalling that $\|X\|_2 = \|\mathbf{1}\|_2 = \sqrt{d}$. Third, any $|\psi'\rangle$ orthogonal to $|\psi\rangle$ can be written as $|\psi'\rangle = Y \otimes \mathbf{1}|\psi\rangle$ for some Y such that $\mathrm{Tr}(Y) = 0$ and $\|Y\|_2 = \sqrt{d}$. Since for any $U \in U(d)$, $U \otimes \bar{U}|\psi\rangle = |\psi\rangle$ and $UY \otimes \bar{U}|\psi\rangle = UYU^* \otimes \mathbf{1}|\psi\rangle$, we then get

$$\begin{aligned}
\|X \otimes \mathbf{1}\, \Delta(|\psi\rangle\langle\psi'|)\, X^* \otimes \mathbf{1}\|_1 &= \left\| X \otimes \mathbf{1}\, |\psi\rangle\langle\psi| \left(\frac{1}{n} \sum_{i=1}^{n} U_i Y^* U_i^* - T^{(1)}(Y^*) \right) \otimes \mathbf{1}\, X^* \otimes \mathbf{1} \right\|_1 \\
&\leqslant \|X \otimes \mathbf{1}\, |\psi\rangle\| \left\| X \otimes \mathbf{1} \left(\frac{1}{n} \sum_{i=1}^{n} U_i Y U_i^* - T^{(1)}(Y) \right) \otimes \mathbf{1}\, |\psi\rangle \right\| \\
&= \left\| X \left(\frac{1}{n} \sum_{i=1}^{n} U_i Y U_i^* - T^{(1)}(Y) \right) \otimes \mathbf{1}\, |\psi\rangle \right\| \\
&\leqslant \left\| X \left(\frac{1}{n} \sum_{i=1}^{n} U_i Y U_i^* - T^{(1)}(Y) \right) \otimes \mathbf{1} \right\|_\infty \\
&= \left\| X \left(\frac{1}{n} \sum_{i=1}^{n} U_i Y U_i^* - T^{(1)}(Y) \right) \right\|_\infty \\
&\leqslant \|X\|_\infty \left\| \frac{1}{n} \sum_{i=1}^{n} U_i Y U_i^* - T^{(1)}(Y) \right\|_\infty
\end{aligned}$$

where the second equality is because $\|X \otimes \mathbf{1}|\psi\rangle\| = \||\varphi\rangle\| = 1$. Now on the one hand $\|X\|_\infty \leqslant \|X\|_2 = \sqrt{d}$. And on the other hand, by Theorem 3.1 for $t = 1$ and $\epsilon/\sqrt{d}$ instead of ϵ, we get that, for $n \geqslant Cd^2(\log d)^6/\epsilon^2$, with probability at least $3/4$, for all Y such that $\|Y\|_2 = \sqrt{d}$,

$$\left\| \frac{1}{n} \sum_{i=1}^{n} U_i Y U_i^* - T^{(1)}(Y) \right\|_\infty \leqslant \frac{\epsilon}{d\sqrt{d}} \|Y\|_1 \leqslant \frac{\epsilon}{d} \|Y\|_2 = \frac{\epsilon}{\sqrt{d}}.$$

And thus, with probability at least $3/4$, for any $|\psi'\rangle$ orthogonal to $|\psi\rangle$,

$$\|X \otimes \mathbf{1}\, \Delta(|\psi\rangle\langle\psi'|)\, X^* \otimes \mathbf{1}\|_1 \leqslant \frac{3\epsilon}{2}.$$

Putting everything together, we obtain that, with probability at least $1/2$, for any state σ (in particular for $\sigma = \mathrm{id} \otimes \mathcal{N}(|\psi\rangle\langle\psi|)$),

$$\|X \otimes \mathbf{1}\, \Delta(\sigma)\, X^* \otimes \mathbf{1}\|_1 \leqslant \frac{3\epsilon}{2}.$$

Inserting this into equation (9), and re-labelling $3\epsilon/2$ in ϵ, yields exactly the claimed result. □

3.4 An alternative formulation

Given a linear map acting on a normed vector space (that of either operators or super-operators in our case), it is natural to define its so-called *induced norms*. For a linear map $\mathcal{M} : L(d) \to L(d)$, the most relevant induced norm is the one-to-one norm, as well as its completely bounded counterpart (also known as diamond norm). These are defined as

$$\|\mathcal{M}\|_{1\to 1} = \sup_{X \in L(d),\ \|X\|_1 \leqslant 1} \|\mathcal{M}(X)\|_1 \text{ and } \|\mathcal{M}\|_\diamond = \sup_{k \in \mathbf{N}} \|\mathrm{id}_k \otimes \mathcal{M}\|_{1\to 1},$$

where $\mathrm{id}_k : L(k) \to L(k)$ denotes the identity map on $L(k)$. By extension, for a linear map $\Xi : \mathcal{L}(d) \to \mathcal{L}(d)$, the most relevant induced norm is the diamond-to-diamond norm, as well as its completely bounded counterpart (which we denote with a double diamond). These are defined as

$$\|\Xi\|_{\diamond\to\diamond} = \sup_{\mathcal{M} \in \mathcal{L}(d),\ \|\mathcal{M}\|_\diamond \leq 1} \|\Xi(\mathcal{M})\|_\diamond \text{ and } \|\Xi\|_{\diamond\diamond} = \sup_{k \in \mathbf{N}} \|\mathrm{id}_k \otimes \Xi\|_{\diamond\to\diamond},$$

where $\mathrm{id}_k : \mathcal{L}(k) \to \mathcal{L}(k)$ denotes the identity map on $\mathcal{L}(k)$.

Using these definitions, we can reformulate Theorems 3.1 and 3.11 as follows:

Corollary 3.12. *Let $0 < \epsilon < 1$. Assume that the probability measure μ on $U(d)$ is a t-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. There exists a universal constant $C > 0$ such that, if $n \geqslant C(td)^t(t \log d)^6/\epsilon^2$, then with probability at least $1/2$, we have*

$$\left\| T^{(t)} - T^{(t)}_{\mu,n} \right\|_{1\to 1} \leqslant \epsilon,$$

where

$$T^{(t)}_{\mu,n} : X \in L(d^t) \mapsto \frac{1}{n} \sum_{i=1}^{n} U_i^{\otimes t} X U_i^{*\otimes t}.$$

Corollary 3.13. *Let $0 < \epsilon < 1$. Assume that the probability measure μ on $U(d)$ is a 2-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. There exists a universal constant $C > 0$ such that, if $n \geqslant Cd^2(\log d)^6/\epsilon^2$, then with probability at least $1/2$, we have*

$$\|\Theta - \Theta_{\mu,n}\|_{\diamond\to\diamond} \leq \epsilon,$$

where

$$\Theta_{\mu,n} : \mathcal{M} \in \mathcal{L}(d) \mapsto \left(\Theta_{\mu,n}(\mathcal{M}) : X \in L(d) \mapsto \frac{1}{n}\sum_{i=1}^{n} U_i \mathcal{M}(U_i^* X U_i) U_i^*\right).$$

For the applications in the next section, it is natural to define a *k-bounded* variant of the completely bounded one-to-one and diamond-to-diamond norms, i.e.

$$\|\mathcal{M}\|_{\diamond,k} = \|\mathrm{id}_k \otimes \mathcal{M}\|_{1\to 1} \text{ and } \|\Xi\|_{\diamond\diamond,k} = \|\mathrm{id}_k \otimes \Xi\|_{\diamond\to\diamond}.$$

By the Schmidt decomposition it is clear that $\|\cdot\|_{\diamond,d} = \|\cdot\|_\diamond$ and $\|\cdot\|_{\diamond\diamond,d^2} = \|\cdot\|_{\diamond\diamond}$. What is more, it is well-known that, for any $k \leqslant d$, $\|\cdot\|_{\diamond,k} \leqslant k\|\cdot\|_{1\to 1}$. We now prove a similar upper bound for $\|\cdot\|_{\diamond\diamond,k}$ in terms of $\|\cdot\|_{\diamond\to\diamond}$.

Lemma 3.14. *For any linear map $\Xi : \mathcal{L}(d) \to \mathcal{L}(d)$, we have*

$$\|\Xi\|_{\diamond\diamond,k} \leq k^2 \|\Xi\|_{\diamond\to\diamond}.$$

Proof. Let $\mathcal{M} \in \mathcal{L}(kd)$ with $\|\mathcal{M}\|_\diamond = 1$ be such that

$$\|\Xi\|_{\diamond\diamond,k} = \|\mathrm{id}_k \otimes \Xi\|_{\diamond\to\diamond} = \|\left(\mathrm{id}_k \otimes \Xi\right)(\mathcal{M})\|_\diamond.$$

By concavity of the diamond norm, we can assume that $\mathcal{M}$ has a Choi matrix $\eta_\mathcal{M} \in L(k^2d^2)$ of rank one, i.e. $\eta_\mathcal{M} = |\varphi_\mathcal{M}\rangle\langle\phi_\mathcal{M}|$ for some $|\varphi_\mathcal{M}\rangle, |\phi_\mathcal{M}\rangle \in \mathbf{C}^{k^2d^2}$. Let us now write $|\varphi_\mathcal{M}\rangle, |\phi_\mathcal{M}\rangle \in \mathbf{C}^{k^2} \otimes \mathbf{C}^{d^2}$ in their Schmidt decomposition:

$$|\varphi_\mathcal{M}\rangle = \sum_{i=1}^{k^2} \sqrt{p_i}|\alpha_i\beta_i\rangle \text{ and } |\phi_\mathcal{M}\rangle = \sum_{i=1}^{k^2} \sqrt{q_i}|\gamma_i\zeta_i\rangle,$$

where $\{|\alpha_i\rangle\}_{1\leqslant i\leqslant k^2}$, $\{|\gamma_i\rangle\}_{1\leqslant i\leqslant k^2}$ and $\{|\beta_i\rangle\}_{1\leqslant i\leqslant k^2}$, $\{|\zeta_i\rangle\}_{1\leqslant i\leqslant k^2}$ are orthonormal families in $\mathbf{C}^{k^2}$ and $\mathbf{C}^{d^2}$ respectively, and where $\{p_i\}_{1\leqslant i\leqslant k^2}$, $\{q_i\}_{1\leqslant i\leqslant k^2}$ are subnormalized probability distributions. We can hence write

$$\mathcal{M} = \sum_{i,j=1}^{k^2} \sqrt{p_i p_j}\mathcal{M}_{ij}^k \otimes \mathcal{M}_{ij}^d,$$

where $\mathcal{M}_{ij}^k : \mathcal{L}(k) \to \mathcal{L}(k)$ is defined as having Choi matrix $\eta_{\mathcal{M}_{ij}^k} = |\alpha_i\rangle\langle\gamma_j| \in L(k^2)$ and $\mathcal{M}_{ij}^d : \mathcal{L}(d) \to \mathcal{L}(d)$ is defined as having Choi matrix $\eta_{\mathcal{M}_{ij}^d} = |\beta_i\rangle\langle\zeta_j| \in L(d^2)$. We then have by the triangle inequality

$$\|\Xi\|_{\diamond\diamond,k} = \|\left(\mathrm{id}_k \otimes \Xi\right)(\mathcal{M})\|_\diamond \leqslant \sum_{i,j=1}^{k^2} \sqrt{p_i q_j}\|\mathcal{M}_{ij}^k \otimes \Xi(\mathcal{M}_{ij}^d)\|_\diamond = \sum_{i,j=1}^{k^2} \sqrt{p_i q_j}\|\Xi(\mathcal{M}_{ij}^d)\|_\diamond.$$

To finish the proof, we then simply have to observe that

$$\sum_{i,j=1}^{k^2} \sqrt{p_i q_j}\|\Xi(\mathcal{M}_{ij}^d)\|_\diamond \leqslant \|\Xi\|_{\diamond\to\diamond} \sum_{i,j=1}^{k^2} \sqrt{p_i q_j} = \|\Xi\|_{\diamond\to\diamond} \left(\sum_{i=1}^{k^2} \sqrt{p_i}\right)\left(\sum_{j=1}^{k^2} \sqrt{q_j}\right) \leqslant k^2\|\Xi\|_{\diamond\to\diamond},$$

where the first inequality is by definition of the diamond-to-diamond norm and the second inequality is due to the Cauchy-Schwarz inequality.

□

4 Application: Quantum non-malleable encryption against adversaries with small quantum memory

Information-theoretically secure quantum encryption has been studied extensively. In particular, the one-time variants of security goals such as confidentiality, authenticity and non-malleability have been defined for quantum encryption. When assessing the efficiency of a symmetric-key encryption scheme, there are three main figures of merit, the running time of the encryption and decryption algorithms, the ciphertext length and the key length. Here, we focus on the latter two figures of merit. Protocols have been designed which achieve the optimal scaling with respect to key length (up to log factors). More precisely, the results are as follows. The quantum one-time pad scheme, that encrypts a quantum system by applying a random element of the Pauli group, requires $2\log d$ bits of key [AMTD00]. The quantum authentication scheme presented in [BCG+02] uses $2\log d + O(s)$ bits of key and $\log d + O(s)$ bits of ciphertext to achieve s bits of security. And non-malleable encryption with unitaries (hence with plaintext space and ciphertext space being the same) can be done with $(4 + o(1))\log d$ bits of key [ABW09]. Here we describe a construction for non-malleable encryption *without adversarial side information* with unitaries using $2\log d + O(\log\log d)$ bits of key. In addition, our scheme has confidentiality against adversaries *with side information*. In other words, it is an alternative to the standard quantum one-time pad with the added property of non-malleability without side information at only an additive logarithmic cost in terms of key length.

One-time-secure quantum encryption

We begin by defining more rigorously the different cryptographic notions mentioned above. In the following, given a a finite set $\mathcal{X}$, the notation $\mathbf{E}_{x\in\mathcal{X}}$ is used to denote the expectation value of a random variable x distributed uniformly on $\mathcal{X}$.

Definition 4.1 (Quantum encryption scheme)**.** *A triple* $(\mathcal{X}, \mathrm{Enc}, \mathrm{Dec})$ *where*

i) $\mathcal{X}$ *is a finite set,*

ii) $\{\mathrm{Enc}_x\}_{x\in\mathcal{X}}$ *is a family of quantum channels* $\mathrm{Enc}_x : L(d_M) \to L(d_C)$*,*

iii) $\{\mathrm{Dec}_x\}_{x\in\mathcal{X}}$ *is a family of quantum channels* $\mathrm{Dec}_x : L(d_C) \to L(d_M)$

is called quantum encryption scheme if

$$\forall\ x \in \mathcal{X},\ \mathrm{Dec}_x \circ \mathrm{Enc}_x = \mathrm{id}.$$

The parameters $\log_2 |\mathcal{X}|$, $\log_2 d_M$ and $\log_2 d_C$ are called key length, message length and ciphertext length, respectively.

Given a quantum state $\sigma \in D(d)$, define the quantum channel $\langle\sigma\rangle \in \mathcal{C}(d)$ by $\langle\sigma\rangle(X) = \mathrm{Tr}(X)\sigma$.

Definition 4.2 (Indistinguishability of ciphertexts)**.** *A quantum encryption scheme has* ϵ*-indistinguishable ciphertexts, if there exists a quantum state* $\sigma \in D(d_C)$ *such that*

$$\|\mathbf{E}_{x\in\mathcal{X}}[\mathrm{Enc}_x] - \langle\sigma\rangle\|_\diamond \leq \epsilon.$$

A quantum encryption scheme has ϵ*-indistinguishable ciphertexts against adversaries without side information if the above inequality holds with the diamond norm replaced by the one-to-one norm.*

Definition 4.3 (Non-malleability)**.** *A quantum encryption scheme is ϵ-non-malleable, if there exists a quantum state $\sigma \in D(d_C)$ such that for all side information dimension d_E and all $\Lambda \in \mathcal{C}(d_C d_E)$ there exist completely positive maps $\Lambda_=, \Lambda_{\neq} \in \mathcal{L}(d_E)$ whose sum is trace-preserving and $p \in [0,1]$ such that*

$$\|\mathbf{E}_{x\in\mathcal{X}} [(\mathrm{Dec}_x \otimes \mathrm{id}) \circ \Lambda \circ (\mathrm{Enc}_x \otimes \mathrm{id})] - (p\,\mathrm{id} \otimes \Lambda_= + (1-p)\langle\sigma\rangle \otimes \Lambda_{\neq})\|_\diamond \leq \epsilon.$$

A quantum encryption scheme is ϵ-non-malleable against adversaries without side information, if there exists a quantum state $\sigma \in D(d_C)$ such that for all $\Lambda \in \mathcal{C}(d_C)$ there exists $p \in [0,1]$ such that

$$\|\mathbf{E}_{x\in\mathcal{X}} [\mathrm{Dec}_x \circ \Lambda \circ \mathrm{Enc}_x] - (p\,\mathrm{id} + (1-p)\langle\sigma\rangle)\|_\diamond \leq \epsilon.$$

Non-stabilized norms and adversaries without quantum side information

Any family of unitary matrices $\{U_x\}_{x\in\mathcal{X}}$ defines a quantum encryption scheme via the encryption channels $\mathrm{Enc}_x(X) = U_x X U_x^*$ and the decryption channels $\mathrm{Dec}_x = \mathrm{Enc}_x^*$. For such unitary quantum encryption schemes, it is easy to see that ϵ-indistinguishability of ciphertexts implies that the family of unitaries is a 2ϵ-approximate 1-design in diamond norm, and any ϵ-approximate 1-design in diamond norm gives rise to a quantum encryption scheme with ϵ-indistinguishable ciphertexts. The weaker property of ϵ-indistinguishability of ciphertexts against adversaries without side information and the ϵ-approximate 1-design property measured in one-to-one norm have the same relationship.

Similarly, if a unitary quantum encryption scheme is ϵ-non-malleable, then it is a 2ϵ-approximate channel twirl in completely bounded diamond-to-diamond norm, and an ϵ-approximate channel twirl in completely bounded diamond-to-diamond norm gives rise to a quantum encryption scheme that is ϵ-non-malleable [ABW09, AM17]. Again, the weaker ϵ-non-malleability against adversaries without side information and the ϵ-approximate channel twirl property measured in diamond-to-diamond norm have the same relationship.

The results in the previous section thus immediately imply the following for random unitary encryption schemes:

Theorem 4.4. *Let $0 < \epsilon < 1$. Assume that the probability measure μ on $U(d)$ is a 2-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. There exists a universal constant $C > 0$ such that, if $n \geqslant Cd^2(\log d)^6/\epsilon^2$, then with probability at least $1/2$, the quantum encryption scheme defined by the family of unitaries $\{U_1, \ldots, U_n\}$ has $\epsilon/\sqrt{d}$-indistinguishable ciphertexts and is ϵ-non-malleable against adversaries without side information.*

Proof. Let us define

$$T^{(1)}_{\mu,n} : X \in L(d) \mapsto \frac{1}{n}\sum_{i=1}^n U_i X U_i^* \text{ and } \Theta_{\mu,n} : \mathcal{M} \in \mathcal{L}(d) \mapsto \frac{1}{n}\sum_{i=1}^n U_i \mathcal{M}(U_i^* \cdot U_i) U_i^*. \quad (10)$$

To begin with notice that, if μ is a 2-design then it is a fortiori a 1-design. Hence by Corollary 3.12 (for $t = 1$ and $\epsilon/\sqrt{d}$ instead of ϵ) and Corollary 3.13, the probability that $T^{(1)}_{\mu,n}$ is an $\epsilon/\sqrt{d}$-approximate 1-design in one-to-one norm and the probability that $\Theta_{\mu,n}$ is an ϵ-approximate channel twirl in diamond-to-diamond norm are both at least $3/4$ for $n \geqslant Cd^2(\log d)^6/\epsilon^2$. By the union bound, both properties hold simultaneously with probability at least $1/2$. And as explained before, if this is so then the corresponding unitary quantum encryption scheme has $\epsilon/\sqrt{d}$-indistinguishable ciphertexts and is ϵ-non-malleable against adversaries without side information. □

Using the result of Lemma 3.14, relating the k-bounded diamond norm to the one-to-one norm and the k-bounded double diamond norm to the diamond-to-diamond norm, we can immediately derive from Theorem 4.4 a generalisation of it that applies to the case where the adversary has side information, but in bounded quantity.

Corollary 4.5. *Let $0 < \epsilon < 1$. Assume that the probability measure μ on $U(d)$ is a 2-design, and let $U_1, \ldots, U_n$ be sampled independently from μ. There exists a universal constant $C > 0$ such that, if $n \geqslant Cd^2(\log d)^6 k^4/\epsilon^2$, then with probability at least $1/2$, the quantum encryption scheme defined by the family of unitaries $\{U_1, \ldots, U_n\}$ has $\epsilon/k\sqrt{d}$-indistinguishable ciphertexts and is ϵ-non-malleable against adversaries with k-bounded side information.*

Proof. Let $T^{(1)}_{\mu,n}, \Theta_\mu$ be defined as in equation (10). We have shown in the proof of Theorem 4.4 that, for $n \geqslant Cd^2(\log d)^6/\epsilon^2$, with probability larger than $1/2$,

$$\left\|T^{(1)} - T^{(1)}_{\mu,n}\right\|_{1\to 1} \leqslant \frac{\epsilon}{\sqrt{d}} \text{ and } \|\Theta - \Theta_{\mu,n}\|_{\diamond\to\diamond} \leqslant \epsilon.$$

Now by Lemma 3.14, we know that this implies that

$$\left\|T^{(1)} - T^{(1)}_{\mu,n}\right\|_{\diamond,k} \leqslant \frac{\epsilon k}{\sqrt{d}} \text{ and } \|\Theta - \Theta_{\mu,n}\|_{\diamond\diamond,k} \leqslant \epsilon k^2.$$

Hence redefining ϵ as ϵk^2, we get that, for $n \geqslant Cd^2(\log d)^6 k^4/\epsilon^2$, with probability larger than $1/2$, $T^{(1)}_{\mu,n}$ is an $\epsilon/k\sqrt{d}$-approximate 1-design in k-bounded diamond norm and $\Theta_{\mu,n}$ is an ϵ-approximate channel twirl in k-bounded double diamond norm. And if this is so then the corresponding unitary quantum encryption scheme has $\epsilon/k\sqrt{d}$-indistinguishable ciphertexts and is ϵ-non-malleable against adversaries with k-bounded side information. □

A note on efficiency

While our scheme is more efficient in terms of key length and in terms of encryption and decryption given the element of the design that needs to be applied (if instantiated with an efficiently implementable 2-design, such as e.g. the Clifford group), specifying the randomly chosen subset of the exact 2-design is inefficient. This is a problem shared by all schemes based on the sub-sampling technique, i.e. in particular by the ones constructed in [HLSW04] and [ABW09]. To construct *efficiently specifiable* approximate designs in the weak norms we consider, that are still smaller than approximate designs in the diamond norm, additional new techniques seem to be necessary. A possible approach would be to, e.g., analyse random quantum circuits with respect to these norms.

Acknowledgements

C.M. thanks David Gross for discussions on t-designs. C.L. acknowledges financial support from the French CNRS (project PEPS JCJC). C.M. was supported by a NWO VIDI grant (Project No. 639.022.519) and a NWO VENI grant (Project No. VI.Veni.192.159).

References

[ABW09] Andris Ambainis, Jan Bouda, and Andreas Winter. Nonmalleable encryption of quantum information. *Journal of Mathematical Physics*, 50(4):042106, 2009.

[ADHW09] Anura Abeyesinghe, Igor Devetak, Patrick Hayden, and Andreas Winter. The mother of all protocols: Restructuring quantum information's family tree. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, page rspa20090202. The Royal Society, 2009.

[AM17] Gorjan Alagic and Christian Majenz. Quantum non-malleability and authentication. In Jonathan Katz and Hovav Shacham, editors, *Advances in Cryptology – CRYPTO 2017*, pages 310–341, Cham, 2017. Springer International Publishing.

[AMR19] Gorjan Alagic, Christian Majenz, and Alexander Russell. Efficient simulation of random states and random unitaries. Cryptology ePrint Archive, Report 2019/1204, 2019. https://eprint.iacr.org/2019/1204.

[AMTD00] A. Ambainis, M. Mosca, A. Tapp, and R. De Wolf. Private quantum channels. In *Proceedings 41st Annual Symposium on Foundations of Computer Science*, pages 547–553, Nov 2000.

[Aub09] Guillaume Aubrun. On almost randomizing channels with a short kraus decomposition. *Communications in Mathematical Physics*, 288(3):1103–1116, Jun 2009.

[BCG+02] H. Barnum, C. Crepeau, D. Gottesman, A. Smith, and A. Tapp. Authentication of quantum messages. In *The 43rd Annual IEEE Symposium on Foundations of Computer Science, 2002. Proceedings.*, pages 449–458, Nov 2002.

[BCR11] Mario Berta, Matthias Christandl, and Renato Renner. The quantum reverse shannon theorem based on one-shot information theory. *Communications in Mathematical Physics*, 306(3):579–615, 2011.

[BHH16] Fernando G. S. L. Brandão, Aram W. Harrow, and Michał Horodecki. Local random quantum circuits are approximate polynomial-designs. *Communications in Mathematical Physics*, 346(2):397–434, Sep 2016.

[DBWR14] Frédéric Dupuis, Mario Berta, Jürg Wullschleger, and Renato Renner. One-shot decoupling. *Communications in Mathematical Physics*, 328(1):251–284, 2014.

[DCEL09] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80:012304, Jul 2009.

[HHWY08] Patrick Hayden, Michał Horodecki, Andreas Winter, and Jon Yard. A decoupling approach to the quantum capacity. *Open Systems & Information Dynamics*, 15(01):7–19, 2008.

[HLSW04] Patrick Hayden, Debbie Leung, Peter W. Shor, and Andreas Winter. Randomizing quantum states: Constructions and applications. *Communications in Mathematical Physics*, 250(2):371–391, Sep 2004.

[HOW07] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Quantum state merging and negative information. *Communications in Mathematical Physics*, 269(1):107–136, 2007.

[JLS18] Zhengfeng Ji, Yi-Kai Liu, and Fang Song. Pseudorandom quantum states. In Hovav Shacham and Alexandra Boldyreva, editors, *Advances in Cryptology – CRYPTO 2018*, pages 126–152, Cham, 2018. Springer International Publishing.

[Kan15] Daniel Kane. Small designs for path-connected spaces and path-connected homogeneous spaces. *Transactions of the American Mathematical Society*, 367(9):6387–6414, 2015.

[LW17] Cécilia Lancien and Andreas Winter. Approximating quantum channels by completely positive maps with small kraus rank. *Preprint*, 2017. https://arxiv.org/abs/1711.00697.

[MBD+17] Christian Majenz, Mario Berta, Frédéric Dupuis, Renato Renner, and Matthias Christandl. Catalytic decoupling of quantum information. *Phys. Rev. Lett.*, 118:080503, Feb 2017.

[MSvW19] Christian Majenz, Christian Schaffner, and Jeroen van Wier. Non-malleability for quantum public-key encryption. Cryptology ePrint Archive, Report 2019/496, 2019. https://eprint.iacr.org/2019/496.

[SDTR13] Oleg Szehr, Frédéric Dupuis, Marco Tomamichel, and Renato Renner. Decoupling with unitary approximate two-designs. *New Journal of Physics*, 15(5):053022, 2013.