

Epsilon-nets, unitary designs and random quantum circuits

Michał Oszmaniec

*Center for Theoretical Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668 Warszawa, Poland and
International Centre for Theory of Quantum Technologies,
University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland*

Adam Sawicki

*Center for Theoretical Physics, Polish Academy of Sciences,
Al. Lotników 32/46, 02-668 Warszawa, Poland*

Michał Horodecki

*International Centre for Theory of Quantum Technologies,
University of Gdańsk, Wita Stwosza 63, 80-308 Gdańsk, Poland*

Epsilon-nets and approximate unitary t -designs are natural notions that capture properties of unitary operations relevant for numerous applications in quantum information and quantum computing. The former constitute subsets of unitary channels that are epsilon-close to any unitary channel. The latter are ensembles of unitaries that (approximately) recover Haar averages of polynomials in entries of unitary channels up to order t .

In this work we establish quantitative connections between these two seemingly different notions. Specifically, we prove that, for a fixed dimension d of the Hilbert space, unitaries constituting δ -approximate t -expanders form ϵ -nets in the set of unitary channels for $t \simeq \frac{d^{5/2}}{\epsilon}$ and $\delta \simeq (\epsilon/C)^{\frac{d^2}{2}}$, where C is a numerical constant. Conversely, we show that ϵ -nets with respect to this metric can be used to construct δ -approximate unitary t -expanders for $\delta \simeq \epsilon t$.

We further apply our findings in conjunction with the recent results of [1] in the context of quantum computing. First, we show that approximate t -designs can be generated by shallow random circuits formed from a set of universal two-qudit gates in the parallel and sequential local architectures considered in [2]. Importantly, our gate sets need not to be symmetric (i.e. contains gates together with their inverses) or consist of gates with algebraic entries. Second, we consider a problem of compilation of quantum gates and prove a non-constructive version of the Solovay-Kitaev theorem for general universal gate sets. Our main technical contribution is a new construction of efficient polynomial approximations to the Dirac delta in the space of quantum channels, which can be of independent interest.

I. INTRODUCTION

Approximate t -designs and ϵ -nets are natural proxies of the set of all unitary transformations of a finite-dimensional Hilbert space. They capture complementary aspects of unitary channels. We start by reviewing here relevance and contexts in which they appear in quantum information theory.

Unitary approximate t -designs [3] are tailored to reproduce statistical moments of degree at most t of the Haar measure on the unitary group. As such, they find numerous applications throughout quantum information, including randomized benchmarking [4], efficient estimation of properties of quantum states [5], decoupling [6], information transmission [7] and quantum state discrimination [8]. Pseudo-random unitaries are also used to model equilibration of quantum systems [2, 9], quantum metrology with random bosonic states [10] and in order to model scrambling inside black holes [11–13]. Recently, approximate unitary designs got a lot of attention in the context of proposals for attaining the so-called quantum computational advantage [14], especially random circuit sampling [15] that was recently realized experimentally by Google [16]. The reason for this is the anticoncentration property [17, 18], which seems essential in the proofs of quantum speedup.

Recently, there was a lot of interest in efficient implementations of pseudo-random quantum unitaries. First, it is known that the multi-qubit Clifford group forms an exact unitary 3-design but fails to be a unitary 4-design [19]. Second, in [2] it was shown that random circuits built from Haar-random 2-qubit gates acting (according to the specified layout) on N -qubit systems of the depth polynomial in N form approximate t -designs. This result holds also if the random two-qubit gate set is replaced by a universal gate set that is symmetric (i.e. contains gates together with their inverses) and consists of gates with algebraic entries. Importantly, both of these requirements are crucial as the arguments of [2] heavily rely on the work by Bourgain and Gamburd [20]. These results were later improved in 2018

in [21], where even faster convergence in n was proved using specially design layouts in which random two-qubit gates were placed. Additionally, recent work [22] (partially) lifted these stringent requirements. Moreover, the authors of [23] showed that random circuits constructed from Clifford gates and a small number of non-Clifford can be used to efficiently generate approximate designs. Finally, there exist proposals for efficient generation of approximate t -designs using diagonal gates [24] and via Hamiltonian [12] and stochastic [25] dynamics.

Epsilon nets form (often discrete) subsets of the set of unitary channels that approximate every unitary operation up to some accuracy. They appear naturally in the context of *compilation* of quantum gates, i.e. the task is to approximate a target unitary gate via the sequence of elementary gates belonging to some "simple" gate-set \mathcal{G} . Traditionally, compilation of quantum gates is carried out using the celebrated Solovay-Kitaev algorithm [26, 27] which states that for any universal and *symmetric* gate-set \mathcal{G} and any target quantum gate U , there exist a sequence of gates from \mathcal{G} that ϵ -approximates U and has length $l \sim \log(\frac{1}{\epsilon})^c$ for $c \approx 3.97$. Moreover, the aforementioned sequence can be found efficiently. Importantly, the Solovay-Kitaev algorithm requires the gate-set to be symmetric as in the course of the compilation it is necessary to perform *group commutators*. There have recently appeared works which partially lifted this restriction by assuming that the gate-set in question contains an irreducible representation of a group [28, 29]. We also note that the relation between efficient gate approximations and spectral gaps (here we study spectral gaps on restricted spaces rather than on the full space of functions on the unitary group) have been previously used in [30].

The notions of approximate t -designs and epsilon-nets seem to be intuitively related but, according to our best knowledge, the quantitative connection between them has not been systematically studied before. We would like to remark however that analysis of the proof of Theorem 5 in [31] allows to infer that approximate t -designs define ϵ -nets for $t \simeq d^3/\epsilon^2$ and $\delta \simeq (\epsilon/\sqrt{d})^{2d^2}$. Moreover, a related problem was recently studied in the context of harmonic analysis on Lie groups. Specifically, recent work in [1] established quantitative relation between spectral gaps on groups and epsilon nets on these manifolds $t \simeq \epsilon^{-2d^2}$ and $\delta \simeq \epsilon^{(d^4+d^2)/2-1}$. We will comment on the relation of these findings to our Result 1 in Section VII.

Overview of the results and their significance— In our work we aim to provide quantitative relation between ϵ -nets and approximate unitary designs. We follow closely the approach that was put forward in [1], where for a semi-simple compact connected Lie group G it was shown that ϵ -nets follow from spectral gaps of certain "transition operators" (defined via the gate-set of interest, and acting on the function spaces built from the irreps of G). We translate these to the quantum information language and observe that when a Lie group G is a group of quantum channels $U(d)$ (isomorphic to the projective unitary group), spectral gaps of the aforementioned transition operators are in one to one correspondence with the parameter δ in the definition of δ -approximate t -expanders [32] (see Eq. (2.3)). Making use of this correspondence, we show that δ -approximate t -designs can form ϵ -nets. We modify the construction proposed in [1] and attain better dependence of t on ϵ and d , the dimension of the Hilbert space (see Section VII for a detailed discussion). Moreover, our arguments do not depend on the detailed knowledge of the representation theory and are instead based solely on the geometry of quantum channels (Result 1). We also prove a converse result i.e. that ϵ -nets can be used to define approximate t -expanders.

These general results are then applied to different problems in quantum computation. First, we give a necessary and sufficient criterion for universality of *any* collection of quantum gates. Second, Result 3 shows a (nonconstructive) variant of Solovay-Kitaev theorem for gate-sets \mathcal{G} that, in contrasts do the existing results, does not require inverses (see the discussion on Solovay-Kitaev theorem above). Finally, we prove in Result 4 that short random quantum circuits generated from two-qubit universal gate-sets \mathcal{G} placed in the parallel and sequential layouts considered in [2] form approximate t -designs. Crucially, compared to previous approaches (see the discussion above) we do not require \mathcal{G} to be symmetric or to have algebraic entries.

Structure of the paper— In Section II we introduce basic concepts and notation. In particular, we describe t -designs and approximate t -designs, and introduce a notion of distance with respect to which we define epsilon nets. This allows us to formulate our main results in Section III. Then, in Section IV we discuss open problems and possible further applications of our results. In Section V we introduce notion of mixing operator T_μ defined on functions acting on unitary channels, and its gap. We relate the operator to the moment operators $T_{\mu,t}$ introduced in Section II.

After these preliminary sections, we are in position to give formal statements and proofs of our findings. The first group of results concerns arbitrary measures on unitary channels. And so, in Section VI we prove that for t large enough, an exact t -design forms an ϵ -net, in Section VII we show that approximate t -design also does. In that part we also prove the result in converse direction, namely that from a ϵ -net with small enough ϵ one can construct a t -design. In Then we move to measures obtained from uniform distribution on sequences of gates from some gate set. In section VIII we apply the above results (employing some additional results from [1]) to prove a nonconstructive version of Solovay-Kitaev theorem which does not require assumption that gate set contains inverses.

Finally, we consider much more structured measures - namely random circuits on n qudits. In Section IX we prove that random circuits (local and parallel) form approximate t -designs without assuming that the gate set contains inverses, and the gates have algebraic entries. We conclude the main part of the article with Section X where we

outline the construction of the polynomial approximation of the Dirac delta on the group of unitary channels. This polynomial function plays a crucial role in the proofs of the results from Sections VI and VII.

The Appendix is largely devoted to technical results needed in the construction of the aforementioned polynomial approximation of the Dirac delta. Some of the results presented there can be of independent interest because of the intriguing connection with the random matrix theory (specifically, Tracy-Widom distribution [33] and distribution of operator norm of GUE matrices).

II. MAIN CONCEPTS AND NOTATIONS

Throughout this work we will be concerned with unitary channels acting on a d -unitary dimensional Hilbert space $\mathcal{H} \simeq \mathbb{C}^d$. A unitary channel is a CPTP map defined by $\mathbf{U}[\rho] = U\rho U^\dagger$, where ρ is a quantum state and $U \in \mathbb{U}(d)$ is a unitary operator on \mathbb{C}^d . In what follows we will denote by $\mathbf{U}(d)$ the set of all unitary quantum channels on \mathbb{C}^d . Note that every unitary operator U uniquely defines a quantum channel but the converse is not true: a quantum channel \mathbf{U} specifies a unitary U up to a global phase. For this reason we can identify $\mathbf{U}(d)$ with the projective unitary group $\text{PU}(d) = \mathbb{U}(d)/\mathbb{U}(1)$. Therefore, $\mathbf{U}(d)$ is a compact connected semi-simple Lie group [34] (we will use this observation in what follows).

In order to define the notion of ϵ -net we need to first specify the distance in the set of unitary channels. We will consider the metric induced by the diamond norm $D(\mathbf{U}, \mathbf{V}) := \|\mathbf{U} - \mathbf{V}\|_\diamond$. This notion of distance has strong operational interpretation of in terms of maximal statistical distinguishability of quantum channels [35]. We will use the following equivalent expression for the diamond norm (see Theorem 26 in [36])

$$D(\mathbf{U}, \mathbf{V}) = \min_{\varphi \in [0, 2\pi)} \|U - \exp(i\varphi)V\|_\infty, \quad (2.1)$$

where $\|\cdot\|_\infty$ denotes the operator norm and U, V are unitaries representing channels \mathbf{U} and \mathbf{V} respectively. We say that a subset $\mathcal{S} \subset \mathbf{U}(d)$ is an ϵ -net (with respect to the metric D), if for every $\mathbf{U} \in \mathbf{U}(d)$ there exist $\mathbf{V} \in \mathcal{S}$ such that $D(\mathbf{U}, \mathbf{V}) \leq \epsilon$. A set of gates $\mathcal{G} \subset \mathbf{U}(d)$ is called *universal* if sequences $\mathbf{V}_n \mathbf{V}_{n-1} \dots \mathbf{V}_1$ of gates from \mathcal{G} form ϵ -nets in $\mathbf{U}(d)$ for arbitrary small ϵ .

The set of unitary channels $\mathbf{U}(d)$ inherits the unique invariant normalized measure from the unitary group $\mathbb{U}(d)$ according to the following prescription. For $\mathcal{S} \subset \mathbf{U}(d)$ we set $\mu_P(\mathcal{S}) = \mu(\varphi^{-1}(\mathcal{S}))$, where μ_P, μ are Haar measures on $\mathbf{U}(d)$ and $\mathbb{U}(d)$ respectively, and $\varphi^{-1}(\mathcal{S})$ is the set of all unitary operators that define quantum channels belonging to \mathcal{S} . Haar measure on $\mathbf{U}(d)$ can be also defined via the action on functions of unitaries that are invariant under the global phase (i.e. $F(\exp(i\alpha)U) = F(U)$, for arbitrary $U \in \mathbb{U}(d)$ and $\alpha \in \mathbb{R}$), $\int_{\mathbf{U}(d)} d\mu_P(\mathbf{U}) F(\mathbf{U}) = \int_{\mathbb{U}(d)} d\mu(U) F(U)$. In what follows we will not differentiate between unitary channels and unitary operators, as well as Haar measures defined on these sets, unless it leads to ambiguity. In particular, will denote by $\text{Vol}(\mathcal{S})$ the Haar measure of a subset \mathcal{S} of unitary channels $\mathbf{U}(d)$ or unitary group $\mathbb{U}(d)$, depending on the context. We will also use the notation $d\mu(\mathbf{U})$ and $d\mu(U)$ for "densities" of Haar measures on $\mathbf{U}(d)$ and $\mathbb{U}(d)$ respectively.

An ensemble of unitaries \mathcal{E} characterized by the probability measure ν is called a t -design [3] iff

$$\int_{\mathbf{U}(d)} d\nu(U) G_t(U) = \int_{\mathbb{U}(d)} d\mu(U) G_t(U), \quad (2.2)$$

where G_t is arbitrary *balanced* polynomial in $\mathbb{U}(d)$, i.e. a function of the form $G_t = \text{tr}(AU^{\otimes t} \otimes \bar{U}^{\otimes t})$, where A is an operator on $(\mathbb{C}^d)^{\otimes 2t}$. Note that balanced polynomials on $\mathbb{U}(d)$ are well defined functions on $\mathbf{U}(d)$. In this work we will be predominantly interested in *discrete* ensembles, i.e. the ones that take the form $\mathcal{E} = \{\nu_i, U_i\}$, for which $\int_{\mathbf{U}(d)} d\nu(U) F(U) = \sum_i \nu_i F(U_i)$. Approximate unitary t -designs (see for example [2, 37]) are ensembles ν of unitaries that satisfy (2.2) up to some desired accuracy. In this work we will focus on a version of approximate t -designs called δ -approximate t -expanders defined as ensembles ν satisfying

$$\|T_{\nu,t} - T_{\mu,t}\|_\infty \leq \delta, \quad (2.3)$$

where for any measure ν (in particular for the Haar measure μ) we define a *moment operator*

$$T_{\nu,t} := \int_{\mathbf{U}(d)} d\nu(U) U^{\otimes t} \otimes \bar{U}^{\otimes t}. \quad (2.4)$$

The quantity $\delta(t, \nu) := \|T_{\nu,t} - T_{\mu,t}\|_\infty$ is sometimes called *expander norm* of ν . There exist other related definitions of approximate designs that use different quantifiers to gauge how well ν approximates the properties of Haar measure

μ (see for example [32]).

III. SUMMARY OF MAIN RESULTS

Here we present our main findings regarding the relation between approximate designs (expanders) and epsilon-nets.

Result 1 (Approximate t -expanders define ϵ -nets). *Consider an ensemble $\mathcal{E} = \{\nu_i, U_i\}$ of unitaries described by the discrete measure ν on $\mathbf{U}(d)$. Let $\epsilon \in [0, d)$ and assume that ensemble \mathcal{E} is a δ -approximate t -expander with $t \simeq \frac{d^{5/2}}{\epsilon}$ (up to logarithmic factors in d and $1/\epsilon$) and $\delta \simeq (\epsilon/C)^{\frac{d^2}{2}}$, where $C = 10\pi$ is a numerical constant. Then, the channels $\{U_i\}$ defined via the elements of \mathcal{E} form an ϵ -net in $\mathbf{U}(d)$ with respect to the distance D induced by the diamond norm.*

We note that setting $\delta = 0$ gives the connection between *exact* t -designs and ϵ -nets. We give the technical formulation of the above result in Theorems 2 and 3. There we state the explicit dependence of t and δ on the dimension of the Hilbert space d and generalise the above statements to arbitrary probability measures (ensembles) on $\mathbf{U}(d)$. Our proofs follow the method presented in [1]. Our technical contributions are twofold. First, we simplify the original arguments making them largely independent of the machinery of group theory and thus more accessible for the broader audience. Second, in Theorem 1 we construct an efficient polynomial approximation of the Dirac delta on $\mathbf{U}(d)$ which allows us to attain better dependence of t on the dimension d and ϵ . Our construction can be of independent interests and its details are provided in Section X.

Result 1 can be used to find out how many times one needs to iterate gates comprising the δ -approximate t -design so that they form an ϵ -net. Specifically in Proposition 2 we prove that that it is enough to iterate them $l \simeq \frac{d^2 \log(\frac{1}{\epsilon})}{\log(\frac{1}{\delta})}$ times. This establishes intriguing connection between complexity of quantum gates and the property of being approximate t -design.

We also prove the connection in the opposite direction. Namelly, we show that epsilon nets can be used to construct approximate designs (see Theorem 4 for the formal statement).

Result 2 (ϵ -nets define approximate t -designs). *Consider a gate-set \mathcal{S} forming an ϵ -net in $\mathbf{U}(d)$. Then, there exists an ensemble of quantum gates from \mathcal{S} which forms an $(2\epsilon t)$ -approximate t -expander.*

We apply the results established above in the context of quantum computing. To this end we use additional ingredient which follows from Theorem 6 of [1] (see Section V and Theorem 5 for the translation of representation-theoretic concepts to the formalism of tensor expanders). Specifically, the spectral gap of the moment operator $T_{\nu,t}$ associated to a measure ν supported on a universal gate-set \mathcal{G} closes not faster than $\frac{A}{\log(t)^2}$. Note that the above relies solely on universality of \mathcal{G} so that the assumptions made e.g. in [20] on algebraic entries of gates and the property that \mathcal{G} is symmetric (i.e. $V \in \mathcal{G}$ implies $V^{-1} \in \mathcal{G}$) are not relevant. Leveraging this and the recent results of [38, 39], it is possible to prove that universality of *any* gate-set \mathcal{G} is equivalent to being δ -approximate t_* -expander, where $\delta < 1$ and t_* depends solely on d . This finding complements recent results [40, 41] that classified semi-simple compact Lie subgroups of $\mathbf{U}(d)$ in terms of their second order commutants).

Finally, we use the above strong spectral gap results of Varju to show the following two results which are relevant to theoretical underpinnings of quantum computing.

Result 3 (Non-constructive inverse-free Solovay-Kitaev). *Let $\mathcal{G} \subset \mathbf{U}(d)$ be a universal gate-set in $\mathbf{U}(d)$ (not necessarily symmetric i.e. $V \in \mathcal{G}$ does not imply $V^{-1} \in \mathcal{G}$). Then, every unitary channel U can be approximated by sequences of gates from \mathcal{G} of length $l \approx \log(\frac{1}{\epsilon})^3$.*

The formal proof can be found in Section VIII. We note that Result 3 does not give a constructive algorithm to find the approximating sequence of gates. Our last result shows that approximate t -designs can be generated efficiently by local random circuits without assuming inverses and algebraic entries. The formal proof is given Section IX.

Result 4. *Let \mathcal{G} be a set of universal two-qudit gates. Consider two types of random circuits on line of n qudits [2].*

- Local random circuits: *we pick uniformly at random two neighboring qudits, and apply gate chosen from \mathcal{G} according to uniform measure $\nu_{\mathcal{G}}$. We denote the resulting distribution by $\nu_{\text{loc}}(\mathcal{G})$.*
- Parallel random circuits: *we apply with probability $1/2$ either $U_{12} \otimes U_{34} \otimes \dots \otimes U_{n-1,n}$ or $U_{23} \otimes U_{45} \otimes \dots \otimes U_{n-2,n-1}$, where each U_{ij} is picked independently from \mathcal{G} according to $\nu_{\mathcal{G}}$. We denote the resulting distribution by $\nu_{\text{par}}(\mathcal{G})$.*

Let $l_{loc, Haar}$ ($l_{par, Haar}$) be lengths of random local (parallel) circuits which are δ -approximate t -expanders, where instead of $\nu_{\mathcal{G}}$ we take Haar measure over two-qudit gates. There exist a constant $C(\mathcal{G})$ such that if

$$l_{loc} \geq n \log^2(t) C(\mathcal{G}) l_{loc, Haar}, \quad l_{par} \geq 2 \log^2(t) C(\mathcal{G}) l_{par, Haar} . \quad (3.1)$$

then, the corresponding random circuits $(\nu_{loc}(\mathcal{G})^{*l_{loc}})$ and $(\nu_{par}(\mathcal{G})^{*l_{par}})$ are δ -approximate t -expanders.

Note that in [2] it was shown that local (parallel) random quantum circuits with Haar distributed gates of lengths satisfying

$$\begin{aligned} l_{loc, Haar} &\geq 42500n [\log_d(4t)]^2 d^2 t^{5+3.1 \log d} (2nt \log d + \log(1/\epsilon)) \\ l_{par, Haar} &\geq 523000 [\log_d(4t)]^2 d^2 t^{5+3.1 \log d} (2nt \log d + \log(1/\epsilon)) \end{aligned} \quad (3.2)$$

are δ -approximate t -expanders. It then follows that circuits constructed from \mathcal{G} scale efficiently with n , too.

Remark 1. It is straightforward to derive analogous bounds for other notions of approximate t -designs (based, for example, on the diamond norm). The conclusions are analogous. Let us stress, however, that our proof technique does not immediately apply to the scenarios considered in [21] and hence we cannot use it to get convergence faster than n for $\sqrt{n} \times \sqrt{n}$ - qubits square lattice. We however believe that this technical problem can be overcome with some effort.

Acknowledgements We are grateful to Stanisław Szarek for explaining to us the intricacies of computing volumes of balls in the unitary group and related manifolds. AS acknowledges financial support from National Science Centre, Poland under the grant SONATA BIS: 2015/18/E/ST1/00200. MH acknowledges support from the Foundation for Polish Science through IRAP project co-financed by EU within the Smart Growth Operational Programme (contract no.2018/MAB/5). MO acknowledges the financial support by TEAM-NET project (contract no. POIR.04.04.00-00-17C1/18-00).

IV. OPEN PROBLEMS

We conclude the introductory part of our work with a list of interesting problems which we left for further research.

- *Optimal scaling of $t(\epsilon, d)$ and $\delta(t, d)$:* Can one improve scaling in the results connecting ϵ -nets with t -designs? We conjecture that the dependence δ is essentially optimal. The same concerns scaling of t with ϵ . We conjecture that with some work it should be possible to obtain $t \simeq d^2$ (for fixed ϵ).
- *Explicit constant in SK theorem:* Unlike in all other results, our version of Solovay-Kitaev theorem contains an unknown constant depending on the dimension and the gate set. To what extent we can determine it (at least to leading order in the dimension)?
- *Improve constants:* dependence on dimension in our estimates is not necessarily optimal, as we have applied quite crude estimate of volume of projective ball. There is also much room for improvement in other places - e.g. estimates such as in Lemma 2 can be sharpened at least for small d . It is not however excluded, that dependence on dimension is, at least asymptotically, optimal.
- *Termination of the universality checking algorithm:* The explicit value of constant in our version of Solovay-Kitaev theorem and the connection between approximate t -designs and ϵ -nets can shed a new light on complexity of universality checking algorithms proposed in [38].
- *Connection with black hole dynamics and complexity growth:* Recently, there were some interesting works connecting complexity of random circuits with black hole dynamics (cf. [13, 42, 43]). It is conceivable that our findings may provide some useful tools, especially in the high complexity regime. In this context it is also natural to explore the possible generalizations of our results to *approximate projective designs* and ϵ -nets in the set of pure quantum states.

V. MIXING OPERATORS ON UNITARY GROUP, THEIR GAP AND APPROXIMATE DESIGNS

In this section we establish the connection between spectral gaps of mixing operators on unitary channels and approximate unitary t -designs (expanders). Let $L^2(\mathbb{U}(d))$ be the Hilbert space space of square-integrable functions on $\mathbb{U}(d)$, i.e. functions satisfying $\int_{\mathbb{U}(d)} d\mu(U) |F(U)|^2 < \infty$, where μ denotes the Haar measure on $\mathbb{U}(d)$. For every $V \in \mathbb{U}(d)$ we introduce a shift operator $T_V : L^2(\mathbb{U}(d)) \rightarrow L^2(\mathbb{U}(d))$ defined via $(T_V(F))(U) = F(V^{-1}U)$. For every

measure ν on $\mathbb{U}(d)$ we can consider an operator $T_\nu : L^2(\mathbb{U}(d)) \rightarrow L^2(\mathbb{U}(d))$ which is defined as a convex combination of operators T_V according to measure ν , $T_\nu = \int_{\mathbb{U}(d)} d\nu(V) T_V$. Its action on functions on $\mathbb{U}(d)$ can be explicitly written as

$$(T_\nu F)(U) = \int_{\mathbb{U}(d)} d\nu(V) F(V^{-1}U) . \quad (5.1)$$

The operator T_ν can be understood it as a transition operator of a random walk on $\mathbb{U}(d)$ in which at every step a unitary is applied at random according to the measure ν .

We shall also consider restriction $T_\nu|_{\mathcal{H}_t}$ of T_ν to the subspace \mathcal{H}_t spanned by balanced polynomials of degree up to t in U as well as in \bar{U} i.e. subspace of functions on $\mathbb{U}(d)$ of the form $G_t(U) = \text{tr}(AU^{\otimes t} \otimes \bar{U}^{\otimes t})$. In particular, if we choose ν to be the Haar measure μ on $\mathbb{U}(d)$ then the operators T_μ and $T_\mu|_{\mathcal{H}_t}$ are projectors - they project onto the space of constant functions on $\mathbb{U}(d)$. Let us denote the space orthogonal to the constant functions on $\mathbb{U}(d)$ by $L_0^2(\mathbb{U}(d))$. We define the gap of T_ν as:

$$g(T_\nu) := 1 - \|T_\nu|_{L_0^2(\mathbb{U}(d))}\|_\infty . \quad (5.2)$$

We note that the so-defined function is a gap, when the support of the measure includes a set of universal gates. We are only interested in such situation, so we will keep name gap for g .

We define a gap for $T_\nu|_{\mathcal{H}_t}$ analogously as for T_ν and denote it by $g(\nu, t)$. By straightforward calculations we get

$$\|T_\mu|_{\mathcal{H}_t} - T_\nu|_{\mathcal{H}_t}\|_\infty = 1 - g(\nu, t) \quad (5.3)$$

The following proposition establishes a very useful connection between $T_\nu|_{\mathcal{H}_t}$ and moment operator $T_{\nu,t}$ introduced in Eq.(2.4).

Proposition 1. *For any measure ν on $\mathbb{U}(d)$ we have*

$$\|T_\mu|_{\mathcal{H}_t} - T_\nu|_{\mathcal{H}_t}\|_\infty = \|T_{\mu,t} - T_{\nu,t}\|_\infty \quad (5.4)$$

and consequently we have $\delta(\nu, t) = 1 - g(\nu, t)$, where $\delta(\nu, t)$ is the expander norm of ν .

Proof. The action of T_ν and T_μ on \mathcal{H}_t is determined by the left regular representation. Under this action \mathcal{H}_t decomposes into irreducible components $\mathcal{H}_t = \bigoplus_\lambda \mathcal{K}^\lambda$ and we have

$$T_\nu|_{\mathcal{H}_t} \approx \bigoplus_\lambda \int_{\mathbb{U}(d)} d\nu(U) \Pi^\lambda(U) , \quad (5.5)$$

where $\Pi^\lambda(U)$ is the matrix corresponding to U via the irreducible representation with the highest weight λ and the symbol \approx denotes unitary equivalence. On the other hand the representation $U \mapsto U^{\otimes t} \otimes \bar{U}^{\otimes t}$ is reducible and decomposes into

$$U^{\otimes t} \otimes \bar{U}^{\otimes t} \approx \bigoplus_{\lambda'} \Pi^{\lambda'}(U) . \quad (5.6)$$

Thus the operator $T_{\nu,t}$ can be written as

$$T_{\nu,t} \approx \bigoplus_{\lambda'} \int_{\mathbb{U}(d)} d\nu(V) \Pi^{\lambda'}(U) . \quad (5.7)$$

We notice, however, that the space \mathcal{H}_t is spanned by the matrix elements of the representation $U^{\otimes t} \otimes \bar{U}^{\otimes t}$ and hence, by the decomposition (5.6), by matrix elements of irreducible representations $\Pi^{\lambda'}$. Let $\mathcal{W}^{\lambda'}$ be the linear span of functions $F_{ij}^{\lambda'}(U) = \langle i | \Pi^{\lambda'}(U) | j \rangle$. It can be verified by direct computation that for every $V \in \mathbb{U}(d)$ we have

$$T_V|_{\mathcal{W}^{\lambda'}} \approx \Pi^{\lambda'}(V) \otimes \mathbb{I}_{m^{\lambda'}} , \quad (5.8)$$

where $m^{\lambda'}$ is the dimension of the multiplicity space equal to $|\mathcal{K}^{\lambda'}|$, the dimension of carrier space of representation $\Pi^{\lambda'}$. Thus it follows that that collection of weights $\{\lambda\}$ and $\{\lambda'\}$ agree, up to multiplicities. The theorem now follows from comparing decompositions (5.5) and (5.7). \square

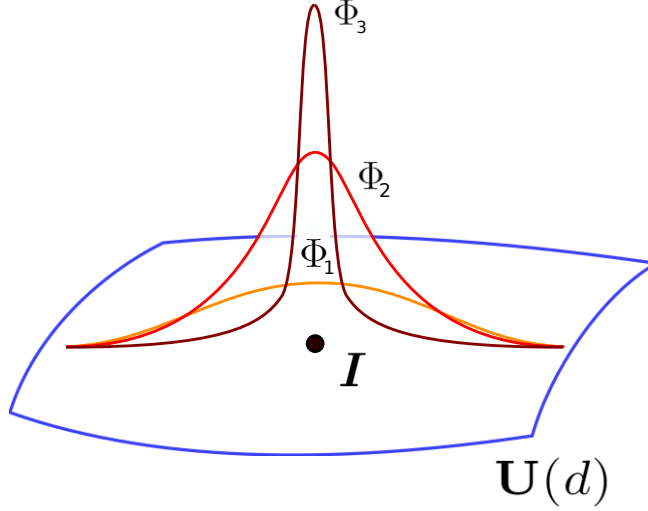


FIG. 1. A graphical presentation of a sequence of polynomial approximations $\Phi_k(\mathbf{U})$ of Dirac delta at \mathbf{I} in $\mathbf{U}(d)$. As the degree k increases the functions Φ_k are more and more peaked in the vicinity of \mathbf{I} , while retaining the normalisation $\int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \Phi_k(\mathbf{U}) = 1$.

As an immediate consequence we get that for a t -design, the gap $g(\nu, t)$ is equal to 1. We conclude this part by noting that composition of operator T_ν is compatible with taking convolutions in the sense that for all l we have $T_{\nu^{*l}} = (T_\nu)^l$. This implies the following well-known result.

Fact 1. *If ν is a δ -approximate t -expander, then ν^{*l} is a δ^l -approximate t -expander.*

VI. EXACT T-DESIGNS AND EPSILON-NETS

In this part we will show that elements of exact t -designs form ϵ -nets with respect to the diamond norm distance provided $t \simeq \frac{d^{5/2}}{\epsilon}$ (up to logarithmic factors in d and $1/\epsilon$). We follow the ideas from [1] with two important differences. First, we significantly reduce the usage of representation theory. Second, we construct a new polynomial approximation of the Dirac delta on the group of quantum channels (see Theorem 1 and Section X where we provide details of the construction). This allows us to obtain improved dependence of t on d and ϵ in Theorem 2.

We start with giving the intuition beyond the proof of our result. We consider a family of real-valued balanced polynomials $\Phi_k \in \mathcal{H}_k$ (i.e. polynomials of degree at most k) that has the following properties:

- Normalisation: $\int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \Phi_k(\mathbf{U}) = 1$, for all k .
- Vanishing integrals on balls sufficiently far from identity \mathbf{I} : for every $\epsilon \in [0, 2]$ and for every \mathbf{V}_0 such that $D(\mathbf{V}_0, \mathbf{I}) \geq \epsilon$ we have

$$\int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U}) \Phi_k(\mathbf{U}) \rightarrow 0 \text{ as } k \rightarrow \infty, \quad (6.1)$$

where $B(\mathbf{V}_0, \epsilon) = \{\mathbf{U} \in \mathbf{U}(d) \mid D(\mathbf{U}, \mathbf{V}_0) \leq \epsilon\}$.

Functions Φ_k can be regarded as polynomial approximation of the Dirac delta localized at \mathbf{I} , the identity channel (see Fig. 1).

We then consider the following integral,

$$I(\nu, \epsilon, k, \mathbf{V}_0) := \int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U}) (T_\nu \Phi_k)(\mathbf{U}), \quad (6.2)$$

where $\mathbf{V}_0 \in \mathbf{U}(d)$, and for any measure $\nu = \{\nu_i, V_i\}$ and a function \mathcal{F} on $\mathbf{U}(d)$ we define (analogously as before for functions on $\mathbb{U}(d)$) $(T_\nu \mathcal{F})(\mathbf{U}) := \sum_i \nu_i \mathcal{F}(\mathbf{V}_i^{-1} \mathbf{U})$. Next, under the assumption that ν is an exact k -design we show in

Lemma 1 that $I(\nu, \epsilon, k, \mathbf{V}_0)$ equals the Haar measure of $B(\mathbf{V}_0, \epsilon/2)$. On the other hand if channels from the support of ν do not form an ϵ -net in $\mathbf{U}(d)$ we can use Eq.(6.1) to prove that $I(\nu, \epsilon, k, \mathbf{V}_0)$ vanishes as $k \rightarrow \infty$ (see Lemma 2 and Theorem 1). We finally look for k such that $I(\nu, \epsilon, k, \mathbf{V}_0)$ is smaller than $\text{Vol}(B(\mathbf{V}_0, \epsilon/2))$, value of which is controlled by Fact 2. This number gives a degree of exact t -design that is ensured to form ϵ -net. The graphical presentation of this general reasoning is given in Fig. 2 while technical details are given below. The main result connecting exact t designs with ϵ nets is Theorem 2.

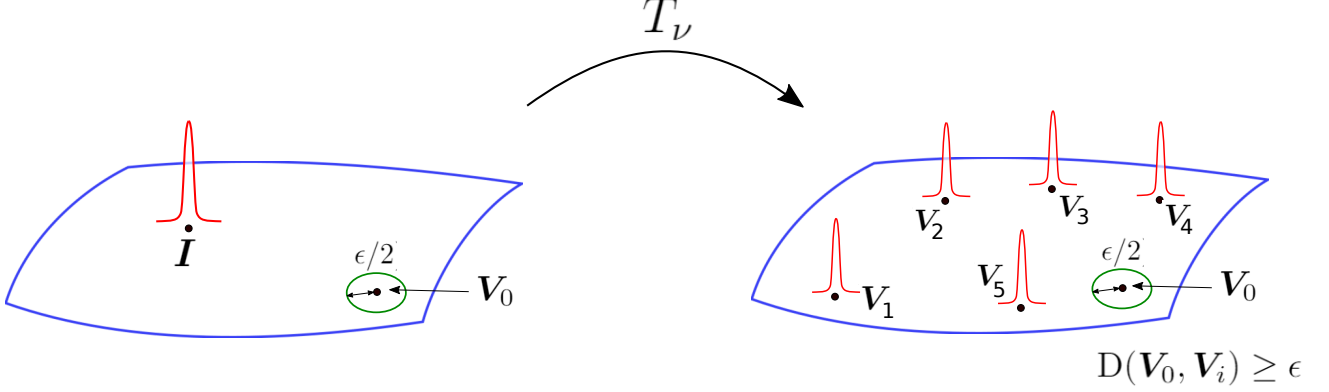


FIG. 2. A visualisation of the general argument that allows to connect t -designs with ϵ -nets. From the t -design property we know that for $k \leq t$ the integral $\int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U})(T_\nu \Phi_k)(\mathbf{U})$ equals $\text{Vol}(B(\mathbf{V}_0, \epsilon/2))$, the volume of the Ball of the radius $\epsilon/2$ centered around \mathbf{V}_0 (interior of the green cycle). On the other hand action of the transition operator T_ν transforms the function Φ_k , initially localized around \mathbf{I} (red peak in the left part of the figure) into a convex combination of functions $\Phi_k^i(\mathbf{U}) = \Phi_k(\mathbf{V}_i^{-1}\mathbf{U})$ localized around points $\mathbf{V}_i \in \text{supp}(\nu)$ (smaller red peaks in the right part of the figure). Assuming that $\text{supp}(\nu)$ does not form an ϵ -net we know that there exist \mathbf{V}_0 that satisfies $D(\mathbf{V}_0, \mathbf{V}_i) \geq \epsilon$. By increasing k and keeping $D(\mathbf{V}_0, \mathbf{V}_i) \geq \epsilon$ we get $\int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U})(T_\nu \Phi_k)(\mathbf{U}) \rightarrow 0$ since the integral is over the ball $B(\mathbf{V}_0, \epsilon/2)$ whose points are far away from unitaries \mathbf{V}_i and functions Φ_k^i approximate the Dirac delta localized at \mathbf{V}_i as k increases. Therefore, there must exist t such that elements of a t -design form an ϵ -net.

Lemma 1. *Let ν be a measure on $\mathbf{U}(d)$ which is an exact unitary t -design. Then for arbitrary function $\Phi \in \mathcal{H}_t$ (i.e. a balanced polynomial of degree at most t in U and in \bar{U}) satisfying*

$$\int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \Phi(\mathbf{U}) = 1, \quad (6.3)$$

and for any $V \in \mathbf{U}(d)$, we have

$$\int_{B(\mathbf{V}, \epsilon)} d\mu(\mathbf{U}) (T_\nu \Phi)(\mathbf{U}) = \text{Vol}(B(\mathbf{V}, \epsilon)), \quad (6.4)$$

where $\epsilon \in [0, 2]$.

Proof. Since ν is an exact t -design, and $\Phi \in \mathcal{H}_t$, the moment operator T_ν projects function Φ onto a constant function. Therefore, due to Eq.(6.3) $T_\nu \Phi = 1$ (a constant function equal to 1). As a result we get Eq.(6.4). \square

As explained above, our goal is to upper bound the integral defined in Eq.(6.2) in terms of k . To this aim we will use the following technical Lemma.

Lemma 2. *Let ν be a measure on $\mathbf{U}(d)$. Suppose that the support of ν is not an ϵ -net in $\mathbf{U}(d)$. Then, there exists \mathbf{V}_0 such that for any function Φ on $\mathbf{U}(d)$, and any κ satisfying $0 \leq \kappa \leq \epsilon$ we have*

$$\int_{B(\mathbf{V}_0, \kappa)} d\mu(\mathbf{U})(T_\nu \Phi)(\mathbf{U}) \leq \max_{\mathbf{V}: D(\mathbf{V}, \mathbf{I}) \geq \epsilon} \int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \Phi(\mathbf{U}). \quad (6.5)$$

Proof. For simplicity we assume that the measure ν is discrete i.e. $\nu = \{\nu_i, V_i\}$. The proof is analogous in the general case. Let \mathbf{V}_0 be a unitary channel that cannot be ϵ -approximated by elements from the support of ν : $D(\mathbf{V}_0, \mathbf{V}_i) \geq \epsilon$.

From the definition of the moment operator (see Eq. (5.1)) we have

$$\int_{B(\mathbf{V}_0, \kappa)} d\mu(\mathbf{U})(T_\nu \Phi)(\mathbf{U}) = \sum_i \nu_i \int_{B(\mathbf{V}_0, \kappa)} d\mu(\mathbf{U}) \Phi(\mathbf{V}_i^{-1} \mathbf{U}) . \quad (6.6)$$

By changing the variables in each summand $\mathbf{U}' = \mathbf{V}_i^{-1} \mathbf{U}$ and denoting $\mathbf{V}_i' = \mathbf{V}_i^{-1} \mathbf{V}_0$ we get

$$\int_{B(\mathbf{V}_0, \kappa)} d\mu(\mathbf{U})(T_\nu \Phi)(\mathbf{U}) = \sum_i \nu_i \int_{B(\mathbf{V}_i', \kappa)} d\mu(\mathbf{U}') \Phi(\mathbf{U}') . \quad (6.7)$$

Finally, using the defining property of \mathbf{V}_0 and employing the unitary invariance of the diamond norm we obtain

$$D(\mathbf{V}_i', \mathbf{I}) = D(\mathbf{V}_i^{-1} \mathbf{V}_0, \mathbf{I}) = D(\mathbf{V}_0, \mathbf{V}_i) \geq \epsilon . \quad (6.8)$$

We conclude the proof by using the above inequality in each summand of Eq.(6.7). \square

The following statement about the volume of the Ball in the space of unitary channels is known as folklore in quantum information community. Here we adapt a rigorous result of [44].

Fact 2 (Estimates for the volume of Ball in the manifold of quantum channels [44]). *Let $B(\mathbf{V}, \epsilon) = \{\mathbf{U} \in \mathbf{U}(d) \mid D(\mathbf{U}, \mathbf{V}) \leq \epsilon\}$ be a ball centered around $\mathbf{V} \in \mathbf{U}(d)$, where D is the diamond norm distance from Eq.(2.1). There exist absolute constants $c, C > 0$ such that for all $\epsilon \in [0, 2]$*

$$\left(\frac{\epsilon}{C}\right)^{d^2-1} \leq \text{Vol}(B(\mathbf{V}, \epsilon)) \leq \left(\frac{\epsilon}{c}\right)^{d^2-1} , \quad (6.9)$$

where $C = 5\pi$ and $c = \frac{\pi}{10}$.

Remark 2. In the original work of Szarek [44] considered general homogenous spaces of $\mathbf{U}(d)$ equipped with the metric induced from the operator norm. By the virtue of the variational characterization of the distance D given in Eq.(2.1) the results presented there apply directly to our scenario.

The last necessary element in our proof strategy is the existence of efficient polynomial approximation of the Dirac δ in the space of unitary channels. Here we present only the final result, while details of the construction and the necessary technical details are presented in Section X and the appendix.

Theorem 1 (Efficient polynomial approximation of the Dirac δ on unitary channels). *Consider a set of Unitary channels $\mathbf{U}(d)$ on d -dimensional quantum system equipped with a metric D induced from the diamond norm (see Eq.(2.1)). Let κ, ϵ, σ be positive numbers satisfying $\epsilon \in [0, 1]$, $\kappa \leq \epsilon$, $\sigma \leq \frac{\epsilon - \kappa}{6\sqrt{d}}$. There exist a function $\mathcal{F}_k^\sigma : \mathbf{U}(d) \rightarrow \mathbb{R}$ with the following properties*

1. *Normalisation:* $\int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \mathcal{F}_k^\sigma(\mathbf{U}) = 1$.
2. *Vanishing integrals on balls sufficiently far from identity channel \mathbf{I} :* for every \mathbf{V} such that $D(\mathbf{V}, \mathbf{I}) \geq \epsilon$ we have

$$\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \mathcal{F}_k^\sigma(\mathbf{U}) \leq 9 \exp\left(-\frac{(\epsilon - \kappa)^2}{4\sigma^2}\right) \left(\frac{\pi}{2}\right)^{d(d-1)} . \quad (6.10)$$

3. *Low degree polynomial:* $\mathcal{F}_k^\sigma(\mathbf{U})$ can be represented as a balanced polynomial in U and \bar{U} of degree

$$k = 5 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{\frac{1}{8} \frac{(\epsilon - \kappa)^2}{d^2 \sigma^2} + \log\left(\frac{1}{\sigma}\right)} . \quad (6.11)$$

4. *Bounded L^2 -norm:*

$$\|\mathcal{F}^\sigma\|_2^2 = \int_{\mathbf{U}(d)} d\mu(\mathbf{U}) |\mathcal{F}_k^\sigma(\mathbf{U})|^2 \leq 9 . \quad (6.12)$$

We are now ready to prove the main result of this section. In the course of the proof we will make use of properties 1, 2 and 3 listed above. Property 4, which bounds the second norm of \mathcal{F}_σ will be used in the subsequent section while discussing connection between approximate designs and ϵ -nets.

Theorem 2 (Exact t -expanders define ϵ -nets for sufficiently large t). *Let $\epsilon \leq 1$ and let ν be a measure on $\mathbb{U}(d)$ which is an exact t -design with*

$$t \geq 5 \frac{d^{5/2}}{\epsilon} \tau(\epsilon, d) , \quad (6.13)$$

for $\tau(\epsilon, d) = \log(6C/\epsilon)^{\frac{1}{2}} \sqrt{\frac{1}{32} \log(6C/\epsilon)^{\frac{1}{2}} + \log\left(\frac{d}{\epsilon} \log(6C/\epsilon)^{\frac{1}{2}}\right)}$, where $C = 5\pi$ is the constant appearing in Fact 2.

Then, the set of unitary channels from the support of ν , $\{\mathbf{V}\}_{V \in \text{supp}(\nu)}$ forms an ϵ -net in $\mathbb{U}(d)$ with respect to the distance D defined in Eq.(2.1).

Recall that for a discrete measure ν we have simply $V \in \text{supp}(\nu)$ iff $\nu(V) > 0$.

Proof. Assume that ν is an exact t -design and that the set of unitary channels from the support of ν , $\{\mathbf{V}\}_{V \in \text{supp}(\nu)}$ is not an ϵ -net. Let \mathbf{V}_0 be a unitary channel that cannot be ϵ -approximated by elements from the support of ν : $D(\mathbf{V}_0, \mathbf{V}_i) \geq \epsilon$. Let \mathcal{F}_k^σ be the function satisfying conditions described in Theorem 1. By Fact 2 and Lemma 1 we have

$$\left(\frac{\epsilon}{2C}\right)^{d^2-1} \leq \text{Vol}(B(\mathbf{V}_0, \epsilon/2)) = \int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U}) (T_\nu \mathcal{F}_k^\sigma)(\mathbf{U}) , \quad (6.14)$$

where $k \leq t$. On the other hand by Lemma 2 we have

$$\int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U}) (T_\nu \mathcal{F}_k^\sigma)(\mathbf{U}) \leq \max_{\mathbf{V}: D(\mathbf{V}, \mathbf{I}) \geq \epsilon} \int_{B(\mathbf{V}, \epsilon/2)} d\mu(\mathbf{U}) \mathcal{F}_k^\sigma(\mathbf{U}) . \quad (6.15)$$

Using Theorem 1 for

$$k = 5 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{\frac{1}{32} \frac{\epsilon^2}{d^2 \sigma^2} + \log\left(\frac{1}{\sigma}\right)} , \quad (6.16)$$

we get

$$\max_{\mathbf{V}: D(\mathbf{V}, \mathbf{I}) \geq \epsilon} \int_{B(\mathbf{V}, \epsilon/2)} d\mu(\mathbf{U}) \mathcal{F}_k^\sigma(\mathbf{U}) \leq 9 \exp\left(-\frac{\epsilon^2}{16\sigma^2}\right) \left(\frac{\pi}{2}\right)^{d(d-1)} . \quad (6.17)$$

As σ decreases (and the degree k increases according to (6.16)) eventually the right-hand side of (6.17) becomes smaller than the lower bound $\left(\frac{\epsilon}{2C}\right)^{d^2-1}$ from (6.14). In particular, by inserting σ which satisfies

$$\frac{1}{2} \left(\frac{\epsilon}{10\pi}\right)^{d^2-1} \geq 9 \exp\left(-\frac{\epsilon^2}{16\sigma^2}\right) \left(\frac{\pi}{2}\right)^{d(d-1)} . \quad (6.18)$$

to Eq.(6.16) we get (by contradiction) the degree k_* such that that unitaries from the support of an exact $k \geq k_*$ design form an ϵ net in $\mathbb{U}(d)$. It is easy to see that taking

$$\sigma \leq \sigma_*(d, \epsilon) = \frac{\epsilon}{d \log(6C/\epsilon)^{\frac{1}{2}}} , \quad (6.19)$$

suffices to satisfy (6.18). Inserting σ_* to (6.16) gives k_* equal to the right-hand side of inequality (6.13). Finally, we remark that dropping the factor of $1/2$ in inequality (6.18) yields essentially identical scaling. \square

VII. APPROXIMATE T-DESIGNS AND EPSILON-NETS

In this section, we will establish even a closer connection between approximate t - designs and ϵ -nets. Specifically, we prove that under suitable conditions approximate t -expanders define ϵ -nets and vice versa.

We first extend the reasoning established in the preceding section. Namely, for *approximate* t -designs the integral $I(\nu, \epsilon, k, \mathbf{V}_0)$ from Eq. (6.2) is not anymore equal $\text{Vol}(B(\mathbf{V}_0, \epsilon))$. Therefore, we need to argue that the integral is not

too small with respect to the volume. This is expected, as ν is almost a k -design. The following Lemma expresses this intuition quantitatively.

Lemma 3. *Let ν be an arbitrary measure on $\mathbb{U}(d)$ which is a δ -approximate unitary k -expander. Then for arbitrary $\mathbf{V} \in \mathbf{U}(d)$, $\epsilon \in [0, 2]$ and a function $\Phi \in \mathcal{H}_k$ (i.e. a balanced polynomial of degree at most k in U and in \bar{U}) satisfying*

$$\int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \Phi(\mathbf{U}) = 1, \quad (7.1)$$

we have the following inequality

$$\int_{B(\mathbf{V}, \epsilon)} d\mu(\mathbf{U}) (T_\nu \Phi)(\mathbf{U}) \geq \text{Vol}(B(\mathbf{V}, \epsilon)) - \delta \sqrt{\text{Vol}(B(\mathbf{V}, \epsilon))} \|\Phi\|_2. \quad (7.2)$$

Proof. We start with the following identity

$$\int_{B(\mathbf{V}, \epsilon)} d\mu(\mathbf{U}) - \int_{B(\mathbf{V}, \epsilon)} d\mu(\mathbf{U}) T_\nu(\Phi)(\mathbf{U}) = \langle 1 - T_\nu \Phi, I_{B(\mathbf{V}, \epsilon)} \rangle, \quad (7.3)$$

where $\langle \cdot, \cdot \rangle$ is the inner product in $L^2(\mathbf{U}(d))$, and I_A is the indicator function of a set $A \subset \mathbf{U}(d)$. Using Cauchy-Schwartz inequality we obtain

$$|\langle 1 - T_\nu \Phi, I_{B(\mathbf{V}, \epsilon)} \rangle| \leq \|1 - T_\nu \Phi\|_2 \|I_{B(\mathbf{V}, \epsilon)}\|_2 = \|1 - T_\nu \Phi\|_2 \sqrt{\text{Vol}(B(\mathbf{V}, \epsilon))}. \quad (7.4)$$

Furthermore, condition (7.1), the assumption $\Phi \in \mathcal{H}_k$ and definitions of T_μ and the infinity norm allows us to write an estimate

$$\|1 - T_\nu \Phi\|_2 = \|(T_\mu - T_\nu)\Phi\|_2 \leq \|(T_\mu - T_\nu)_{\mathcal{H}_k}\|_\infty \|\Phi\|_2 = \delta \|\Phi\|_2, \quad (7.5)$$

where in the last equality we used Proposition 1. By combining bounds (7.5) and (7.4) with (7.3) we obtain the desired result, i.e. we get (7.2). \square

With the help of the above Lemma and due to properties of a carefully chosen polynomial approximation to the Dirac delta given in Theorem 1 we are in the position to prove the main result of this section.

Theorem 3 (δ -approximate t -expanders define ϵ -nets). *Suppose that a measure ν on $\mathbb{U}(d)$ is a δ -approximate unitary t -expander with*

$$t \geq 5 \frac{d^{5/2}}{\epsilon} \tau(\epsilon, d), \quad \delta \leq \frac{1}{6} \left(\frac{\epsilon}{2C} \right)^{\frac{d^2}{2}}, \quad (7.6)$$

and $\tau(\epsilon, d) = \log(6C/\epsilon)^{\frac{1}{2}} \sqrt{\frac{1}{32} \log(6C/\epsilon)^{\frac{1}{2}} + \log\left(\frac{d}{\epsilon} \log(6C/\epsilon)^{\frac{1}{2}}\right)}$, where $C = 5\pi$ is the constant appearing in Fact 2.

Then, the set of unitary channels $\{\mathbf{V}\}_{\mathbf{V} \in \text{supp}(\nu)}$ forms an ϵ -net in $\mathbf{U}(d)$ with respect to the distance D defined in (2.1).

Remark 3. *A similar result follows from arguments given in the proof of Theorem 5 in [31]. From careful analysis of the arguments presented there it can be shown that δ -approximate t -expanders with $t \simeq d^3/\epsilon^2$ and $\delta \simeq (\epsilon/\sqrt{d})^{2d^2}$ define ϵ -nets with respect to the distance between unitary channels induced from the Hilbert-Schmidt norm*

$$\tilde{D}(\mathbf{U}, \mathbf{V}) := \min_{\varphi \in [0, 2\pi)} \|U - \exp(i\varphi)V\|_{\text{HS}}. \quad (7.7)$$

Our result gives a more favorable scaling of t and δ in both d and ϵ . This is because the inequality $D(\mathbf{U}, \mathbf{V}) \leq \tilde{D}(\mathbf{U}, \mathbf{V})$ implies that ϵ -net with respect to distance \tilde{D} is automatically ϵ -net with respect to distance D . In order to attain the scaling claimed above we used tight bounds on volumes of Hilbert-Schmidt balls in $\mathbb{U}(d)$ (cf. [33] Theorem 5.11).

Proof. We proceed analogously as in the proof of Theorem 2. Assume that ν is a δ -approximate t -design and that the set of unitary channels from the support of ν , $\{\mathbf{V}\}_{\mathbf{V} \in \text{supp}(\nu)}$, is not an ϵ -net. We choose \mathbf{V}_0 to be a unitary channel that cannot be ϵ -approximated by elements from the support of ν : $D(\mathbf{V}_0, \mathbf{V}_i) \geq \epsilon$. Moreover, we take \mathcal{F}_k^σ to be the polynomial function described in Theorem 1 for $\sigma = \sigma_*(d, \epsilon)$ (c.f Eq.(6.19)) and $k = 5 \frac{d^{5/2}}{\epsilon} \tau(\epsilon, d)$.

From Lemma 3 and Eq.(6.12) it follows that

$$\text{Vol}(B(\mathbf{V}_0, \epsilon/2)) - 3\delta \text{Vol}(B(\mathbf{V}_0, \epsilon/2))^{\frac{1}{2}} \leq \int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U}) (T_\nu \mathcal{F}_k^\sigma)(\mathbf{U}) , \quad (7.8)$$

for any $k \leq t$. On the other hand, by repeating the same arguments as in the proof of Theorem 2 we have

$$\int_{B(\mathbf{V}_0, \epsilon/2)} d\mu(\mathbf{U}) (T_\nu \mathcal{F}_k^\sigma)(\mathbf{U}) \leq \exp\left(-\frac{\epsilon^2}{16\sigma^2}\right) \left(\frac{\pi}{2}\right)^{d(d-1)} . \quad (7.9)$$

It is now clear that if δ is such that

$$3\delta \text{Vol}(B(\mathbf{V}_0, \epsilon/2))^{\frac{1}{2}} \leq \frac{1}{2} \text{Vol}(B(\mathbf{V}_0, \epsilon/2)) , \quad (7.10)$$

then we obtain inequality (6.18). However, already from the proof of Theorem 2 we know that this inequality cannot be satisfied for $\sigma = \sigma_*(d, \epsilon)$ and $k = 5\frac{d^{5/2}}{\epsilon}\tau(\epsilon, d)$ and hence unitaries from the support of ν must form an ϵ -net. We conclude the proof observing that $\delta \leq \frac{1}{6}\left(\frac{\epsilon}{2C}\right)^{\frac{d^2}{2}}$ is a sufficient condition for validity (7.10). This can be verified easily using Fact 2. \square

We now prove the statement in the opposite direction to the one given in Theorem 3.

Theorem 4 (ϵ -nets in the set of unitary channels can be used to define $(2\epsilon t)$ -approximate unitary t -expanders). *Consider a subset of unitary channels $\mathcal{S} \subset \mathbf{U}(d)$ which forms an ϵ -net in $\mathbf{U}(d)$ with respect to the distance D defined in (2.1). Then, there exists an ensemble $\mathcal{E} = \{\nu_i, \mathbf{V}_i\}$, with $\mathbf{V}_i \in \mathcal{S}$, which forms an $(2\epsilon t)$ -approximate t -expander.*

Proof. We present an explicit (although possibly computationally inefficient) construction of an ensemble of gates form \mathcal{S} which will (ϵt) -approximate t -design for $\delta = \epsilon t$. First, we note that, by definition of ϵ -net, elements for \mathcal{S} define a cover of $\mathbf{U}(d)$ via balls of radius *at most* ϵ :

$$\mathbf{U}(d) = \bigcup_{\mathbf{V} \in \mathcal{S}} B(\mathbf{V}, \epsilon) . \quad (7.11)$$

Since $\mathbf{U}(d)$ is a compact space, we can take a finite collection of gates $\{\mathbf{V}_i\}_{i=1}^K \subset \mathcal{S}$ such that

$$\mathbf{U}(d) = \bigcup_{i=1}^K B(\mathbf{V}_i, \epsilon) . \quad (7.12)$$

We note that K in the above equation is some, in general unknown, but *finite* number. We now use Eq. (7.12) to define a disjoint collection of subsets \mathcal{V}_i that cover $\mathbf{U}(d)$. We set

$$\mathcal{V}_1 = B(\mathbf{V}_1, \epsilon) , \quad \mathcal{V}_{k+1} = B(\mathbf{V}_{k+1}, \epsilon) \setminus \bigcup_{i=1}^k B(\mathbf{V}_i, \epsilon) , \quad k = 1, \dots, K-1 . \quad (7.13)$$

By the construction we have $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset$ whenever $i \neq j$. Moreover, $\bigcup_{i=1}^K \mathcal{V}_i = \mathbf{U}(d)$, while it might also happen that $\mathcal{V}_i = \emptyset$ for $i > K'$, where $K' = \min\{k \mid \bigcup_{i=1}^k B(\mathbf{V}_i, \epsilon) = \mathbf{U}(d)\}$. From the definition we have $\mathcal{V}_i \subset B(\mathbf{V}_i, \epsilon)$ and hence for $i \leq K'$

$$\mathbf{U} \in \mathcal{V}_i \implies D(\mathbf{U}, \mathbf{V}_i) \leq \epsilon . \quad (7.14)$$

We are now ready to define a discrete ensemble that is a desired approximate t -design. We set V_i to be *any* unitary operator which is compatible with channel \mathbf{V}_i and define

$$\mathcal{E} = \{\nu_i, V_i\}_{i=1}^{K'} , \quad \text{where } \nu_i = \mu_P(\mathcal{V}_i) . \quad (7.15)$$

The normalization of the probability ν_i follows from the construction. Let us finally upper bound $\|T_{\nu, t} - T_{\mu, t}\|_\infty$. We

have

$$T_{\nu,t} - T_{\mu,t} = \sum_{i=1}^{K'} \mu_P(\mathcal{V}_i) V_i^{\otimes t} \otimes \bar{V}_i^{\otimes t} - \int_{\mathbb{U}(d)} d\mu(U) U^{\otimes t} \otimes \bar{U}^{\otimes t} = \sum_{i=1}^{K'} \int_{\varphi^{-1}(\mathcal{V}_i)} d\mu(U) [V_i^{\otimes t} \otimes \bar{V}_i^{\otimes t} - U^{\otimes t} \otimes \bar{U}^{\otimes t}] . \quad (7.16)$$

We have the following chain of inequalities

$$\|T_{\nu,t} - T_{\mu,t}\|_{\infty} \leq \sum_{i=1}^{K'} \int_{\varphi^{-1}(\mathcal{V}_i)} d\mu(U) \|V_i^{\otimes t} \otimes \bar{V}_i^{\otimes t} - U^{\otimes t} \otimes \bar{U}^{\otimes t}\| \leq \sum_{i=1}^{K'} \int_{\varphi^{-1}(\mathcal{V}_i)} d\mu(U) \|V_i \otimes \bar{V}_i - U \otimes \bar{U}\|_{\infty} t \leq 2\epsilon t . \quad (7.17)$$

The second inequality follows from the well-known telescopic bound (see for example page 27 in [2])

$$\|A^{\otimes t} - B^{\otimes t}\|_{\infty} \leq \|A - B\|_{\infty} t , \quad (7.18)$$

applied for $A = V_i \otimes \bar{V}_i$ and $B = U \otimes \bar{U}$. The third inequality in Eq.(7.17) follows from Eq.(2.1) and the fact that numbers $\{\mu_P(\mathcal{V}_i)\}_{i=1}^{K'}$ sum up to 1. To see this we first chose the relative phase between U and V_i in such a way that these operators saturate Eq.(2.1) and arrive at the bound $\|V_i \otimes \bar{V}_i - U \otimes \bar{U}\|_{\infty} \leq 2D(U, V)$. Second, we use Eq.(7.14) which ensures $D(V_i, U) \leq \epsilon$. \square

VIII. SEQUENCES OF GATES AND EPSILON-NETS

Let $\mathcal{G} \subset \mathbb{U}(d)$ be the support of a probability measure $\nu_{\mathcal{G}}$ on $\mathbb{U}(d)$. Then the words of length l composed of gates from the set \mathcal{G} , we denote them by \mathcal{G}_l , constitute the support of $\nu_{\mathcal{G}}^{*l}$. In this section we explore properties of sets \mathcal{G}_l and formulate an inverse-free version of the Solovay-Kitaev theorem. Our first result is the following

Proposition 2. *Let ν be an arbitrary measure on $\mathbb{U}(d)$ which is a δ -approximate unitary t -expander with*

$$t \geq 5 \frac{d^{5/2}}{\epsilon} \tau(\epsilon, d) , \quad (8.1)$$

and $\tau(\epsilon, d) = \log(6C/\epsilon)^{\frac{1}{2}} \sqrt{\frac{1}{32} \log(6C/\epsilon)^{\frac{1}{2}} + \log\left(\frac{d}{\epsilon} \log(6C/\epsilon)^{\frac{1}{2}}\right)}$, where $C = 5\pi$ is the constant appearing in Fact 2. Then for

$$l \geq \frac{\log(6) + \frac{d^2}{2} \log\left(\frac{2C}{\epsilon}\right)}{1 - \delta} , \quad (8.2)$$

the set of unitary channels in the support of ν^{*l} , $\{\mathbf{V}\}_{\mathbf{V} \in \text{supp}(\nu^{*l})}$, forms an ϵ -net in $\mathbf{U}(d)$ with respect to the distance D defined in (2.1).

Proof. By Fact 1 it follows that the support of ν^{*l} defines δ^l -approximate t -expander. Next, if l satisfies (8.2) one easily checks that

$$\delta^l \leq \frac{1}{6} \left(\frac{\epsilon}{2C} \right)^{\frac{d^2}{2}} , \quad (8.3)$$

Therefore by Theorem 3 the support of ν^{*l} is an ϵ -net. \square

We now reformulate a strong result by Peter Varju (Theorem 6 in [1]) in the language of approximate t -expanders (see part XIA of Appendix for details)

Theorem 5 (Slow decay of the spectral gap). *Let ν be arbitrary probability measure on $\mathbf{U}(d)$. Then, there exist a natural number t_0 and a constant $B > 0$ (depending only on d) such that for all natural $t > t_0$ we have*

$$\|T_{\nu,t} - T_{\mu,t}\|_{\infty} \leq 1 - \frac{1 - \delta(\nu, t_0)}{B \log(t)^2} , \quad (8.4)$$

where $\delta(\nu, t_0) = \|T_{\nu, t_0} - T_{\mu, t_0}\|_\infty$. In other words the gap of the random walk generated by ν cannot decrease faster than $\log^{-2}(t)$ for large t .

The above Theorem in conjunction with Proposition 2 allows us to state an inverse-free version of the Solovay-Kitaev theorem. We note that an equivalent result has already appeared in [1]. It was, however, obscured by the mathematical character of that work.

Theorem 6 (Non-constructive inverse-free Solovay-Kitaev). *Let $\mathcal{G} \subset \mathbf{U}(d)$ be a universal gate-set in $\mathbf{U}(d)$ (not necessarily symmetric i.e. $V \in \mathcal{G}$ does not imply $V^{-1} \in \mathcal{G}$). Let $\nu_{\mathcal{G}}$ be a uniform measure on \mathcal{G} . Then, there exists absolute constant $C > 0$ (depending on d), such that for*

$$l \geq C(d) \frac{\log^3(\frac{1}{\epsilon})}{1 - \delta(\nu_{\mathcal{G}}, t_0)} \quad (8.5)$$

the set \mathcal{G}_l forms an ϵ -net in $\mathbf{U}(d)$.

Proof. Proposition 2 tells us that if \mathcal{G} is a $\delta(\nu_{\mathcal{G}}, t)$ -approximate t -expander, with t given by (8.1), then for

$$l \geq \frac{\log(6) + \frac{d^2}{2} \log(\frac{2C}{\epsilon})}{1 - \delta(\nu_{\mathcal{G}}, t)}, \quad (8.6)$$

\mathcal{G}_l is an ϵ -net. On the other hand, Theorem 5 allows us to bound $1 - \delta(\nu_{\mathcal{G}}, t)$ as follows

$$1 - \delta(\nu_{\mathcal{G}}, t) \geq \frac{1 - \delta(\nu_{\mathcal{G}}, t_0)}{B \log^2(t)}. \quad (8.7)$$

Combining (8.7) with (8.6) and making use of (8.1) we obtain (8.5). \square

Remark 4. Another application of our results that relate ϵ -nets and δ -approximate t -designs is connected to universality of gate-sets \mathcal{G} . In [38, 39] it was shown that the necessary condition for universality of a gate-set $\mathcal{G} \subset \mathbf{U}(d)$ is $\dim(\text{Comm}(U \otimes \bar{U} | U \in \mathcal{G})) = 2$. Moreover, sets \mathcal{G} that satisfy the necessary condition are either universal or they generate finite subgroups of $\mathbf{U}(d)$. Furthermore, in order to verify universality of a set \mathcal{G} that satisfies the necessary condition one has to check that there is l such that \mathcal{G}_l forms an ϵ -net with $\epsilon \leq \frac{1}{2\sqrt{2}}$. Thus using Proposition 2 a gate-set \mathcal{G} satisfying the necessary condition is universal iff it is a δ -approximate t -expander with $\delta < 1$ and t given by (8.1) with $\epsilon = \frac{1}{2\sqrt{2}}$. Otherwise \mathcal{G} generates a finite group (this follows from Lemma 4.8 of [38]). Therefore checking universality of \mathcal{G} can be reduced to two steps 1) checking if $\dim(\text{Comm}(U \otimes \bar{U} | U \in \mathcal{G})) = 2$ and 2) checking if $\delta(\nu_{\mathcal{G}}, t) < 1$ for t given by (8.1) with $\epsilon = \frac{1}{2\sqrt{2}}$.

IX. RANDOM CIRCUITS AND APPROXIMATE DESIGNS

In this section we shall prove that random circuits composed of universal gates are approximate t -designs *without* any assumptions on the set of gates (i.e. unlike in [2] we shall not assume that the set contains inverses or that the unitaries have algebraic entries). Importantly, we are not using result due to Bourgain and Gomburd [20] who proved that universal set of gates has a gap that does not diverge with growing t under the assumption of algebraic entries. However, from the results of Varju [1] is possible to prove a lower bound on the the gap, which vanishes very slowly with t , yet without any assumptions (see Theorem 5).

Let $\nu_{\mathcal{G}}$ be uniform measure on set of gates \mathcal{G} , and let \mathcal{G}^\dagger be the set of inverses of gates from \mathcal{G} , and $\nu_{\mathcal{G}^\dagger}$ uniform measure on \mathcal{G}^\dagger . Note that $\nu_{\mathcal{G}\mathcal{G}^\dagger} = \nu_{\mathcal{G}} * \nu_{\mathcal{G}^\dagger}$. We shall also employ the following (well-known in the mathematics community, see e.g. [1]) Lemma in order to remove the assumption that the set of gates contains inverses.

Lemma 4 (Bounds on the spectral gap without assuming a symmetric gate-set). *Let \mathcal{G} be arbitrary finite gateset in $\mathbf{U}(d)$. Let $\nu_{\mathcal{G}}$ and $\nu_{\mathcal{G}^\dagger}$ be two measures uniformly supported on \mathcal{G} and \mathcal{G}^\dagger respectively. We have the following inequalities*

$$\delta(\nu_{\mathcal{G}}, t)^2 = \delta(\nu_{\mathcal{G}\mathcal{G}^\dagger}, t) \quad (9.1)$$

$$\frac{1}{2}g(\nu_{\mathcal{G}} * \nu_{\mathcal{G}^\dagger}, t) \leq g(\nu_{\mathcal{G}}, t) \leq g(\nu_{\mathcal{G}} * \nu_{\mathcal{G}^\dagger}, t) \quad (9.2)$$

Proof. Recall that $\delta(\nu_{\mathcal{G}}, t) = \|T_{\nu_{\mathcal{G}}, t} - T_{\mu, t}\|_{\infty}$. From definition of $T_{\nu, t}$ we have for any measures ν, ν'

$$T_{\nu * \nu', t} = T_{\nu, t} T_{\nu', t}, \quad T_{\nu, t} = T_{\mu, t} \oplus T_{\nu, t}^{\perp}, \quad T_{\mu, t} = T_{\mu, t}^2 = T_{\mu, t}^{\dagger}, \quad (9.3)$$

so that $\delta(\nu, t) = \|T_{\nu_{\mathcal{G}}, t}^{\perp}\|$. It also immediately follows that $T_{\nu * \nu', t}^{\perp} = T_{\nu, t}^{\perp} T_{\nu', t}^{\perp}$. Further, from definition of moment operators $T_{\nu_{\mathcal{G}}, t}$ we have

$$T_{\nu_{\mathcal{G}^{\dagger}}, t} = T_{\nu_{\mathcal{G}}, t}^{\dagger} \quad (9.4)$$

which gives

$$T_{\nu_{\mathcal{G}^{\dagger}}, t}^{\perp} = (T_{\nu_{\mathcal{G}}, t}^{\perp})^{\dagger}. \quad (9.5)$$

We then also get

$$T_{\nu_{\mathcal{G}\mathcal{G}^{\dagger}}, t}^{\perp} = T_{\nu_{\mathcal{G}} * \nu_{\mathcal{G}^{\dagger}}, t}^{\perp} = T_{\nu_{\mathcal{G}}, t}^{\perp} T_{\nu_{\mathcal{G}^{\dagger}}, t}^{\perp} = T_{\nu_{\mathcal{G}}, t}^{\perp} (T_{\nu_{\mathcal{G}}, t}^{\perp})^{\dagger} \quad (9.6)$$

and hence

$$\delta(\nu_{\mathcal{G}\mathcal{G}^{\dagger}}, t) = \|T_{\nu_{\mathcal{G}\mathcal{G}^{\dagger}}, t}^{\perp}\| = \|T_{\nu_{\mathcal{G}}, t}^{\perp} (T_{\nu_{\mathcal{G}}, t}^{\perp})^{\dagger}\| = \|T_{\nu_{\mathcal{G}}, t}^{\perp}\|^2 = \delta(\nu_{\mathcal{G}}, t)^2. \quad (9.7)$$

We have thus proved the formula (9.1). Now, the first inequality of (9.2) follows by definition of g : $g = 1 - \delta$ and the use of $\sqrt{x} \leq \frac{1}{2} + \frac{1}{2}x$ for $x \geq 0$, while the second one follows from $x^2 \leq x$ for $x \leq 1$, \square

We shall now consider two layouts for random circuits acting on n qudits, composed of two qudit gates from set \mathcal{G} : (i) local random circuits and (ii) parallel random circuits. Local random circuits are the following. We pick uniformly at random two neighboring qudits, and apply gate chosen from \mathcal{G} according to uniform measure. The resulting measure we shall denote by $\nu_{loc}^n(\mathcal{G})$. Let us also denote by $\nu_{loc}^n(\mu)$ similarly defined measure, but with $\nu_{\mathcal{G}}$ replaced with Haar measure on two qudits μ . Regarding parallel random circuits, we apply with probability $1/2$ either unitary $U_{12} \otimes U_{34} \otimes \dots \otimes U_{n-1, n}$ or $U_{23} \otimes U_{45} \otimes \dots \otimes U_{n-2, n-1}$ where each U_{ij} is picked independently from \mathcal{G} according to $\nu_{\mathcal{G}}$. The resulting measure we shall denote by $\nu_{par}^n(\mathcal{G})$ and if $\nu_{\mathcal{G}}$ is replaced by Haar, by $\nu_{par}^n(\mu)$. We shall now relate the gaps of two steps of such circuits to the gap of one step of circuit with measure $\nu_{\mathcal{G}\mathcal{G}^{\dagger}}$. In this way we shall reduce the problem to gate sets with inverses so that we then can invoke results on such circuits from [2].

Lemma 5 (Bound on the gap of random local quantum circuits for non symmetric gate-set). *Let $\nu_{loc}^{(n)}(\mathcal{G})$ be a measure describing random local quantum circuits generated two qudit gate-set \mathcal{G} (not necessarily symmetric). We have the following lower bound*

$$g\left(\nu_{loc}^{(n)}(\mathcal{G}) * \nu_{loc}^{(n)}(\mathcal{G}^{\dagger})\right) \geq \frac{1}{n-1} g(\nu_{loc}^{(n)}(\mathcal{G}\mathcal{G}^{\dagger})) \quad (9.8)$$

Proof. By definition of $\nu_{loc}^{(n)}(\mathcal{G})$ we have

$$T_{\nu_{loc}^{(n)}(\mathcal{G}), t} = \frac{1}{n-1} \sum_{i=1}^{n-1} A_{ii+1}. \quad (9.9)$$

where $A_{ii+1} = T_{\nu_{\mathcal{G}}, t}$ with \mathcal{G} acting on qudits i and $i+1$. Note that A_{ii+1} is not necessarily Hermitian. Denoting for clarity by $P_{H_{aar}}^{\perp}$ the complement of $T_{\mu, t}$ we write

$$\begin{aligned} 1 - g\left(\nu_{loc}^{(n)}(\mathcal{G}) * \nu_{loc}^{(n)}(\mathcal{G}^{\dagger})\right) &= \|P_{H_{aar}}^{\perp} T_{\nu_{loc}^{(n)}(\mathcal{G}), t} T_{\nu_{loc}^{(n)}(\mathcal{G}^{\dagger}), t}^{\dagger} P_{H_{aar}}^{\perp}\| = \\ &\|P_{H_{aar}}^{\perp} \left(\frac{1}{(n-1)^2} \sum_i A_{ii+1} A_{ii+1}^{\dagger} + \frac{1}{(n-1)^2} \sum_{i \neq j} A_{ii+1} A_{jj+1}^{\dagger} \right) P_{H_{aar}}^{\perp}\| \leq \\ &\leq \|P_{H_{aar}}^{\perp} \left(\frac{1}{(n-1)^2} \sum_i A_{ii+1} A_{ii+1}^{\dagger} \right) P_{H_{aar}}^{\perp}\| + \frac{(n-1)^2 - (n-1)}{(n-1)^2} = \frac{1}{n-1} (1 - g(\nu_{loc}^{(n)}(\mathcal{G}\mathcal{G}^{\dagger})) + \frac{(n-1)^2 - (n-1)}{(n-1)^2}). \end{aligned}$$

The first equality is definition of gap. The second uses Eq.(9.9). The inequality comes from triangle inequality and $\|A_{ii+1}\| \leq 1$ (since A 's are moment operators). The last equality follows from $T_{\nu * \nu', t} = T_{\nu, t} T_{\nu', t}$, applied to operators

A. Hence we obtain the claimed result

$$g\left(\nu_{loc}^{(n)}(\mathcal{G}) * \nu_{loc}^{(n)}(\mathcal{G}^\dagger)\right) \geq \frac{1}{n-1} g(\nu_{loc}^{(n)}(\mathcal{G}\mathcal{G}^\dagger)) . \quad (9.10)$$

□

In exactly analogous way one proves

Lemma 6. [Bound on the gap of random parallel quantum circuits for non symmetric gate-set] Let $\nu_{par}^{(n)}(\mathcal{G})$ be a measure describing random parallel quantum circuits generated two qudit gate-set \mathcal{G} (not necessarily symmetric). We have the following lower bound

$$g\left(\nu_{par}^{(n)}(\mathcal{G}) * \nu_{par}^{(n)}(\mathcal{G}^\dagger)\right) \geq \frac{1}{2} g(\nu_{par}^{(n)}(\mathcal{G}\mathcal{G}^\dagger)) . \quad (9.11)$$

Next we need Lemma proved in [2] (it is not formulated as a separate Lemma, but it is a contents of the proof of Corollary 7 of [2]).

Lemma 7 (local circuits). For a set $\tilde{\mathcal{G}}$ of gates containing inverses we have

$$g(\nu_{loc}^{(n)}(\tilde{\mathcal{G}}), t) \geq g(\nu_{\tilde{\mathcal{G}}}^{(n)}, t) g(\nu_{loc}^{(n)}(\mu), t), \quad g(\nu_{par}^{(n)}(\tilde{\mathcal{G}})) \geq g(\nu_{\tilde{\mathcal{G}}}^{(n)}, t) g(\nu_{par}^{(n)}(\mu), t) \quad (9.12)$$

I.e. we have the same relation for local as well as parallel circuits.

We will also make use of theorem by Peter Varju, which we reformulated in Theorem 5). It states that

$$g(\nu, t) \geq \frac{g(\nu, t_0)}{B \log^2(t)} \quad (9.13)$$

for any measure ν on $\mathbb{U}(d)$, where B depends only on d . Finally we shall need an estimate for the length of circuits that is needed to produce an δ -approximate t -design, provided that single step has gap $g(\nu, t)$:

Proposition 3. For any measure ν on $\mathbb{U}(d)$, the measure ν^{*l} is δ -approximate t -design, if

$$l \geq \frac{1}{g(\nu, t)} \log \frac{1}{\delta} . \quad (9.14)$$

Proof. In order ν^{*l} to be t -design we need $(1 - g(\nu, t))^l \leq \delta$. Taking logarithm of both sides, and using $1 - g \leq e^{-g}$ proves the estimate. □

We are now ready to prove the main result of this section

Theorem 7. A random local circuit composed of two-qudit local gates from universal set of G given by $\nu_{loc}^{(n)}(\mathcal{G})$ is an δ -approximate t -design provided its length l_{loc} satisfies

$$l_{loc} \geq n \log^2(t) C(\mathcal{G}) l_{loc, Haar} \quad (9.15)$$

where $C(\mathcal{G}) > 0$ is a constant depending only on set of gates and dimension d , while $l_{loc, Haar}$ is the length of the local Haar circuit that is δ -approximate t -design. For random parallel circuit, with analogous notation we have

$$l_{par} \geq 2 \log^2(t) C(\mathcal{G}) l_{par, Haar} \quad (9.16)$$

Proof. Let us prove the result for local circuits first. We have the following chain of inequalities:

$$\begin{aligned} g(\nu_{loc}^{(n)}(\mathcal{G}), t) &\geq \frac{1}{2} g(\nu_{loc}^{(n)}(\mathcal{G}) * \nu_{loc}^{(n)}(\mathcal{G}^\dagger), t) \geq \frac{1}{2} \frac{1}{n-1} g(\nu_{loc}^{(n)}(\mathcal{G}\mathcal{G}^\dagger), t) \geq \\ &\frac{1}{2} \frac{1}{n-1} g(\nu_{loc}^{(n)}(\mathcal{G}\mathcal{G}^\dagger), t) g(\nu_{loc}^{(n)}(\mu), t) \geq \frac{1}{2} \frac{1}{n} \frac{g(\nu_{\mathcal{G}\mathcal{G}^\dagger}, t_0)}{B(d) \log^2 t} g(\nu_{loc}^{(n)}(\mu), t) \geq \frac{1}{2} \frac{1}{n} \frac{g(\nu_{\mathcal{G}}, t_0)}{B(d) \log^2 t} g(\nu_{loc}^{(n)}(\mu), t) = \\ &= \frac{1}{n} \frac{C(\mathcal{G})}{\log^2 t} g(\nu_{loc}^{(n)}(\mu), t) \end{aligned} \quad (9.17)$$

Here the first inequality comes from Lemma 4, the second from Lemma 5, the third from Lemma 7, the fourth from Eq. (9.13) and the fifth from Lemma 4. Now, since \mathcal{G} is universal $g(\nu_{\mathcal{G}}, t_0)$ and hence $C(\mathcal{G})$ is nonzero. Then applying Lemma 3 and inserting $l_{H_{aar}}$ in place of

$$\frac{\ln \frac{1}{\delta}}{g(\nu_{loc}^{(n)}(\mu))} \quad (9.18)$$

ends the proof. In the case of parallel circuits the proof is exactly the same with just one difference: instead of $1/(n-1)$ there is factor $1/2$ after second inequality, since instead of Lemma 5 concerning local circuits we apply Lemma 6 concerning parallel circuits. \square

X. POLYNOMIAL APPROXIMATION OF DIRAC DELTA ON UNITARY CHANNELS

In this section we present the main idea behind the construction of the polynomial approximation of the Dirac delta in the manifold of unitary channels. The features of this particular approximation were stated without a proof in Theorem 1. In our exposition we will follow a ‘bottom up’ approach. The polynomial function on $\mathbf{U}(d)$ will be constructed from the Fourier series truncation (denoted by $f_{p,k}^\sigma$) of a suitable symmetric function f_p^σ on d dimensional torus $\mathbb{T}^d = \{\boldsymbol{\varphi} = (\varphi_1, \dots, \varphi_d) \mid \varphi_i \in [-\pi, \pi]\}$. This Fourier truncation will be then used to define a function $F_{p,k}^\sigma$ which is a “class extension” of $f_{p,k}^\sigma$, i.e. is a function on $\mathbf{U}(d)$ defined via $F_{p,k}^\sigma(U) = f_{p,k}^\sigma(\text{Eig}(U))$, where $\text{Eig}(U)$ denotes a diagonal matrix of eigenvalues of a unitary operator U . Finally, the function $F_{p,k}^\sigma$ will be averaged over the global phase, resulting in a well-defined polynomial function $\tilde{F}_{p,k}^\sigma$ on $\mathbf{U}(d)$. This function will define (up to a normalisation constant) a polynomial approximation of Dirac δ , denoted by \mathcal{F}_k^σ , whose existence is claimed in Theorem 1.

We will adopt the following convention when referring to elements from the sets relevant in our considerations: $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$, $\mathbf{n}, \mathbf{k} \in \mathbb{Z}^d$. Moreover, we denote by $\mathbf{x} \cdot \mathbf{y}$ the standard inner product in \mathbb{R}^d (note that we can apply it to elements of $\mathbb{Z}^d \subset \mathbb{R}^d$). Finally, we will denote by $|\mathbf{x}|$ and $|\mathbf{x}|_1$ respectively euclidean and 1-norm of $\mathbf{x} \in \mathbb{R}^d$.

We begin by introducing a number of useful functions on \mathbb{R}^d and \mathbb{T}^d . The standard Gaussian distribution on \mathbb{R}^d is defined by

$$f^\sigma : \mathbb{R}^d \rightarrow \mathbb{R}^d, \quad f^\sigma(\mathbf{x}) := \frac{1}{(\sqrt{2\pi}\sigma)^d} \exp\left(-\frac{\mathbf{x}^2}{2\sigma^2}\right). \quad (10.1)$$

It will be convenient for us to introduce a periodized version of this function

$$f_p^\sigma : \mathbb{T}^d \rightarrow \mathbb{T}^d, \quad f_p^\sigma(\boldsymbol{\varphi}) := \sum_{\mathbf{k} \in \mathbb{Z}^d} f^\sigma(\boldsymbol{\varphi} + 2\pi\mathbf{k}). \quad (10.2)$$

By the virtue of the Poisson summation formula [45] we know that for any function $f \in L^1(\mathbb{R}^d)$ that satisfies:

$$|f(\mathbf{x})| \leq \frac{C}{(1 + |\mathbf{x}|)^{d+\alpha}}, \quad (10.3)$$

for some positive constants C and α , we have

$$f_p(\boldsymbol{\varphi}) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n} \in \mathbb{Z}^d} \hat{f}(\mathbf{n}) \exp(\mathbf{i}\mathbf{n} \cdot \boldsymbol{\varphi}), \quad \text{where } \hat{f}(\mathbf{n}) = \int_{\mathbb{R}^d} d\mathbf{y} f(\mathbf{y}) \exp(-\mathbf{i}\mathbf{n} \cdot \mathbf{y}), \quad (10.4)$$

is the standard Fourier transform of f computed at point $\mathbf{n} \in \mathbb{Z}^d$. Using the fact that $\hat{f}^\sigma(\mathbf{n}) = e^{-\frac{1}{2}\sigma^2\mathbf{n}^2}$ we obtain

$$f_p^\sigma(\boldsymbol{\varphi}) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n} \in \mathbb{Z}^d} e^{-\frac{1}{2}\sigma^2\mathbf{n}^2} \exp(\mathbf{i}\mathbf{n} \cdot \boldsymbol{\varphi}). \quad (10.5)$$

Analogously we define a truncated version of $f_p^\sigma(\boldsymbol{\varphi})$,

$$f_{p,k}^\sigma(\boldsymbol{\varphi}) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n} \in S_k} e^{-\frac{1}{2}\sigma^2\mathbf{n}^2} \exp(\mathbf{i}\mathbf{n} \cdot \boldsymbol{\varphi}), \quad (10.6)$$

where $S_k = \{\mathbf{n} \mid |\mathbf{n}|_1 \leq k\}$.

Finally we define a ‘phase averaged’ versions of functions f_p^σ and $f_{p,k}^\sigma$

$$f_p^{\sigma,a}(\boldsymbol{\varphi}) := \frac{1}{2\pi} \int_0^{2\pi} d\phi f_p^\sigma(\boldsymbol{\varphi} + (\phi, \dots, \phi)) , \quad f_{p,k}^{\sigma,a}(\boldsymbol{\varphi}) := \frac{1}{2\pi} \int_0^{2\pi} d\phi f_{p,k}^\sigma(\boldsymbol{\varphi} + (\phi, \dots, \phi)) . \quad (10.7)$$

We use the fact that functions f_p^σ and $f_{p,k}^\sigma$ are functions on \mathbb{T}^d that are invariant under the permutation of angles. Therefore, we can define class functions F^σ and F_k^σ on $\mathbb{U}(d)$ that recover f_p^σ and $f_{p,k}^\sigma$ when restricted to \mathbb{T}^d . In other words

$$F^\sigma(U) := f_p^\sigma(\text{Eig}(U)) , \quad F_k^\sigma(U) := f_{p,k}^\sigma(\text{Eig}(U)) , \quad (10.8)$$

where $\text{Eig}(U) = \text{diag}(\exp(i\phi_1), \dots, \exp(i\phi_d))$ is a diagonal matrix formed by eigenvalues of U . When we average the above functions over the global phase we get well-defined functions on the group of unitary channels $\mathbf{U}(d)$.

$$\tilde{\mathcal{F}}^\sigma = \mathbb{P}_{\text{phase}} F^\sigma , \quad \tilde{\mathcal{F}}_k^\sigma = \mathbb{P}_{\text{phase}} F_k^\sigma , \quad (10.9)$$

where linear operator $\mathbb{P}_{\text{phase}} : L^2(\mathbb{U}(d)) \rightarrow L^2(\mathbb{U}(d))$ is defined by

$$(\mathbb{P}_{\text{phase}} F)(U) = \frac{1}{2\pi} \int_0^{2\pi} d\phi F(\exp(i\phi)U) . \quad (10.10)$$

We note that $\mathbb{P}_{\text{phase}}$ is an orthonormal projector in $L^2(\mathbb{U}(d))$ that projects onto functions in $L^2(\mathbb{U}(d))$ that are invariant under a global phase transformation. As explained in Section II we can interpret such functions as functions defined on $\mathbf{U}(d)$. The normalised version of $\tilde{\mathcal{F}}_k^\sigma$,

$$\mathcal{F}_k^\sigma := \tilde{\mathcal{F}}_k^\sigma / \mathcal{N}_k^\sigma , \quad \mathcal{N}_k^\sigma := \int_{\mathbf{U}(d)} d\mu(U) \tilde{\mathcal{F}}_k^\sigma(U) \quad (10.11)$$

is our candidate for a ‘low degree’ approximation of the Dirac δ at \mathbf{I} . We shall prove Theorem 1 via a sequence of technical Lemmas that will eventually cover all the properties stated in Theorem 1. It will be also convenient to introduce auxiliary function of $\mathbf{U}(d)$,

$$\mathcal{F}^\sigma := \tilde{\mathcal{F}}^\sigma / \mathcal{N}^\sigma , \quad \mathcal{N}^\sigma := \int_{\mathbf{U}(d)} d\mu(U) \tilde{\mathcal{F}}^\sigma(U) \quad (10.12)$$

that will serve as a reference function that \mathcal{F}_k^σ approximates as $k \rightarrow \infty$. We begin with the following Lemma.

Lemma 8. *The function \mathcal{F}_k^σ defined in the preceding paragraphs satisfies $\mathcal{F}_k^\sigma \in \mathcal{H}_k$ i.e. is a balanced polynomial of degree k in U and \bar{U} .*

The proof of this result, which seems intuitive at the first sight, turns out to surprisingly complex. We present it in Part XIB of the Appendix.

We proceed with giving a number of properties of function $\tilde{\mathcal{F}}^\sigma$. The relevant properties of $\tilde{\mathcal{F}}_k^\sigma$ will be derived latter by controlling the error resulting from the truncation. In order to facilitate the computations involved, our proof strategy effectively shifts the considerations from $\mathbf{U}(d)$ to $\mathbb{U}(d)$. In particular, for class functions defined on the unitary group $\mathbb{U}(d)$ we often make use of the Weyl integration formula [46], which ensures that for any class function F on $\mathbb{U}(d)$ we have

$$\int_{\mathbb{U}(d)} d\mu(U) F(U) = \int_{\mathbb{T}^d} d\mu(\boldsymbol{\varphi}) F(\text{diag}(\exp(i\varphi_1), \dots, \exp(i\varphi_d))) , \quad (10.13)$$

where the measure on \mathbb{T}^d is the *push-forward* of a Haar measure on $\mathbb{U}(d)$ and is given by

$$d\mu(\boldsymbol{\varphi}) = \frac{1}{(2\pi)^d d!} \prod_{1 \leq i < j \leq d} |e^{i\varphi_i} - e^{i\varphi_j}|^2 d\varphi_1 \dots d\varphi_d \quad (10.14)$$

Although the corresponding formula is guaranteed to exist in principle also for class functions on $\mathbf{U}(d)$, we are not aware of any explicit expressions analogous to Eq.(10.13).

Lemma 9 (Lower bound on the normalization constant \mathcal{N}^σ). *Let \mathcal{N}_σ be defined as in Eq.(10.12) and let $\sigma \leq \frac{\pi}{4\sqrt{d}}$. We have the following inequality*

$$\mathcal{N}^\sigma \geq \frac{1}{2} C_d \sigma^{d(d-1)} \left(\frac{2}{\pi} \right)^{d(d-1)}, \quad (10.15)$$

where $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^{d d!}}$.

Sketch of the proof. Observe first that due to the definition of the Haar measure on $\mathbf{U}(d)$ (see Section II) the normalisation constant \mathcal{N}^σ can be expressed via the integral from function F^σ (defined in Eq.(10.8))

$$\mathcal{N}^\sigma = \int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U}) = \int_{\mathbb{U}(d)} d\mu(U) \left(\frac{1}{2\pi} \int_0^{2\pi} d\varphi F^\sigma(\exp(i\varphi)U) \right) = \int_{\mathbb{U}(d)} d\mu(U) F^\sigma(U), \quad (10.16)$$

where in the last equality we used invariance of the Haar measure on $\mathbb{U}(d)$ under the translations by unitary operations (in this case $\exp(i\varphi)I$). Importantly, by the virtue of Weyl integration formula (cf. Eq.(10.13)) the integral appearing in the right-hand side of Eq.(10.16) can be expressed via the integral of the periodized Gaussian f_p^σ defined on \mathbb{T}^d . This allows us to write

$$\int_{\mathbb{U}(d)} d\mu(U) F^\sigma(U) = \int_{\mathbb{T}^d} d\mu(\varphi) f_p^\sigma(\varphi) \geq \int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi), \quad (10.17)$$

where the inequality follows from $f_p^\sigma(\varphi) \geq f^\sigma(\varphi)$. The function f^σ turns out to be closely related to the GUE ensemble of random Hermitian matrices [33] which ultimately allows us to establish the following bound

$$\int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi) \geq \frac{1}{2} C_d \sigma^{d(d-1)} \left(\frac{2}{\pi} \right)^{d(d-1)}, \quad (10.18)$$

where the dimension-dependant constant $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^{d d!}}$ appears because of the usage of the *Mehta integral* [47]. Combining the above inequality with Eq.(10.16) and Eq.(10.17) concludes proofs of Lemma 9. The detailed reasoning justifying Eq. (10.18) is given in Lemma 15 in Appendix XIC. \square

The following result allows us to upper bound the rate of decay of integrals of the form $\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U})$, where $D(\mathbf{V}, \mathbf{I}) \geq \epsilon$ in terms of the integrals on the unitary group $\mathbb{U}(d)$. The latter turn out to be simpler to analyze.

Lemma 10. *Let F^σ and $\tilde{\mathcal{F}}^\sigma$ be functions on $\mathbb{U}(d)$ and $\mathbf{U}(d)$ defined in Eq.(10.8) and Eq.(10.9) respectively. Let $\epsilon \geq \kappa \geq 0$. Then, for any \mathbf{V} satisfying $D(\mathbf{V}, \mathbf{I}) \geq \epsilon$ we have the following inequality*

$$\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U}) \leq \int_{B(\mathbf{I}, \epsilon - \kappa)^c} d\mu(U) F^\sigma(U), \quad (10.19)$$

where $B(\mathbf{I}, r)^c = \{U \in \mathbb{U}(d) \mid \|U - \mathbf{I}\| > r\}$ is the complement of the ball with respect to the operator norm in $\mathbb{U}(d)$.

Proof. We begin by noting that for $D(\mathbf{V}, \mathbf{I}) \geq \epsilon$ and $\kappa \leq \epsilon$ we have $B(\mathbf{V}, \kappa) \subset B(\mathbf{I}, \epsilon - \kappa)^c$ and consequently

$$\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U}) \leq \int_{B(\mathbf{I}, \epsilon - \kappa)^c} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U}). \quad (10.20)$$

Next, from the characterization of the diamond norm given Eq.(2.1) we get that for all $r > 0$

$$\{U \in \mathbb{U}(d) \mid D(\mathbf{U}, \mathbf{I}) \leq r\} = \bigcup_{\phi} \{U \in \mathbb{U}(d) \mid \|e^{i\phi} \mathbf{I} - U\| \leq r\}. \quad (10.21)$$

The connection between the Haar measures on $\mathbb{U}(d)$ and $\mathbf{U}(d)$ and the definition of $\tilde{\mathcal{F}}^\sigma$ gives

$$\int_{B(\mathbf{I}, r)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U}) = \int_{\bigcup_{\phi} B(\exp(i\phi)\mathbf{I}, r)} d\mu(U) \mathbb{P}_{\text{phase}} F^\sigma(U) = \int_{\bigcup_{\phi} B(\exp(i\phi)\mathbf{I}, r)} d\mu(U) F^\sigma(U), \quad (10.22)$$

where in the last equality we used the invariance of the set $\bigcup_{\phi} B(\exp(i\phi)I, r)$ with respect to the multiplication by the global phase. Next, the condition $F^\sigma(U) \geq 0$ implies

$$\int_{B(I, r)} d\mu(U) \tilde{\mathcal{F}}^\sigma(U) \geq \int_{B(I, r)} d\mu(U) F^\sigma(U) . \quad (10.23)$$

By setting $r = \epsilon - \kappa$ in the above inequality and combining this with Eq.(10.16) we get

$$\int_{B(I, \epsilon - \kappa)^c} d\mu(U) \tilde{\mathcal{F}}^\sigma(U) \leq \int_{B(I, \epsilon - \kappa)^c} d\mu(U) F^\sigma(U) . \quad (10.24)$$

Inserting this inequality to Eq. (10.20) concludes the proof. \square

We now want to control the rate of decay of the integral appearing in the right-hand side of Eq.(10.19). To this end we use Weyl integration formula (cf. Eq.(10.13)) which gives

$$\int_{B(I, r)^c} d\mu(U) F^\sigma(U) = \int_{B_\infty(0, r)^c} d\mu(\varphi) f_p^\sigma(\varphi) , \quad (10.25)$$

where $B_\infty(0, r) = \{\varphi \in \mathbb{T}^d \mid |\varphi_i| \leq r\}$ and $B_\infty(0, r)^c$ is its complement in \mathbb{T}^d . Next, in Lemma 16 given in part XI C of the Appendix we establish upper bounds on the right-hand side of (10.25). This result is proven by (i) establishing appropriate upper bounds on the norm $\|f^\sigma - f_p^\sigma\|_1$ where $\|\cdot\|_1$ denotes L^1 norm of the space of integrable functions on \mathbb{T}^d equipped with the measure $d\mu(\varphi)$, and (ii) connecting $\int_{B_\infty(0, r)^c} d\mu(\varphi) f_p^\sigma(\varphi)$ to the tail behaviour of the operator norm of GUE matrices [33]. In this way we obtain the following lemma:

Lemma 11. *Let F^σ be the function on $\mathbb{U}(d)$ defined in Eq.(10.8). Let $\sigma \leq \frac{r}{4\sqrt{d}}$ and $r \leq 2/3$. We have the following inequality*

$$\int_{B(I, r)^c} d\mu(U) F^\sigma(U) \leq \frac{3}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}} \quad (10.26)$$

where $B(I, r)^c = \{U \in \mathbb{U}(d) \mid \|U - I\| > r\}$ and $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$.

We conclude our characterisation of the functions $\tilde{\mathcal{F}}^\sigma$ and \mathcal{F}^σ by the following proposition, which follows easily from the well-known inequality between L^2 and L^1 norms

$$\|\tilde{\mathcal{F}}^\sigma\|_2 \leq \|\tilde{\mathcal{F}}^\sigma\|_1 , \quad (10.27)$$

and the positivity of $\tilde{\mathcal{F}}^\sigma$, which ensures that $\|\tilde{\mathcal{F}}^\sigma\|_1 = \mathcal{N}^\sigma$.

Proposition 4. *Let $\tilde{\mathcal{F}}^\sigma$ be a function on $\mathbb{U}(d)$ defined in Eq. (10.9) and let \mathcal{N}^σ be a constant defined in Eq. (10.12). We have the following inequality*

$$\|\tilde{\mathcal{F}}^\sigma\|_2 \leq \mathcal{N}^\sigma . \quad (10.28)$$

Consequently, L^2 norm of $\mathcal{F}^\sigma = \tilde{\mathcal{F}}^\sigma / \mathcal{N}^\sigma$ satisfies

$$\|\mathcal{F}^\sigma\|_2 \leq 1 . \quad (10.29)$$

Combining the results stated in Lemmas 9, 10, 11 with Proposition 4 establishes that the function \mathcal{F}^σ is normalized, satisfies $\|\mathcal{F}^\sigma\|_2 \leq 1$, and for $\sigma \leq \frac{\epsilon - \kappa}{4\sqrt{d}}$ we have

$$\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \mathcal{F}^\sigma(\mathbf{U}) \leq 3 \exp\left(-\frac{(\epsilon - \kappa)^2}{4\sigma^2}\right) \left(\frac{\pi}{2}\right)^{d(d-1)} , \quad (10.30)$$

whenever $D(\mathbf{V}, \mathbf{I}) \geq \epsilon$. Looking at the above conditions we see that they closely resemble properties required from \mathcal{F}_k^σ in Theorem 1. The following key Lemma controls the rate of approximation of $\tilde{\mathcal{F}}^\sigma$ by $\tilde{\mathcal{F}}_k^\sigma$ in L^2 norm.

Lemma 12 (Approximation of $\tilde{\mathcal{F}}^\sigma$ by $\tilde{\mathcal{F}}_k^\sigma$). *Let $k \geq d/\sigma$ and let $\sigma \leq 1/2$. Let $\tilde{\mathcal{F}}_k^\sigma$ and $\tilde{\mathcal{F}}^\sigma$ be functions defined in Eq.(10.9). We have the following upper bound on the L^2 distance between these functions*

$$\left\| \tilde{\mathcal{F}}^\sigma - \tilde{\mathcal{F}}_k^\sigma \right\|_2 \leq 10 C_d \frac{e^{-\frac{1}{4}(\frac{k}{\sqrt{d}} - \sqrt{d})^2 \sigma^2}}{\sigma} . \quad (10.31)$$

Proof sketch. Similary as before we reduce the problem to consideration of functions on \mathbb{T}^d . First, due to the fact that $\mathbb{P}_{\text{phase}}$ is an orthonormal projector in $L^2(\mathbb{U}(d))$ we get

$$\left\| \tilde{\mathcal{F}}^\sigma - \tilde{\mathcal{F}}_k^\sigma \right\|_2 = \left\| \mathbb{P}_{\text{phase}} (F^\sigma - F_k^\sigma) \right\|_2 \leq \|F^\sigma - F_k^\sigma\|_2 . \quad (10.32)$$

Using the Weyl integration formula and definitions of class functions F^σ, F_k^σ (cf. Eq.(10.8)) we obtain $\|F^\sigma - F_k^\sigma\|_2 = \|f_p^\sigma - f_{p,k}^\sigma\|_2$, where the L^2 distance between f_p^σ and $f_{p,k}^\sigma$ is computed using the measure $d\mu(\varphi)$ on \mathbb{T}^d (this is a consequence of Eq.(10.13)). The claimed result follows now from the inequality (valid for $k \geq d/\sigma$ and $\sigma \leq 1/2$).

$$\|f_p^\sigma - f_{p,k}^\sigma\|_2 \leq 10 C_d \frac{e^{-\frac{1}{4}(\frac{k}{\sqrt{d}} - \sqrt{d})^2 \sigma^2}}{\sigma} \quad (10.33)$$

which we prove in Lemma 18 in Part XID of the Appendix using trigonometric expansions (10.5) and (10.6). \square

The following proposition asserts that for sufficiently large degree k the L^2 -distance $\|\tilde{\mathcal{F}}^\sigma - \tilde{\mathcal{F}}_k^\sigma\|_2$ is comparable with \mathcal{N}^σ (cf. Lemma 9) and the upper bound on $\int_{B(I,r)^c} d\mu(U) F^\sigma(U)$ from Lemma 11.

Proposition 5. *Let $\tilde{\mathcal{F}}_k^\sigma$ and $\tilde{\mathcal{F}}^\sigma$ be functions defined in Eq.(10.9). Moreover let*

$$\sigma \leq \min\{1/8, \frac{\pi}{4\sqrt{d}}\} , \quad k \geq 5 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{\frac{1}{8} \frac{r^2}{d^2 \sigma^2} + \ln \frac{1}{\sigma}} , \quad r \leq \frac{2}{3} . \quad (10.34)$$

Then we have

$$\|\tilde{\mathcal{F}}^\sigma - \tilde{\mathcal{F}}_k^\sigma\|_2 \leq \frac{1}{2} \max \left\{ \mathcal{N}^\sigma, \frac{3}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}} \right\} , \quad (10.35)$$

where \mathcal{N}^σ is a constant defined in Eq.(10.12) and $\frac{3}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}}$ is an upper bound on the integral $\int_{B(I,r)^c} d\mu(U) F^\sigma(U)$ from Lemma 11.

The proof of the above result follows from comparison of upper bound (10.31) with the bounds given in Lemmas 9 and 11. The comparison is provided by the (technical) Lemma 20 proved in Part XID of the Appendix. We have now all the necessary ingredients to justify that the function \mathcal{F}_k^σ satisfies all the properties required by Theorem 1.

Proof of Theorem 1. For $\epsilon \geq \kappa$ we set $r = \epsilon - \kappa$. According to assumptions of the Theorem we will assume $\sigma \leq \frac{\epsilon - \kappa}{4\sqrt{d}}$ and $k \geq 5 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{\frac{1}{8} \frac{r^2}{d^2 \sigma^2} + \ln \frac{1}{\sigma}}$. Then, Lemma 8 ensures that \mathcal{F}_k^σ is a polynomial of a suitable degree in U and \bar{U} as claimed in property 3 in Theorem 1. We proceed with proofs of the remaining three properties \mathcal{F}_k^σ .

We start by establishing the normalization of \mathcal{F}_k^σ (condition 1 in Theorem 1). By definition of \mathcal{F}_k^σ (cf. Eq.(10.11)) this is equivalent to showing

$$\mathcal{N}_k^\sigma = \int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}_k^\sigma(\mathbf{U}) \neq 0 . \quad (10.36)$$

By simple manipulations we get

$$\int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}_k^\sigma(\mathbf{U}) = \int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U}) - \int_{\mathbf{U}(d)} d\mu(\mathbf{U}) (\tilde{\mathcal{F}}^\sigma(\mathbf{U}) - \tilde{\mathcal{F}}_k^\sigma(\mathbf{U})) \geq \int_{\mathbf{U}(d)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}^\sigma(\mathbf{U}) - \|\tilde{\mathcal{F}}^\sigma - \tilde{\mathcal{F}}_k^\sigma\|_2 , \quad (10.37)$$

where the inequality comes from applying the Cauchy-Schwartz inequality in $L^2(\mathbf{U}(d))$ to functions $G_1(\mathbf{U}) = \tilde{\mathcal{F}}^\sigma(\mathbf{U}) -$

$\tilde{\mathcal{F}}_k^\sigma(\mathbf{U})$ and $G_2(\mathbf{U}) = 1$. Using (10.35) in the above inequality we obtain

$$\mathcal{N}_k^\sigma \geq \frac{1}{2} \mathcal{N}^\sigma > 0 . \quad (10.38)$$

The proof of the second property in Theorem 1 (decay of integrals over balls $B(\mathbf{V}, \kappa)$ when $D(\mathbf{V}, \mathbf{I}) \geq \epsilon$) follows the similar logic. Specifically, using the Cauchy-Schwartz inequality leads, as previously, to

$$\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}_k^\sigma \leq \int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}_k^\sigma + \|\tilde{\mathcal{F}} - \tilde{\mathcal{F}}_k^\sigma\|_2 . \quad (10.39)$$

Application of the bound (10.35) from Proposition 5 and results of Lemmas 10 and 11 gives

$$\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}_k^\sigma \leq \frac{3}{2} \times \frac{3}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{(\epsilon - \kappa)^2}{\sigma^2}} . \quad (10.40)$$

Using the definition of \mathcal{F}_k^σ and employing (10.38) together with the lower bound for \mathcal{N}^σ from Lemma 9 we finally obtain the desired result

$$\int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \mathcal{F}_k^\sigma = \frac{1}{\mathcal{N}_k^\sigma} \int_{B(\mathbf{V}, \kappa)} d\mu(\mathbf{U}) \tilde{\mathcal{F}}_k^\sigma \leq 3 \times 3 \exp\left(-\frac{(\epsilon - \kappa)^2}{4\sigma^2}\right) \left(\frac{\pi}{2}\right)^{d(d-1)} . \quad (10.41)$$

We conclude the proof by giving an upper bound on $\|\mathcal{F}^\sigma\|_2$. We proceed analogously as before by relating \mathcal{F}_k^σ with $\tilde{\mathcal{F}}_k^\sigma$:

$$\|\mathcal{F}_k^\sigma\|_2 = \frac{1}{\mathcal{N}_k^\sigma} \|\tilde{\mathcal{F}}_k^\sigma\|_2 \leq \frac{2}{\mathcal{N}^\sigma} \left(\|\tilde{\mathcal{F}}^\sigma\|_2 + \|\tilde{\mathcal{F}}_k^\sigma - \tilde{\mathcal{F}}^\sigma\|_2 \right) , \quad (10.42)$$

where the inequality follows from (10.38) (applied to the denominator) and triangle inequality for $\|\cdot\|_2$ (applied to the numerator). Next, using Proposition 4 ($\|\tilde{\mathcal{F}}^\sigma\|_2 \leq \mathcal{N}^\sigma$) and again employing (10.35) ($\|\tilde{\mathcal{F}}_k^\sigma - \tilde{\mathcal{F}}^\sigma\|_2 \leq \mathcal{N}^\sigma/2$), we finally obtain $\|\mathcal{F}_k^\sigma\|_2 \leq 3$, which concludes the proof. \square

-
- [1] P. P. Varju, *Doc. Math.* **18**, 1137 (2013).
 - [2] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, *Communications in Mathematical Physics* **346**, 397 (2016), [arXiv:1208.0692 \[quant-ph\]](#).
 - [3] C. Dankert, R. Cleve, J. Emerson, and E. Livine, *Physical Review A* **80**, 012304 (2009), [arXiv:quant-ph/0606161 \[quant-ph\]](#).
 - [4] J. M. Epstein, A. W. Cross, E. Magesan, and J. M. Gambetta, *Phys. Rev. A* **89**, 062321 (2014), [arXiv:1308.2928 \[quant-ph\]](#).
 - [5] H.-Y. Huang, R. Kueng, and J. Preskill, *arXiv e-prints*, [arXiv:2002.08953 \(2020\)](#), [arXiv:2002.08953 \[quant-ph\]](#).
 - [6] O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner, *New Journal of Physics* **15**, 053022 (2013), [arXiv:1109.4348 \[quant-ph\]](#).
 - [7] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, *Proceedings of the Royal Society of London Series A* **465**, 2537 (2009), [arXiv:quant-ph/0606225 \[quant-ph\]](#).
 - [8] P. Sen, *arXiv e-prints*, [quant-ph/0512085 \(2005\)](#), [arXiv:quant-ph/0512085 \[quant-ph\]](#).
 - [9] L. Masanes, A. J. Roncaglia, and A. Acín, *arXiv e-prints*, [arXiv:1108.0374 \(2011\)](#), [arXiv:1108.0374 \[quant-ph\]](#).
 - [10] M. Oszmaniec, R. Augusiak, C. Gogolin, J. Kołodyński, A. Acín, and M. Lewenstein, *Phys. Rev. X* **6**, 041044 (2016).
 - [11] D. A. Roberts and B. Yoshida, *Journal of High Energy Physics* **2017**, 121 (2017), [arXiv:1610.04903 \[quant-ph\]](#).
 - [12] Y. Nakata, C. Hirche, M. Koashi, and A. Winter, *Phys. Rev. X* **7**, 021006 (2017), [arXiv:1609.07021 \[quant-ph\]](#).
 - [13] AAAAA, *Doc. Math.* **18**, 1137 (2013).
 - [14] A. W. Harrow and A. Montanaro, *Nature (London)* **549**, 203 (2017), [arXiv:1809.07442 \[quant-ph\]](#).
 - [15] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Nature Physics* **14**, [arXiv:1608.00263 \(2018\)](#), [arXiv:1608.00263 \[quant-ph\]](#).
 - [16] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostitsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant,

- X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, *Nature (London)* **574**, 505 (2019).
- [17] D. Hangleiter, J. Bermejo-Vega, M. Schwarz, and J. Eisert, *Quantum* **2**, 65 (2018).
- [18] M. Yoganathan, R. Jozsa, and S. Strelchuk, *Proceedings of the Royal Society of London Series A* **475**, 20180427 (2019), [arXiv:1806.03200 \[quant-ph\]](#).
- [19] H. Zhu, R. Kueng, M. Grassl, and D. Gross, arXiv e-prints, arXiv:1609.08172 (2016), [arXiv:1609.08172 \[quant-ph\]](#).
- [20] J. Bourgain and A. Gamburd, arXiv e-prints, arXiv:1108.6264 (2011), [arXiv:1108.6264 \[math.GR\]](#).
- [21] A. Harrow and S. Mehraban, arXiv e-prints, arXiv:1809.06957 (2018), [arXiv:1809.06957 \[quant-ph\]](#).
- [22] R. Mezher, J. Ghalbouni, J. Dgheim, and D. Markham, *Entropy* **22**, 92 (2020), [arXiv:1911.03704 \[quant-ph\]](#).
- [23] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth, arXiv e-prints, arXiv:2002.09524 (2020), [arXiv:2002.09524 \[quant-ph\]](#).
- [24] Y. Nakata, C. Hirche, C. Morgan, and A. Winter, *Journal of Mathematical Physics* **58**, 052203 (2017), [arXiv:1502.07514 \[quant-ph\]](#).
- [25] L. Banchi, D. Burgarth, and M. J. Kastoryano, *Physical Review X* **7**, 041015 (2017), [arXiv:1704.03041 \[quant-ph\]](#).
- [26] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (American Mathematical Society, Boston, MA, USA, 2002).
- [27] C. M. Dawson and M. A. Nielsen, arXiv e-prints, quant-ph/0505030 (2005), [arXiv:quant-ph/0505030 \[quant-ph\]](#).
- [28] I. S. B. Sardharwalla, T. S. Cubitt, A. W. Harrow, and N. Linden, arXiv e-prints, arXiv:1602.07963 (2016), [arXiv:1602.07963 \[quant-ph\]](#).
- [29] A. Bouland and M. Ozols, arXiv e-prints, arXiv:1712.09798 (2017), [arXiv:1712.09798 \[quant-ph\]](#).
- [30] A. W. Harrow, B. Recht, and I. L. Chuang, *Journal of Mathematical Physics* **43**, 4445 (2002), [arXiv:quant-ph/0111031 \[quant-ph\]](#).
- [31] M. B. Hastings and A. W. Harrow, *Quantum Info. Comput.* **9**, 336360 (2009).
- [32] R. A. Low, *Pseudo-randomness and Learning in Quantum Computation*, Ph.D. thesis, - (2010).
- [33] G. Aubrun and S. J. Szarek, *Alice and Bob meet Banach: the interface of asymptotic geometric analysis and quantum information theory*, Mathematical surveys and monographs (American Mathematical Society, Providence, RI, 2017).
- [34] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction* (Springer, 2000).
- [35] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*, 10th ed. (Cambridge University Press, USA, 2011).
- [36] N. Johnston, D. W. Kribs, and V. I. Paulsen, *Quantum Info. Comput.* **9**, 1635 (2009).
- [37] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, *Phys. Rev. Lett.* **116**, 170502 (2016), [arXiv:1605.00713 \[quant-ph\]](#).
- [38] A. Sawicki and K. Karnas, *Annales Henri Poincaré* **18**, 3515 (2017), [arXiv:1609.05780 \[quant-ph\]](#).
- [39] A. Sawicki and K. Karnas, *Phys. Rev. A* **95**, 062303 (2017), [arXiv:1610.00547 \[quant-ph\]](#).
- [40] Z. Zimborás, R. Zeier, T. Schulte-Herbrüggen, and D. Burgarth, *Phys. Rev. A* **92**, 042309 (2015).
- [41] M. Oszmaniec and Z. Zimborás, *Phys. Rev. Lett.* **119**, 220502 (2017), [arXiv:1705.11188 \[quant-ph\]](#).
- [42] D. A. Roberts and B. Yoshida, *Journal of High Energy Physics* **2017**, 121 (2017), [arXiv:1610.04903 \[quant-ph\]](#).
- [43] L. Susskind, arXiv e-prints, arXiv:1810.11563 (2018), [arXiv:1810.11563 \[hep-th\]](#).
- [44] S. J. Szarek, *QUANTUM PROBABILITY* **43**, 395 (1998).
- [45] C. Gasquet and P. Witomski, *Fourier Analysis and Applications* (Springer, 1999).
- [46] R. W. Goodman and N. R. Wallach, *Symmetry, Representations, and Invariants*, Graduate Texts in Mathematics (Springer, Dordrecht, 2009).
- [47] I. G. Macdonald, *SIAM Journal on Mathematical Analysis* **13**, 988 (1982).
- [48] I. N. Bronshtein, K. A. Semendiaev, G. Musiol, and H. Muhlig, *Handbook of Mathematics* (Springer, 2007).

XI. APPENDIX

In the appendix we provide technical details not included in the main text. Specifically, in Part [XIA](#) we prove the relation between expander norms and restricted gaps on projective unitary group. This allows us to complete the proof of Theorem [5](#). In Part [XIB](#) we give the proof of Lemma [8](#). In the rest of the Appendix we technical results important for the construction of a polynomial approximation of Dirac delta on $\mathbf{U}(d)$. Part [XIC](#) contains proofs of certain properties of of periodized Gaussian in \mathbb{T}^d (and its extension to the unitary group $\mathbf{U}(d)$). The latter Part [XID](#) presents estimates for the convergence of the polynomial truncation of this function in suitable norms. Finally, in Part [XIE](#) we gather auxiliary results and facts that are used in ealier sections of the Appendix.

A. Proof of Theorem [5](#)

Before we use the results of [\[1\]](#) we need to introduce a couple of concepts from representation theory (we refer the reader to [\[34\]](#) for the comprehensive introduction to representation theory of semisimple Lie groups and Lie algebras).

Let G be a compact semisimple Lie group and let Π be a representation of G in a finite-dimensional Hilbert space \mathcal{K} . Let π be the associated representation of the Lie algebra of G denoted by \mathfrak{g} . Let \mathfrak{t} be a Lie algebra of the maximal torus in G . Weights $\alpha \in \mathfrak{t}$ encode joint eigenvalues of elements $X \in \mathfrak{t}$ in irreducible representations of G . In other words for every weight α there exist a representation π and a vector $|\psi_\alpha\rangle$ such that for all $X \in \mathfrak{t}$

$$\pi(X) |\psi_\alpha\rangle = i\langle\alpha, X\rangle |\psi_\alpha\rangle , \quad (11.1)$$

where $\langle\cdot, \cdot\rangle$ is a non-degenerate inner product in \mathfrak{g} induced by the Killing form. Recall that irreducible representations of compact connected semisimple Lie Groups are finite-dimensional labelled by the so-called highest weights λ , which are weights that satisfy some additional technical properties. Moreover, for every finite-dimensional representation \mathcal{K} of G it is possible to choose a basis consisting of weight vectors.

In [1] the author considers the decay of spectral gap of T_ν for semisimple compact Lie groups, G . In particular he focuses on restriction, $T_\nu|_{\mathcal{K}_r}$, of the operator T_ν to the space \mathcal{K}_r defined by

$$\mathcal{K}_r = \bigoplus_{\lambda: 0 < \|\lambda\|_G \leq r} \mathcal{K}_\lambda , \quad (11.2)$$

where \mathcal{K}_λ denotes the irreducible representation of highest weight λ and $\|\lambda\|_G = \sqrt{\langle\lambda, \lambda\rangle}$ is the norm of λ induced by the Killing form in \mathfrak{g} . In what follows we will be extensively using the notation $\mathcal{K}_\lambda \subset_G \mathcal{K}$ to denote the situation in which irreducible representation \mathcal{K}_λ of G appears in the decomposition of \mathcal{K} onto irreducible components.

The spectral gap of $T_\nu|_{\mathcal{K}_r}$ is then defined as

$$\text{gap}_r(G, \nu) = 1 - \|T_\nu|_{\mathcal{K}_r}\|_\infty . \quad (11.3)$$

In this setting the following theorem holds

Theorem 8 (Theorem 6 in [1]). *For every semisimple compact connected Lie group G , there are numbers c and r_0 such that the following holds. Let ν be an arbitrary probability measure on G . Then*

$$\text{gap}_r(G, \nu) \geq c \text{gap}_{r_0}(G, \nu) \log^{-2}(r) . \quad (11.4)$$

In what follows we apply the above result for $G = \mathbf{U}(d)$. In order to see how Theorem 5 follows from Theorem 8 let us first note that inequality (11.4) can be written as

$$\|T_\nu|_{\mathcal{K}_r}\|_\infty \leq 1 - \frac{1 - \|T_\nu|_{\mathcal{K}_{r_0}}\|_\infty}{(1/c) \log^2(r)} . \quad (11.5)$$

We thus aim to find the relation between r and t . To this end we first notice that representation $\Pi^{1,1}(\mathbf{U}) := U \otimes \bar{U}$ decomposes as the direct sum of the adjoint representation Ad and the trivial representation. Furthermore the adjoint representation is the faithful representation of $\mathbf{U}(d)$. Thus every irreducible representation \mathcal{H}_λ will appear in the decomposition of $(\Pi^{1,1})^{\otimes t}$ into irreducible components for sufficiently large t . The same can be said about the decomposition of \mathcal{H}_t (on which $\mathbf{U}(d)$ acts via its regular representation) into irreducible representations of $\mathbf{U}(d)$. This follows from considerations given in Section V and the fact that representations $\Pi_A(U) = U^{\otimes t} \otimes \bar{U}^{\otimes t}$ is equivalent to $\Pi_B(U) = (\Pi^{1,1}(U))^{\otimes t}$. Using this observation we conclude that there exist t_0 such that for all λ satisfying $0 < \|\lambda\|_G \leq r_0$ we have $\mathcal{K}_\lambda \subset_{\mathbf{U}(d)} \mathcal{H}_{t_0}$ (for some suitable t_0 depending on r_0). Therefore we have

$$\|T_\nu|_{\mathcal{K}_{r_0}}\|_\infty \leq \|T_\nu|_{\mathcal{H}_{t_0}} - T_\mu|_{\mathcal{H}_{t_0}}\|_\infty = \|T_{\nu, t_0} - T_{\mu, t_0}\|_\infty , \quad (11.6)$$

where $T_\mu|_{\mathcal{H}_{t_0}}$ is a projector onto a trivial representation (space of constant functions) in \mathcal{H}_{t_0} . The first inequality in the above equation comes from the fact that (c.f. Eq.(11.2)) \mathcal{K}_r does not contain trivial representations of $\mathbf{U}(d)$. The equality follows from Proposition 1 in Section V.

In the second step we observe that all irreducible representations appearing in $(\Pi^{1,1})^{\otimes t}$ have highest weights of magnitude $\|\lambda\|_G \leq at$, where a depends only on d . This follows from the fact that weight vectors associated with the representation $(\Pi^{1,1})^{\otimes t}$ can be chosen to have tensor product structure i.e.

$$|\psi_\beta\rangle = |\psi_{\alpha_1}\rangle \otimes |\psi_{\alpha_2}\rangle \otimes \dots \otimes |\psi_{\alpha_t}\rangle , \quad (11.7)$$

where $|\psi_{\alpha_2}\rangle \in \mathbb{C}^d \otimes \mathbb{C}^d$ is a weight vector in the representation $\Pi^{1,1}$. Consequently, all weights β occurring in the representation $(\Pi^{1,1})^{\otimes t}$ (or equivalently the regular representation of $\mathbf{U}(d)$ restricted to the function space \mathcal{H}_t) are

sums of weights associated with $\Pi^{1,1}$: $\beta = \sum_{i=1}^t \alpha_i$. Using triangle inequality we obtain that for all λ such that $K_\lambda \subset_{\mathbf{U}(d)} \mathcal{H}_t$

$$\|\lambda\|_{\mathbf{U}(d)} \leq t \max_{\alpha} \|\alpha\|_{\mathbf{U}(d)}, \quad (11.8)$$

where the optimization is over all weights α appearing in $\Pi^{1,1}$. It is clear that $\max_{\alpha} \|\alpha\|_{\mathbf{U}(d)} = a$ depends only on the dimension d .

Inequalities (11.6) and (11.8) allow us to employ Eq.(11.5) in the context of expander norms. First, from (11.8) it follows that every nontrivial representation $K_\lambda \subset_{\mathbf{U}(d)} \mathcal{H}_t$ satisfies also $K_\lambda \subset_{\mathbf{U}(d)} K_{at}$ and consequently

$$\|T_{\nu,t} - T_{\mu,t}\|_{\infty} \leq \|T_{\nu}|K_{at}\|_{\infty}. \quad (11.9)$$

Applying to the above first (11.5) and then (11.6) gives

$$\|T_{\nu,t} - T_{\mu,t}\|_{\infty} \leq 1 - \frac{1 - \|T_{\nu}|K_{r_0}\|_{\infty}}{(1/c) \log^2(at)} \leq 1 - \frac{1 - \|T_{\nu,t_0} - T_{\mu,t_0}\|_{\infty}}{(1/c) \log^2(at)}. \quad (11.10)$$

It is now easy to verify that for $t \geq 2 \log(at) \leq C \log(t)$ for $C = 1 + \frac{\log(a)}{\log(2)}$ and therefore by setting $B = C^2/c$ we obtain the claimed result: for $t \geq t_0$ (t_0 is defined above Eq.(11.6)) and for any probability measure ν on $\mathbf{U}(d)$ we have

$$\|T_{\nu,t} - T_{\mu,t}\|_{\infty} \leq 1 - \frac{1 - \|T_{\nu,t_0} - T_{\mu,t_0}\|_{\infty}}{B \log^2(t)}. \quad (11.11)$$

B. Proof of Lemma 8

In order to prove Lemma 8 we introduce the following notation. Let $\mathbf{n} = (n_1, \dots, n_d) \in \mathbb{Z}^d$ be such that $\sum_{i=1}^d n_i = 0$ and $\sum_{i=1}^d |n_i| = 2k$. Let S_d denote the symmetric group on d symbols. We define action of $\sigma \in S_d$ on $\mathbf{n} \in \mathbb{Z}^d$ by:

$$\sigma(\mathbf{n}) = (n_{\sigma(1)}, \dots, n_{\sigma(d)}). \quad (11.12)$$

Recall that a partition of a set X is a set of non-empty subsets of X such that every element $x \in X$ is in exactly one of these subsets. Let P_d be a partition of \mathbf{n} , where we view \mathbf{n} as an d -element set. Let $|P_d|$ be the number of subsets in partition P_d . By the abuse of notation we define $P_d(\mathbf{n})$ to be a vector whose first $|P_d|$ coefficients are sums of elements in the corresponding partition subsets and the remaining coefficients are equal to zero. For example, for $d = 3$ and partition P_3 of \mathbf{n} given by $\{\{n_1, n_2\}, n_3\}$ we have $P_3(\mathbf{n}) = (n_1 + n_2, n_3, 0)$. Finally, recall that $\mathcal{H}_k = \text{Span} \left\{ \text{tr}(A_t U^{\otimes t} \otimes \bar{U}^{\otimes t}) : A \in \text{End}(\mathbb{C}^{d^{2t}}), t \in \{0, 1, \dots, k\} \right\}$. Our aim is to prove the following Lemma whose corollary is Lemma 8.

Lemma 13. \mathcal{H}_k contains all functions of the form:

$$\sum_{\sigma \in S_d} e^{i\sigma(\mathbf{n})\varphi}, \quad (11.13)$$

where \mathbf{n} is such that $\sum_{i=1}^d n_i = 0$ and $\sum_{i=1}^d |n_i| = 2t$, $0 \leq t \leq k$. In particular $\mathcal{F}_k^\sigma \in \mathcal{H}_k$.

In order to prove the above Lemma we first observe that if C_m is the natural matrix representation of the cyclic permutation $(1, \dots, m)$ on $\mathbb{C}^{d^m} = \mathbb{C}^d \otimes \dots \otimes \mathbb{C}^d$ then

$$\text{tr}(C_m A_1 \otimes \dots \otimes A_m) = \text{tr}(A_1 \cdot \dots \cdot A_m) \quad (11.14)$$

Thus if we choose $A_t \in \text{End}(\mathbb{C}^{d^{2t}})$ to be

$$A_t = C_{n_1} \otimes C_{n_2} \otimes \dots \otimes C_{n_\alpha} \otimes C_{n_{\alpha+1}} \otimes \dots \otimes C_{n_d}, \quad (11.15)$$

where $\sum_{i=1}^\alpha n_i = t = \sum_{i=\alpha+1}^d n_i$ the resulting function is:

$$f_{\mathbf{n}}(\varphi) := \text{tr} (A_t U^{\otimes t} \otimes \bar{U}^{\otimes t}) = \text{tr} U^{n_1} \text{tr} U^{n_2} \dots \text{tr} U^{n_\alpha} \text{tr} \bar{U}^{n_{\alpha+1}} \dots \text{tr} \bar{U}^{n_d} = \quad (11.16)$$

$$= \prod_{j=1}^{\alpha} \left(\sum_{k=1}^d e^{in_j \phi_k} \right) \prod_{j=\alpha+1}^d \left(\sum_{k=1}^d e^{-in_j \phi_k} \right), \quad (11.17)$$

which can be reduced to

$$f_{\mathbf{n}}(\varphi) = \sum_{P_d} \alpha(P_d) \sum_{\sigma \in S_d} e^{i\sigma(P_d(\mathbf{n}))\varphi}, \quad \mathbf{n} = (n_1, n_2, \dots, n_\alpha, -n_{\alpha+1}, \dots, -n_d). \quad (11.18)$$

We are now ready to give a proof of Lemma 13.

Proof. By direct calculations one checks that Lemma 13 is valid for $k = 0$ and $k = 1$. We follow by induction, i.e. we assume Lemma 13 is valid for $k \geq 1$ and our aim is to show that this implies its validity for $k + 1$. Let \mathbf{n} be such that $\sum_{i=1}^d n_i = 0$ and $\sum_{i=1}^d |n_i| = 2(k + 1)$. Consider the function $f_{\mathbf{n}} \in \mathcal{H}_{k+1}$. The summand corresponding to the full partition P_d is:

$$\sum_{\sigma \in S_d} e^{i\sigma(\mathbf{n})\varphi} \quad (11.19)$$

Thus to show that (11.19) belongs to \mathcal{H}_{k+1} it suffices to show that for other partitions P_d the corresponding summands appearing in (11.18) are either in \mathcal{H}_k or can be easily proved to be in \mathcal{H}_{k+1} . The latter happens only when $\sum |P_d(\mathbf{n})_i| = 2k + 2$ that is P_d respects division of \mathbf{n} into two parts $\{n_1, \dots, n_\alpha\}$ and $\{n_{\alpha+1}, \dots, n_d\}$. Note however that such $P_d(\mathbf{n})$ has at least one zero entry. We can perform the same reasoning with function $f_{P_d(\mathbf{n})}$ and select vectors $P_{d'}(P_d(\mathbf{n}))$ satisfying $\sum |P_{d'}(P_d(\mathbf{n}))_i| = 2k + 2$. They necessarily have at least two zero coefficients. Following this procedure we always arrive at the vector $\mathbf{n}_1 = (k + 1, 0, \dots, 0, -k - 1)$ whose corresponding function $f_{\mathbf{n}_1} \in \mathcal{H}_{k+1}$ is

$$\sum_{\sigma \in S_d} e^{i\sigma(\mathbf{n}_1)\varphi} + a_1, \quad (11.20)$$

where $a_1 \in \mathcal{H}_k$. Thus $\sum_{\sigma \in S_d} e^{i\sigma(\mathbf{n}_1)\varphi} \in \mathcal{H}_k$. Next, reversing the path of the above reasoning we obtain the desired result. \square

C. Estimates for Gaussian functions on a torus

In this part of the Appendix we complete proofs of Lemmas 9 and 11 from Section X. For reader's convenience we collect here concepts and notations that will be used in the reminder of the Appendix. We will use the following measures defined on \mathbb{T}^d :

$$d\varphi = d\varphi_1 \dots d\varphi_d, \quad d\mu(\varphi) = \frac{1}{(2\pi)^d d!} \prod_{1 \leq i < j \leq d} |e^{i\varphi_i} - e^{i\varphi_j}|^2 d\varphi. \quad (11.21)$$

We also introduce the counterparts of these measures on \mathbb{R}^d :

$$d\mathbf{x} = dx_1 \dots dx_d, \quad d\mu_G(\mathbf{x}) = \frac{1}{(2\pi)^d d!} \prod_{1 \leq i < j \leq d} (x_i - x_j)^2 d\mathbf{x}. \quad (11.22)$$

Recall the functions on \mathbb{T}^d that were used to define polynomial approximation to the Dirac delta at $\mathbf{I} \in \mathbf{U}(d)$:

$$f^\sigma(\varphi) = \frac{1}{(\sqrt{2\pi}\sigma)^d} e^{-\frac{1}{2} \frac{\varphi^2}{\sigma^2}}, \quad f_p^\sigma(\varphi) = \frac{1}{(\sqrt{2\pi}\sigma)^d} \sum_{\mathbf{k} \in \mathbb{Z}^d} e^{-\frac{1}{2} \frac{(\varphi + 2\pi\mathbf{k})^2}{\sigma^2}}. \quad (11.23)$$

Poisson summation formula implies

$$f_p^\sigma(\varphi) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n} \in \mathbb{Z}^d} e^{-\frac{1}{2}\mathbf{n}^2 \sigma^2} e^{-i\mathbf{n}\varphi} . \quad (11.24)$$

The truncation of this function to trigonometric polynomials of degree at most k is given by

$$f_{p,k}^\sigma(\varphi) = \frac{1}{(2\pi)^d} \sum_{\mathbf{n} \in S_k} e^{-\frac{1}{2}\mathbf{n}^2 \sigma^2} e^{-i\mathbf{n}\varphi} , \quad (11.25)$$

where $S_k = \{\mathbf{n} \mid |\mathbf{n}|_1 \leq k\}$. The above define class functions on $\mathbb{U}(d)$ which were used in Section X:

$$F^\sigma(U) = f_p^\sigma(\text{Eig}(U)) , \quad F_k^\sigma(U) = f_{p,k}^\sigma(\text{Eig}(U)) . \quad (11.26)$$

As we explained in Section X, the Weyl integration formula for class functions in $\mathbb{U}(D)$

$$\int_{\mathbb{U}(d)} d\mu(U) F(U) = \int_{\mathbb{T}^d} d\mu(\varphi) f(\varphi) , \quad (11.27)$$

enables to compute numerous quantities relevant for functions F^σ, F_k^σ solely in terms of the functions $f_p^\sigma, f_{p,k}^\sigma$ defined on \mathbb{T}^d .

In what follows it will be expedient to bound integrals with respect to the measure $d\mu(\varphi)$ by integrals with respect to the measure $d\mu_G(\mathbf{x})$. To this end we establish the following technical result.

Lemma 14. *Then for any non-negative integrable function $s : \mathbb{T}^d \rightarrow \mathbb{R}$ we have:*

$$\int_A d\mu(\varphi) s(\varphi) \leq \int_A d\mu_G(\mathbf{x}) s(\mathbf{x}) \quad \text{for all } A \subset \mathbb{T}^d \quad (11.28)$$

and

$$\int_A d\mu(\varphi) s(\varphi) \geq \left(\frac{2}{\pi}\right)^{d(d-1)} \int_A d\mu_G(\mathbf{x}) s(\mathbf{x}) \quad \text{for all } A \subset \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]^{\times d} . \quad (11.29)$$

Proof. The proof follows from comparing densities of measures $d\mu(\varphi)$ and $d\mu_G(\mathbf{x})$. We first observe $d\mu(\varphi) = \prod_{1 \leq i < j \leq d} 4 \sin^2(\frac{\varphi_i - \varphi_j}{2}) d\varphi$. The upper bound in Eq.(11.28) follows then from the estimate

$$\prod_{1 \leq i < j \leq d} 4 \sin^2(\frac{\varphi_i - \varphi_j}{2}) \leq \prod_{1 \leq i < j \leq d} (\varphi_i - \varphi_j)^2 , \quad (11.30)$$

where we used the fact that for all $x \in \mathbb{R}$ $|\sin(x)| \leq |x|$. On the other hand using the bound $\frac{2}{\pi}|x| \leq |\sin(x)|$ valid for $x \in [-\pi/2, \pi/2]$ we get that for all $\varphi_i \in [-\pi/2, \pi/2]$

$$\prod_{1 \leq i < j \leq d} 4 \sin^2(\frac{\varphi_i - \varphi_j}{2}) \leq \prod_{1 \leq i < j \leq d} \frac{2}{\pi^2} (\varphi_i - \varphi_j)^2 = \left(\frac{2}{\pi}\right)^{d(d-1)} \prod_{1 \leq i < j \leq d} (\varphi_i - \varphi_j)^2 , \quad (11.31)$$

which completes the proof of Eq.(11.29). \square

1. Lower bound on normalization constant \mathcal{N}^σ

In the sketch of the proof of Lemma 9 in Section X we argued that $\mathcal{N}^\sigma \geq \int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi)$. The following result completes the proof of Eq.(10.18).

Lemma 15. *For σ satisfying $\sigma \leq \frac{\pi}{4\sqrt{d}}$ we have*

$$\int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi) \geq \frac{1}{2} C_d \sigma^{d(d-1)} \left(\frac{2}{\pi}\right)^{d(d-1)} , \quad (11.32)$$

where $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$.

Proof. We first employ Eq.(11.29) from Lemma 14 to $A = [-\pi/2, \pi/2]^{\times d}$ and function f^σ obtaining

$$\int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi) \geq \left(\frac{2}{\pi}\right)^{d(d-1)} \int_{[-\pi/2, \pi/2]^{\times d}} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) . \quad (11.33)$$

The integral over $[-\pi/2, \pi/2]^{\times d}$ can be further decomposed as

$$\int_{[-\pi/2, \pi/2]^{\times d}} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) = \int_{\mathbb{R}^d} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) - \int_{([- \pi/2, \pi/2]^{\times d})^c} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) . \quad (11.34)$$

Both terms in the above expression can be connected to respectively: *statistical sum* of GUE ensemble, and the tail probability of the maximal eigenvalue distribution of a GUE random matrix. Concretely, employing estimates from Lemma 21 (given in Part XIE) we get that for $\sigma \leq \pi/(4\sqrt{d})$

$$\int_{\mathbb{R}^d} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) = C_d \sigma^{d(d-1)} , \quad \int_{([- \pi/2, \pi/2]^{\times d})^c} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) \leq \frac{1}{2} C_d \sigma^{d(d-1)} , \quad (11.35)$$

where $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$. Inserting the above expressions into expressions into Eq.(11.35) and Eq.(11.34) concludes the proof. \square

2. Upper bound for integral of F^σ over the complement of a ball

The following result is an effective restatement of Lemma 11 from Section X. The equivalence of both results follows from the Weyl integration formula which implies $\int_{B(I,r)} d\mu(U) F^\sigma(U) = \int_{B_\infty(0,r)} d\mu(\varphi) f_p^\sigma(\varphi)$.

Lemma 16 (Restatement of Lemma 11 from Section X). *For σ and r satisfying $\sigma \leq \frac{r}{4\sqrt{d}}$, $r \leq 2/3$ we have*

$$\int_{B_\infty(0,r)^c} d\mu(\varphi) f_p^\sigma(\varphi) \leq \frac{3}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{\epsilon^2}{\sigma^2}} , \quad (11.36)$$

where $B_\infty(0,r) = \{\varphi \in \mathbb{T}^d \mid |\varphi_i| \leq r\}$ and $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$.

Proof. From the positivity of f^σ and the definition of f_p^σ it follows that

$$\int_{B_\infty(0,r)^c} f_p^\sigma(\varphi) d\mu(\varphi) = \int_{B_\infty(0,r)^c} f^\sigma(\varphi) d\mu(\varphi) + \|f_p^\sigma - f^\sigma\|_1 . \quad (11.37)$$

Using Lemma 14 allows us to write an estimate $\int_{B(0,r)^c} f^\sigma(\varphi) d\mu(\varphi) \leq \int_{([-r,r]^{\times d})^c} f^\sigma(\mathbf{x}) d\mu_G(\mathbf{x})$, relating the integrals on the torus with integrals on \mathbb{R}^d . Next, using Lemma 17 (bounding $\|f_p^\sigma - f^\sigma\|_1$) and Lemma 21 (bounding the tail of Gaussian integral) we obtain that for $\sigma \leq 1/(6\sqrt{d})$ and $\sigma \leq r/(4\sqrt{d})$

$$\int_{B_\infty(0,r)^c} f_p^\sigma(\varphi) d\mu(\varphi) \leq \frac{1}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{\epsilon^2}{\sigma^2}} + C_d \sigma^{d(d-1)} e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} \leq \frac{3}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{\epsilon^2}{\sigma^2}} , \quad (11.38)$$

where in the last inequality we assumed $r \leq \sqrt{2}\pi$. Finally, all the assumptions made on σ and r in the above reasoning will be satisfied provided $\sigma \leq \frac{r}{4\sqrt{d}}$, $r \leq 2/3$. \square

Lemma 17 (L^1 -norm difference between f^σ and f_p^σ). *For σ satisfying $\sigma \leq \frac{1}{6\sqrt{d}}$ we have*

$$\|f^\sigma - f_p^\sigma\|_1 \leq C_d \sigma^{d(d-1)} e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} , \quad (11.39)$$

where $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$.

Proof. As f_p^σ is the periodization of f^σ we have

$$f_p^\sigma(\varphi) = f^\sigma(\varphi) + \sum_{\mathbf{k} \in \mathbb{Z}^d, \mathbf{k} \neq 0} f^\sigma(\varphi + 2\pi\mathbf{k}). \quad (11.40)$$

Combining this with the positivity of both f^σ we get

$$\|f^\sigma - f_p^\sigma\|_1 = \sum_{\mathbf{k} \in \mathbb{Z}^d, \mathbf{k} \neq 0} \int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi + 2\pi\mathbf{k}). \quad (11.41)$$

We can write it as

$$\|f^\sigma - f_p^\sigma\|_1 = \sum_{m=1}^{\infty} \sum_{\mathbf{k} \in A_m} \int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi + 2\pi\mathbf{k}), \quad (11.42)$$

where $A_m = \{\mathbf{k} : -m \leq k_i \leq m, i = 1, \dots, d; \exists_i |k_i| = m\}$ (i.e. A_m is the boundary of a regular d -cube of in \mathbb{Z}^d centered at the origin and having diameter $2m$ in L^∞ distance). The function $f_{\mathbf{k}}^\sigma(\varphi) := f^\sigma(\varphi + 2\pi\mathbf{k})$ is centered around point $-2\pi\mathbf{k}$ and it is positive. Thus we can apply Lemma 14 to get

$$\int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi + 2\pi\mathbf{k}) \leq \int_{\mathbb{T}^d} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x} + 2\pi\mathbf{k}). \quad (11.43)$$

Now, for $\mathbf{k} \in A_m$, $\mathbf{x} \in \mathbb{T}^d$ we have $|\mathbf{x} + 2\pi\mathbf{k}|_\infty \geq \pi(2m-1)$ (where we used the L^∞ norm: $|\mathbf{x}|_\infty = \max_i |x_i|$). Therefore we can bound

$$\int_{\mathbb{T}^d} d\mu(\varphi) f^\sigma(\varphi + 2\pi\mathbf{k}) \leq \int_{B_\infty(0, \pi(2\pi-1))^c} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) \quad (11.44)$$

where $B_\infty(\mathbf{y}, r) = \{\mathbf{x} \in \mathbb{R}^d \mid |x_i - y_i| \leq r, i = 1, \dots, d\}$. Next we apply tail bound from Lemma 21 which gives

$$\int_{B(0, 2\pi m - \pi)^c} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) \leq \frac{1}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}}. \quad (11.45)$$

which holds for $\sigma \leq \frac{\pi(2m-1)}{2\sqrt{d}}$ so that it is enough to assume that $\sigma \leq \frac{\pi}{2\sqrt{d}}$. Going back to Eq.(11.41) we get

$$\|f^\sigma - f_p^\sigma\|_1 \leq \frac{1}{2} C_d \sigma^{d(d-1)} \sum_{m=1}^{\infty} \sum_{\mathbf{k} \in A_m} e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}} = \frac{1}{2} C_d \sigma^{d(d-1)} \sum_{m=1}^{\infty} |A_m| e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}}. \quad (11.46)$$

Next we use an estimate for the number of elements in A_m : $|A_m| \leq 2d(2m+1)^{d-1}$ which gives (11.47)

$$\|f^\sigma - f_p^\sigma\|_1 \leq d C_d \sigma^{d(d-1)} \sum_{m=1}^{\infty} (2m+1)^{d-1} e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}} \leq d C_d \sigma^{d(d-1)} 3^{d-1} \sum_{m=1}^{\infty} (2m-1)^{d-1} e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}}, \quad (11.47)$$

where in the last inequality we used $\left(\frac{2m+1}{2m-1}\right)^{d-1} \leq 3^{d-1}$ for $m \geq 1$. In order to bound the series

$$\sum_{m=1}^{\infty} (2m-1)^{d-1} e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}}, \quad (11.48)$$

we present each summand as a product of two terms

$$\sum_{m=1}^{\infty} (2m-1)^{d-1} e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}} \leq \max_{m \geq 1} \left((2m-1)^{d-1} e^{-\frac{1}{8} \frac{\pi^2 (2m-1)^2}{\sigma^2}} \right) \sum_{m=1}^{\infty} e^{-\frac{1}{8} \frac{\pi^2 (2m-1)^2}{\sigma^2}} \quad (11.49)$$

For the first term we have

$$\max_{m \geq 1} \left((2m-1)^{d-1} e^{-\frac{1}{8} \frac{\pi^2 (2m-1)^2}{\sigma^2}} \right) \leq \max_{x \geq 0} \left(x^{d-1} e^{-\frac{1}{8} \frac{\pi^2 (x-2)^2}{\sigma^2}} \right) \leq \left(\frac{4(d-1)\sigma^2}{e\pi^2} \right)^{\frac{d-1}{2}}, \quad (11.50)$$

where we used Proposition 10 Next we consider the second term of (11.49). We have

$$\begin{aligned} \sum_{m=1}^{\infty} e^{-\frac{1}{8} \frac{\pi^2 (2m-1)^2}{\sigma^2}} &\leq 2e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} + \sum_{m=3}^{\infty} e^{-\frac{1}{8} \frac{\pi^2 (2m-2)^2}{\sigma^2}} = 2e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} + \sum_{l=2}^{\infty} e^{-\frac{1}{8} \frac{\pi^2 l^2}{\sigma^2}} \leq \\ &\leq 2e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} + \int_1^{\infty} dx e^{-\frac{1}{8} \frac{\pi^2 x^2}{\sigma^2}} dx \leq 2e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} + \frac{\sigma\sqrt{8}}{\sqrt{\pi}} e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} \leq 3e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}}, \end{aligned} \quad (11.51)$$

where to bound the integral we have used the Hoeffding bound of Proposition 9 and in the last step we assumed that $\sigma \leq \sqrt{\pi/8}$. Inserting (11.51) and (11.50) into (11.49) we get

$$\sum_{m=1}^{\infty} (2m+1)^{d-1} e^{-\frac{1}{4} \frac{\pi^2 (2m-1)^2}{\sigma^2}} \leq 3 \left(\frac{4(d-1)\sigma^2}{e\pi^2} \right)^{\frac{d-1}{2}} e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}}. \quad (11.52)$$

Coming back to (11.47) we obtain

$$\|f^\sigma - f_p^\sigma\|_1 \leq dC_d \sigma^{d(d-1)} 3^{d-1} 3 \left(\frac{4(d-1)\sigma^2}{e\pi^2} \right)^{\frac{d-1}{2}} e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}} = B_d(\sigma) C_d \sigma^{d(d-1)} e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}}, \quad (11.53)$$

where $B_d(\sigma) = d3^d \left(\frac{4(d-1)\sigma^2}{e\pi^2} \right)^{\frac{d-1}{2}}$. It is easy to see that $B_d(\sigma)$ is non-increasing function of σ for $d \geq 2$ and that $B_d(1/(6\sqrt{d})) \leq 1$. Therefore for $\sigma \leq \frac{1}{6\sqrt{d}}$ we have

$$\|f^\sigma - f_p^\sigma\|_1 \leq C_d \sigma^{d(d-1)} e^{-\frac{1}{8} \frac{\pi^2}{\sigma^2}}. \quad (11.54)$$

During the above considerations we have assumed

$$\sigma \leq \sqrt{\pi/8}, \quad \sigma \leq \frac{\pi}{2\sqrt{d}}, \quad (11.55)$$

and therefore $\sigma \leq \frac{1}{6\sqrt{d}}$ is the the strongest constraint we had to impose. \square

D. Polynomial truncation of periodized Gaussian

This part of the Appendix is devoted to technical results on approximation of $\tilde{\mathcal{F}}^\sigma$ by its polynomial truncation \tilde{F}_k^σ . As explained in Section X the crucial number of interest, the L^2 -distance $\|\tilde{\mathcal{F}}^\sigma - \tilde{\mathcal{F}}_k^\sigma\|_2$ can be upper bounded by L^2 -distance between the corresponding functions on \mathbb{T}^d (computed with respect to the measure $d\mu(\varphi)$), $\|f_{p,k}^\sigma - f_p^\sigma\|_2$. This is the language that will be used in what follows.

Lemma 18 (Restatement of Lemma 12 from Section X). *Assume that*

$$k \geq \frac{d}{\sigma}, \quad \sigma \leq \frac{1}{2}. \quad (11.56)$$

Let $f_p^\sigma, f_{p,k}^\sigma$ be functions on \mathbb{T}^d defined in Eq.(11.24). We have the following estimate on their L^2 distance

$$\|f_{p,k}^\sigma - f_p^\sigma\|_2 \leq 10 C_d \frac{e^{-\frac{1}{4}(\frac{k}{\sqrt{d}} - \sqrt{d})^2 \sigma^2}}{\sigma}, \quad (11.57)$$

where $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$.

Proof. Using the definition of the L^2 norm on \mathbb{T}^d we get

$$\|f_{p,k}^\sigma - f_p^\sigma\|_2^2 = \int_{\mathbb{T}^d} d\mu(\varphi) |f_{p,k}^\sigma(\varphi) - f_p^\sigma(\varphi)|^2 \leq \frac{2^{d(d-1)}}{(2\pi)^d d!} \int_{\mathbb{T}^d} d\varphi |f_{p,k}^\sigma(\varphi) - f_p^\sigma(\varphi)|^2, \quad (11.58)$$

where we have used definition of $d\mu(\varphi)$ (cf. Eq.(11.21)) and the inequality $\prod_{1 \leq i < j \leq d} |e^{i\varphi_i} - e^{i\varphi_j}|^2 \leq 2^{d(d-1)}$. By expanding $f_{p,k}^\sigma(\varphi) - f_p^\sigma(\varphi)$ in a trigonometric series and using orthogonality of functions $\{\exp(i\mathbf{n}\varphi)\}_{\mathbf{n} \in \mathbb{Z}^d}$ on \mathbb{T}^d (equipped with measure $d\varphi$) we obtain

$$\|f_{p,k}^\sigma - f_p^\sigma\|_2^2 \leq \frac{2^{d(d-1)}}{(2\pi)^d d!} \int_{\mathbb{T}^d} d\varphi \left| \frac{1}{(2\pi)^d} \sum_{\mathbf{n}: |\mathbf{n}|_1 > k} e^{-\frac{1}{2}\mathbf{n}^2 \sigma^2} e^{-i\mathbf{n}\varphi} \right|^2 = \frac{1}{(2\pi)^d} \frac{2^{d(d-1)}}{(2\pi)^d d!} \sum_{\mathbf{n}: |\mathbf{n}|_1 > k} e^{-\mathbf{n}^2 \sigma^2}, \quad (11.59)$$

where the summation in second and third expression above is over $\mathbf{n} \in \mathbb{Z}^d$ corresponding to trygonometric polynomials of degree exceeding k (i.e $\sum_{i=1}^d |n_i| > k$). Using the well-known bound $\mathbf{n}^2 \geq |\mathbf{n}|_1^2/d$, valid for $\mathbf{n} \in \mathbb{Z}^d$, we obtain

$$\sum_{\mathbf{n}: |\mathbf{n}|_1 > k} e^{-\mathbf{n}^2 \sigma^2} \leq \sum_{\mathbf{n}: \mathbf{n}^2 > \frac{k^2}{d}} e^{-\mathbf{n}^2 \sigma^2}. \quad (11.60)$$

The second sum in the above expression can be bounded using Lemma 19 given below. For this result it follows that for $k \geq d/\sigma$ and $\sigma \leq 1/2$ we have

$$\sum_{\mathbf{n}: \mathbf{n}^2 > \frac{k^2}{d}} e^{-\sigma^2 \mathbf{n}^2} \leq \frac{\pi^{\frac{d}{2}} \sqrt{8\pi}}{\Gamma(\frac{d}{2}) \sigma^2} e^{-\frac{1}{2}(\frac{k}{\sqrt{d}} - \sqrt{d})^2 \sigma^2}. \quad (11.61)$$

Inserting this inequality to (11.60) and using the result in (11.59) yields

$$\|f_{p,k}^\sigma - f_p^\sigma\|_2 \leq C_d A_d \frac{e^{-\frac{1}{4}(\frac{k}{\sqrt{d}} - \sqrt{d})^2 \sigma^2}}{\sigma}, \quad (11.62)$$

where $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$ and

$$A_d = \frac{1}{\prod_{k=1}^d k!} \sqrt{\frac{2^{d(d-1)} \pi^{\frac{d}{2}} \sqrt{8\pi} d!}{\Gamma(\frac{d}{2})}}. \quad (11.63)$$

In Proposition 8 we prove that for all d we have $A_d \leq 10$. Making use of this result proves Eq.(11.57). \square

We now prove the result on the upper bound of the series appearing in Eq.(11.60).

Lemma 19. For $r \geq \frac{\sqrt{d}}{\sigma}$ and $\sigma \leq 1/2$ we have

$$\sum_{\mathbf{n}: \mathbf{n}^2 > r^2} e^{-\sigma^2 \mathbf{n}^2} \leq \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \frac{2\sqrt{2\pi}}{\sigma^2} e^{-\frac{1}{2}(r - \sqrt{d})^2 \sigma^2}, \quad (11.64)$$

where the summation is over $\mathbf{n} \in \mathbb{Z}^d$ corresponding satisfying $\mathbf{n}^2 \geq r^2$.

Proof. We write

$$\sum_{\mathbf{n}^2 \geq r^2} e^{-\sigma^2 \mathbf{n}^2} = \sum_{l=0}^{\infty} \sum_{\mathbf{n} \in D_l} e^{-\sigma^2 \mathbf{n}^2} \leq \sum_{l=0}^{\infty} |D_l| e^{-\sigma^2 (r + l\sqrt{d})^2}, \quad (11.65)$$

with $|D_l|$ being the number of elements of the set D_l defined by $D_l = \{\mathbf{n} \in \mathbb{Z}^d \mid (r + l\sqrt{d})^2 \leq \mathbf{n}^2 \leq (r + l\sqrt{d} + \sqrt{d})^2\}$. To evaluate $|D_l|$ we note that

$$|D_l| \leq |B_l| \leq \text{vol}(B(r + l\sqrt{d} + \sqrt{d})) , \quad (11.66)$$

where $B_l = \{\mathbf{n} \in \mathbb{Z}^d \mid \|\mathbf{n}\|^2 \leq (r + l\sqrt{d})^2\}$, and $B(r)$ denotes the Euclidean ball in \mathbb{R}^d of radius r . Indeed, consider all the points \mathbf{n} contained in B_l . These are all points \mathbf{n} contained in the Euclidean ball of radius $r + l\sqrt{d}$. We now note, that each such point is in the middle of the unit d -dimensional cube containing only this ball. The diameter of such cube is \sqrt{d} , hence if we enlarge the radius of the ball by \sqrt{d} , the number of the points in the $(r + l\sqrt{d})$ -ball will be no smaller than the volume of the enlarged ball. Since

$$\text{vol}(B(r)) = c_d r^d \quad \text{for} \quad c_d = \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \quad (11.67)$$

we have

$$|D_l| \leq c_d (r + l\sqrt{d} + \sqrt{d})^d. \quad (11.68)$$

Therefore we obtain the following upper bound

$$\sum_{\mathbf{n}: \|\mathbf{n}\|^2 > r^2} e^{-\sigma^2 \|\mathbf{n}\|^2} \leq c_d \sum_{l=0}^{\infty} (r + l\sqrt{d} + \sqrt{d})^d e^{-\sigma^2 (r + l\sqrt{d})^2}. \quad (11.69)$$

In what follows we estimate this series by an integral. To this end we need to choose such r that the function will be nonincreasing for $x \geq -1$. We find that the function $g_r(x) = (r + x\sqrt{d} + \sqrt{d})^d e^{-\frac{1}{2}(r + x\sqrt{d})^2 \sigma^2}$ has three critical points:

$$x_0 = -1 - \frac{r}{\sqrt{d}}, \quad x_{\pm} = \frac{1}{2} \left(-1 - \frac{2r}{\sqrt{d}} \pm \sqrt{\frac{2}{\sigma^2} + 1} \right). \quad (11.70)$$

It follows that the function is nonincreasing for $x \geq x_+$, so that we need r such that $x_+ \leq -1$. We rewrite the inequality $x_+ \leq -1$ as follows:

$$r \geq \frac{\sqrt{d}}{2} \left(1 + \sqrt{\frac{2}{\sigma^2} + 1} \right) \quad (11.71)$$

Note that assuming $\sigma \leq 1/2$ we have

$$1 + \sqrt{\frac{2}{\sigma^2} + 1} \leq \frac{2}{\sigma} \quad (11.72)$$

so that it is enough to take

$$r \geq \frac{\sqrt{d}}{\sigma}. \quad (11.73)$$

Since for $g_r(-1) = r^d e^{-(r-\sqrt{d})^2 \sigma^2} > 0$ and for $x \rightarrow \infty$ it goes to zero, we obtain that $g_r(x) > 0$ for $x \geq -1$ and we can bound the sum by an integral

$$\sum_{l=0}^{\infty} (r + l\sqrt{d} + \sqrt{d})^d e^{-\sigma^2 (r + l\sqrt{d})^2} \leq \int_{-1}^{\infty} dx (r + x\sqrt{d} + \sqrt{d})^d e^{-\sigma^2 (r + x\sqrt{d})^2} = \frac{1}{\sqrt{d}} \int_{r-\sqrt{d}}^{\infty} dy (y + \sqrt{d})^d e^{-y^2 \sigma^2}. \quad (11.74)$$

Using positivity of the integrand within the integration limits, we now bound this integral as follows

$$\begin{aligned} \int_{r-\sqrt{d}}^{\infty} dy (y + \sqrt{d})^d e^{-y^2 \sigma^2} &= \int_{r-\sqrt{d}}^{\infty} dy \left((y + \sqrt{d}) e^{-\frac{y^2 \sigma^2}{d}} \right)^d \leq \\ &\leq \max_{y \geq r-\sqrt{d}} \left((y + \sqrt{d})^d e^{-\frac{y^2 \sigma^2}{2}} \right) \int_{r-\sqrt{d}}^{\infty} dy (y + \sqrt{d}) e^{-\frac{y^2 \sigma^2}{2}}. \end{aligned} \quad (11.75)$$

Let $u(y) = (y + \sqrt{d})^d e^{-\frac{y^2 \sigma^2}{2}}$. The function u has three critical points: minimum y_- and maximum y_+ given by

$$y_{\pm} = \frac{\sqrt{d}}{2} \left(-1 \pm \sqrt{\frac{4}{\sigma^2} + 1} \right), \quad y_0 = -\sqrt{d}. \quad (11.76)$$

The maximum y_+ is a global maximum, and $y_+ \geq 0$. Therefore

$$\max_{y \geq r - \sqrt{d}} u(y) \leq u(y_+) \leq y_+ + \sqrt{d} = \frac{\sqrt{d}}{2} \left(1 + \sqrt{\frac{4}{\sigma^2} + 1} \right) \leq \frac{2\sqrt{d}}{\sigma}, \quad (11.77)$$

where the last inequality holds for $\sigma \leq \frac{3}{2}$. The second term of the right-hand-side of inequality (11.75) we bound by Hoeffding-type bound from Proposition 9:

$$\int_{r - \sqrt{d}}^{\infty} (y + \sqrt{d}) e^{-\frac{y^2 \sigma^2}{2}} dy \leq \frac{\sqrt{2\pi}}{\sigma} e^{-\frac{1}{2}(r - \sqrt{d})^2 \sigma^2} \quad (11.78)$$

valid for positive lower limit, i.e. for $r \geq \sqrt{d}$. Inserting this into (11.75) we get

$$\int_{r - \sqrt{d}}^{\infty} (y + \sqrt{d})^d e^{-y^2 \sigma^2} dy \leq \frac{2\sqrt{2\pi}\sqrt{d}}{\sigma^2} e^{-\frac{1}{2}(r - \sqrt{d})^2 \sigma^2}. \quad (11.79)$$

which gives

$$\sum_{l=0}^{\infty} (r + l\sqrt{d} + \sqrt{d})^d e^{-\sigma^2(r + l\sqrt{d})^2} \leq \frac{2\sqrt{2\pi}}{\sigma^2} e^{-\frac{1}{2}(r - \sqrt{d})^2 \sigma^2}. \quad (11.80)$$

Finally, inserting this into (11.69) we get

$$\sum_{\mathbf{n}: \mathbf{n}^2 > r^2} e^{-\sigma^2 \mathbf{n}^2} \leq c_d \frac{2\sqrt{2\pi}}{\sigma^2} e^{-\frac{1}{2}(r - \sqrt{d})^2 \sigma^2}. \quad (11.81)$$

This ends the proof. \square

We will now state a technical lemma which will allow to prove Proposition 5 from Section X. The lemma sets k for which $\|f_{p,k}^{\sigma} - f_p^{\sigma}\|_2$ (actually its upper bound given in (11.57)) is smaller than both lower bound on \mathcal{N}^{σ} (given in Lemma 15) and the upper bound on tails of the periodized Gaussian function (given in Lemma 16). Since $\|\tilde{\mathcal{F}}^{\sigma} - \tilde{\mathcal{F}}_k^{\sigma}\|_2 \leq \|f_{p,k}^{\sigma} - f_p^{\sigma}\|_2$, this is enough to prove Proposition 5.

Lemma 20. *For $\sigma \leq 1/8$ and*

$$k \geq 5 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{\frac{1}{8} \frac{r^2}{d^2 \sigma^2} + \ln \frac{1}{\sigma}} \quad (11.82)$$

we have

$$10 C_d \frac{e^{-\frac{1}{4}(\frac{k}{\sqrt{d}} - \sqrt{d})^2 \sigma^2}}{\sigma} \leq \frac{1}{2} \min \left\{ \frac{3}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}}, \frac{1}{2} C_d \sigma^{d(d-1)} \left(\frac{2}{\pi} \right)^{d(d-1)} \right\} \quad (11.83)$$

where $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$.

Proof. It is enough to prove a bit stronger estimate:

$$10 C_d \frac{e^{-\frac{1}{4}(\frac{k}{\sqrt{d}} - \sqrt{d})^2 \sigma^2}}{\sigma} \leq \frac{1}{4} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}} \left(\frac{2}{\pi} \right)^{d(d-1)}. \quad (11.84)$$

We thus need to find how large should be k to ensure the above inequality.

Using $\pi \leq 4$ and $40 \times 2^{d(d-1)} \leq 2^{3d^2}$ (valid for $d \geq 2$) we get that the inequality (11.84) is implied by the following one

$$2^{3d^2} e^{-\frac{1}{4} \frac{(k-d)^2}{d} \sigma^2} \leq \sigma^{d^2} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}}. \quad (11.85)$$

Assuming now $\sigma \leq 1/8$ (so that $2^{-3d^2} \geq \sigma^{d^2}$), we get that (11.84) is implied by

$$e^{-\frac{1}{4} \frac{(k-d)^2}{d} \sigma^2} \leq \sigma^{2d^2} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}}. \quad (11.86)$$

Taking logarithm of both sides, we can rewrite this as follows

$$k \geq 2 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{2 \ln \frac{1}{\sigma} + \frac{1}{4} \frac{r^2}{d^2 \sigma^2}} + d \quad (11.87)$$

For $\sigma \leq 1/8$ we have

$$2 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{2 \ln \frac{1}{\sigma} + \frac{1}{4} \frac{r^2}{d^2 \sigma^2}} \geq d, \quad (11.88)$$

so to fulfill the inequality (11.87) (and hence (11.84)) it is enough that

$$k \geq 3 \frac{d^{\frac{3}{2}}}{\sigma} \sqrt{2 \ln \frac{1}{\sigma} + \frac{1}{4} \frac{r^2}{d^2 \sigma^2}} \quad (11.89)$$

Of course we can take a bit larger but better looking k as in (11.82). In the course of the proof we have assumed that $k \geq d/\sigma$, $\sigma \leq 1/2$ and $\sigma \leq 1/8$. These constraints are fulfilled if $\sigma \leq 1/8$ and k satisfies (11.82). This ends the proof. \square

E. Auxiliary technical results and facts

1. Estimates of integrals of Gaussian-Vandermonde on \mathbb{R}^d

In this section we shall use the following notation

$$\Delta(\mathbf{x})^2 = \prod_{1 \leq i < j \leq d} (x_i - x_j)^2. \quad (11.90)$$

Lemma 21 (Upper bounds on Gaussian integrals). *Let f^σ be a standard Gaussian function given by Eq.(10.1) and let $([-r, r]^{\times d})^c$ be the complement of $[-r, r]^{\times d}$ in \mathbb{R}^d , and $C_d = \frac{\prod_{k=1}^d k!}{(2\pi)^d d!}$. Then we have the following upper bounds for the integrals*

$$\int_{([-r, r]^{\times d})^c} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) \leq \frac{1}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{2} d \left(\frac{r}{\sigma \sqrt{d}} - 2 \right)^2}, \text{ for } \sigma \leq \frac{r}{2\sqrt{d}}, \quad (11.91)$$

$$\int_{([-r, r]^{\times d})^c} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) \leq \frac{1}{2} C_d \sigma^{d(d-1)} e^{-\frac{1}{4} \frac{r^2}{\sigma^2}}, \text{ for } \sigma \leq \frac{r}{4\sqrt{d}}. \quad (11.92)$$

Moreover,

$$\int_{\mathbb{R}^d} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) = C_d \sigma^{d(d-1)}. \quad (11.93)$$

Proof. By the definition of $\mu_G(\mathbf{x})$ (see Eq.(11.22)) we have

$$\int_{([-r, r]^{\times d})^c} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) = \frac{1}{(2\pi)^d d!} \frac{1}{(\sqrt{2\pi}\sigma)^d} \int_{([-r, r]^{\times d})^c} d\mathbf{x} \Delta(\mathbf{x})^2 e^{-\frac{1}{2} \frac{\mathbf{x}^2}{\sigma^2}} \quad (11.94)$$

Introducing the variable $\mathbf{y} = \mathbf{x}/\sigma$ and making use of the probability measure for Gaussian Unitary Ensemble reduces (11.94) to

$$\begin{aligned} \frac{1}{(2\pi)^d d!} \frac{\sigma^{d(d-1)}}{(2\pi)^{\frac{d}{2}}} \int_{([-r/\sigma, r/\sigma]^{\times d})^c} d\mathbf{y} \Delta(\mathbf{y})^2 e^{-\frac{1}{2}\mathbf{y}^2} &= C_d \sigma^{d(d-1)} \left(\frac{1}{\prod_{k=1}^d k!} \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{([-r/\sigma, r/\sigma]^{\times d})^c} d\mathbf{y} \Delta(\mathbf{y})^2 e^{-\frac{1}{2}\mathbf{y}^2} \right) = \\ &= C_d \sigma^{d(d-1)} \Pr_{A \sim \text{GUE}} \left(\|A\|_\infty \geq \frac{r}{\sigma} \right). \end{aligned} \quad (11.95)$$

Next we make use of Proposition 6

$$\Pr_{A \sim \text{GUE}} \left(\|A\|_\infty \geq \frac{r}{\sigma} \right) = \Pr_{A \sim \text{GUE}} \left(d^{-1/2} \|A\|_\infty \geq \frac{r}{\sigma\sqrt{d}} \right) \leq \frac{1}{2} e^{-\frac{1}{2}d \left(\frac{r}{\sigma\sqrt{d}} - 2 \right)^2}. \quad (11.96)$$

Combining (11.96) with (11.95) we obtain (11.91). We next note that for $\sigma \leq \frac{r}{4\sqrt{d}}$ we have

$$\frac{r}{\sigma\sqrt{d}} - 2 \geq \frac{r}{2\sigma\sqrt{d}}, \quad (11.97)$$

therefore

$$\Pr_{A \sim \text{GUE}} \left(\|A\|_\infty \geq \frac{r}{\sigma} \right) \leq \frac{1}{2} e^{-\frac{r^2}{4\sigma^2}}. \quad (11.98)$$

Inserting this estimate into (11.95) we obtain estimate (11.92). Proceeding similarly we have

$$\int_{[-r, r]^{\times d}} d\mu_G(\mathbf{x}) f^\sigma(\mathbf{x}) = C_d \sigma^{d(d-1)} \Pr_{A \sim \text{GUE}} \left(\|A\|_\infty \leq \frac{r}{\sigma} \right), \quad (11.99)$$

which, taking $r \rightarrow \infty$, proves (11.93). \square

2. Auxiliary facts

In this part we provide a number of auxiliary facts that we used in previous sections of the Appendix

Proposition 6 (Tail bounds for spectral norm of GUE matrices [33]). *Let $\Pr_{A \sim \text{GUE}}$ be the probability measure on Hermitian matrices given by the Gaussian Unitary Ensemble. For $a > 0$ we have*

$$\Pr_{A \sim \text{GUE}} \left(\|d^{-1/2} A\|_\infty \geq 2 + a \right) \leq \frac{1}{2} e^{-\frac{1}{2}da^2}. \quad (11.100)$$

The constant C_d that we introduced in Lemma 9 follows from

Proposition 7 (Mehta integral [47]). *The following integral has the analytical form*

$$\frac{1}{(2\pi)^{\frac{d}{2}}} \int_{\mathbb{R}^d} d\mathbf{x} \Delta(\mathbf{x})^2 e^{-\frac{1}{2}\mathbf{x}^2} = \prod_{k=1}^d k! \quad (11.101)$$

Bellow we provide upper bound on the constant A_d introduced in the proof of Lemma 18.

Proposition 8. *For all positive integers d we have*

$$A_d = \frac{1}{\prod_{k=1}^d k!} \sqrt{\frac{2^{d(d-1)} \pi^{\frac{d}{2}} 2\sqrt{2\pi} d!}{\Gamma(\frac{d}{2})}} \leq 10. \quad (11.102)$$

Proof. We will show that for $d \geq 6$ A_d is decreasing sequence. Then the proof follows by directly verifying that numerical value of A_d for $d = 2, 3, 4, 5, 6$ is smaller than 10. To prove monotonicity we write:

$$\frac{A_{d+1}^2}{A_d^2} = \frac{1}{(d!)^2 (d+1)^2} 2^{2d} 2\sqrt{2\pi} \frac{\Gamma(d/2)}{\Gamma(d/2 + 1/2)} \quad (11.103)$$

Now, using that $\Gamma(x)$ is increasing for $x \geq 2$, and dropping $(d+1)^2$ in denominator, as well as bounding $d!$ from below by Stirling-type inequality [48]: $k! \geq \sqrt{2\pi} \sqrt{k} \left(\frac{k}{e}\right)^k$, we get

$$\frac{A_{d+1}^2}{A_d^2} \leq \sqrt{2} \left(\frac{2e}{d}\right)^{2d} \quad (11.104)$$

The right hand side is decreasing function of d and is less than 1 for $d \geq 6$, which proves that A_d is monotonically decreasing for $d \geq 6$. \square

Next, we recall a well known bound for a tail of Gaussian integral.

Proposition 9 (Hoeffding-type bound [48]). *For positive r we have*

$$\int_r^\infty dx e^{-b^2 x^2} \leq \frac{\sqrt{\pi}}{b} e^{-b^2 r^2} \quad (11.105)$$

We also give a useful fact regarding maximum of function $x^k e^{-px^2}$

Proposition 10. *For a positive p we have*

$$\max_{x \geq 0} x^k e^{-px^2} = \left(\frac{k}{2pe}\right)^{\frac{k}{2}}. \quad (11.106)$$