Commun. Math. Phys. 289, 597–652 (2009)
Digital Object Identifier (DOI) 10.1007/s00220-009-0787-3

Communications in
Mathematical
Physics

Local Asymptotic Normality for Finite Dimensional Quantum Systems

Jonas Kahn[1], **Mădălin Guţă**[2]

[1] Département de Mathématiques, Université Paris-Sud 11, Bât 425, 91405 Orsay Cedex, France
[2] School of Mathematical Sciences, University of Nottingham, University Park, NG7 2RD, Nottingham, UK. E-mail: Madalin.Guta@nottingham.ac.uk

Received: 7 July 2008 / Accepted: 16 January 2009
Published online: 28 March 2009 – © Springer-Verlag 2009

Abstract: Previous results on local asymptotic normality (LAN) for qubits [16,19] are extended to quantum systems of arbitrary finite dimension d. LAN means that the quantum statistical model consisting of n identically prepared d-dimensional systems with joint state $\rho^{\otimes n}$ converges as $n \to \infty$ to a statistical model consisting of classical and quantum Gaussian variables with fixed and known covariance matrix, and unknown means related to the parameters of the density matrix ρ. Remarkably, the limit model splits into a product of a classical Gaussian with mean equal to the diagonal parameters, and independent harmonic oscillators prepared in thermal equilibrium states displaced by an amount proportional to the off-diagonal elements. As in the qubits case [16], LAN is the main ingredient in devising a general two step adaptive procedure for the optimal estimation of completely unknown d-dimensional quantum states. This measurement strategy shall be described in a forthcoming paper [18].

1. Introduction

Quantum statistics deals with problems of statistical inference arising in quantum mechanics. The first significant results in this area appeared in the seventies and tackled issues such as quantum Cramér-Rao bounds for unbiased estimators, optimal estimation for families of states possessing a group symmetry, estimation of Gaussian states, optimal discrimination between non-commuting states. It is impossible to list all contributions but the following references may give the flavour of these developments [7,8,27,28,47,48]. The more recent theoretical advances [2,3,6,23,24,35] are closely related to the rapid development of quantum information and quantum engineering, and are often accompanied by practical implementations [1,20,39,40].

An important topic in quantum statistics is that of optimal estimation of an unknown state using the results of measurements performed on n identically prepared quantum systems [4,5,9,13,14,25,26,30,32,45]. In the case of two dimensional systems, or qubits, the problem has been solved explicitly in the *Bayesian set-up*, in the particular case

of an invariant prior and figure of merit based on the fidelity distance between states [5]. However the method used there does not work for more general priors, loss functions, or higher dimensions. In the *pointwise approach*, Hayashi and Matsumoto [26] showed that the Holevo bound [28] for the variance of locally unbiased estimators can be achieved asymptotically, and provided a sequence of measurements with this property. Their results, building on earlier work [21,22], indicate for the first time the emergence of a Gaussian limit in the problem of optimal state estimation for qubits. The extension to d-dimensional case is analysed by Matsumoto in [33].

In [16,19] we performed a detailed analysis of this phenomenon (again for qubits), and showed that we deal with the quantum generalization of an important concept in mathematical statistics called *local asymptotic normality*. As a corollary, we devised a two steps adaptive measurement strategy for state estimation which is asymptotically optimal for a large class of loss functions and priors, and could be practically implemented using continuous-time measurements. In 'classical statistics', the idea of approximating a sequence of statistical models by a family of Gaussian distributions appeared in [46], and was fully developed by Le Cam [31] who coined the term "local asymptotic normality". Among the many applications we mention its role in asymptotic optimality theory and in proving the asymptotic normality of certain estimators such as the maximum likelihood estimator. The aim of this paper is to extend the results of [16,19] to systems of *arbitrary dimension* $d < \infty$, and thus provide the main tool for solving the *open problem* of optimal state estimation for d-dimensional quantum systems [18].

Before stating the main result of the paper we shall explain briefly the meaning of local asymptotic normality for two dimensional systems [16,19]. We are given n qubits identically prepared in an unknown state ρ. Asymptotic normality means that for large n we can encode the statistical information contained in the state $\rho^{\otimes n}$ into a Gaussian model consisting of a classical random variable with distribution $N(u, I^{-1})$, and a quantum harmonic oscillator prepared in a (Gaussian) displaced thermal state ϕ_ζ. The term *local* refers to how ρ is related to the parameters $\theta = (u, \zeta)$, as explained below.

For a more precise formulation let us parametrise the qubit states by their Bloch vectors $\rho(\vec{r}) = \frac{1}{2}(\mathbf{1} + \vec{r}\,\vec{\sigma})$, where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ are the Pauli matrices. The neighbourhood of the state ρ_0 with $\vec{r_0} = (0, 0, 2\mu - 1)$ and $1/2 < \mu < 1$, is a three-dimensional ball parametrised by the deviation $u \in \mathbb{R}$ of diagonal elements and $\zeta \in \mathbb{C}$ of the off-diagonal ones,

$$\rho_\theta = \begin{pmatrix} \mu + u & \zeta^* \\ \zeta & 1 - \mu - u \end{pmatrix}, \qquad \theta = (u, \zeta) \in \mathbb{R} \times \mathbb{C}. \tag{1.1}$$

Note that ρ_0 is to be considered fixed and known but otherwise arbitrary, and can be taken to be diagonal without any loss of generality. Consider now n identically prepared qubits whose individual states are in a neighbourhood of ρ_0 of size $1/\sqrt{n}$, so that their joint state is $\rho_\theta^n := \left[\rho_{\theta/\sqrt{n}}\right]^{\otimes n}$ for some unknown θ. We would like to understand the structure of the family (statistical experiment)

$$\mathcal{Q}_n := \{\rho_\theta^n : \|\theta\| \le C\}, \tag{1.2}$$

as a whole, more precisely what is its asymptotic behavior as $n \to \infty$?

For this we consider a quantum harmonic oscillator with position and momentum operators satisfying the commutation relations $[\mathbf{Q}, \mathbf{P}] = i\mathbf{1}$. We denote by $\{|k\rangle, k \ge 0\}$ the eigenbasis of the number operator and define the thermal equilibrium state

$$\phi = (1 - e^{-\beta}) \sum_{k=0}^{\infty} e^{-k\beta} |k\rangle\langle k|, \qquad e^{-\beta} = \frac{1-\mu}{\mu},$$

which has centered Gaussian distributions for both $\mathbf{Q}$ and $\mathbf{P}$ with variance $1/(4\mu - 2) > 1/2$. We define a family of displaced thermal equilibrium states

$$\phi^{\zeta} := D^{\zeta}(\phi) := W(\zeta/\sqrt{2\mu - 1})\, \phi\, W(\zeta/\sqrt{2\mu - 1})^*, \tag{1.3}$$

where $W(\zeta) := \exp(\zeta a^* - \bar{\zeta} a)$ is the unitary displacement operator with $\zeta \in \mathbb{C}$. Additionally we consider a classical *Gaussian shift model* consisting of the family of normal distributions $N(u, \mu(1-\mu))$ with unknown center u and fixed known variance. The classical-quantum statistical experiment to which we alluded above is defined by the family of densities

$$\mathcal{R} := \{\phi^{\theta} := \mathcal{N}(u, \mu(1-\mu)) \otimes \phi^{\zeta} : \|\theta\| \leq C\}, \tag{1.4}$$

where the unknown parameters $\theta = (u, \zeta) \in \mathbb{R} \times \mathbb{C}$ are the same as those of $\mathcal{Q}_n$.

Theorem 1.1 [16,19]. *Let $\mathcal{Q}_n$ be the quantum statistical experiment* (1.2) *and let $\mathcal{R}$ be the classical-quantum experiment* (1.4)*. Then for each n there exist quantum channels (normalized completely positive maps)*

$$T_n : M\left((\mathbb{C}^2)^{\otimes n}\right) \to L^1(\mathbb{R}) \otimes \mathcal{T}(L^2(\mathbb{R})),$$
$$S_n : L^1(\mathbb{R}) \otimes \mathcal{T}(L^2(\mathbb{R})) \to M\left((\mathbb{C}^2)^{\otimes n}\right),$$

with $\mathcal{T}(L^2(\mathbb{R}))$ the trace-class operators, such that

$$\lim_{n\to\infty} \sup_{\|\theta\| \leq C} \|\phi_\theta - T_n\left(\rho_\theta^n\right)\|_1 = 0,$$
$$\lim_{n\to\infty} \sup_{\|\theta\| \leq C} \|\rho_\theta^n - S_n\left(\phi_\theta\right)\|_1 = 0,$$

for an arbitrary constant $C > 0$. The norm of trace class operators is $\|\tau\|_1 := \mathrm{Tr}(|\tau|)$.

The theorem shows that from a statistical point of view the joint qubits states are asymptotically indistinguishable from the limit Gaussian system. At the first sight one might object that the local nature of the result prevents us from drawing any conclusions for the original model of a completely unknown state ρ. However this is not a limitation, but reflects the correct normalisation of the parameters with $n \to \infty$. Indeed as n grows we have more information about the state which can be pinned down to a region of size slightly larger than $1/\sqrt{n}$ by performing rough measurements on a small proportion of the systems. After this 'localisation' step, we can use more sophisticated techniques to better estimate the state within the local neighbourhood of the first step estimator, and it is here where we use the local asymptotic normality result. Indeed, since locally the states are uniformly close to displaced Gaussian states we can pull back the optimal (heterodyne) measurement for estimating the latter to get an asymptotically optimal measurement for the former. Based on this insight we have proposed a realistic measurement set-up for this purpose using an atom-field interaction and continuous measurements in the field [16].

This paper deals with the extension of the previous result to d-dimensional systems. Like in the two-dimensional case we parametrise the neighbourhood of a fixed (diagonal) state ρ_0 by a vector $\vec{u} \in \mathbb{R}^{d-1}$ of diagonal parameters and $d(d-1)/2$ complex parameters $\vec{\zeta} = (\zeta_{j,k} : j < k)$, one for each off-diagonal matrix element (cf. (4.2) and (4.4)). We consider the same $1/\sqrt{n}-$scaling and look at the family

$$\mathcal{Q}_n = \left\{ \left[\rho_{\theta/\sqrt{n}} \right]^{\otimes n} : \theta = (\vec{u}, \vec{\zeta}) \in \Theta_n \subset \mathbb{R}^{d-1} \otimes \mathbb{C}^{d(d-1)/2} \right\},$$

where Θ_n is a ball of local parameters whose size is allowed to grow slowly with n.

As in the 2-dimensional case, the limit model is the product of a classical statistical model depending on the parameters $\vec{u}$ and a quantum model depending on $\vec{\zeta}$. Moreover the quantum part splits into a tensor product of displaced thermal states of quantum oscillators, one for each off-diagonal matrix element $\zeta_{j,k}$ with $j < k$. Thus

$$\phi^\theta = \mathcal{N}(\vec{u}, I_{\rho_0}^{-1}) \otimes \bigotimes_{j<k} \phi_{j,k}^{\zeta_{j,k}}, \qquad \theta = (\vec{u}, \vec{\zeta}).$$

Here, I_{ρ_0} is the Fisher information matrix of the multinomial model with parameters $(\mu_1, \ldots, \mu_d)$ described in Example 3.4, and $\phi_{j,k}^{\zeta_{j,k}}$ is the displaced thermal equilibrium state defined in (4.15) with inverse temperature $\beta = \ln(\mu_j/\mu_k)$.

Theorem 4.3 is the main result of the paper and shows the convergence of $\mathcal{Q}_n$ to the Gaussian model

$$\mathcal{R}_n = \left\{ \phi^\theta : \theta \in \Theta_n \subset \mathbb{R}^{d-1} \otimes \mathbb{C}^{d(d-1)/2} \right\},$$

in the spirit of Theorem 1.1. On the technical side, the uniform convergence holds over local neighbourhoods Θ_n which are allowed to grow with n rather than being fixed balls. This is essential for constructing the two stage optimal measurement: first localise within a neighbourhood Θ_n, and then apply the optimal Gaussian measurement. The details of this construction are similar to the two dimensional case and will be given in a subsequent paper [18].

Despite the similarity to the two dimensional case, the proof of the d-dimensional result has additional features which may be responsible for the fact that the optimal estimation problem has remained unsolved until now. The proof is based on the following observations:

- the n systems space $(\mathbb{C}^d)^{\otimes n}$ decomposes into a direct sum of irreducible representations of $SU(d)$, each representation being labelled by a Young diagram λ (cf. Theorem 4.1);
- the joint state $\rho_{\theta/\sqrt{n}}^{\otimes n}$ has the block diagonal form (4.8), the block weights $\lambda \to p_\lambda^{\theta,n}$ depend only on the diagonal parameters $\vec{u}$ and are closely related to the multinomial distribution of Example 3.4. This classical statistical model converges to the $(d-1)$-dimensional Gaussian shift model $N(\vec{u}, I_{\rho_0}^{-1})$;
- there exists an isometry V_λ mapping basis vectors $|\mathbf{m}, \lambda\rangle$ of the irreducible representation $\mathcal{H}_\lambda$ *almost* into number vectors $|\mathbf{m}\rangle$ of the multimode Fock space, where $\mathbf{m} = \{m_{j,k} : j < k\}$ is the collection of eigenvalues of the number operators for all oscillators.

- given a typical λ, the conditional block-state $\rho_\lambda^{\theta,n}$ can be mapped with V_λ into a multimode state which is close (in trace norm) to the Gaussian product state $\otimes_{j<k}\phi_{j,k}^{\zeta_{j,k}}$. This can be done *uniformly* over the typical diagrams whose normalised shapes have $1/\sqrt{n}$ fluctuations around $(\mu_1, \mu_2, \ldots, \mu_d)$, and over parameters $\theta \in \Theta_n$.

The first item is the well known Weyl duality which is extensively used in quantum statistics for i.i.d. states. The probability distribution of the second point has also been analysed in the context of large deviations [30] for the estimation of the state eigenvalues. The third point shows that the basis $|\mathbf{m}, \lambda\rangle$ is almost orthogonal for indices $\mathbf{m}$ which are not too big. This basis is obtained by projecting tensors of the form $f_\mathbf{a} := f_{a(1)} \otimes \cdots \otimes f_{a(n)}$ onto a subspace of $(\mathbb{C}^d)^{\otimes n}$ which is isomorphic to $\mathcal{H}_\lambda$ (cf. Theorem 5.2). Let us place the indices $\{a(i) : i = 1 \ldots n\}$ in the boxes of the diagram λ along rows, starting from the left end of the first row, to obtain a tableau $t_\mathbf{a}$. It turns out that we only need to consider $f_\mathbf{a}$ for which $t_\mathbf{a}$ is a semistandard tableau (nondecreasing along rows, increasing along columns). Then the label $\mathbf{m} := \{m_{i,j} : j > i\}$ is the collection of integers $m_{i,j}$ equal to the number of $j's$ on the row i, and is in one to one correspondence with $\mathbf{a}$. The following is an example of such semistandard tableau

$$t_\mathbf{m} = \begin{array}{|c|c|c|c|c|c|c|c|c|c|c|c|c|}\hline 1&1&1&1&1&1&1&1&2&2&2&3&3\\\hline 2&2&2&2&2&3\\\cline{1-6} 3&3&3\\\cline{1-3}\end{array}, \qquad \text{with } m_{1,2} = 3, m_{1,3} = 2, m_{2,3} = 1.$$

The relatively large number of i's in the row i is intentional, since it turns out that the 'relevant' vectors, i.e. those carrying the states $\rho_\lambda^{\theta,n}$, have indices $m_{i,j}$ small compared with the length of the rows ($\lambda_i \approx n\mu_i$ for typical representations λ). More precisely, in Sect. 7.2 we prove the following quasi-orthogonality result which allows us to carry the block states over to the oscillator space: if $\mathbf{m} \neq \mathbf{l}$ and $|\mathbf{l}| \leq |\mathbf{m}| \leq n^\eta$, then

$$|\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda \rangle| = O(n^{(9\eta-2)|\mathbf{m}-\mathbf{l}|/12}) \xrightarrow[n\to\infty]{} 0 \qquad \text{for } \eta < 2/9.$$

The proof of the fourth point involves a detailed analysis of the state $\rho_\lambda^{\theta,n}$ through its coefficients in the basis $|\mathbf{m}, \lambda\rangle$ of $\mathcal{H}_\lambda$. When $\theta = 0$ the state is diagonal and its coefficients approach uniformly those of the multidimensional thermal state $\phi^{\vec{0}} = \otimes_{j<k}\phi_{j,k}$ as shown in Lemma 6.3. The next step is to apply $SU(d)$ rotations and obtain the states $\rho_\lambda^{\theta,n}$. In Lemmas 6.4 and 6.5 it is shown that the unitary operations $\mathrm{Ad}[U_\lambda(\zeta/\sqrt{n})]$ act on $\rho_\lambda^{0,n}$ in the same way as the displacement operator $D^{\vec{\zeta}}$ acts on the thermal state $\phi^{\vec{0}}$. A remarkable fact is that in the limit the different off-diagonal parameters 'separate' into a product of shift experiments for quantum oscillators, one for each off-diagonal index $(j < k)$. This could be guessed from the Quantum Central Limit Theorem 4.4 which is related to the restriction of our result to $\theta = 0$.

Due to the apparent intricacy of the main result, the paper is organised according to the 'onion peeling' principle. We start in Sect. 2 with general classical statistical notions which motivate our investigation in quantum statistics. In particular we explain the relevance of the Le Cam distance between statistical models as a statistically meaningful way to describe convergence. Section 3 presents the classical version of local asymptotic normality with the multinomial model as an example.

In Sect. 4 we introduce the quantum statistical model consisting of n identical quantum systems with joint state $\rho^{\theta,n}$ described by diagonal and rotation parameters. We also introduce the multimode Gaussian states appearing in the limit. With this we can formulate the main result, Theorem 4.3.

In Sect. 5 we introduce the basis $|\mathbf{m}, \lambda\rangle$ and the isometry V_λ allowing us to define the channels T_n and S_n connecting the two statistical models.

In Sect. 6 we break the proof of the main theorem into manageable lemmas, essentially by using triangle inequalities. Each lemma deals with a different aspect of the convergence and has an interest in its own.

Finally, the technical proofs are collected in Sect. 7. Notably, Subsect. 7.1 and Lemma 7.1 contain the combinatorial substance of the paper.

Our investigation relies on the theory of representations of $SU(d)$. We refer to [11, 12, 15] for proofs of standard results and more details.

As in the two-dimensional case [16], local asymptotic normality provides a two stage adaptive measurement strategy which is asymptotically optimal for both Bayesian and pointwise viewpoints, and for a large range of 'distances' on the state space [18].

Our result is derived under the assumption that the density matrix ρ_0 is strictly positive and has no degenerate eigenvalues. Although we believe that a limit experiment exists for *all* states, the limit model may have a different structure. For example in the case $d = 2$, when $\mu_1 = \mu_2 = 1/2$ the limit is a completely classical Gaussian shift experiment as it can be gleaned from the quantum Central Limit Theorem [37] and from the alternative version of local asymptotic normality based on 'weak convergence' [17]. More drastically, when the state ρ_0 is pure (but we do not have this information), the classical Gaussian shift describing the eigenvalue parameter is replaced by a Poisson experiment and the correct scaling is $1/n$. A detailed analysis of these remaining cases will be pursued in a separate publication.

Throughout, we will use the following symbols: φ, ψ for states, ϕ, ρ for density matrices, T, S, M for channels (randomisations), $\mathcal{E}, \mathcal{P}, \mathcal{Q}$ for statistical models (experiments), θ, ζ, u for parameters, $\alpha, \beta, \gamma, \delta, \epsilon$ for positive constants, λ for Young diagrams.

2. Classical and Quantum Statistical Experiments

In this section we introduce some basic notions from classical statistics with the aim of defining the Le Cam distance between statistical models and local asymptotic normality. In parallel, we shall define the quantum analogues and point out their relevance in quantum statistics. The reader may find the conceptual framework helpful in understanding the quantum version of the result.

Let X be a random variable with values in the measure space $(\mathcal{X}, \Sigma_\mathcal{X})$, and let us assume that its probability distribution P belongs to some family $\{P_\theta : \theta \in \Theta\}$ where the parameter θ is unknown. Statistical inference deals with the question of how to use the available data X in order to draw conclusions about some property of θ. We shall call the family

$$\mathcal{E} := \{P_\theta : \theta \in \Theta\}, \tag{2.1}$$

a *statistical experiment or statistical model* over $(\mathcal{X}, \Sigma_\mathcal{X})$ [31].

In quantum statistics the data is replaced by a quantum system prepared in a state φ which belongs to a family $\{\varphi_\theta : \theta \in \Theta\}$ of states over an algebra of observables. In order to make a statistical inference about θ one first has to measure the system, and then apply statistical techniques to draw conclusions from the data consisting of the measurement outcomes. An important difference with the classical case is that the experimenter has the possibility to choose the measurement set-up M, and each set-up will lead to a different classical model $\{P_\theta^{(M)} : \theta \in \Theta\}$, where $P_\theta^{(M)}$ is the distribution of outcomes when performing the measurement M on the system prepared in state φ_θ.

The guiding idea of this paper is to investigate the structure of the family of quantum states

$$\mathcal{Q} := \{\varphi_\theta : \theta \in \Theta\},$$

which will be called a *quantum statistical experiment*. We shall show that in an important asymptotic set-up, namely that of a large number of identically prepared systems, the joint state can be approximated by a multidimensional quantum Gaussian state, for *all* possible preparations of the individual systems. This will bring a drastic simplification in the problem of optimal estimation for d-dimensional quantum systems, which can then be solved in the asymptotic framework [18].

2.1. Classical and quantum randomisations. Any statistical decision (e.g. estimator, test) can be seen as data processing using a *Markov kernel*. Suppose we are given a random variable X taking values in $(\mathcal{X}, \Sigma_\mathcal{X})$ and we want to produce a 'decision' $y \in \mathcal{Y}$ based on the data X. The space $\mathcal{Y}$ may be for example the parameter space Θ in the case of estimation, or just the set $\{0, 1\}$ in the case of testing between two hypotheses. For every value $x \in \mathcal{X}$ we choose y randomly with probability distribution given by $K_x(dy)$. Assuming that $K : \mathcal{X} \times \Sigma_\mathcal{Y} \to [0, 1]$ is measurable with respect to x for all fixed $A \in \Sigma_\mathcal{Y}$, we can regard K as a map from probability distributions over $(\mathcal{X}, \Sigma_\mathcal{X})$ to probability distributions over $(\mathcal{Y}, \Sigma_\mathcal{Y})$ with

$$K(P)(A) = \int K_x(A)P(dx), \quad A \in \Sigma_\mathcal{Y}. \tag{2.2}$$

A *statistic* $S : \mathcal{X} \to \mathcal{Y}$ is a particular example of such a procedure, where K_x is simply the delta measure at $S(x)$.

Besides statistical decisions, there is another important reason why one would like to apply such treatment to the data, namely to summarize it in a more convenient and informative way for future purposes as illustrated in the following simple example. Consider n independent identically distributed random variables $X_1, \ldots, X_n$ with values in $\{0, 1\}$ and distribution $P_\theta := (1 - \theta, \theta)$ with $\theta \in \Theta := (0, 1)$. The associated statistical experiment is $\mathcal{E}_n := \{P_\theta^n : \theta \in \Theta\}$.

It is easy to see that $\bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$ is an unbiased estimator of θ and moreover it is a *sufficient statistic* for $\mathcal{E}_n$, *i.e.* the conditional distribution $P_\theta^n(\cdot|\bar{X}_n = \bar{x})$ does not depend on θ! In other words the dependence on θ of the total sample $(X_1, X_2, \ldots, X_n)$ is completely captured by the statistic $\bar{X}_n$ which can be used as such for any statistical decision problem concerning $\mathcal{E}_n$. If we denote by $\bar{P}_\theta^n$ the distribution of $\bar{X}_n$ then the experiment $\bar{\mathcal{E}}_n = \{\bar{P}_\theta^n : \theta \in \Theta\}$, is statistically equivalent to $\mathcal{E}_n$.

The Markov kernel K defined in (2.2) maps the experiment $\mathcal{E}$ of Eq. (2.1) into the experiment

$$\mathcal{F} := \{Q_\theta : \theta \in \Theta\},$$

over $(\mathcal{Y}, \Sigma_\mathcal{Y})$ with $Q_\theta = K(P_\theta)$. For mathematical convenience it is useful to represent such transformations in terms of linear maps between linear spaces.

Definition 2.1. *A positive linear map*

$$T_* : L^1(\mathcal{X}, \Sigma_\mathcal{X}, P) \to L^1(\mathcal{Y}, \Sigma_\mathcal{Y}, Q)$$

is called a stochastic operator or transition if $\|T_*(g)\|_1 = \|g\|_1$ *for every* $g \in L^1_+(\mathcal{X})$. *A positive linear map*

$$T : L^\infty(\mathcal{Y}, \Sigma_\mathcal{Y}, Q) \to L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P)$$

is called a Markov operator if $T\mathbf{1} = \mathbf{1}$, *and if for any* $f_n \downarrow 0$ *in* $L^\infty(\mathcal{Y})$ *we have* $Tf_n \downarrow 0$.

A pair (T_*, T) as above is called a dual pair if

$$\int fT(g)dP = \int T_*(f)gdQ,$$

for all $f \in L^1(\mathcal{X}, \Sigma_\mathcal{X}, P)$ and $g \in L^\infty(\mathcal{Y}, \Sigma_\mathcal{Y}, Q)$. It is a theorem that for any stochastic operator T_* there exists a unique dual Markov operator T and vice versa.

What is the relation between Markov operators and Markov kernels ? Roughly speaking, any Markov kernel defines a Markov operator when we restrict to families of dominated probability measures. Let us assume that all distributions P_θ of the experiment $\mathcal{E}$ defined in (2.1) are absolutely continuous with respect to a fixed probability distribution P, such that there exist densities $p_\theta := dP_\theta/dP : \mathcal{X} \to \mathbb{R}_+$. Such an experiment is called *dominated* and in concrete situations this condition is usually satisfied. Let $K_x(dy)$ be a Markov kernel (2.2) such that $Q_\theta = K(P_\theta)$, then we define associated Markov operator $(T(f))(x) := \int f(y)k_x(dy)$ and have

$$Q_\theta = P_\theta \circ T, \quad \forall\theta. \tag{2.3}$$

When the probability distributions of two experiments are related to each other as in (2.3), we say that $\mathcal{F}$ is a *randomisation* of $\mathcal{E}$. From the duality between T and T_* we obtain an equivalent characterization in terms of the stochastic operator $T_* : L^1(\mathcal{X}, \Sigma_\mathcal{X}, P) \to L^1(\mathcal{Y}, \Sigma_\mathcal{Y}, Q)$ such that

$$T_*(dP_\theta/dP) = dQ_\theta/dQ, \quad \forall\theta\,.$$

The concept of randomisation is weaker than that of Markov kernel transformation, but under the additional condition that $(\mathcal{Y}, \Sigma_\mathcal{Y})$ is locally compact space with countable base and Borel σ-field, it can be shown that any randomisation can be implemented by a Markov kernel [41].

What is the analogue of randomisations in the quantum case ? In the language of operator algebras $L^\infty(\mathcal{X}, \Sigma_\mathcal{X}, P)$ is a commutative von Neumann algebra and $L^1(\mathcal{X}, \Sigma_\mathcal{X}, P)$ is the space of (densities of) *normal* linear functionals on it. The stochastic operator T_* is the classical version of *quantum channel*, *i.e.* a completely positive normalized (trace-preserving) map

$$T_* : \mathcal{A}_* \to \mathcal{B}_*,$$

where $\mathcal{A}_*, \mathcal{B}_*$ are the spaces of normal states on the von Neumann algebra $\mathcal{A}$ and respectively $\mathcal{B}$. Any normal state φ on $\mathcal{A}$ has a density ρ with respect to the trace such that $\varphi(A) = \mathrm{Tr}(\rho A)$ for all $A \in \mathcal{A}$. The dual of T_* is

$$T : \mathcal{B} \to \mathcal{A},$$

which is a unital completely positive map and has the property that $T_*(\varphi)(b) = \varphi(T(b))$ for all $b \in \mathcal{B}$ and $\varphi \in \mathcal{A}_*$. We interpret such quantum channels as possible physical transformations from input to output states.

A particular class of channels is that of measurements. In this case the input is the state of a quantum system described by an algebra $\mathcal{A}$, and the output is a probability distribution over the space of outcomes $(\mathcal{X}, \Sigma_{\mathcal{X}})$. Any measurement is described by a positive linear map

$$M : L^{\infty}(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{A},$$

which is completely specified by the image of characteristic functions of measurable sets, also called *positive operator valued measure* (POVM).

The corresponding channel acting on states is a positive map $M_* : \mathcal{A}_* \to L^1(\mathcal{X}, \Sigma_{\mathcal{X}}, P)$ given by

$$M_*(\varphi)(A) = \varphi(M(A)) = \mathrm{Tr}(\rho M(A)),$$

where ρ is the density matrix of φ. By applying the channel M to the quantum statistical experiment consisting of the family of states $\mathcal{Q} = \{\varphi_\theta : \theta \in \Theta\}$ on $\mathcal{A}$ we obtain a classical statistical experiment

$$\mathcal{Q}_M := \{M_*(\varphi_\theta) : \theta \in \Theta\},$$

over the outcomes space $(\mathcal{X}, \Sigma_{\mathcal{X}})$.

As in the classical case, quantum channels can be seen as ways to compare quantum experiments. The first steps in this direction were made by Petz [34,36,38] who developed the theory of *quantum sufficiency* dealing with the problem of characterizing when a sub-algebra of observables contains the same statistical information about a family of states, as the original algebra. More generally, two experiments $\mathcal{Q} := \{\mathcal{A}, \varphi_\theta : \theta \in \Theta\}$ and $\mathcal{R} := \{\mathcal{B}, \psi_\theta : \theta \in \Theta\}$ are called *statistically equivalent* if there exist channels $T : \mathcal{A} \to \mathcal{B}$ and $S : \mathcal{B} \to \mathcal{A}$ such that

$$\psi_\theta \circ T = \varphi_\theta \quad \text{and} \quad \varphi_\theta \circ S = \psi_\theta \quad \forall \theta.$$

As consequence, for any measurement $M : L^{\infty}(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{A}$ there exists a measurement $T \circ M : L^{\infty}(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{B}$ such that the resulting classical experiments coincide $\mathcal{Q}_M = \mathcal{R}_{T\circ M}$. Thus for any statistical problem, and any procedure concerning the experiment $\mathcal{Q}$ there exists a procedure for $\mathcal{R}$ with the same risk (average error), and vice versa.

2.2. The Le Cam distance and its statistical meaning. We have seen that two experiments are statistically equivalent when they can be transformed into each other by means of quantum channels. When this cannot be done exactly, we would like to have a measure of how close the two experiments are when we allow any channel transformation. We define the *deficiency* of $\mathcal{R}$ with respect to $\mathcal{Q}$ as

$$\delta(\mathcal{R}, \mathcal{Q}) = \inf_T \sup_\theta \|\varphi_\theta - \psi_\theta \circ T\|, \tag{2.4}$$

where the infimum is taken over all channels $T : \mathcal{A} \to \mathcal{B}$. The norm distance between two states on $\mathcal{A}$ is defined as

$$\|\varphi_1 - \varphi_2\| := \sup\{|\varphi_1(a) - \varphi_2(a)| : a \in \mathcal{A}, \|a\| \leq 1\},$$

and for $\mathcal{A} = \mathcal{B}(\mathcal{H})$ it is equal to $\|\rho_1 - \rho_2\|_1 := \mathrm{Tr}(|\rho_1 - \rho_2|)$, where ρ_i is the density matrix of the state φ_i. When $\delta(\mathcal{R}, \mathcal{Q}) = 0$ we say that $\mathcal{R}$ is more informative than

$\mathcal{Q}$. Note that $\delta(\mathcal{R}, \mathcal{Q})$ is not symmetric but satisfies a triangle inequality of the form $\delta(\mathcal{R}, \mathcal{Q}) + \delta(\mathcal{Q}, \mathcal{T}) \geq \delta(\mathcal{R}, \mathcal{T})$. By symmetrizing we obtain a proper distance over the space of equivalence classes of experiments, called Le Cam's distance [31],

$$\Delta(\mathcal{Q}, \mathcal{R}) := \max\left(\delta(\mathcal{Q}, \mathcal{R}),\ \delta(\mathcal{R}, \mathcal{Q})\right). \tag{2.5}$$

What is the statistical meaning of the Le Cam distance ? We shall show that if $\delta(\mathcal{R}, \mathcal{Q}) \leq \epsilon$, then for any statistical decision problem with loss function between 0 and 1, any measurement procedure for $\mathcal{Q}$ can be matched by a measurement procedure for $\mathcal{R}$ whose risk will be at most ϵ larger than the previous one.

A decision problem is specified by a *decision space* $(\mathcal{X}, \Sigma_{\mathcal{X}})$ and a *loss function* $W_\theta : \mathcal{X} \to [0, 1]$ for each $\theta \in \Theta$. We are given a quantum system prepared in the state $\varphi_\theta \in \mathcal{A}_*$ with unknown parameter $\theta \in \Theta$ and would like to perform a measurement with outcomes in $\mathcal{X}$ such that the expected value of the loss function W_θ is small. Let

$$M : L^\infty(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{A},$$

be such a measurement, and $P_\theta^{(M)} = \varphi_\theta \circ M$, then the *risk* at θ is

$$R(M, \theta) := \int_{\mathcal{X}} W_\theta(x) P_\theta^{(M)}(dx).$$

Since the point θ is unknown one would like to obtain a small risk over all possible realizations

$$R_{max}(M) = \sup_{\theta \in \Theta} R(M, \theta).$$

The *minimax risk* is then $R_{minmax} := \inf_M R_{max}(M)$.

Coming back to the experiments $\mathcal{Q}$ and $\mathcal{R}$ we shall compare their achievable risks for a given decision problem as above. Consider the measurement $N : L^\infty(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{B}$ given by $N = T \circ M$, where $T : \mathcal{A} \to \mathcal{B}$ is the channel which achieves the infimum in (2.4). Then

$$\begin{aligned} R(N, \theta) &= \int_{\mathcal{X}} W_\theta(x) P_\theta^{(N)}(dx) = \psi_\theta(T \circ M(W_\theta)) \\ &\leq \|\psi_\theta \circ T - \varphi_\theta\| + \varphi_\theta(M(W_\theta)) \leq \delta(\mathcal{R}, \mathcal{Q}) + R(M, \theta), \end{aligned}$$

where we have used the fact that $0 \leq W_\theta \leq 1$.

Lemma 2.2. *For every achievable risk $R(M, \theta)$ for $\mathcal{Q}$ there exists a measurement $N : L^\infty(\mathcal{X}, \Sigma_{\mathcal{X}}, P) \to \mathcal{B}$ for $\mathcal{R}$ such that*

$$R(N, \theta) \leq R(M, \theta) + \delta(\mathcal{R}, \mathcal{Q}).$$

As a consequence,

$$R_{minmax}(\mathcal{R}) \leq R_{minmax}(\mathcal{Q}) + \delta(\mathcal{R}, \mathcal{Q}).$$

A similar inequality holds in the Bayesian framework where one optimises the average risk with respect to a prior distribution π over Θ: $R_\pi(M) = \int_\Theta R(M, \theta)\pi(d\theta)$.

3. Local Asymptotic Normality in Statistics

In this section we describe the notion of local asymptotic normality and its significance in statistics [31,41–43]. Suppose that we observe $X_1, \ldots, X_n$, where X_i take values in a measurable space $(\mathcal{X}, \Sigma_{\mathcal{X}})$ and are independent, identically distributed with distribution P_θ indexed by a parameter θ belonging to an open subset $\Theta \subset \mathbb{R}^m$. The full sample is a single observation from the product P_θ^n of n copies of P_θ on the sample space (Ω^n, Σ^n). Local asymptotic normality means that for large n such statistical experiments can be approximated by Gaussian experiments after a suitable reparametrisation. Let θ_0 be a fixed point and define a local parameter $u = \sqrt{n}(\theta - \theta_0)$ characterizing points in a small neighbourhood of θ_0, and rewrite P_θ^n as $P^n_{\theta_0+u/\sqrt{n}}$ seen as a distribution depending on the parameter u. Local asymptotic normality means that for large n the experiments

$$\left\{P^n_{\theta_0+u/\sqrt{n}} : u \in \mathbb{R}^m\right\} \quad \text{and} \quad \left\{N(u, I_{\theta_0}^{-1}) : u \in \mathbb{R}^m\right\},$$

have the same statistical properties when the models $\theta \mapsto P_\theta$ are sufficiently 'smooth'. The point of this result is that while the original experiment may be difficult to analyse, the limit one is a tractable *Gaussian shift experiment* in which we observe a single sample from the normal distribution with unknown mean u and fixed variance matrix $I_{\theta_0}^{-1}$. Here, $\left[I_{\theta_0}\right]_{ij} = \mathbb{E}_{\theta_0}\left[\ell_{\theta_0,i}\ell_{\theta_0,j}\right]$ is the Fisher information matrix at θ_0, with $\ell_{\theta,i} := \partial \log p_\theta / \partial\theta_i$ the score function and p_θ is the density of P_θ with respect to a reference probability distribution P.

There exist two formulations of the result depending on the notion of convergence which one uses. In this paper we only discuss the *strong* version based on convergence with respect to the Le Cam distance, and we refer to [43] for another formulation using the so-called weak convergence (convergence in distribution of finite dimensional marginals of the likelihood ratio process), and to [17] for its generalization to quantum statistical experiments.

Before formulating the theorem, we explain what sufficiently smooth means. The least restrictive condition is that p_θ is *differentiable in quadratic mean*, *i.e.* there exists a measurable function $\ell_\theta : \mathcal{X} \to \mathbb{R}$ such that

$$\lim_{u\to 0} \int \left[p_{\theta+u}^{1/2} - p_\theta^{1/2} - u^t \ell_\theta p_\theta^{1/2}\right]^2 dP = 0.$$

Note that ℓ_θ must still be interpreted as a score function since under some regularity conditions we have $\partial p_\theta^{1/2}/\partial\theta_i = \frac{1}{2}(\partial \log p_\theta/\partial\theta_i)p_\theta^{1/2}$.

Theorem 3.1. *Let $\mathcal{E} := \{P_\theta : \theta \in \Theta\}$ be a statistical experiment with $\Theta \subset \mathbb{R}^d$ and $P_\theta \ll P$ such that the map $\theta \to p_\theta$ is differentiable in quadratic mean. Define*

$$\mathcal{E}_n = \{P^n_{\theta_0+u/\sqrt{n}} : \|u\| \le C\}, \quad \mathcal{F} = \{N(u, I_{\theta_0}^{-1}) : \|u\| \le C\},$$

with I_{θ_0} the Fisher information matrix of $\mathcal{E}$ at point θ_0, and C a positive constant. Then $\Delta(\mathcal{E}_n, \mathcal{F}) \to 0$. In other words, there exist sequences of randomisations T_n and S_n such that:

$$\lim_{n\to\infty} \sup_{\|u\|\le C} \|T_n(P^n_{\theta_0+u/\sqrt{n}}) - N(u, I_{\theta_0}^{-1})\| = 0,$$

$$\lim_{n\to\infty} \sup_{\|u\|\le C} \|P^n_{\theta_0+u/\sqrt{n}} - S_n(N(u, I_{\theta_0}^{-1}))\| = 0.$$

Remark 3.2. Note that the statement of the theorem is not of Central Limit type which typically involves convergence *in distribution* to a Gaussian distribution at a *single* point θ_0. Local asymptotic normality states that the convergence is *uniform* in a $1/sqrtn$-neighbourhood of θ_0, and moreover the variance of the limit Gaussian is fixed whereas the variance obtained from the Central Limit Theorem depends on the point θ. Additionally, the randomisation transforming the data $(X_1, \ldots, X_n)$ into the Gaussian variable is the same for all $\theta = \theta_0 + u/\sqrt{n}$ and thus does not require *a priori* the knowledge of θ.

Remark 3.3. Local asymptotic normality is the basis of many important results in asymptotic optimality theory and explains the asymptotic normality of certain estimators such as the maximum likelihood estimator. The quantum version introduced in the next section plays a similar role for the case of the quantum statistical model. An asymptotically optimal estimation strategy based on local asymptotic normality was derived in [16] for two-dimensional systems.

Example 3.4. Let $P_\mu = (\mu_1, \ldots, \mu_d)$ be a probability distribution with unknown parameters $(\mu_1, \ldots, \mu_{d-1}) \in \mathbb{R}_+^{d-1}$ satisfying $\mu_i > 0$ and $\sum_{i \le d-1} \mu_i < 1$. The Fisher information at a point μ is

$$I(\mu)_{ij} = \sum_{k=1}^{d-1} \mu_k (\delta_{ik}\mu_i^{-1} \cdot \delta_{jk}\mu_j^{-1}) + (1 - \sum_{l=1}^{d-1} \mu_l)^{-1} = \delta_{ij}\mu_i^{-1} + (1 - \sum_{l=1}^{d-1} \mu_l)^{-1}, \quad (3.1)$$

and its inverse is

$$V(\mu)_{ij} := [I(\mu)^{-1}]_{ij} = \delta_{ij}\mu_i - \mu_i\mu_j. \quad (3.2)$$

The multinomial experiment P_μ^n consisting of n i.i.d. date with distribution P_μ converges locally to $\mathcal{F} := (N(u, V(\mu)) : u \in \mathbb{R}^{d-1}, \|u\| \le C)$.

The experiment $\mathcal{F}$ will appear again in Theorem 4.3, as the classical part of the limit Gaussian shift experiment.

4. Local Asymptotic Normality in Quantum Statistics

In this section we present the main result of the paper. Local asymptotic normality for d-dimensional quantum systems means roughly the following: the sequence $\mathcal{Q}_n$ of experiments consisting of joint states $\rho^{\otimes n}$ of n identical quantum systems prepared independently in the same state ρ, converges to a limit experiment $\mathcal{R}$ which is a quantum-classical Gaussian model involving displaced thermal equilibrium states of $d(d-1)/2$ oscillators and a $(d-1)$-dimensional classical Gaussian shift model. As in the classical case, the result has a local nature reflecting the $1/\sqrt{n}$ rate of convergence of state estimation. A neighbourhood of a fixed diagonal state $\rho_0 = \mathrm{Diag}(\mu_1, \ldots, \mu_d)$ is parametrised by (changes in the) diagonal parameters $\vec{u} \in \mathbb{R}^{d-1}$ and off-diagonal parameters $\vec{\zeta} \in \mathbb{C}^{d(d-1)/2}$. The latter can be implemented by small unitary rotations. The limit Gaussian model has a classical part $N(\vec{u}, V(\mu))$ with fixed known variance $V(\mu)$, and a quantum part $\otimes_{j<k} \phi_{j,k}^{\zeta_{j,k}}$ with each $\phi_{j,k}^{\zeta_{j,k}}$ being a thermal equilibrium state with $\beta_{j,k} = \ln(\mu_j/\mu_k)$, displaced in phase space by an amount proportional to $\zeta_{j,k}$.

The reason for choosing the above parametrisation is twofold. Firstly, it unveils the important separation between 'classical' and 'quantum' parameters, and the further separation among the different off-diagonal parameters. Secondly, it is very convenient for

the proof. However as we shall see in Sect. 4.5, the limit experiment can be formulated in a 'coordinate-free' way in terms of quasifree states on a CCR-algebra. Although it is not needed in the main theorem, we include this formulation linking our result to the Quantum Central Limit Theorem. We stress again that local asymptotic normality is not a consequence of the Central Limit Theorem, indeed the latter is not even an ingredient in the proof but gives an indication as to what is the limit state when all parameters are zero.

4.1. The n-tuple of d-dimensional systems. As explained in Sect. 3 for the classical case, our theory will be local in nature, so we shall be interested in a (shrinking) neighbourhood of an arbitrary but fixed faithful state with density given by the diagonal matrix

$$\rho_0 = \mathrm{Diag}(\mu_1, \mu_2, \ldots, \mu_d) \quad \text{with } \mu_1 > \mu_2 > \cdots > \mu_d > 0. \tag{4.1}$$

A sufficiently small neighbourhood of ρ_0 in the state space can be parametrised by $\theta := (\vec{u}, \vec{\zeta})$ as follows:

$$\tilde{\rho}_\theta := \begin{bmatrix} \mu_1 + u_1 & \zeta_{1,2}^* & \cdots & \zeta_{1,d}^* \\ \zeta_{1,2} & \mu_2 + u_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \zeta_{d-1,d}^* \\ \zeta_{1,d} & \cdots & \zeta_{d-1,d} & \mu_d - \sum_{i=1}^{d-1} u_i \end{bmatrix}, \qquad u_i \in \mathbb{R},\ \zeta_{j,k} \in \mathbb{C}. \tag{4.2}$$

Indeed, note that if θ is small enough then $\tilde{\rho}_\theta$ is a density matrix.

Let $\delta := \inf_{1 \le i \le d} \mu_i - \mu_{i+1}$, with $\mu_{d+1} = 0$, be the minimal separation between the eigenvalues. Throughout the paper we restrict to states satisfying (4.1), for which the minimal separation δ is strictly positive.

In the first order in $\theta/\sqrt{\delta}$, the family $\tilde{\rho}_\theta$ is obtained by first perturbing the diagonal elements of ρ_0 with $\vec{u}$ and then performing a small unitary transformation with

$$U(\vec{\zeta}) := \exp\left[i \left(\sum_{1 \le j < k \le d} \frac{\mathrm{Re}(\zeta_{j,k}) T_{j,k} + \mathrm{Im}(\zeta_{j,k}) T_{k,j}}{\sqrt{\mu_j - \mu_k}} \right) \right], \tag{4.3}$$

where $T_{j,k}$ are generators of the Lie algebra of $SU(d)$ defined in (7.2). The advantage of the latter parametrisation is that we can fully exploit the machinery of irreducible group representations. For this reason, in all subsequent computations we shall work with the 'unitary' family

$$\rho_\theta := U(\vec{\zeta}) \mathrm{Diag}\left(\mu_1 + u_1, \mu_2 + u_2, \ldots, \mu_d - \sum_{i=1}^{d-1} u_i \right) U^*(\vec{\zeta}), \quad u_i \in \mathbb{R},\ \zeta_{j,k} \in \mathbb{C}. \tag{4.4}$$

but we keep in mind the relationship with (4.2).

As in the classical case, the parameter θ will be scaled by the factor $1/\sqrt{n}$ meaning that we zoom in around ρ_0 with the rate equal to the typical estimation rate based on n samples. Let $\rho^{\theta,n} := \rho_{\theta/\sqrt{n}}^{\otimes n}$ and let $\mathcal{Q}_n$ be the sequence of statistical experiments

$$\mathcal{Q}_n := \left\{ \rho^{\theta,n} : \theta \in \Theta_n \right\}, \tag{4.5}$$

consisting of n systems, each one prepared in a state $\rho_{\theta/\sqrt{n}}$ situated in a local neighbourhood of ρ_0. The local parameter $\theta = (\vec{u}, \vec{\zeta})$ belongs to a neighbourhood Θ_n of the origin of $\mathbb{R}^{d-1} \times \mathbb{C}^{d(d-1)/2}$ which is allowed to grow slowly with n in a way that will be made precise later.

One of the principal tools in our result is the representation theory of the special unitary group $SU(d)$. Due to lack of space we shall not include any proofs and refer to [11,12,15] for details. In particular we shall be working with the well known tensor representation which will be analysed in increasing depth across the following sections.

The space $(\mathbb{C}^d)^{\otimes n}$ carries two commuting group representations: that of $SU(d)$ given by

$$\pi_n(U) : f_1 \otimes \cdots \otimes f_n \mapsto Uf_1 \otimes \cdots \otimes Uf_n, \quad U \in SU(d), \tag{4.6}$$

and that of the permutation group $S(n)$ given by

$$\tilde{\pi}_d(\tau) : f_1 \otimes \cdots \otimes f_n \mapsto f_{\tau^{-1}(1)} \otimes \cdots \otimes f_{\tau^{-1}(n)}, \quad \tau \in S(n). \tag{4.7}$$

Since the two group representations commute with each other, the representation space decomposes into a direct sum of tensor products of irreducible representations. It turns out that the irreducible representations of $SU(d)$ and $S(n)$ are indexed by *Young diagrams* with d rows for the former and n boxes for the latter. A Young diagram is defined by a tuple of ordered integers $\lambda = (\lambda_1 \geq \lambda_2 \cdots \geq \lambda_k)$ with λ_i the number of boxes on row i (see Fig. 1). As we shall see later this pictorial representation will be very useful in understanding the structure of the irreducible representations $(\mathcal{H}_\lambda, \pi_\lambda)$ of $SU(d)$.

The following theorem called *Schur-Weyl duality* shows that the only tensor products appearing in the above mentioned direct sum are those of irreducible representations indexed by the same λ, and in particular the algebras generated by $\pi_n(u)$ and respectively $\tilde{\pi}_d(\tau)$ are each other's commutant!

Theorem 4.1. *Let π_n and $\tilde{\pi}_d$ be the representations of $SU(d)$ and respectively $S(n)$ on $(\mathbb{C}^d)^{\otimes n}$. Then the representation space decomposes into a direct sum of tensor products of irreducible representations of $SU(d)$ and $S(n)$ indexed by Young diagrams with d lines and n boxes:*

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_\lambda \mathcal{H}_\lambda \otimes \mathcal{K}_\lambda,$$

$$\pi_n \equiv \bigoplus_\lambda \pi_\lambda \otimes \mathbf{1}_{\mathcal{K}_\lambda}, \quad \tilde{\pi}_d \equiv \bigoplus_\lambda \mathbf{1}_{\mathcal{H}_\lambda} \otimes \tilde{\pi}_\lambda.$$

In particular $\rho^{\theta,n} = \rho_{\theta/\sqrt{n}}^{\otimes n}$ and $\tilde{\pi}_d(\tau)$ commute for all τ. Hence we have the block diagonal form for the joint states

$$\rho^{\theta,n} = \bigoplus_\lambda p_\lambda^{\theta,n} \rho_\lambda^{\theta,n} \otimes \frac{\mathbf{1}_{\mathcal{K}_\lambda}}{M_n(\lambda)}, \tag{4.8}$$

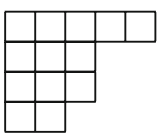

Fig. 1. Young diagram with $\lambda = (5, 3, 3, 2)$

where $M_n(\lambda)$ is the dimension of $\mathcal{K}_\lambda$, $p_\lambda^{\theta,n}$ is a probability distribution over the Young diagrams, and $\rho_\lambda^{\theta,n}$ is a density matrix on $\mathcal{H}_\lambda$. From (4.4) and the Schur-Weyl duality, we get the expression of the block states

$$\rho_\lambda^{\theta,n} = U_\lambda(\vec{\zeta}/\sqrt{n})\, \rho_\lambda^{u,0,n}\, U_\lambda(\vec{\zeta}/\sqrt{n})^*. \tag{4.9}$$

We interpret the decomposition (4.8) as follows: by doing a 'which block' measurement we obtain information about θ through the probability density $p_\lambda^{\theta,n}$. In fact it is easy to see that $p_\lambda^{\theta,n}$ does not depend on $\vec{\zeta}$, so it only gives information about the diagonal parameters $\vec{u}$. Later on we shall see that the model $p^{\theta,n}$ has the same limit as the classical multinomial model described in Example 3.4. Once this information has been obtained, one still possesses a conditional quantum state $\rho_\lambda^{\theta,n}$. It turns out that this state carries information about the rotation parameters $\vec{\zeta}$, and we shall show that the statistical model described by the conditional state converges to a 'purely quantum' Gaussian shift experiment.

4.2. Displaced thermal equilibrium states of a harmonic oscillator. The ground state of a quantum harmonic oscillator or the laser state of a monochromatic light pulse are well known examples of quantum Gaussian states. Both physical systems are described by the same algebra of observables generated by the canonical 'position' and 'momentum' observables $\mathbf{Q}$ and $\mathbf{P}$ satisfying the Heisenberg commutation relation: $\mathbf{QP} - \mathbf{PQ} = i\mathbf{1}$. These observables can be represented on the Hilbert space $L^2(\mathbb{R})$ as

$$(\mathbf{Q}f)(x) = xf(x), \quad (\mathbf{P}f)(x) = -i\frac{df}{dx}(x), \quad f \in L^2(\mathbb{R}). \tag{4.10}$$

The space $L^2(\mathbb{R})$ has a special orthonormal basis $\{|0\rangle, |1\rangle, \dots\}$ with the vector $|m\rangle$ given by

$$H_m(x)e^{-x^2/2}/(\sqrt{\pi}2^m m!)^{1/2},$$

where H_m are the Hermite polynomials [10]. These are the eigenvectors of the number operator $\mathbf{N} := \frac{1}{2}(\mathbf{Q}^2 + \mathbf{P}^2 - \mathbf{1})$ counting the number of 'excitations' of the oscillator or the number of photons in the case of the light beam, such that $\mathbf{N}|m\rangle = m|m\rangle$.

The creation and annihilation operators $\mathbf{a}^* := (\mathbf{Q} - i\mathbf{P})/\sqrt{2}$ and respectively $\mathbf{a} := (\mathbf{Q} + i\mathbf{P})/\sqrt{2}$, satisfy $[\mathbf{a}, \mathbf{a}^*] = \mathbf{1}$ and act as 'ladder' operators on the number basis:

$$\mathbf{a}|m\rangle = \sqrt{m}\,|m-1\rangle, \quad \mathbf{a}^*|m\rangle = \sqrt{m+1}\,|m+1\rangle.$$

It can be easily checked that both $\mathbf{Q}$ and $\mathbf{P}$ have Gaussian distribution with respect to the vacuum state $|0\rangle$. In fact they are 'jointly Gaussian'

$$\langle 0|\exp(iu\mathbf{Q} + iv\mathbf{P})|0\rangle = \exp\left(-\frac{1}{4}(u^2+v^2)\right).$$

We shall often use the complex form of the unitary *Weyl operators*

$$W(z) := \exp(z\mathbf{a}^* - \bar{z}\mathbf{a}) = \exp(ip_0\mathbf{Q} - iq_0\mathbf{P}), \quad z = (q_0 + ip_0)/\sqrt{2} \in \mathbb{C},$$

which satisfy the Weyl relations $W(z)^* W(z') W(z) = \exp\left(2i\mathrm{Im}(\bar{z}' z)\right) W(z')$. The coherent (vector) states $|z\rangle$ are obtained by displacing the vacuum state with Weyl operators

$$|z\rangle := W(z)\,|0\rangle = \exp(-|z|^2/2) \sum_{m=0}^{\infty} \frac{z^m}{\sqrt{m!}}\,|m\rangle\,. \tag{4.11}$$

They are Gaussian states with the same variance as the vacuum, and means $\langle z|\,\mathbf{Q}\,|z\rangle = \sqrt{2}\mathrm{Re}(z)$ and $\langle z|\,\mathbf{P}\,|z\rangle = \sqrt{2}\mathrm{Im}(z)$:

$$\langle z|\,W(z')\,|z\rangle = \exp\left(-\frac{1}{2}|z - z'|^2 + 2i\,\mathrm{Im}(\bar{z}' z)\right).$$

Besides, coherent states, an important role in our discussion will be played by the thermal equilibrium states. For every $\beta > 0$ we define the Gaussian state

$$\varphi_\beta(W(z)) = \exp\left(-\frac{|z|^2}{2\tanh(\beta/2)}\right). \tag{4.12}$$

Its density matrix consists of a mixture of k-photon states with geometrical weights

$$\phi_\beta = (1 - e^{-\beta}) \sum_{k=0}^{\infty} e^{-k\beta}\,|k\rangle\,\langle k|\,, \tag{4.13}$$

and can also be obtained by 'smearing' the coherent states with a Gaussian kernel:

$$\phi_\beta = \frac{e^\beta - 1}{\pi} \int_{\mathbb{C}} \exp\left(-(e^\beta - 1)|z|^2\right) |z\rangle\,\langle z|\,dz. \tag{4.14}$$

The thermal equilibrium states can be shifted in 'phase space' by means of displacement operations D^z which act by adjoining with unitaries $W(z)$, i.e. $D^z(\cdot) := \mathrm{Ad}[W(z)](\cdot) = W(z)^* \cdot W(z)$. The result is a Gaussian state φ_β^z with the same variance as φ_β and the same means as $|z\rangle\langle z|$:

$$\begin{aligned} \varphi_\beta^z(W(z')) &:= \exp\left(-\frac{|z|^2}{2\tanh(\beta/2)} + 2i\mathrm{Im}(\bar{z}' z)\right), \\ \phi_\beta^z &:= D^z(\phi_\beta) := W(z)^* \phi_\beta W(z). \end{aligned} \tag{4.15}$$

4.3. The multimode Fock space and the limit Gaussian shift experiment. We now consider $d(d-1)/2$ commuting harmonic oscillators, with a joint state consisting of independent Gaussian states. Let us define the *multimode Fock space*,

$$\mathcal{F} := \bigotimes_{1 \le j < k \le d} L^2(\mathbb{R}),$$

in which we identify the number basis

$$|\mathbf{m}\rangle = \bigotimes_{j<k} \left|m_{j,k}\right\rangle, \quad \mathbf{m} = \left\{m_{j,k} \in \mathbb{N} : j < k\right\}. \tag{4.16}$$

For each of the oscillators we define the thermal equilibrium state $\phi_{j,k} := \phi_{\beta_{j,k}}$, where $\beta_{j,k} = \ln(\mu_j/\mu_k)$, and $\{\mu_1, \ldots, \mu_d\}$ are the eigenvalues of the density matrix ρ_0 (cf. (4.1)). We now use the Weyl operators to displace these states by an amount proportional to the off-diagonal elements $\zeta_{j,k}$ of ρ^θ (cf. (4.2) and (4.4))

$$\phi_{j,k}^{\zeta_{j,k}} := W\left(\frac{\zeta_{j,k}}{2\sqrt{\mu_j - \mu_k}}\right)^* \phi_{j,k}\, W\left(\frac{\zeta_{j,k}}{2\sqrt{\mu_j - \mu_k}}\right).$$

Next we define the joint state $\varphi^{\vec{\zeta}}$ of the oscillators with density matrix

$$\phi^{\vec{\zeta}} = \bigotimes_{j<k} \phi_{j,k}^{\zeta_{j,k}} \in \mathcal{T}_1(\mathcal{F}), \tag{4.17}$$

where $\mathcal{T}_1(\mathcal{F})$ is the space of trace-class operators on $\mathcal{F}$.

The states $\phi^{\vec{\zeta}}$ form the quantum part of the limit Gaussian experiment. The classical part is identical to the $(d-1)$-dimensional Gaussian shift model $N(\vec{u}, V(\mu))$ of Example 3.4, where $\mu = \{\mu_1, \ldots, \mu_d\}$.

Definition 4.2. *On the algebra $L^\infty(\mathbb{R}^{d-1}) \otimes \mathcal{B}(\mathcal{F})$ we define normal state φ^θ with density*

$$\phi^\theta := \mathcal{N}(\vec{u}, V(\mu)) \otimes \phi^{\vec{\zeta}} \in L^1(\mathbb{R}^{d-1}) \otimes \mathcal{T}_1(\mathcal{F}), \tag{4.18}$$

where $\mathcal{N}(\vec{u}, V(\mu))$ is the Gaussian density of Example 3.4. The quantum-classical Gaussian experiment $\mathcal{R}$ is defined by

$$\mathcal{R} = \{\phi^\theta : \theta = (\vec{u}, \vec{\zeta}) \in \mathbb{R}^{d-1} \times \mathbb{C}^{d(d-1)/2}\}.$$

4.4. The main theorem. We are now ready to formulate the main result of the paper. In view of subsequent application to optimal state estimation, it is essential to consider (slowly) growing domains of the local parameters. For given $\beta, \gamma > 0$ we define

$$\Theta_{n,\beta,\gamma} = \left\{(\vec{\zeta}, \vec{u}) : \|\vec{\zeta}\|_\infty \le n^\beta, \|\vec{u}\|_\infty \le n^\gamma\right\}.$$

Recall that δ is the separation between the eigenvalues of ρ_0 given by Eq. (4.1). Though we use parametrisation (4.4) for density matrices ρ_θ, recall that in the first order this is approximated by $\tilde{\rho}_\theta$ defined in (4.2). In fact it can be shown that the same theorem holds for the latter parametrisation.

Theorem 4.3. *Let $\delta > 0$, let $\beta < 1/9$ and $\gamma < 1/4$. Let the quantum experiments*

$$\mathcal{Q}_n = \left\{\rho^{\theta,n} : \theta \in \Theta_{n,\beta,\gamma}\right\}, \quad \mathcal{R}_n = \left\{\phi^\theta : \theta \in \Theta_{n,\beta,\gamma}\right\},$$

where $\rho^{\theta,n} = \rho_{\theta/\sqrt{n}}^{\otimes n}$ is the state on $M\left((\mathbb{C}^d)^{\otimes n}\right)$ given by Eq. (4.4), and ϕ^θ is given by (4.18).

Then, there exist channels (completely positive, normalised maps)

$$T_n : M(\mathbb{C}^d)^{\otimes n} \to L^1(\mathbb{R}^{d-1}) \otimes \mathcal{T}_1(\mathcal{F}), \tag{4.19}$$

$$S_n : L^1(\mathbb{R}^{d-1}) \otimes \mathcal{T}_1(\mathcal{F}) \to M(\mathbb{C}^d)^{\otimes n} \tag{4.20}$$

with $\mathcal{T}_1(\mathcal{F})$ the space of trace-class operators on $\mathcal{F}$, such that

$$\sup_{\theta\in\Theta_{n,\beta,\gamma}} \left\|\phi^\theta - T_n(\rho^{\theta,n})\right\|_1 = O(n^{-\kappa}), \tag{4.21}$$

$$\sup_{\theta\in\Theta_{n,\beta,\gamma}} \left\|S_n(\phi^\theta) - \rho^{\theta,n}\right\|_1 = O(n^{-\kappa}), \tag{4.22}$$

where $\kappa > 0$ depends only on δ, β and γ. In particular we have

$$\lim_{n\to\infty} \Delta(\mathcal{Q}_n, \mathcal{R}_n) = 0,$$

where $\Delta(\cdot,\cdot)$ is the Le Cam distance defined in (2.5).

In other words, we get polynomial speed of convergence of the approximation, which is enough to build two-step evaluation strategies in the finite experiments globally asymptotically equivalent to strategies in the limit experiment [18]. The main steps of the proof are given in a sequence of lemmas in Sect. 6 assembled into Theorem 6.7. The bound (4.22) follows easily from (4.21) as shown in Sect. 6.2.

4.5. The relation between LAN and CLT. One way to think of local asymptotic normality is the following: we would like to understand the asymptotic behaviour of the collective (fluctuation) observables (4.25) with respect to a *whole neighbourhood* of the state ρ, how the limit distribution changes as we change the reference state $\rho^{\otimes n}$.

The quantum Central Limit Theorem [37] describes the asymptotic behaviour of the same observables with respect to a *fixed* state, and is one of the ingredients in the proof of a different version of LAN based on *weak convergence* [17]. However, in the case of strong convergence, which is the object of this paper, CLT does not play any role since we are interested in convergence in norm rather than in distribution, and uniformly over a range of parameters.

The purpose of the section is to derive a 'coordinate free' version of the limit Gaussian experiment using the Central Limit Theorem and the notion of symmetric logarithmic derivative. The reader interested in the proof of the main theorem may skip the following pages and continue with Sect. 5.

Let ρ be the density matrix of a fixed faithful state on $M(\mathbb{C}^d)$. To ρ we associate an algebra of canonical commutation relations carrying a Gaussian state φ. The Quantum Central Limit Theorem [37] says that φ is the limit distribution of certain multi-particle observables with respect to product states $\rho^{\otimes n}$.

Let

$$(A, B)_\rho := \mathrm{Tr}(\rho\, A \circ B), \quad \text{where } A \circ B := \frac{AB + BA}{2},$$

be a positive inner product on the real linear space of *selfadjoint operators* $M(\mathbb{C}^d)_{sa}$. We define the Hilbert space with inner product $(\cdot,\cdot)_\rho$,

$$L^2(\rho) := \{A \in M(\mathbb{C}^d)_{sa} : \mathrm{Tr}(A\rho) = 0\}.$$

Let σ be the *symplectic form* on $L^2(\rho)$,

$$\sigma(A, B) := \frac{i}{2}\mathrm{Tr}(\rho\,[A, B]).$$

The C^*-algebra of canonical commutation relations $CCR(L^2(\rho), \sigma)$ is generated by the Weyl operators $W(A)$ satisfying the relations

$$W(A)^* = W(-A), \quad W(A)W(B) = W(A+B)\exp(-i\sigma(A,B)), \quad A, B \in L^2(\rho).$$

On $CCR(L^2(\rho), \sigma)$ we define the Gaussian (quasifree) state

$$\varphi(W(A)) := \exp\left(-\frac{1}{2}\|A\|_\rho^2\right), \quad \|A\|_\rho^2 = (A,A)_\rho. \tag{4.23}$$

The state φ is regular, i.e. there exists a representation $(\pi, \mathcal{H})$ of the algebra $CCR(L^2(\rho), \sigma)$ such that the one parameter family $t \mapsto \pi(W(tA))$ is weakly continuous and φ extends to a normal state on the von Neumann algebra generated by $\pi(CCR(L^2(\rho), \sigma))$. This means that there exist selfadjoint 'field operators' $B(A)$ such that $\pi(W(tA)) = \exp(itB(A))$, and there exists a density matrix $\phi_\pi \in \mathcal{T}_1(\mathcal{H})$ such that

$$\varphi(W(A)) = \mathrm{Tr}\left(\exp(iB(A))\phi_\pi\right), \quad A \in L^2(\rho).$$

The representation $(\pi, \mathcal{H})$ can be obtained through the GNS construction, or by 'diagonalising' the CCR algebra as we shall see in a moment. From (4.23) we deduce that the distribution of $B(A)$ with respect to φ is a centred normal distribution with variance $\|A\|_\rho^2$. From the Weyl relations it follows that the fields satisfy the following canonical commutation relations $[B(A), B(C)] = 2i\sigma(A,C)\mathbf{1}$.

Consider now the tensor product $\bigotimes_{k=1}^n M(\mathbb{C}^d)$ which is generated by elements of the form

$$A^{(k)} = \mathbf{1} \otimes \cdots \otimes A \otimes \cdots \otimes \mathbf{1}, \tag{4.24}$$

with A acting on the k^{th} position of the Hilbert space tensor product $\left(\mathbb{C}^d\right)^{\otimes n}$. We are interested in the asymptotics as $n \to \infty$ of the joint distribution under the state $\rho^{\otimes n}$, of 'fluctuation' elements of the form

$$F_n(A) := \frac{1}{\sqrt{n}} \sum_{k=1}^n A^{(k)}. \tag{4.25}$$

Theorem 4.4 [Quantum CLT]. *Let $A_1, \ldots, A_s \in L^2(\rho)$. Then the following holds*:

$$\lim_{n\to\infty} \mathrm{Tr}\left(\rho^{\otimes n}\left(\prod_{l=1}^s F_n(A_l)\right)\right) = \varphi\left(\prod_{l=1}^s (B(A_l))\right),$$
$$\lim_{n\to\infty} \mathrm{Tr}\left(\rho^{\otimes n}\left(\prod_{l=1}^s \exp(iF_n(A_l))\right)\right) = \varphi\left(\prod_{l=1}^s W(A_l)\right).$$

Although the algebra $CCR(L^2(\rho), \sigma)$ may look rather abstract, its structure can be easily understood by 'diagonalising' it. Let us assume that ρ is a diagonal matrix $\rho_0 = \mathrm{Diag}(\mu_1, \ldots, \mu_d)$. The Hilbert space $L^2(\rho_0)$ decomposes as direct sum of orthogonal subspaces $\mathcal{H}_{\rho_0} \oplus \mathcal{H}_{\rho_0}^\perp$, where

$$\mathcal{H}_{\rho_0} := \mathrm{Lin}\{A : [A, \rho_0] = 0, \mathrm{Tr}(A\rho_0) = 0\}, \quad \text{and} \quad \mathcal{H}_{\rho_0}^\perp = \mathrm{Lin}\{T_{j,k}, j \neq k\}, \tag{4.26}$$

with $T_{j,k}$ the generators of the $\mathfrak{su}(d)$ algebra defined in (7.2).

The elements $W(A)$ with $A \in \mathcal{H}_{\rho_0}$ generate the center of the algebra which is isomorphic to the algebra of bounded continuous functions $C_b(\mathbb{R}^{d-1})$. Explicitly, we identify the coordinates in $\mathbb{R}^{d-1}$ with the basis $\{d_i = -\mu\mathbf{1} + E_{i,i} : i = 1, \dots d-1\}$ of $\mathcal{H}_{\rho_0}$, (see (7.2) for the definition of $E_{i,i}$). Then the covariance matrix for the basis vectors is

$$(d_i, d_j)_{\rho_0} = \mathrm{Tr}(\rho_0 d_i d_j) = \delta_{i,j}\mu_i - \mu_i\mu_j = [V(\mu)]_{i,j},$$

where V_μ is the covariance matrix (3.2). Moreover

$$t_{j,k} := T_{j,k}/\sqrt{2(\mu_j - \mu_k)}, \quad j \neq k, \tag{4.27}$$

form an *orthogonal and symplectic basis* of $\mathcal{H}_{\rho_0}^{\perp}$, i.e.

$$\sigma(t_{j,k}, t_{k,j}) = -1/2, \quad j < k, \quad \text{and } \sigma(t_{j,k}, t_{l,m}) = 0 \quad \text{for } \{j,k\} \neq \{l,m\},$$

which means that $\{t_{j,k}, t_{k,j}\}$ generate isomorphic algebras of quantum harmonic oscillator which we denote by $CCR(\mathbb{C})$. From

$$\|t_{j,k}\|^2_{\rho_0} = \mathrm{Tr}(\rho_0 t^2_{j,k}) = \frac{\mu_j + \mu_k}{2(\mu_j - \mu_k)}$$

and (4.12) we conclude that each of the oscillators is prepared independently in the thermal equilibrium state $\varphi_{j,k} = \varphi_{\beta_{j,k}}$ with $\beta_{j,k} = \ln(\mu_j/\mu_k)$.

Based on the discussion of Sects. 4.2 and 4.3 we can choose $\mathcal{H} := L^2(\mathbb{R}^{d-1}) \otimes \mathcal{F}$ and define the regular representation π of $CCR(L^2(\rho_0), \sigma)$ on this space in a straightforward way and its von Neumann completion is $L^\infty(\mathbb{R}^{d-1}) \otimes \mathcal{B}(\mathcal{F})$. The state φ decomposes as

$$\varphi \cong N(0, V_\mu) \otimes \bigotimes_{j<k} \varphi_{j,k}, \tag{4.28}$$

which is precisely the state φ^θ for $\theta = (\vec{u}, \vec{\xi}) = (\vec{0}, \vec{0})$, defined in (4.18).

Remark. The family of states φ^θ of the experiment $\mathcal{R}$ is obtained by shifting φ^0 with the help of symmetric logarithmic derivatives. Since this falls outside the scope of this paper we refer the reader to [26] for more details.

5. Explicit Form of the Channels and First Steps of the Proof

5.1. Second look at the irreducible representations of $SU(d)$. Before explaining the steps involved in the proof, let us take a closer look at the block states (4.9). Recall that we have the decomposition of Theorem 4.1 over Young diagrams with n boxes and

$$\rho^{\theta,n} = \bigoplus_\lambda \rho_\lambda^{\theta,n} \otimes \frac{\mathbf{1}_{\mathcal{K}_\lambda}}{M_n(\lambda)}.$$

Let $\{f_1, \dots, f_d\}$ be the eigenvectors of ρ_0, i.e. the standard basis vectors of $\mathbb{C}^d$. Then the eigenvectors of $\rho_0^{\otimes n} = \rho^{0,n}$ are tensor products

$$f_{\mathbf{a}} := f_{a(1)} \otimes f_{a(2)} \otimes \cdots \otimes f_{a(n)},$$

and the eigenvalues $\prod_k \lambda_{a(k)}$ do not depend on the order of the vectors in the product.

5.1.1. Projecting onto a copy of $\mathcal{H}_\lambda$. Our aim is to 'project' to an irreducible representation $\mathcal{H}_\lambda$ and obtain an explicit expression for the eigenvectors of the block components $\rho_\lambda^{\theta,n}$. Such a projection is not unique, in fact for any rank one operator $|v\rangle\langle u| \in \mathcal{B}(\mathcal{K}_\lambda)$ with $\langle u|v\rangle = 1$ we can define a (not necessarily orthogonal) projection $y = y^2$ on a copy of $\mathcal{H}_\lambda$,

$$y_\lambda(u, v) := \mathbf{1}_{\mathcal{H}_\lambda} \otimes |v\rangle\langle u| : (\mathbb{C}^d)^{\otimes n} \to \mathcal{H}_\lambda \otimes |v\rangle.$$

However the action of $y_\lambda(u, v)$ on basis vectors $f_{\mathbf{a}}$ depends on a particular identification between $(\mathbb{C}^d)^{\otimes n}$ and the direct sum in Theorem 4.1. Therefore we need a direct way of defining such a projection and the key observation is that $y_\lambda(u, v)$ is a *minimal* projection in the algebra $\mathrm{Alg}(\tilde{\pi}_d(\tau) : \tau \in S(n))$, i.e. it cannot be decomposed into a sum of non-zero projections, and vice-versa any minimal projection is of this form. The following recipe (given without proof) shows how to construct minimal projections in the $S(n)$ group algebra. We recall that the group *-algebra $\mathcal{A}(S(n))$ is the linear space spanned by the group elements endowed with a product stemming from the group product

$$a = \sum_{\tau \in S(n)} a(\tau)\tau, \quad b = \sum_{\varrho \in S(n)} b(\varrho)\varrho \implies ab = \sum_{\tau,\varrho \in S(n)} a(\tau)b(\varrho)\tau\varrho$$

$$= \sum_{\sigma \in S(n)} \left(\sum_{s \in S(n)} a(\sigma s^{-1})b(s) \right) \sigma,$$

and with adjoint $a^* = \sum_{\tau \in S(n)} a(\tau)\tau^{-1}$.

Let λ be a Young diagram with n boxes and consider the (standard) Young tableau t in which the boxes are filled with the numbers $\{1, \ldots, n\}$ in increasing order from left to right along rows, starting with the top row and down columns, as shown in the left-side tableau of Fig. 2.

Define the group algebra elements

$$P_\lambda = \sum_{\sigma \in \mathcal{R}_\lambda} \sigma, \quad Q_\lambda = \sum_{\tau \in \mathcal{C}_\lambda} \mathrm{sgn}(\tau)\tau,$$

where $\mathcal{R}_\lambda$ is the $S(n)$-subgroup of permutation leaving the rows of t invariant, and $\mathcal{C}_\lambda$ is the subgroup of permutations leaving the columns of t invariant. Note that P_λ and Q_λ are self-adjoint elements of the $S(n)$ group algebra satisfying

$$P_\lambda P_\lambda = |\mathcal{R}_\lambda| P_\lambda = (\prod_{i=1}^{d} \lambda_i!) P_\lambda, \qquad Q_\lambda Q_\lambda = |\mathcal{C}(\lambda)| Q_\lambda = (\prod_{i=1}^{d} i^{\lambda_i - \lambda_{i+1}}) Q_\lambda. \tag{5.1}$$

The *Young symmetriser* is defined as

$$Y_\lambda := Q_\lambda P_\lambda.$$

1	3	4	6
2	5		
7			

1	1	2	2	3
2	3	3		
3				

Fig. 2. Left: a standard Young tableaux. Right: a semi-standard Young tableau for $d = 3$

Theorem 5.1. *Up to a scalar normalising factor, the Young symmetriser Y_λ is minimal projection in $\mathcal{A}(S(n))$ and $y_\lambda := q_\lambda p_\lambda = \tilde{\pi}_d(Q_\lambda)\tilde{\pi}_d(P_\lambda)$ projects onto a copy of $\mathcal{H}_\lambda \subset (\mathbb{C}^d)^{\otimes n}$.*

The action of the Young symmetriser y_λ on basis vectors $f_{\mathbf{a}} \in (\mathbb{C}^d)^{\otimes n}$ follows easily from the definition of Y_λ. For each $f_{\mathbf{a}}$ we fill the boxes of λ with the indices $a(k)$ going along rows from left to right, starting with the top row and finishing with the bottom one. For example, if $\lambda = \ydiagram{2,2}$ and $f_{\mathbf{a}} = f_2 \otimes f_2 \otimes f_1 \otimes f_2 \otimes f_1$ then $t_{\mathbf{a}} = \begin{array}{|c|c|c|}\hline 2&2&1\\\hline 2&1\\\cline{1-2}\end{array}$. $S(n)$ has an obvious action on the set of tableaux by permuting the *content* of the boxes which are numbered from 1 to n in the standard way as in Fig. 2. The action of the Young symmetriser $y_\lambda = q_\lambda p_\lambda$ on $f_{\mathbf{a}}$ is deduced from the action on the tableau $t_{\mathbf{a}}$: one first symmetrises with respect to components which are in the same row, and then antisymmetrises with respect to components in the same column. For example if $\lambda = \ydiagram{2,1}$ then

$$y_\lambda(f_2 \otimes f_1 \otimes f_3) = f_2 \otimes f_1 \otimes f_3 + f_1 \otimes f_2 \otimes f_3 - f_3 \otimes f_1 \otimes f_2 - f_3 \otimes f_2 \otimes f_1.$$

5.1.2. Finding a basis in $\mathcal{H}_\lambda$. By the previous theorem the vectors $y_\lambda f_{\mathbf{a}}$ span $\mathcal{H}_\lambda$, but are not linearly independent. We show now how to select a basis (subset of linearly independent vectors spanning $\mathcal{H}_\lambda$). A *semistandard* Young tableau is a diagram filled with numbers in $\{1, \ldots, d\}$ such that the entries are non-decreasing along rows from left to right and increasing along columns from top to bottom, as in the right-side of Fig. 2.

Theorem 5.2. *The vectors $y_\lambda f_{\mathbf{a}}$ for which $t_{\mathbf{a}}$ is a semistandard Young tableau form a (non-orthogonal) basis of the irreducible representation $(\pi_\lambda, \mathcal{H}_\lambda)$.*

Since the values in the rows are nondecreasing, there is a one-to-one correspondence between Young tableaux $t_{\mathbf{a}}$ and vectors $\mathbf{m} = (m_{i,j})_{1 \le i < j \le d}$, where $m_{i,j}$ is the number of j's appearing in line i of the Young tableau $t_{\mathbf{a}}$. Note that we need only consider $m_{i,j}$ for $j > i$, as there is no j in line i if $j < i$ (the columns are increasing), and the number of i in line i is $\lambda_i - \sum_{j=i+1}^d m_{i,j}$. For example, if $t_{\mathbf{a}} = \begin{array}{|c|c|c|c|c|}\hline 1&1&2&3&3\\\hline 2&3\\\cline{1-2}3\\\cline{1-1}\end{array}$ then $\mathbf{m} = \{m_{1,2} = 1, m_{1,3} = 2, m_{2,3} = 1\}$.

By a slight abuse of notation we shall denote the corresponding vectors by $y_\lambda f_{\mathbf{m}}$ and the normalised vectors

$$|\mathbf{m}, \lambda\rangle := \mathcal{N}(\mathbf{m}, \lambda) y_\lambda f_{\mathbf{m}}, \tag{5.2}$$

where $\mathcal{N}(\mathbf{m}, \lambda) = 1/\|y_\lambda f_{\mathbf{m}}\|$. This constant is in general not easy to compute but we shall describe its asymptotic properties in Sect. 7.3.

Using (5.1) we have

$$\langle y_\lambda f_{\mathbf{a}} | y_\lambda f_{\mathbf{b}} \rangle = \langle q_\lambda p_\lambda f_{\mathbf{a}} | q_\lambda p_\lambda f_{\mathbf{b}} \rangle = \langle p_\lambda f_{\mathbf{a}} | q_\lambda^2 p_\lambda f_{\mathbf{b}} \rangle = (\prod_{i=1}^d i^{\lambda_i - \lambda_{i+1}}) \langle p_\lambda f_{\mathbf{a}} | y_\lambda f_{\mathbf{b}} \rangle. \tag{5.3}$$

In order to get further simplifications, we examine some special vector states, that we shall call by analogy with the Fock spaces *finite-dimensional coherent states*.

The first is the special vector $|\mathbf{0}, \lambda\rangle$, the *highest weight vector* of the representation $(\pi_\lambda, \mathcal{H}_\lambda)$, which later on will play the role of the *finite-dimensional vacuum*. This vector,

as we have seen, corresponds to the semi-standard Young tableau where all the entries in row i are i. An immediate consequence is that

$$p_\lambda |f_{\mathbf{0}}\rangle = (\prod_{i=1}^{d} \lambda_i !)|f_{\mathbf{0}}\rangle. \tag{5.4}$$

Moreover $\langle f_{\mathbf{0}}|q_\lambda f_{\mathbf{0}}\rangle = 1$ since any column permutation produces a vector orthogonal to $f_{\mathbf{0}}$. Thus the normalised vector is:

$$|\mathbf{0}, \lambda\rangle = \frac{1}{\prod_{i=1}^{d} \lambda_i ! \sqrt{i^{\lambda_i - \lambda_{i+1}}}} y_\lambda |f_{\mathbf{0}}\rangle. \tag{5.5}$$

The finite-dimensional coherent states are defined as $\pi_\lambda(U)|\mathbf{0}_\lambda\rangle$ for $U \in SU(d)$. From $[p_\lambda, \pi_\lambda(U)] = 0$ and (5.4), we get $p_\lambda \pi_\lambda(U)|\mathbf{0}_\lambda\rangle = (\prod_{i=1}^{d} \lambda_i !)U|\mathbf{0}_\lambda\rangle$, thus

$$\langle y_\lambda f_{\mathbf{m}}|\pi_\lambda(U)|\mathbf{0}, \lambda\rangle = \sqrt{\prod_{i=1}^{d} i^{\lambda_i - \lambda_{i+1}}} \langle p_\lambda f_{\mathbf{m}}|q_\lambda \pi_\lambda(U) f_{\mathbf{0}}\rangle. \tag{5.6}$$

The latter expression holds for any linear combination of $f_{\mathbf{m}}$ on the left-hand side, in particular $\pi_\lambda(V) f_{\mathbf{0}}$ for another unitary operator V. In Lemma 7.1, we shall examine asymptotics of (5.6) for specific sequences of unitaries U when $n \to \infty$. One of the main tools will be formula (7.6).

The following expressions of the dimensions of $\mathcal{K}_\lambda$ and $\mathcal{H}_\lambda$ are given without proof: Let $g_{l,m}$ be the *hook length* of the box (l, m), defined as one plus the number of boxes under plus the number of boxes to the right. For example the diagram $(5, 3, 3)$ has the hook lengths : $\begin{array}{|c|c|c|c|c|}\hline 7&6&5&2&1\\\hline 4&3&2\\\cline{1-3} 3&2&1\\\cline{1-3}\end{array}$.

The dimension $M_n(\lambda)$ of $\mathcal{K}_\lambda$ is

$$M_n(\lambda) = \frac{n!}{\prod_{\substack{l=1\ldots d \\ m=1\ldots\lambda_l}} g_{l,m}},$$

and can be rewritten in the following form which is more adapted to our needs:

$$M_n(\lambda) = \binom{n}{\lambda_1, \ldots, \lambda_d} \prod_{\substack{l=1\ldots d \\ k=l+1\ldots d}} \frac{\lambda_l - \lambda_k + k - l}{\lambda_l + k - l}. \tag{5.7}$$

To summarise, we have defined a non-orthonormal basis $\{|\mathbf{m}, \lambda\rangle\}$ of $\mathcal{H}_\lambda$ such that $|\mathbf{m}, \lambda\rangle$ are eigenvectors of $\rho^{\vec{0},\vec{u},n}$ for all λ, with eigenvalues:

$$\langle \mathbf{m}, \lambda|\rho^{\vec{0},\vec{u},n}|\mathbf{m}, \lambda\rangle = \prod_{i=1}^{d} (\mu_i^{\vec{u},n})^{\lambda_i} \prod_{j=i+1}^{d} \left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}}, \tag{5.8}$$

where $\mu_i^{\vec{u},n} = \mu_i + u_i/\sqrt{n}$ for $1 \le i \le (d-1)$ and $\mu_d^{\vec{u},n} = \mu_d - (\sum_i u_i)/\sqrt{n}$.

The next step is to take into account the action of the unitary $U(\vec{\zeta})$. We define the automorphism

$$\Delta^{\vec{\zeta},n} : M((\mathbb{C}^d)^{\otimes n}) \to M((\mathbb{C}^d)^{\otimes n}),$$

by

$$\tau \mapsto \Delta^{\vec{\zeta},n}(\tau) = \mathrm{Ad}[U(\vec{\zeta},n)](\tau) := U(\vec{\zeta}/\sqrt{n})^{\otimes n}\,\tau\,U^*(\vec{\zeta}/\sqrt{n})^{\otimes n}. \tag{5.9}$$

Then we have $\rho^{\vec{\zeta},\vec{u},n} = \Delta^{\vec{\zeta},n}(\rho^{\vec{0},\vec{u},n})$. By Theorem 4.1 and using the decomposition (4.8), we get the blockwise action on irreducible components

$$\Delta^{\vec{\zeta},n}(\rho^{\otimes n}) = \bigoplus_{\lambda} \Delta_{\lambda}^{\vec{\zeta},n}(\rho_\lambda) \otimes \mathbf{1}_{\mathcal{K}_\lambda},$$

where $\Delta_{\lambda}^{\vec{\zeta},n} = \mathrm{Ad}[U_\lambda(\vec{\zeta},n)]$. In particular we have

$$\rho_{\lambda}^{\vec{\zeta},\vec{u},n} = \Delta_{\lambda}^{\vec{\zeta},n}(\rho_{\lambda}^{\vec{0},\vec{u},n}). \tag{5.10}$$

With these notations, we can set about building the channels T_n.

5.2. Description of T_n. We look for channels

$$T_n : M((\mathbb{C}^d)^{\otimes n}) \to L^1(\mathbb{R}^{d-1}) \otimes \mathcal{T}_1(\mathcal{F})$$

of the form:

$$T_n : \rho^{\theta,n} \longmapsto \sum_{\lambda} p_{\lambda}^{\theta,n}\,\tau_{\lambda}^{n} \otimes \left(V_\lambda \rho_{\lambda}^{\theta,n} V_\lambda^*\right). \tag{5.11}$$

Here, V_λ is an isometry from $\mathcal{H}_\lambda$ to $\mathcal{F}$, i.e. $V_\lambda^* V_\lambda = \mathbf{1}_{\mathcal{H}_\lambda}$. On the classical side, τ_λ^n is a probability law on $\mathbb{R}^{d-1}$. We may view τ^n as a Markov kernel (2.2) from the set of diagrams λ to $\mathbb{R}^{d-1}$.

The channel T_n can be described by the following sequence of operations. We first perform a 'which block' measurement over the irreducible representations and get a result λ. Then, on the one hand, we apply a classical randomisation to λ, and on the other hand we apply a channel depending on our result λ to the conditional state ρ_λ.

The underlying ideas are the following:

1) The probability distribution $p_\lambda^{\theta,n}$ is essentially a multinomial depending *only* on $\vec{u}$, as it can be deduced from (5.8) and (5.7). As we have seen in Example 3.4, this converges (in Le Cam sense) to a classical Gaussian shift experiment. Here, in order to obtain the strong norm convergence we need to smooth the discrete distribution into a continuous one with respect to the Lebesgue measure. We choose a particular smoothing distribution that will ensure the uniform L^1 convergence to the Gaussian model (Lemma 6.1).

Definition 5.3. *Let τ_λ^n be the probability density on $\mathbb{R}^{d-1}$ defined for all λ such that $\sum \lambda_i = n$, by:*

$$\tau_\lambda^n(\mathrm{d}x) = \tau_\lambda^n(x)\mathrm{d}x = \mathrm{d}x\, n^{(d-1)/2}\chi(A_{\lambda,n}), \tag{5.12}$$

where $A_{\lambda,n} = \{x \in \mathbb{R}^{d-1} : |n^{1/2}x_i + n\mu_i - \lambda_i| \le 1/2, 1 \le i \le d-1\}$. We further denote

$$b_\lambda^{\theta,n} = p_\lambda^{\theta,n}\tau_\lambda^n,$$

depending on θ only through $\vec{u}$.

2) For the quantum part, we map the 'finite-dimensional vacuum' $|\mathbf{0}, \lambda\rangle$ to the Fock space vacuum $|\mathbf{0}\rangle$, and the basis vectors $|\mathbf{m}, \lambda\rangle$ of $\mathcal{H}_\lambda$ 'near' the basis vectors $|\mathbf{m}\rangle$ of the Fock space $\mathcal{F}$ (cf. Definitions (5.2) and respectively (4.16)). Here we need to tackle the problem that $\{|\mathbf{m}, \lambda\rangle\}$ is not an orthonormal basis but only becomes so asymptotically. The following lemma provides the isometry V_λ appearing in (5.11).

Lemma 5.4. *Let $\eta < 2/9$. Suppose that $\lambda_i - \lambda_{i+1} \geq \delta n$ for all $1 \leq i \leq d$, with the convention $\lambda_{d+1} = 0$. Then for $n > n_0(\eta, \delta, d)$ there exists an isometry $V_\lambda : \mathcal{H}_\lambda \to \mathcal{F}$ such that, $V|\mathbf{0}, \lambda\rangle = |\mathbf{0}\rangle$ and for $0 < |\mathbf{m}| \leq n^\eta$,*

$$\langle \mathbf{m}| V_\lambda = \frac{1}{\sqrt{1 + (\tilde{C}n)^{(9\eta-2)/12}/\delta^{1/3}}} \langle \mathbf{m}, \lambda| ,$$

where $\tilde{C} = \tilde{C}(\eta, d)$ is a constant. More precisely, n_0 can be taken of the form $(C(d)/\delta^2)^{1/(1-3\eta)}$.

Proof. See Sect. 7.2. The main tool is Lemma 7.3. □

For Young diagrams which do not satisfy the assumption of the previous lemma, the isometry V_λ can be defined arbitrarily. The reason is that those blocks have vanishing collective weight and can be neglected altogether (cf. Lemma 6.2).

From this operational description we conclude that T_n is a proper channel since τ^n is a Markov kernel and V_λ is an isometry. We then want to prove that $T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n})$ is close to $\phi^{\mathbf{0}}$ and that the finite-dimensional operations $\Delta_\lambda^{\vec{\zeta},n}$ have almost the same action as the displacement operators D^ζ of the Fock space, cf. (4.15). Finite-dimensional coherent states and formula (4.14) will be the stepping stone to those results.

6. Main Steps of the Proof

6.1. Why T_n does the work. We shall break (4.21) in small manageable pieces. The result and brief explanatory remarks, repeating those in the derivation, are given from (6.3) on.

We introduce first a few shorthand notations: the restriction of T_n to the block λ is

$$T_\lambda : \rho_\lambda^{\theta,n} \mapsto V_\lambda \rho_\lambda^{\theta,n} V_\lambda^*,$$

so that

$$T_n : \rho^{\theta,n} \mapsto \sum_\lambda p_\lambda^{\theta,n} \tau_\lambda^n \otimes T_\lambda(\rho_\lambda^{\theta,n}) = \sum_\lambda b_\lambda^{\theta,n} \otimes \phi_\lambda^{\theta,n}.$$

We also define $T_\lambda^* : \phi \mapsto V_\lambda^* \phi V_\lambda$. and note that $T_\lambda^* T_\lambda = \mathrm{Id}_{\mathcal{H}_\lambda}$.

We expand (5.11) as

$$\begin{aligned} T_n(\rho^{\theta,n}) &= \sum_\lambda b_\lambda^{\theta,n} \otimes \phi_\lambda^{\theta,n} \\ &= \mathcal{N}(\vec{u}, V_\mu) \otimes \phi^{\vec{\zeta}} - (\mathcal{N}(\vec{u}, V_\mu) - \sum_\lambda b_\lambda^{\theta,n}) \otimes \phi^{\vec{\zeta}} - \sum_\lambda b_\lambda^{\theta,n} \otimes \phi^{\vec{\zeta}} - \phi_\lambda^{\theta,n}. \end{aligned}$$

Proving (4.21) then amounts to proving

$$\sup_{\theta\in\Omega_{n,\epsilon}} \|(\mathcal{N}(\vec{u},V_\mu)-\sum_\lambda b_\lambda^{\theta,n})\otimes\phi^{\vec{\xi}}+\sum_\lambda b_\lambda^{\theta,n}\otimes(\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n})\|_1\le Cn^{-\epsilon/\delta}.$$

We now use the triangle inequality to upper bound this norm by a sum of "elementary" terms to be treated separately in the following sections:

$$\begin{aligned}
&\|(\mathcal{N}(\vec{u},V_\mu)-\sum_\lambda b_\lambda^{\theta,n})\otimes\phi^{\vec{\xi}}+\sum_\lambda b_\lambda^{\theta,n}\otimes(\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n})\|_1\\
&\le\|(\mathcal{N}(\vec{u},V_\mu)-\sum_\lambda b_\lambda^{\theta,n})\otimes\phi^{\vec{\xi}}\|_1+\sum_\lambda\|b_\lambda^{\theta,n}\otimes(\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n})\|_1\\
&\le\|\phi^{\vec{\xi}}\|_1\|(\mathcal{N}(\vec{u},V_\mu)-\sum_\lambda b_\lambda^{\theta,n})\|_1+\sum_\lambda\|b_\lambda^{\theta,n}\|_1\|(\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n})\|_1.
\end{aligned}$$

Since $\|\phi^{\vec{\xi}}\|_1=\|\mathcal{N}(\vec{u},V_\mu)\|_1=\|\phi_\lambda^{\theta,n}\|=1$, we have $\|\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n}\|_1\le 2$. Similarly $\sum_\lambda\|b_\lambda^{\theta,n}\|_1=1$ because $\|b_\lambda^{\theta,n}\|_1=p_\lambda^{\theta,n}$. We split the sum over λ in two parts, one for which it is expected that $\|\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n}\|_1$ is small, and the other on which the sum of all $\|b_\lambda^{\theta,n}\|_1$ is small. Specifically, define the set of *typical Young diagrams*,

$$\Lambda_{n,\alpha}:=\{\lambda:|\lambda_i-n\mu_i|\le n^\alpha,\,1\le i\le d\},\quad\text{for }\alpha>1/2,\tag{6.1}$$

then

$$\begin{aligned}
&\|T_n(\rho^{\theta,n})-\mathcal{N}(\vec{u},V_\mu)\otimes\phi^{\vec{\xi}}\|\\
&\le\|\mathcal{N}(\vec{u},V_\mu)-\sum_\lambda b_\lambda^{\theta,n}\|_1+\sup_{\lambda\in\Lambda_{n,\alpha}}\|\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n}\|_1+2\sum_{\lambda\notin\Lambda_{n,\alpha}}\|b_\lambda^{\theta,n}\|_1.
\end{aligned}\tag{6.2}$$

The first term corresponds to the convergence of the classical experiment in the Le Cam sense. If the second term is small, then on $\Lambda_{n,\alpha}$, the (purely quantum) family $\rho_\lambda^{\theta,n}$ is near the family $\phi^{\vec{\xi}}$. The last term is small due to the concentration of $p_\lambda^{\theta,n}$ around the representations with shape $\lambda_i=n\mu_i$. In other words, the only representations that matter are those in $\Lambda_{n,\alpha}$.

The hardest term to dominate (notice that the two others are classical) is the second. We transform it until we reach tractable fragments:

$$\begin{aligned}
&\|\phi^{\vec{\xi}}-\phi_\lambda^{\theta,n}\|_1=\|\phi^{\vec{\xi}}-T_\lambda(\rho_\lambda^{\theta,n})\|_1\\
&=\|D^{\vec{\xi}}(\phi^{\vec{0}})-[T_\lambda\Delta_\lambda^{\vec{\xi},n}T_\lambda^*](T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n}))\|_1\\
&=\|D^{\vec{\xi}}(\phi^{\vec{0}})-D^{\vec{\xi}}(T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n}))+D^{\vec{\xi}}(T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n}))-[T_\lambda\Delta_\lambda^{\vec{\xi},n}T_\lambda^*](T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n}))\|_1\\
&\le\|D^{\vec{\xi}}(\phi^{\vec{0}})-D^{\vec{\xi}}(T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n}))\|_1+\|[D^{\vec{\xi}}-T_\lambda\Delta_\lambda^{\vec{\xi},n}T_\lambda^*](T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n})-\phi^{\vec{0}})\|_1\\
&\quad+\|[D^{\vec{\xi}}-T_\lambda\Delta_\lambda^{\vec{\xi},n}T_\lambda^*](\phi^{\vec{0}})\|_1\\
&\le 3\|T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n})-\phi^{\vec{0}}\|_1+\|[D^{\vec{\xi}}-T_\lambda\Delta_\lambda^{\vec{\xi},n}T_\lambda^*](\phi^{\vec{0}})\|_1,
\end{aligned}$$

where in the last inequality we have used the fact that the displacement operators are isometries.

Note that the first term does not depend on $\vec{\zeta}$ and the second term is small if the displacement operators $\Delta_\lambda^{\vec{\zeta},n}$ and $D^{\vec{\zeta}}$ have 'similar action' on an appropriate domain. Using the integral formula (4.14) for gaussian states ϕ_β and the fact that $\phi^{\vec{0}}$ is a tensor product of such states (cf. (4.18)), we bound the second term by

$$\|[D^{\vec{\zeta}} - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^*](\phi^{\vec{0}})\|_1 \leq \int_{\mathbb{C}^{d(d-1)/2}} f(\vec{z}) \|[D^{\vec{\zeta}} - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^*](|\vec{z}\rangle\langle\vec{z}|)\|_1 d\vec{z},$$

where

$$f(\vec{z}) = \prod_{i<j} \frac{\mu_i - \mu_j}{\pi \mu_j} \exp\left(-\frac{\mu_i - \mu_j}{\mu_j}|z_{i,j}|^2\right),$$

and $|\vec{z}\rangle\langle\vec{z}| = D^{\vec{z}}(|\mathbf{0}\rangle\langle\mathbf{0}|)$ is the multimode coherent state, so

$$[D^{\vec{\zeta}} - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^*](|\vec{z}\rangle\langle\vec{z}|) = [D^{\vec{\zeta}} D^{\vec{z}} - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^* D^{\vec{z}}](|\mathbf{0}\rangle\langle\mathbf{0}|).$$

Now, f is a probability density, and the norm in the integrand is dominated by two. By splitting the integral we obtain

$$\begin{aligned}\|[D^{\vec{\zeta}} - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^*](\phi^{\vec{0}})\|_1 \leq 2 \int_{\|\vec{z}\|>n^\beta} f(\vec{z}) \mathrm{d}\vec{z} + \sup_{\|\vec{z}\|\leq n^\beta} \|[D^{\vec{\zeta}} D^{\vec{z}} \\ - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^* D^{\vec{z}}](|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1.\end{aligned}$$

By adding and subtracting additional terms

$$\begin{aligned}D^{\vec{\zeta}} D^{\vec{z}} - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^* D^{\vec{z}} = D^{\vec{\zeta}+\vec{z}} - T_\lambda \Delta_\lambda^{\vec{\zeta}+\vec{z},n} T_\lambda^* \\ + T_\lambda \Delta_\lambda^{\vec{\zeta}+\vec{z},n} T_\lambda^* - T_\lambda \Delta_\lambda^{\vec{\zeta},n} \Delta_\lambda^{\vec{z},n} T_\lambda^* \\ + T_\lambda \Delta_\lambda^{\vec{\zeta},n} \Delta_\lambda^{\vec{z},n} T_\lambda^* - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^* D^{\vec{z}},\end{aligned}$$

we deduce that

$$\begin{aligned}\|[D^{\vec{\zeta}} - T_\lambda \Delta_\lambda^{\vec{\zeta},n} T_\lambda^*](|\vec{z}\rangle\langle\vec{z}|)\|_1 \leq \|[D^{\vec{\zeta}+\vec{z}} - T_\lambda \Delta_\lambda^{\vec{\zeta}+\vec{z},n} T_\lambda^*](|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1 \\ + \|[\Delta_\lambda^{\vec{\zeta}+\vec{z},n} - \Delta_\lambda^{\vec{\zeta},n} \Delta_\lambda^{\vec{z},n}](|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|)\|_{\mathbf{1}} \\ + \|[\Delta_\lambda^{\vec{z},n} T_\lambda^* - T_\lambda^* D^{\vec{z}}](|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1,\end{aligned}$$

where the last two terms on the right side have been simplified using properties of T_λ, T_λ^*, $\Delta_\lambda^{\vec{\zeta},n}$. Notice that the first and third norms are essentially the same and the three terms are small if the action of $\Delta_\lambda^{\vec{\zeta}}$ is mapped into that of the displacement operators $D^{\vec{\zeta}}$.

Putting all this together, our 'expanded' form for (4.21) is

$$\sup_{\theta\in\Omega_{n,\beta,\gamma}} \|T_n(\rho^{\theta,n}) - \phi^{\vec{\zeta}} \otimes \mathcal{N}(\vec{u}, V_\mu)\|_1 \tag{6.3}$$

$$\leq \sup_{\theta\in\Omega_{n,\beta,\gamma}} \|\mathcal{N}(\vec{u}, V_\mu) - \sum_\lambda b_\lambda^{\theta,n}\|_1 \tag{6.4}$$

$$+2 \sup_{\theta\in\Omega_{n,\beta,\gamma}} \sum_{\lambda\notin\Lambda_{n,\alpha}} \|b_\lambda^{\theta,n}\|_1 \tag{6.5}$$

$$+3 \sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|\phi^{\vec{0}} - T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n})\|_1 \tag{6.6}$$

$$+ \sup_{\|\vec{z}\|\leq n^\beta} \sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|[D^{\vec{\zeta}+\vec{z}} - T_\lambda \Delta_\lambda^{\vec{\zeta}+\vec{z},n} T_\lambda^*](|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1 \tag{6.7}$$

$$+ \sup_{\|\vec{z}\|\leq n^\beta} \sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|[D^{\vec{z}} - T_\lambda \Delta_\lambda^{\vec{z},n} T_\lambda^*](|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1 \tag{6.8}$$

$$+ \sup_{\|\vec{z}\|\leq n^\beta} \sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|[\Delta_\lambda^{\vec{\zeta}+\vec{z},n} - \Delta_\lambda^{\vec{\zeta},n}\Delta_\lambda^{\vec{z},n}](|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|)\|_1 \tag{6.9}$$

$$+2\int_{\|\vec{z}\|\geq n^\beta} f(\vec{z})\mathrm{d}\vec{z}. \tag{6.10}$$

The last Gaussian tail term is less than $C\exp(-\delta n^{2\beta})$, where C depends only on the dimension d. Under the hypothesis $n^{2\beta} > 2/\delta$, this can be bounded again by $O(n^{-2\beta})$.

The following lemmas provide upper bounds for each of the terms. Before each lemma we remind the reader what is the significance of the bound. The proofs are gathered in Sect. 7.

The classical part of the channel is a Markov kernel τ (see Definition 5.3) mapping the 'which block' distribution $p_\lambda^{\theta,n}$ into the density $b_\lambda^{\theta,n}$ on $\mathbb{R}^{d-1}$ which approaches uniformly the gaussian shift experiment (6.4). Recall that $b_\lambda^{\theta,n}$ depends only on $\vec{u}$ and not on $\vec{\zeta}$, so that we have the same parameter set for the two classical experiments.

Lemma 6.1. *With the above definitions, for any ϵ, we have*

$$\sup_{\theta\in\Omega_{n,\beta,\gamma}} \|\mathcal{N}(\vec{u}, V_\mu) - \sum_\lambda b_\lambda^{\theta,n}\|_1 = O(n^{-1/4+\epsilon}/\delta, n^{-1/2+\gamma}/\delta).$$

The next lemma deals with (6.5) by showing concentration around Young diagrams λ in the 'typical subset' (6.1). This allows we to restrict to this set of diagrams in further estimates.

Lemma 6.2. *Let $\alpha - \gamma - 1/2 > 0$. Then, with the above definitions we have*

$$\sup_{\theta\in\Omega_{n,\beta,\gamma}} \sum_{\lambda\notin\Lambda_{n,\alpha}} \|b_\lambda^{\theta,n}\|_1 = O\left(n^{d^2}\exp(-n^{2\alpha-1}/2)\right),$$

with the $O(\cdot)$ term converging to zero.

The term (6.6) shows that when the rotation parameter is zero, the block states $\rho_\lambda^{\vec{0},\vec{u},n}$ are essentially thermal equilibrium states, as one would expect from the quantum Central Limit Theorem 4.4. However the convergence here is in norm rather than in distribution, and uniform over the various parameters.

Lemma 6.3. *Let* $0 < \eta < 2/9$. *With the above definitions, we have*

$$\sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|\phi^{\vec{0}} - T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n})\|_1 = O(n^{-1/2+\gamma+\eta}/\delta,\ n^{(9\eta-2)/24}/\delta^{1/6}, \exp(-\delta n^\eta)).$$

The terms (6.7) and (6.8) show that the 'finite dimensional coherent states' obtained by performing small rotations on the 'finite-dimensional vacuum' are uniformly close to their infinite dimensional counterparts, thus justifying the coherent state terminology.

Lemma 6.4. *Let* $\epsilon > 0$ *be such that* $2\beta + \epsilon \leq \eta < 2/9$. *Then,*

$$\sup_{\|\vec{z}\|\leq n^\beta} \sup_{\|\vec{\xi}\|\leq n^{-1/2+2\beta}/\delta} \sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|[D^{\vec{\zeta}+\vec{z}} - T_\lambda \Delta_\lambda^{\vec{\zeta}+\vec{z},\vec{\xi},n} T_\lambda^*](|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1 = R(n)$$

with

$$R(n)^2 = O(n^{(9\eta-2)/12}\delta^{-1/3}, n^{-1+2\beta+\eta}\delta^{-1}, n^{-1/2+3\beta+2\epsilon}\delta^{-3/2}, n^{-1+\alpha+2\beta}\delta^{-1}, n^{-1+\alpha+\eta}\delta^{-1}, n^{-1+3\eta}\delta^{-1}, n^{-\beta}). \tag{6.11}$$

For estimating the terms (6.7, 6.8), *the case when* $\vec{\xi} = \vec{0}$ *is sufficient. This more general form is useful for the proof of Lemma* 6.5. *The unitary operation is defined as* $\Delta_\lambda^{\vec{\zeta},\xi,n} := \mathrm{Ad}[U_\lambda(\vec{\zeta},\xi,n)]$ *with* $U(\vec{\zeta},\xi,n))$ *the general* $SU(d)$ *element of* (7.1).

Finally (6.9) shows that the 'finite-dimensional' displacement operators multiply as the corresponding displacement operators when acting on the vacuum.

Lemma 6.5. *With the above definitions, under the same hypotheses as in Lemma* 6.4, *we have*

$$\sup_{\|\vec{z}\|\leq n^\beta} \sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|[\Delta_\lambda^{\vec{\zeta}+\vec{z},n} - \Delta_\lambda^{\vec{\zeta},n}\Delta_\lambda^{\vec{z},n}](|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|)\|_1 = R(n)$$

with $R(n)$ *given by Eq.* (6.11).

From the last three lemmas, together with the bound on the remainder integral (6.10) we obtain the following lemma which can be plugged into the bound (6.2):

Lemma 6.6. *With the above notations under the same hypotheses as in Lemma* 6.4, *we have*

$$\sup_{\theta\in\Omega_{n,\beta,\gamma}} \sup_{\lambda\in\Lambda_{n,\alpha}} \|\phi^{\vec{\zeta}} - \phi_\lambda^{\theta,n}\| = R(n) + O(n^{-1/2+\gamma+\eta}/\delta)$$

with $R(n)$ *given by Eq.* (6.11).

Gathering all these results and using the inequalities $\alpha - \gamma - 1/2 > 0$, $2\beta + \epsilon \leq \eta < 2/9$ we get the following relations between the error terms: $n^{-1/2+\beta+\eta/2}/\delta^{1/2} = o(n^{-1/2+3\eta/2}/\delta^{1/2})$ and $n^{-1/2+\alpha/2+\beta}/\delta^{1/2} = o(n^{-1/2+\alpha/2+\eta/2}/\delta^{1/2})$.

This yields the next theorem which provides the bound (4.21).

Theorem 6.7. *For any $\delta > 0$, $0 < \gamma < 1/4$, $\epsilon > 0$, $1/2 + \gamma < \alpha < 1$, $\eta < 2/9$, $0 < \beta < (\eta - \epsilon)/2$, the sequence of channels T_n satisfies*

$$\sup_{\theta \in \Omega_{n,\beta,\gamma}} \left\| T_n(\rho^{\theta,n}) - \phi \right\|_1 = O(n^{-1/4+3\beta/2+\epsilon}\delta^{-3/2} + n^{-1/2+\alpha/2+\eta/2}\delta^{-1/2} + n^{-1/2+3\eta/2}\delta^{-1/2} + n^{-\beta/2} + n^{-1/2+\gamma+\eta}/\delta + n^{(9\eta-2)/24}/\delta^{1/6} + \exp(-\delta n^{\eta})). \tag{6.12}$$

For any given $0 < \delta < 1$, $\beta < 1/9$ and $\gamma < 1/4$, we can choose α, η, ϵ satisfying the above conditions, such that the right side is of order $O(n^{-\kappa})$, with $\kappa > 0$ depending on β, γ, δ.

6.2. Definition of S_n and proof of its efficiency. The channel S_n is essentially the inverse of T_n and as we shall see, (4.22) can be deduced from (4.21).

On the classical side we need a Markov kernel completing the equivalence between the family $p_\lambda^{\vec{u},n}$ and $\mathcal{N}(\vec{u}, V_\mu)$. Let σ^n be defined by

$$\sigma^n : x \in \mathbb{R}^{d-1} \mapsto \delta_{\lambda_x}, \tag{6.13}$$

where λ_x is the Young diagram such that $\sum_1^d \lambda_i = n$, and $|n^{1/2}x_i + n\mu_i - \lambda_i| < 1/2$, for $2 \le i \le d$. No such diagram exists, we set λ_x to any admissible value, for example $(n, 0, \ldots, 0)$. Notice that with (5.12), $\sigma^n \circ \tau^n \circ \sigma^n = \sigma^n$. Moreover any probability on the λ such that $\sum_1^d \lambda_i = n$ is in the image of σ^n, so that $\sigma^n \circ \tau^n(p^{\theta,n}) = p^{\theta,n}$.

Lemma 6.8. *With the above definitions, for any ϵ, we have*

$$\sup_{\|\vec{u}\| \le n^\gamma} \left\| \sigma^n \mathcal{N}(\vec{u}, V_\mu) - p^{\vec{u},n} \right\|_1 = O(n^{-1/2+\epsilon}/\delta, n^{-1/4+\gamma}/\delta).$$

Proof. See the end of Sect. 7.5. □

The channel S_n is given by the following sequence of operations acting on the two spaces of the product $L^1(\mathbb{R}^{d-1}) \otimes \mathcal{T}_1(\mathcal{F})$. Given a sample from the probability distribution $N(\vec{u}, V_\mu)$, we use the Markov kernel σ^n to produce a Young diagram λ. Conditional on λ we send the quantum part through the channel

$$S_\lambda : \phi \mapsto \tilde{S}_\lambda(\phi) \otimes \frac{\mathbf{1}_{\mathcal{K}_\lambda}}{M_n(\lambda)}$$

with

$$\tilde{S}_\lambda : \phi \mapsto T_\lambda^* \phi + (1 - \mathrm{Tr}(T_\lambda^*(\phi)))|\mathbf{0}, \lambda\rangle\langle \mathbf{0}, \lambda|.$$

The second term is rather arbitrary and ensures that $\tilde{S}_\lambda$ is trace preserving map. What is important is that for any density operator ρ_λ on the block λ, the operator $\tilde{S}_\lambda$ reverts the action of T_λ:

$$\begin{aligned} \tilde{S}_\lambda T_\lambda(\rho_\lambda) &= T_\lambda^* T_\lambda(\rho_\lambda) + (1 - \mathrm{Tr}(T_\lambda^* T_\lambda(\rho_\lambda)))|\mathbf{0}, \lambda\rangle\langle \mathbf{0}, \lambda| \\ &= \rho_\lambda + (1 - \mathrm{Tr}(\rho_\lambda))|\mathbf{0}, \lambda\rangle\langle \mathbf{0}, \lambda| = \rho_\lambda. \end{aligned}$$

Now

$$S_n(\mathcal{N}(\vec{u}, V_\mu) \otimes \phi^{\vec{\zeta}}) = \bigoplus_\lambda [\sigma^n \mathcal{N}(\vec{u}, V_\mu)](\lambda) \tilde{S}_\lambda(\phi^{\vec{\zeta}}) \otimes \frac{\mathbf{1}_{\mathcal{K}_\lambda}}{M_n(\lambda)},$$

and with the notation $\sigma^n \mathcal{N}_\lambda^{\vec{u}} := [\sigma^n \mathcal{N}(\vec{u}, V_\mu))](\lambda)$ and $q_\lambda^{\vec{u},n} := \min(\sigma^n \mathcal{N}_\lambda^{\vec{u}}, p_\lambda^{\vec{u},n})$ we have

$$S_n(\phi^{\vec{\zeta}} \otimes \mathcal{N}(\vec{u}, V_\mu)) - \rho^{\theta,n} = \bigoplus_\lambda \Big\{ q_\lambda^{\vec{u},n} (\tilde{S}_\lambda(\phi^{\vec{\zeta}}) - \rho_\lambda^{\theta,n}) + (\sigma^n \mathcal{N}_\lambda^{\vec{u}} - q_\lambda^{\vec{u},n}) \tilde{S}_\lambda(\phi^{\vec{\zeta}}) - (p_\lambda^{\vec{u},n} - q_\lambda^{\vec{u},n}) \rho_\lambda^{\theta,n} \Big\} \otimes \frac{\mathbf{1}_{\mathcal{K}_\lambda}}{M_n(\lambda)}.$$

Taking L^1 norms, and using that all ϕ's and ρ's have trace 1 and that channels (such as $\tilde{S}_\lambda$) are trace preserving, we get the bound:

$$\begin{aligned} \|S_n(\phi^{\vec{\zeta}} \otimes \mathcal{N}(\vec{u}, V_\mu)) - \rho^{\theta,n}\|_1 &\leq \sum_\lambda \|q_\lambda^{\vec{u},n} (\tilde{S}_\lambda(\phi^{\vec{\zeta}}) - \rho_\lambda^{\theta,n})\|_1 + \sum_\lambda |\sigma \mathcal{N}_\lambda^{\vec{u}} - p_\lambda^{\vec{u},n}| \\ &\leq 2 \sum_{\lambda \notin \Lambda_{n,\alpha}} q_\lambda^{\vec{u},n} + \sup_{\lambda \in \Lambda_{n,\alpha}} \|\tilde{S}_\lambda(\phi^{\vec{\zeta}}) - \rho_\lambda^{\theta,n}\|_1 + \|\sigma^n \mathcal{N}(\vec{u}, V_\mu) - p^{\vec{u},n}\|_1 \\ &\leq 2 \sum_{\lambda \notin \Lambda_{n,\alpha}} q_\lambda^{\vec{u},n} + \sup_{\lambda \in \Lambda_{n,\alpha}} \|\phi^{\vec{\zeta}} - T_\lambda(\rho_\lambda^{\theta,n})\|_1 + \|\sigma^n \mathcal{N}(\vec{u}, V_\mu) - p^{\vec{u},n}\|_1. \end{aligned}$$

Now the first term is smaller than the remainder term of the gaussian outside a ball whose radius is n^α. Hence this term is going to zero faster than any polynomial, independently on δ and $\vec{u}$ for $\|\vec{u}\| \leq n^\gamma$. The second term is treated in Lemma 6.6 (recalling that $\phi_\lambda^{\theta,n} = T_\lambda(\rho_\lambda^{\theta,n})$), and the third term is treated in Lemma 6.8. This ends the proof of (4.22). □

7. Technical Proofs

7.1. Combinatorial and representation theoretical tools. Here we continue the analysis of the $SU(d)$ irreducible representations $(\pi_\lambda, \mathcal{H}_\lambda)$ started in Sect. 5.1. The purpose of this section is to provide good estimates of quantities of the type $\langle \mathbf{m}, \lambda \mid \pi_\lambda(U) \mid \mathbf{l}, \lambda \rangle$ which will be needed in the proofs of Lemmas 7.3 and 6.4.

We shall use the following form of a general $SU(d)$ element and the shorthand notations:

$$\begin{aligned} U(\vec{\zeta}, \vec{\xi}) &:= \exp\left[i \left(\sum_{i=1}^{d-1} \xi_i H_i + \sum_{1 \leq j < k \leq d} \frac{\mathrm{Re}(\zeta_{j,k}) T_{j,k} + \mathrm{Im}(\zeta_{j,k}) T_{k,j}}{\sqrt{\mu_j - \mu_k}} \right) \right], \\ U(\vec{\zeta}, \vec{\xi}, n) &:= U(\vec{\zeta}/\sqrt{n}, \vec{\xi}/\sqrt{n}), \quad U(\vec{\zeta}) := U(\vec{\zeta}, \vec{0}), \quad U(\vec{\zeta}, n) := U(\vec{\zeta}/\sqrt{n}), \end{aligned} \tag{7.1}$$

where H_i and $T_{i,j}$ are the generators of $SU(d)$ defined by

$$\begin{aligned} H_j &= E_{j,j} - E_{j+1,j+1} \quad \text{for } j \leq d-1, \\ T_{j,k} &= i E_{j,k} - i E_{k,j} \quad \text{for } 1 \leq j < k \leq d, \\ T_{k,j} &= E_{j,k} + E_{k,j} \quad \text{for } 1 \leq j < k \leq d, \end{aligned} \tag{7.2}$$

with $E_{i,j}$ the matrix with entry (i, j) equal to 1, and all others equal to 0.

We first introduce some new notations and remind the reader about the already existing ones:

1) We write $l(c)$ for the length of the column c in the Young diagram λ. There are then $\lambda_i - \lambda_{i+1}$ columns such that $l(c) = i$. An alternative definition is $l(c) = \inf\{i : \lambda_i \geq c\}$.
2) Recall that we denote by $f_{\mathbf{a}}$ the basis vectors $f_{a(1)} \otimes \cdots \otimes f_{a(n)}$, and to each vector we associate a Young tableau $t_{\mathbf{a}}$, where the indices $a(i)$ fill the boxes of a diagram λ in a particular way. We denote by $t^c_{\mathbf{a}}$ the column c of $t_{\mathbf{a}}$, i.e. the function $t^c_{\mathbf{a}} : \{1, \ldots, l(c)\} \to \{1, \ldots, d\}$ that associates to the row number r the value of the entry of that Young tableau in column c, row r. For example, if $t_{\mathbf{a}} = \begin{smallmatrix}\boxed{2}\boxed{2}\boxed{1}\\ \boxed{2}\boxed{1}\end{smallmatrix}$ we get the values:

$$t^1_{\mathbf{a}}(1) = 2, \quad t^1_{\mathbf{a}}(2) = 2, \quad t^2_{\mathbf{a}}(1) = 2, \quad t^2_{\mathbf{a}}(2) = 1, \quad t^3_{\mathbf{a}}(1) = 1.$$

We shall often be interested in the image $t^c_{\mathbf{a}}(\{1, \ldots, l(c)\})$ as an unordered set, or compare $t^c_{\mathbf{a}}$ to Id^c, the identity function on the integers $\{1, \ldots, l(c)\}$.
3) Recall also that $\mathcal{H}_\lambda$ is spanned by the vectors $y_\lambda f_{\mathbf{a}}$ for which $t_{\mathbf{a}}$ is a semistandard Young tableau, and $y_\lambda = q_\lambda p_\lambda$ is the Young symmetriser (cf. Theorem 5.2). If $t_{\mathbf{a}}$ is semistandard then we can use the alternative notation $f_{\mathbf{m}}$ for $f_{\mathbf{a}}$ since $\mathbf{a}$ is in one-to-one correspondence with $\mathbf{m} = \{m_{i,j} : 1 \leq i < j \leq d\}$, where $m_{i,j}$ is the number of j's in the row i of $t_{\mathbf{a}}$. The normalised vectors are

$$|\mathbf{m}, \lambda\rangle := y_\lambda f_{\mathbf{m}} / \| y_\lambda f_{\mathbf{m}} \|.$$

4) Let $\mathcal{O}_\lambda(\mathbf{m})$ be the orbit of $f_{\mathbf{m}}$ under the subgroup $\mathcal{R}_\lambda$ of row permutations. This consists of vectors $f_{\mathbf{b}}$ which have exactly $m_{i,j}$ boxes with j in row i, and the rest are i. In particular, row i has no entries smaller than i. Since the action of permutations is transitive, we have

$$p_\lambda f_{\mathbf{m}} = \sum_{\sigma \in \mathcal{R}_\lambda} f_{\mathbf{a}\circ\sigma} = \sum_{f_{\mathbf{b}} \in \mathcal{O}_\lambda(\mathbf{m})} \frac{\#\mathcal{R}_\lambda}{\#\mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{b}}. \tag{7.3}$$

5) Since we antisymmetrize with q_λ, we are only interested in the $t_{\mathbf{a}}$ (not necessarily semistandard) which do not have two equal entries in the same column. Such tableaux $t_{\mathbf{a}}$ (or vectors $f_{\mathbf{a}}$) shall be called *admissible* and their set is denoted $\mathcal{V}$.
6) For any $f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m})$ we define

$$\Gamma(f_{\mathbf{a}}) := |\mathbf{m}| - \#\{1 \leq c \leq \lambda_1 : t^c_{\mathbf{a}} \neq \mathrm{Id}^c\},$$

and denote by $\mathcal{V}^\Gamma(\mathbf{m})$ the set of vectors $f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m}) \bigcap \mathcal{V}$ with $\Gamma(f_a) = \Gamma$. Then we have

$$\mathcal{O}_\lambda(\mathbf{m}) \bigcap \mathcal{V} = \bigcup_{\Gamma \in \mathbb{N}} \mathcal{V}^\Gamma(\mathbf{m}).$$

Note that $\Gamma(f_{\mathbf{a}}) \geq 0$ and is zero if and only if each column $t^c_{\mathbf{a}}$ is either Id^c or of the form $t^c_{\mathbf{a}}(r) = j\delta_{r=i} + r\delta_{r\neq i}$ for some $i \leq l(c) < j$. A $t^c_{\mathbf{a}}$ of this form will be called an (i, j)-*substitution*.

The following 'algorithm' shows how to build all the possible $f_{\mathbf{a}} \in \mathcal{V}^\Gamma(\mathbf{m})$, thus enabling us to estimate the size of $\mathcal{V}^\Gamma(\mathbf{m})$.

Algorithm. Let $(\mathbf{m}, \lambda)$ be fixed but otherwise arbitrary. In order to generate a particular admissible $f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m})$ we need to select the $m_{i,j}$ boxes on row i which are filled with j, for all $i < j$. The rest of the boxes are filled automatically with i's. The constraint is that no column should have two boxes filled with the same number.

Generating a diagram can be described intuitively as follows. We start with the 'vacuum' vector (tableau) $f_{\mathbf{0}} := f_{\mathbf{m}=\mathbf{0}}$ (row i is filled exclusively with i's), and with a set of $|\mathbf{m}|$ bricks containing $m_{i,j}$ identical bricks labelled (i, j), for each pair $i < j$. To change the content of a box from i into j we place an (i, j)-brick in that box. This procedure is repeated until all bricks have been used, each box being modified at most once.

At this stage each column c may contain several bricks placed in the appropriate boxes, so that its configuration is uniquely defined by the set of bricks κ which shall be called a *column-modifier*. For example if $\kappa = \{(i, j), (f, l)\}$, then the column has entries

$$t^c_{\mathbf{a}}(k) = \begin{cases} j & \text{if } k = i; \\ l & \text{if } k = f; \\ k & \text{otherwise.} \end{cases}$$

Note that a column-modifier is not an arbitrary collection of bricks but one that can be used to produce a column with different entries. In the previous example, if $i < f$ this means either ($j \neq f$ and $j, l > l(c)$) or ($j = f$ and $l > l(c)$). The elementary one-brick column-modifier denoted $\kappa(i, j)$ can only be used in a column with $i \leq l(c) < j$, otherwise the entry j would appear twice.

Now, since the length of a column is at most d and all entries must be different, there are less than $d!$ different types of column-modifiers. Another important remark is that a column-modifier always increases the value of the modified cells, so that in this case $t^c_{\mathbf{a}}(\{1, \ldots, l(c)\}) \neq \{1, \ldots, l(c)\}$.

Alternatively to the above scenario where the bricks are inserted sequentially, we can first cluster them into $|\mathbf{m}| - \Gamma$ column-modifiers, and then apply each column-modifier to a particular column. A given collection of column-modifiers is uniquely determined by $\{m_\kappa : \kappa\}$, where m_κ is the multiplicity of κ. This procedure is detailed in the following 3 stages:

I. Choose Γ bricks among our $|\mathbf{m}|$. As we have $d(d-1)/2$ different types of bricks (recall that $i > j$), and we do not distinguish between identical bricks, there are at most $[d(d-1)/2]^\Gamma$ possibilities. For $\Gamma = 0$, we have only one choice.

II. Consider the remaining bricks as a set of elementary column-modifiers. Starting from these, we sequentially add each of the Γ bricks selected in the first stage, to one of these elementary column-modifiers to form non-elementary ones. At each step we have at most $d!$ different *types* of column modifiers to which we can attach the new brick. Note that we do not distinguish between column modifiers of the same type, but rather consider them as an unordered set. Hence, we have less than $(d!)^\Gamma$ possibilities. If $\Gamma = 0$ there is only one possibility.
Note that at the end of Stage II at least $\max\{0, |\mathbf{m}| - 2\Gamma\}$ of the column-modifiers are elementary, and that $m_{\kappa(i,j)} \leq m_{i,j}$.

III. Apply the column-modifiers to the columns of $f_{\mathbf{0}}$, so that no two modifiers are applied to the same column and the resulting $f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m})$ is admissible. By construction $\Gamma(f_{\mathbf{a}}) = \Gamma$ and all admissible tableaux can be generated in this way.

For counting the number of possibilities for the third stage we apply the column modifiers sequentially, but since some of them may be identical we need to divide by the combinatorial factor $\prod_\kappa m_\kappa!$, where m_κ is the number of column modifiers of type κ.

We distinguish between elementary column modifiers of type $\kappa(i, j)$ and composite ones. There are less than n possibilities of inserting a composite column-modifier κ. An elementary one of type $\kappa(i, j)$ can only be inserted in a column with at least i rows, and since the resulting vector has to be admissible, the column cannot contain another j, so its length is smaller than j. There are $\lambda_i - \lambda_j$ such columns. Hence the number of possibilities at stage three of the algorithm is upper bounded by

$$\prod_{\kappa \neq \kappa(i,j)} \frac{n^{m_\kappa}}{m_\kappa!} \cdot \prod_{i<j} \frac{(\lambda_i - \lambda_j)^{m_{\kappa(i,j)}}}{m_{\kappa(i,j)}!}. \tag{7.4}$$

When $\Gamma = 0$, for each elementary column modifier $\kappa(i, j)$ the number of available columns is at least $(\lambda_i - \lambda_j - |\mathbf{m}|)_+ := \max\{0, \lambda_i - \lambda_j - |\mathbf{m}|\}$. Thus we have the following lower bound:

$$\prod_{i<j} \frac{(\lambda_i - \lambda_j - |\mathbf{m}|)_+^{m_{i,j}}}{m_{i,j}!}. \tag{7.5}$$

Note that the upper bound (7.4) depends on the set of multiplicities $\{m_\kappa\}$.

We now return to our list of notations and definitions:

7) To each column of $t_\mathbf{a}$ we associated a column modifier which completely determines its content. If $m_\kappa^\mathbf{a}$ is the number of columns with column-modifer κ, we collect all multiplicities in $E := \{m_\kappa^\mathbf{a} : \kappa\}$. In particular Γ is a function of E,

$$\Gamma(f_\mathbf{a}) = |\mathbf{m}| - \sum_\kappa m_\kappa^\mathbf{a}.$$

Vectors for which $\Gamma(f_\mathbf{a}) = 0$ have the same multiplicity set E^0, where $m_{\kappa(i,j)} = m_{i,j}$ for all $i < j$ and the other $m_\kappa = 0$. Similarly to $\mathcal{V}^\Gamma(\mathbf{m})$, we denote by $\mathcal{V}^E(\mathbf{m})$ the set of tableaux in $\mathcal{O}_\lambda(\mathbf{m}) \bigcap \mathcal{V}$ with $E(f_\mathbf{a}) = E$, in particular

$$\mathcal{V}^\Gamma(\mathbf{m}) = \bigcup_{E:\Gamma(E)=\Gamma} \mathcal{V}^E(\mathbf{m}).$$

8) To each column c of $t_\mathbf{a}$ we associate two disjoint sets: the added entries $\{t_\mathbf{a}^c(1), \ldots, t_\mathbf{a}^c(l(c))\} \setminus \{1, \ldots, l(c)\}$ and the deleted entries $\{1, \ldots, l(c)\} \setminus \{t_\mathbf{a}^c(1), \ldots, t_\mathbf{a}^c(l(c))\}$. This data is placed into a single set by attaching a $\pm$ sign to each entry, indicating if it is added or deleted. It is easy to verify that if $t_\mathbf{a}$ is admissible, the set of added and deleted entries is uniquely determined by the column-modifer κ associated to c, and hence shall be denoted by $S(\kappa)$. For example $S(\kappa(i, j)) = \{(i, -), (j, +)\}$ and for $\kappa = \{(i, j), (j, k)\}$ we have $S(\kappa) = \{(i, -), (k, +)\}$. We define the multiplicities $m_S^\mathbf{a} = \sum_{\kappa:S(\kappa)=S} m_\kappa^\mathbf{a}$ and $F(f_\mathbf{a}) := \{m_S^\mathbf{a} : S\}$. To summarise, we have defined the maps $f_\mathbf{a} \longmapsto E(f_\mathbf{a}) \longmapsto F(f_\mathbf{a})$.

We now state our estimates. The first point of the following lemma is an exact formula serving as the main tool to prove some of the bounds below.

Lemma 7.1. 1. *For any unitary operator $U \in M(\mathbb{C}^d)$, for any basis vectors $f_\mathbf{a}$ and $f_\mathbf{b}$, we have*

$$\langle f_\mathbf{a} | q_\lambda U^{\otimes n} f_\mathbf{b} \rangle = \prod_{1 \leq c \leq \lambda_1} \det(U^{t_\mathbf{a}^c, t_\mathbf{b}^c}), \tag{7.6}$$

where $U^{t^c_{\mathbf{a}},t^c_{\mathbf{b}}}$ is the $l(c)\times l(c)$ minor of U given by $[U^{t^c_{\mathbf{a}},t^c_{\mathbf{b}}}]_{i,j}=U_{t^c_{\mathbf{a}}(i),t^c_{\mathbf{b}}(j)}$. Under the assumptions

$$\left.\begin{array}{ll} |\mathbf{m}|\le n^{\eta}, & \lambda\in\Lambda_{n,\alpha} \\ \inf_i|\mu_i-\mu_{i+1}|\ge\delta, & \mu_d\ge\delta \\ \|\vec{\zeta}\|_1\le Cn^{\beta}, & \|\vec{\xi}\|_1\le n^{-1/2+2\beta}/\delta, \\ \beta\le 1/2 & n>\left(2\delta^{-1}\right)^{1/(1-\alpha)} \end{array}\right\} \tag{7.7}$$

we have the following estimates with remainder terms uniform in the eigenvalues $\mu_\bullet$:

2. *The number of admissible $f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{m})$ with $\Gamma(f_{\mathbf{a}})=0$ is*

$$\#\mathcal{V}^0(\mathbf{m})=\prod_{j>i}\frac{(\lambda_i-\lambda_j)^{m_{i,j}}}{m_{i,j}!}(1+O(n^{-1+2\eta}/\delta)). \tag{7.8}$$

3. *Let $E:=\{m_\kappa:\kappa\}$ with $\Gamma(E)=\Gamma$. The number of admissible $f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{m})$ with $E(f_{\mathbf{a}})=E$ is bounded by:*

$$\#\mathcal{V}^E(\mathbf{m})\le n^{-\Gamma+\sum_{i<j}(m_{i,j}-m_{\kappa(i,j)})}\prod_{j>i}\frac{(\lambda_i-\lambda_j)^{m_{\kappa(i,j)}}}{m_{\kappa(i,j)}!}. \tag{7.9}$$

4. *The number of admissible $f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{m})$ with $\Gamma(f_{\mathbf{a}})=\Gamma$ is bounded by:*

$$\#\mathcal{V}^\Gamma(\mathbf{m})\le C^\Gamma n^{-\Gamma}\delta^{-2\Gamma}|\mathbf{m}|^{2\Gamma}\prod_{j>i}\frac{(\lambda_i-\lambda_j)^{m_{i,j}}}{m_{i,j}!}, \tag{7.10}$$

   for a constant $C=C(d)$.

5. *Let $f_{\mathbf{a}}\in\mathcal{V}^{\Gamma^a}(\mathbf{l})$, and consider $\mathcal{V}^{\Gamma^b}(\mathbf{m})\subset\mathcal{O}_\lambda(\mathbf{m})$ for some fixed Γ^b. Then:*

$$|\langle f_{\mathbf{a}}|q_\lambda\sum_{f_{\mathbf{b}}\in\mathcal{V}^{\Gamma^b}(\mathbf{m})}f_{\mathbf{b}}\rangle|\le\begin{cases}0 & \text{if } \Gamma^b\ne|\mathbf{m}|-|\mathbf{l}|+\Gamma^a\\ (C|\mathbf{m}|)^{\Gamma^b} & \text{otherwise}\end{cases}, \tag{7.11}$$

   with $C=C(d)$.

6. *If $f_{\mathbf{a}}\in\mathcal{V}^0(\mathbf{m})$, then*

$$\langle f_{\mathbf{a}}|q_\lambda\sum_{f_{\mathbf{b}}\in\mathcal{O}_\lambda(\mathbf{m})}f_{\mathbf{b}}\rangle=1. \tag{7.12}$$

7. *If $f_{\mathbf{a}}\in\mathcal{V}^0(\mathbf{m})$ so that its set of elementary column-modifiers is $E^0=\{m_{\kappa(i,j)}=m_{i,j}\}$, then*

$$\langle f_{\mathbf{a}}|q_\lambda U(\vec{\zeta},\vec{\xi},n)^{\otimes n}f_{\mathbf{0}}\rangle=\exp\left(i\phi-\frac{\|\vec{\zeta}\|_2^2}{2}\right)\prod_{i<j}\left(\frac{\zeta_{i,j}}{\sqrt{n}\sqrt{\mu_i-\mu_j}}\right)^{m_{i,j}}r(n), \tag{7.13}$$

   with the phase and error factor

$$\phi=\sqrt{n}\sum_{i=1}^{d-1}(\mu_i-\mu_{i+1})\xi_i,$$

$$r(n)=1+O\left(n^{-1+2\beta+\eta}\delta^{-1},n^{-1/2+2\beta}\delta^{-1},n^{-1+2\beta+\alpha}\delta^{-1}\right).$$

8. *If $f_{\mathbf{a}} \in \mathcal{V}^E(\mathbf{m})$, so that its set of column-modifiers is $E = \{m_\kappa : \kappa\}$ and $\Gamma(E) = \Gamma$, then*

$$\left|\langle f_{\mathbf{a}}|q_\lambda U(\vec{\zeta},\vec{\xi},n)^{\otimes n} f_{\mathbf{0}}\rangle\right| \leq \exp\left(-\frac{\|\vec{\zeta}\|_2^2}{2}\right)\left(\frac{C\|\vec{\zeta}\|}{\sqrt{n\delta}}\right)^{-\Gamma+\sum_{i<j}(m_{i,j}-m_{\kappa(i,j)})} \prod_{i<j}\left(\frac{\zeta_{i,j}}{\sqrt{n}\sqrt{\mu_i-\mu_j}}\right)^{m_{\kappa(i,j)}} r(n), \tag{7.14}$$

with $C = C(d)$ a constant and $r(n)$ as in point 7 above.

9. *Under the further hypotheses that $\|\vec{z}\| \leq n^\beta$, $m_{i,j} \leq 2|\zeta_{i,j} + z_{i,j}|n^{\beta+\epsilon}$ for some $\epsilon > 0$, we have:*

$$\langle \sum_{f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}}|q_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n) f_{\mathbf{0}}\rangle = \exp\left(i\phi - \frac{\|\vec{\zeta}+\vec{z}\|_2^2}{2}\right)\prod_{i<j}\frac{\left((\zeta_{i,j}+z_{i,j})(\sqrt{n}\sqrt{\mu_i-\mu_j})\right)^{m_{i,j}}}{m_{i,j}!} r(n), \tag{7.15}$$

with

$$r(n) = 1 + O\left(n^{-1+2\beta+\eta}\delta^{-1}, n^{-1+2\beta+\alpha}\delta^{-1}, n^{-1+2\eta}\delta^{-1}, n^{-1+\alpha+\eta}\delta^{-1}, \delta^{-3/2}n^{-1/2+3\beta+2\epsilon}\right).$$

10. *Under the further hypotheses that $|\mathbf{l}| \leq |\mathbf{m}|$ and $n^{1-3\eta} > 2C/\delta^2$, where $C = C(d)$,*

$$|\langle \sum_{f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{l})} f_{\mathbf{a}}|q_\lambda \sum_{f_{\mathbf{b}}\in\mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{b}}\rangle| \leq (C|\mathbf{m}|)^{|\mathbf{m}|-|\mathbf{l}|}\prod_{i<j}\frac{(\lambda_i-\lambda_j)^{l_{i,j}}}{l_{i,j}!}\left(\frac{C|\mathbf{l}|^2|\mathbf{m}|}{n\delta^2}\right)^{\Gamma^a_{\min}(\mathbf{l},\mathbf{m})} \tag{7.16}$$

with

$$\Gamma^a_{\min}(\mathbf{l},\mathbf{m}) \geq \frac{(|\mathbf{l}-\mathbf{m}|+3|\mathbf{l}|-3|\mathbf{m}|)_+}{6}. \tag{7.17}$$

11. *We have*

$$\langle \sum_{f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}}|q_\lambda \sum_{f_{\mathbf{b}}\in\mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{b}}\rangle = \prod_{i<j}\frac{(\lambda_i-\lambda_j)^{m_{i,j}}}{m_{i,j}!}\left(1+O(n^{3\eta-1}/\delta)\right). \tag{7.18}$$

Proof of (7.6). We first express $\langle f_{\mathbf{a}}|U^{\otimes n} f_{\mathbf{b}}\rangle$ as a product of matrix entries of U:

$$\langle f_{\mathbf{a}}|U^{\otimes n} f_{\mathbf{b}}\rangle = \prod_{1\leq c\leq\lambda_1}\prod_{1\leq r\leq l(c)}\langle f_{t^c_{\mathbf{a}}(r)}|U f_{t^c_{\mathbf{b}}(r)}\rangle = \prod_{1\leq c\leq\lambda_1}\prod_{1\leq r\leq l(c)} U_{t^c_{\mathbf{a}}(r),t^c_{\mathbf{b}}(r)}.$$

Since the subgroup of column permutations $\mathcal{C}_\lambda$ is the product of the permutation groups of each column, each $\sigma \in \mathcal{C}_\lambda$ is $\sigma = s_1 \dots s_{\lambda_1}$ with s_c a permutation of column c which transforms $t^c_{\mathbf{b}}(r)$ into $t^c_{\mathbf{b}}(s_c(r))$. Then

$$\begin{aligned}\langle f_{\mathbf{a}}|q_\lambda U^{\otimes n} f_{\mathbf{b}}\rangle = \langle f_{\mathbf{a}}|U^{\otimes n} q_\lambda f_{\mathbf{b}}\rangle &= \sum_{\sigma\in\mathcal{C}_\lambda} \epsilon(\sigma) \prod_{1\le c\le\lambda_1} \prod_{1\le r\le l(c)} U_{t^c_{\mathbf{a}}(r),t^c_{\mathbf{b}}(s_c(r))} \\ &= \prod_{1\le c\le\lambda_1} \sum_{s_c\in S_c} \epsilon(s_c) \prod_{1\le r\le l(c)} U_{t^c_{\mathbf{a}}(r),t^c_{\mathbf{b}}(s_c(r))} \\ &= \prod_{1\le c\le\lambda_1} \det(U^{t^c_{\mathbf{a}},t^c_{\mathbf{b}}}).\end{aligned}$$

□

Proof of (7.8). The number of admissible $f_{\mathbf{a}}$ such that $\Gamma(f_{\mathbf{a}}) = 0$ is given by the products of the possibilities at each stage of the algorithm. For the first two stages, there is exactly one possibility when $\Gamma = 0$. Hence $\#\mathcal{V}^0$ is the number of possibilities at the third stage. Here the upper bound (7.4) reads: as $\prod_{j>i}(\lambda_i - \lambda_j)^{m_{i,j}}/m_{i,j}!$. On the other hand, we may use (7.5) as a lower bound, recalling that $\lambda_i - \lambda_j \ge \delta n/2$ and $|\mathbf{m}| \le n^\eta$ (cf. (7.7)). This yields the result (7.8). □

Proof of (7.9). The number of $f_{\mathbf{a}}$ in $\mathcal{V}^E$ is given by the third stage of the algorithm (the two first stages yield a particular E). We then obtain (7.9) by applying (7.4) and neglecting the $m_\kappa!$ factors, while noticing that $\sum_\kappa m_\kappa = |\mathbf{m}| - \Gamma$. □

Proof of (7.10). The set $\mathcal{V}^\Gamma$ is the union of all $\mathcal{V}^E$ with $\Gamma(E) = \Gamma$. Now the first two stages of the algorithm imply that there are at most C^Γ different E with the latter property, with $C = C(d)$. □

Now we use (7.9) to upper-bound $\mathcal{V}^E$ as follows. Since $\sum m_{\kappa(i,j)} \ge |\mathbf{m}| - 2\Gamma$, we may write $\prod_\kappa m_{\kappa(i,j)}! \ge \prod_{i<j} m_{i,j}! \sup_{i<j} m_{i,j}^{-2\Gamma}$. Moreover $\lambda_i - \lambda_j \ge \delta n/2$. By putting together we obtain

$$\#\mathcal{V}^E \le n^{-\Gamma}\delta^{-2\Gamma}|\mathbf{m}|^{2\Gamma} \prod_{j>i} \frac{(\lambda_i - \lambda_j)^{m_{i,j}}}{m_{i,j}!}, \quad \forall E \text{ with } \Gamma(E) = \Gamma.$$

Multiplying by the number of possible E yields the result.

Proof of (7.11). We are applying (7.6) with $U = \mathbf{1}$. Since both $f_{\mathbf{a}}$ and $f_{\mathbf{b}}$ are products of basis vectors, the scalar product $\langle f_{\mathbf{a}} \mid q_\lambda f_{\mathbf{b}}\rangle$ is equal to -1 or 1 if $t^c_{\mathbf{a}}([1, l(c)]) = t^c_{\mathbf{b}}([1, l(c)])$ for all columns, and 0 otherwise. Here we denote by $t^c_{\mathbf{a}}([1, l(c)])$ the set of entries $\{t^c_{\mathbf{a}}(1), \dots, t^c_{\mathbf{a}}(l(c))\}$.
Now, since a modified column cannot satisfy $t^c_{\mathbf{a}}([1, l(c)]) = [1, l(c)]$ (and the same for $\mathbf{b}$), the vectors $f_{\mathbf{a}}$ and $f_{\mathbf{b}}$ are orthogonal unless they have the same number of modified columns. Finally, that number is $|\mathbf{l}| - \Gamma^a$ for $f_{\mathbf{a}}$ and $|\mathbf{m}| - \Gamma^b$ for $f_{\mathbf{b}}$. This yields the first line of (7.11). □

We now concentrate on the case when $\Gamma^b = |\mathbf{m}| - |\mathbf{l}| + \Gamma^a$. Since $|\langle f_{\mathbf{a}} \mid q_\lambda f_{\mathbf{b}}\rangle| \le 1$, we can bound the sum of scalar products by the number of non-zero inner products. The question is how many diagrams $f_{\mathbf{b}}$ have the same content (seen as an unordered set) in each column as $f_{\mathbf{a}}$: $t^c_{\mathbf{a}}([1, l(c)]) = t^c_{\mathbf{b}}([1, l(c)])$, or equivalently $S(\kappa^c_{\mathbf{a}}) = S(\kappa^c_{\mathbf{b}})$.

For building the relevant $f_{\mathbf{b}}$, we can follow the algorithm with the further condition that, at stage three, all the column-modifiers are applied in such a way that the unordered column content is identical to that of $f_{\mathbf{a}}$.

The first two stages of the algorithm are the same so they yield a C^{Γ^b} factor. We now have a collection $\{m_\kappa\}$ of column modifiers which have to be placed so that they match the column content of $f_{\mathbf{a}}$. For each S we identify the column modifiers $\kappa_1, \ldots, \kappa_{r(S)}$ such that $S(\kappa_i) = S$ for all $1 \le i \le r(S)$. The total number of such objects is $m_S := \sum_{i \le r(S)} m_{\kappa_i}$ and the number of ways in which they can be inserted to produce *distinct* diagrams is

$$\binom{m_S}{m_{\kappa_1} \ldots m_{\kappa_{r(S)}}}.$$

Recall that the number of elementary column-modifiers $\sum_{i<j} m_{\kappa(i,j)}$ is at least $|\mathbf{m}| - 2\Gamma^b$. Moreover, each elementary column-modifier $\kappa(i, j)$ corresponds to a different $S(\kappa(i, j)) = \{(i, -), (j, +)\}$. Thus

$$|\mathbf{m}| - 2\Gamma^b \le \sum_{i<j} m_{\kappa(i,j)} \le \sum_S \max_{\kappa: S(\kappa)=S} m_\kappa.$$

Since

$$\sum_S m_S = \sum_\kappa m_\kappa = |\mathbf{m}| - \Gamma^b,$$

we obtain

$$\sum_S \left(m_S - \max_{\kappa: S(\kappa)=S} m_\kappa \right) \le \Gamma^b.$$

This implies

$$\prod_S \binom{m_S}{m_{\kappa_1} \ldots m_{\kappa_{r(S)}}} \le |\mathbf{m}|^{\Gamma^b}.$$

Multiplying by the C^{Γ^b} of the first stages, we get (7.11).

Proof of (7.12). As shown above the only non-zero contributions come from $f_{\mathbf{b}} \in \mathcal{V}^0 \subset \mathcal{O}_\lambda(\mathbf{m})$.
Since $\Gamma^b = 0$, the constant from the two first stages of the algorithm is 1, $m_S = m_{i,j} = m_{\kappa(i,j)}$ for all S corresponding to an elementary column-modifier, and 0 otherwise. So the combinatorial factor is again one: we do not have any choice in our placement of column-modifiers. In other words, the only $f_{\mathbf{b}}$ such that $\langle f_{\mathbf{a}} \mid q_\lambda f_{\mathbf{b}} \rangle \neq 0$ is $f_{\mathbf{a}}$. Finally, $\langle f_{\mathbf{a}} \mid q_\lambda f_{\mathbf{a}} \rangle = 1$. □

Proof of (7.13). From (7.6) we deduce

$$\langle f_{\mathbf{a}} | q_\lambda U(\vec{\zeta}, \vec{\xi}, n)^{\otimes n} f_{\mathbf{0}} \rangle = \prod_{1 \le c \le \lambda_1} \det(U^{t^c_{\mathbf{a}}, \mathrm{Id}^c}), \quad U = U(\vec{\zeta}, \vec{\xi}, n).$$

We shall use the Taylor expansion of the unitary $U(\vec{\zeta}, \vec{\xi}, n)$ to estimate the above determinants. □

Entry-wise, for all $1 \le i \le d$ on the first line, and all $1 \le i < j \le d$ on the second and third lines:

$$\begin{aligned}
U_{i,i}(\vec{\zeta},\vec{\xi},n) &= 1 + i\frac{\xi_i\delta_{i\neq d} - \xi_{i-1}\delta_{i\neq 1}}{\sqrt{n}} - \frac{1}{2n}\sum_{j\neq i}\frac{|\zeta_{i,j}|^2}{|\mu_i-\mu_j|} \\
&\quad + O(\|\vec{\zeta}\|^3 n^{-3/2}\delta^{-3/2}, \|\vec{\zeta}\|\|\vec{\xi}\|n^{-1}\delta^{-1/2}, \|\vec{\xi}\|^2 n^{-1}); \\
U_{i,j}(\vec{\zeta},\vec{\xi},n) &= -\frac{1}{\sqrt{n}}\frac{\zeta^*_{i,j}}{\sqrt{\mu_i-\mu_j}} + O(\|\vec{\zeta}\|^2 n^{-1}\delta^{-1}, \|\vec{\zeta}\|\|\vec{\xi}\|n^{-1}\delta^{-1/2}); \\
U_{j,i}(\vec{\zeta},\vec{\xi},n) &= \frac{1}{\sqrt{n}}\frac{\zeta_{i,j}}{\sqrt{\mu_i-\mu_j}} + O(\|\vec{\zeta}\|^2 n^{-1}\delta^{-1}, \|\vec{\zeta}\|\|\vec{\xi}\|n^{-1}\delta^{-1/2}).
\end{aligned}$$

If $\vec{\zeta} = O(n^{\beta})$, $\|\vec{\xi}\| \le n^{-1/2+2\beta}/\delta$, and $\beta < 1/2$, the remainder terms are $O(n^{-3/2+3\beta}\delta^{-3/2})$ for the first line and $O(n^{-1+2\beta}\delta^{-1})$ for the last two lines.

Therefore, when our parameters are in this range, we can give precise enough evaluations of the determinants. The idea is to find the dominating terms in the expansion of the determinant

$$\det A = \sum_{\sigma}\prod_{i}\epsilon(\sigma)A_{i,\sigma(i)}.$$

Note that we can use the above Taylor expansions inside the determinant since the number of terms in the product is at most d.

Since $f_{\mathbf{a}} \in \mathcal{V}^0$, all $t^c_{\mathbf{a}}$ are either Id^c, or an (i,j)-*substitution*. If $t^c_{\mathbf{a}} = \mathrm{Id}^c$, the summands with more than two non-diagonal terms are of the same order as the remainder term, so that only the identity and the transpositions count in $\sum_\sigma \prod_i A_{i,\sigma(i)}$. Let $l = l(c)$, then

$$\upsilon(l) := \det(U^{\mathrm{Id}^c,\mathrm{Id}^c}(\vec{\zeta},\vec{\xi},n)) = 1 + i\frac{\xi_l}{\sqrt{n}} - \frac{1}{2n}\sum_{\substack{1\le i\le l\\ l+1\le j\le d}}\frac{|\zeta_{i,j}|^2}{\mu_i-\mu_j} + O(n^{-3/2+3\beta}\delta^{-3/2}).$$

Note that for $l = d$, we get the usual determinant of $U(\vec{\zeta},\vec{\xi},n)$ which is 1.

Consider now the case $t^c_{\mathbf{a}} \neq \mathrm{Id}^c$. Since $t^c_{\mathbf{a}}(r) \ge r$ for all r, there exists a whole column of $U^{t^c_{\mathbf{a}},\mathrm{Id}^c}$ whose entries are smaller in modulus than $O(\|\vec{\zeta}\|/\sqrt{n\delta}) = O(n^{-1/2+\beta}\delta^{-1})$. In particular if $t^c_{\mathbf{a}}$ is an (i,j)-*substitution*, then the only summand that is of this order comes from the identity. So that

$$\upsilon(i,j) := \det(U^{t^c_{\mathbf{a}},\mathrm{Id}^c}(\vec{\zeta},\vec{\xi},n)) = \frac{\zeta_{i,j}}{\sqrt{n}\sqrt{\mu_i-\mu_j}} + O(n^{-1+2\beta}\delta^{-1}). \tag{7.19}$$

Note that this approximation does not depend on $l(c)$, but only on i and j.

We now put together the estimated determinants in the product (7.6). For each $i < j$ there are $m_{i,j}$ columns of the type (i,j)-substitution. Out of the $\lambda_l - \lambda_{l+1}$ columns of length $l = l(c)$ there are $\lambda_l - \lambda_{l+1} - R_l$ of the type Id^c, with $0 \le R_l \le |\mathbf{m}|$.

Hence:

$$\langle f_{\mathbf{a}}|q_\lambda U(\vec{\zeta},\vec{\xi},n)^{\otimes n} f_{\mathbf{0}}\rangle = \prod_{l=1}^{d}(\upsilon(l))^{\lambda_l-\lambda_{l+1}}\prod_{1\le i<j\le d}(\upsilon(i,j))^{m_{i,j}}\prod_{l=1}^{d}(\upsilon(l))^{-R_l}. \tag{7.20}$$

Now $\upsilon(l) = 1 + O(n^{-1+2\beta}\delta^{-1})$ and $R_l \le |\mathbf{m}| \le n^\eta$, so the last product is $1 + O(n^{-1+2\beta+\eta}\delta^{-1})$. Similarly, since $\lambda \in \Lambda_{n,\alpha}$ we have $\lambda_l - \lambda_{l+1} = n(\mu_l - \mu_{l+1}) + O(n^\alpha)$, and we can use Lemma 7.2 given at the end of this section to estimate the first product as follows:

$$\prod_{l=1}^{d} \upsilon(l)^{\lambda_l - \lambda_{l+1}} = \prod_{l=1}^{d} \exp\left(i\phi_l - \frac{1}{2} \sum_{\substack{1\le i\le l \\ l+1\le j\le d}} |\zeta_{i,j}|^2 \frac{\mu_l - \mu_{l+1}}{\mu_i - \mu_j} \right) r(n)$$
$$= \exp\left(i\phi - \frac{\|\vec{\zeta}\|_2^2}{2} \right) r(n),$$

with

$$\tilde{r}(n) = 1 + O(n^{-1+\alpha+2\beta}\delta^{-1}, n^{-1/2+2\beta}\delta^{-1}),$$
$$\phi_l = \delta_{l\neq d}\sqrt{n}(\mu_l - \mu_{l+1})\xi_l,$$
$$\phi = \sqrt{n}\sum_{l=1}^{d-1}(\mu_l - \mu_{l+1})\xi_l.$$

We now turn our attention to the middle product on the right side of (7.20),

$$\upsilon(i,j)^{m_{i,j}} = \left(\frac{\zeta_{i,j}}{\sqrt{n}\sqrt{\mu_i - \mu_j}} \right)^{m_{i,j}} \left(1 + O\left(n^{-1+2\beta+\eta}\delta^{-1} \right) \right),$$

where we have used that $|\mathbf{m}| \le n^\eta$. Inserting into (7.20) yields (7.13).

Note that $\langle f_{\mathbf{a}} | q_\lambda U(\vec{\zeta}, \vec{\xi}, n)^{\otimes n} f_{\mathbf{0}} \rangle = 0$ if there exist $i < j$ such that $\zeta_{i,j} = 0$ and $m_{i,j} \neq 0$.

Proof of (7.14). We may write, much like in (7.20),

$$\langle f_{\mathbf{a}} | q_\lambda U(\vec{\zeta}, \vec{\xi}, n)^{\otimes n} f_{\mathbf{0}} \rangle = \prod_{l=1}^{d} (\upsilon(l)))^{\lambda_l - \lambda_{l+1}} \prod_{\kappa} (\upsilon(\kappa))^{m_\kappa} \prod_{l=1}^{d} (\upsilon(l))^{-R_l},$$

where $0 \le R_l \le |\mathbf{m}| - \Gamma$ and $\upsilon(\kappa)$ is the determinant of the minor of U corresponding to having applied the column-modifier κ. We can further split the column-modifers into elementary ones $\kappa(i,j)$ and non-elementary ones κ'. □

Then $\langle f_{\mathbf{a}} | q_\lambda U(\vec{\zeta}, \vec{\xi}, n)^{\otimes n} f_{\mathbf{0}} \rangle$ can be written as

$$\prod_{l=1}^{d} (\upsilon(l)))^{\lambda_l - \lambda_{l+1}} \prod_{i<j} (\upsilon(i,j))^{m_{\kappa(i,j)}} \prod_{l=1}^{d} (\upsilon(l))^{-R_l} \prod_{\kappa'} \left(\upsilon(\kappa') \right)^{m_{\kappa'}}.$$

The first three products on the right side can be treated as above. For the fourth product we give a rough upper bound based on the following observation. If the entries in the column have been modified in an admissible way, then $t_{\mathbf{a}}^c(i) = j > l(c)$ for some i, so that $|\upsilon(\kappa)| \le C\|\vec{\zeta}\|/\sqrt{n\delta}$ for any κ, with some constant $C = C(d)$.

Thus by using the previous point

$$\left|\langle f_{\mathbf{a}}|q_\lambda U(\vec{\zeta},\vec{\xi},n)^{\otimes n} f_{\mathbf{0}}\rangle\right| \leq \exp\left(-\frac{\|\vec{\zeta}\|_2^2}{2}\right)\left(\frac{C\|\vec{\zeta}\|}{\sqrt{n\delta}}\right)^{\sum_{\kappa'} m_{\kappa'}} \prod_{i<j}\left(\frac{|\zeta_{i,j}|}{\sqrt{n}\sqrt{\mu_i-\mu_j}}\right)^{m_{\kappa(i,j)}} r(n). \quad (7.21)$$

We obtain (7.14) by noting that the number of non-elementary modifiers is

$$\sum_{\kappa'} m_{\kappa'} = -\Gamma + \sum_{i<j}(m_{i,j} - m_{\kappa(i,j)}).$$

Proof of (7.15). Note that only admissible vectors in $\mathcal{O}_\lambda(\mathbf{m})$ can bring non-zero contributions. We shall split the sum into sub-sums using $\mathcal{O}_\lambda(\mathbf{m}) \bigcap \mathcal{V} = \bigcup_E \mathcal{V}^E(\mathbf{m})$, and compare each sub-sum against the benchmark $\mathcal{V}^0 = \mathcal{V}^{E^0}$.
From the bounds on $\vec{\zeta}$ and $\vec{z}$ we obtain $\|\vec{\zeta}+\vec{z}\| = O(n^\beta)$, so we can apply the previous points with $\vec{\zeta}+\vec{z}$ instead of $\vec{\zeta}$.

Using (7.8) and (7.13) and recalling that $\lambda \in \Lambda_{n,\alpha}$, we get:

$$\langle \sum_{f_{\mathbf{a}}\in\mathcal{V}^0} f_{\mathbf{a}}|q_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n)^{\otimes n} f_{\mathbf{0}}\rangle = \exp\left(i\phi - \|\vec{\zeta}+\vec{z}\|_2^2/2\right)\prod_{i<j}\frac{\left((\zeta_{i,j}+z_{i,j})\sqrt{n}\sqrt{\mu_i-\mu_j}\right)^{m_{i,j}}}{m_{i,j}!} r(n)$$

with error factor

$$r(n) = 1 + O\left(n^{-1+2\beta+\eta}\delta^{-1}, n^{-1/2+2\beta}\delta^{-1}, n^{-1+2\beta+\alpha}\delta^{-1}, n^{-1+2\eta}\delta^{-1}, n^{-1+\alpha+\eta}\delta^{-1}\right).$$

For $E \neq E^0$ we combine (7.14) and (7.9) to obtain

$$\begin{aligned}
&|\langle \sum_{f_{\mathbf{a}}\in\mathcal{V}^E} f_{\mathbf{a}}|q_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n) f_{\mathbf{0}}\rangle| \cdot |\langle \sum_{f_{\mathbf{a}}\in\mathcal{V}^0} f_{\mathbf{a}}|q_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n) f_{\mathbf{0}}\rangle|^{-1} \\
&\leq n^{-\Gamma}\prod_{i<j}\left(\frac{\lambda_i-\lambda_j}{n}\right)^{m_{\kappa(i,j)}-m_{i,j}} \frac{m_{i,j}!}{m_{\kappa(i,j)}!}\left(\frac{\|\vec{\zeta}+\vec{z}\|}{\sqrt{\delta n}}\right)^{-\Gamma} \\
&\quad\times\prod_{i<j}\left(\frac{\sqrt{\delta n}|\zeta_{i,j}+z_{i,j}|}{\|\vec{\zeta}+z\|\sqrt{n}\sqrt{\mu_i-\mu_j}}\right)^{m_{\kappa(i,j)}-m_{i,j}} r(n) \\
&\leq O(n^{-\Gamma(1/2+\beta)})\delta^{-\Gamma/2}\prod_{i<j:m_{i,j}\neq 0}\left(\frac{|\zeta_{i,j}+z_{i,j}|\sqrt{\mu_i-\mu_j}}{m_{i,j}\|\vec{\zeta}+\vec{z}\|}\right)^{m_{\kappa(i,j)}-m_{i,j}} \\
&\leq O\left((2\delta^{-3/2}n^{-1/2+3\beta+2\epsilon})^\Gamma\right),
\end{aligned}$$

with $O(\cdot)$ uniform in Γ. In the second inequality we used

$$m_{i,j}!/m_{\kappa(i,j)}! \leq m_{i,j}^{m_{i,j}-m_{\kappa(i,j)}}, \qquad \sum_{i<j}(m_{\kappa(i,j)} - m_{i,j}) \geq -2\Gamma, \qquad \lambda \in \Lambda_{n,\alpha},$$

and in the third inequality we used

$$m_{i,j} \leq 2|\zeta_{i,j} + z_{i,j}|n^{\beta+\epsilon}, \qquad \frac{|\zeta_{i,j} + z_{i,j}|\sqrt{\mu_i - \mu_j}}{m_{i,j}\|\vec{\zeta} + \vec{z}\|} \leq 1.$$

Furthermore, for a given Γ, there are at most C^Γ different E such that $\Gamma(E) = \Gamma$, corresponding to the possible choices in the first two stages of the algorithm, where $C = C(d)$. Hence, if n is large enough, so that $2C\delta^{-3/2}n^{-1/2+3\beta+2\epsilon} < 1$, we have:

$$\begin{aligned}
\langle \sum_{f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}}|q_\lambda U(\vec{\zeta} + z, \vec{\xi}, n) f_{\mathbf{0}}\rangle &= \sum_\Gamma \langle \sum_{f_{\mathbf{a}} \in \mathcal{V}^\Gamma} f_{\mathbf{a}}|q_\lambda U(\vec{\zeta} + z, \vec{\xi}, n) f_{\mathbf{0}}\rangle \\
&= \left(1 + O(\delta^{-3/2}n^{-1/2+3\beta+2\epsilon})\right) \exp\left(i\phi - \|\vec{\zeta} + z\|_2^2/2\right) \\
&\quad \times \prod_{i<j} \frac{\left((\vec{\zeta} + z)_{i,j}(\sqrt{n}\sqrt{\mu_i - \mu_j})\right)^{m_{i,j}}}{m_{i,j}!} r(n) \\
&= \exp\left(i\phi - \|\vec{\zeta} + z\|_2^2/2\right) \prod_{i<j} \frac{\left((\vec{\zeta} + z)_{i,j}(\sqrt{n}\sqrt{\mu_i - \mu_j})\right)^{m_{i,j}}}{m_{i,j}!} r_2(n),
\end{aligned}$$

where the sum over Γ was bounded using a geometric series and

$$\begin{aligned}
r_2(n) &= 1 \\
&\quad + O\left(n^{-1+2\beta+\eta}\delta^{-1}, n^{-1+\alpha+\beta}\delta^{-1}, n^{-1+2\eta}\delta^{-1}, n^{-1+\alpha+\eta}\delta^{-1}, \delta^{-3/2}n^{-1/2+3\beta+2\epsilon}\right).
\end{aligned}$$

This is exactly (7.15). □

Proof of (7.16). We choose Γ^a and Γ^b satisfying the condition $\Gamma^b - \Gamma^a = |\mathbf{m}| - |\mathbf{l}|$ under which the inner products in (7.11) are non-zero. By multiplying (7.10) and (7.11), we see that:

$$\begin{aligned}
|\langle \sum_{f_{\mathbf{a}} \in \mathcal{V}^{\Gamma^a}(\mathbf{l})} f_{\mathbf{a}}|q_\lambda \sum_{f_{\mathbf{b}} \in \mathcal{V}^{\Gamma^b}(\mathbf{m})} f_{\mathbf{b}}\rangle| &\leq (C|\mathbf{m}|)^{\Gamma^b} \prod_{i<j} \frac{(\lambda_i - \lambda_j)^{l_{i,j}}}{l_{i,j}!} \left(\frac{C|\mathbf{l}|^2}{n\delta^2}\right)^{\Gamma^a} \\
&= (C|\mathbf{m}|)^{|\mathbf{m}|-|\mathbf{l}|} \prod_{i<j} \frac{(\lambda_i - \lambda_j)^{l_{i,j}}}{l_{i,j}!} \left(\frac{C|\mathbf{l}|^2|\mathbf{m}|}{n\delta^2}\right)^{\Gamma^a}.
\end{aligned} \tag{7.22}$$

It remains to sum up the upper bounds over all relevant pairs (Γ^a, Γ^b). If $n^{1-3\eta} > 2C/\delta^2$, the dominating term in the sum of bounds is that corresponding to the smallest possible Γ^a. The question is, what is the smallest possible value of Γ^a leading to non-zero inner products?

A necessary condition for $f_{\mathbf{a}}$ not to be orthogonal to $f_{\mathbf{b}}$ is that for each set S of suppressed and added values, the two vectors have the same multiplicities $m^{\mathbf{a}}_S = m^{\mathbf{b}}_S$.

The following argument provides a lower bound for $\Gamma(f_{\mathbf{a}}) + \Gamma(f_{\mathbf{b}})$. The idea is to count the minimum number of 'horizontal box shuffling' operations necessary in order to transform a Young tableau $t_{\mathbf{a}'} \in \mathcal{O}_\lambda(\mathbf{m})$ into the tableau $t_{\mathbf{a}}$. Since $|\mathbf{m}| \leq n^\eta$ and $\lambda_d \geq \delta n + O(n^\alpha)$, the tableau $t_{\mathbf{a}'}$ can be chosen to have at most one modified box per

column (thus $\Gamma(f_{\mathbf{a}'}) = 0$), and such that each of the modified columns of $t_{\mathbf{a}}$ are also modified in $t_{\mathbf{a}'}$. We also choose $t_{\mathbf{b}'}$ in a similar fashion.

Now at each step we horizontally move one elementary column modifier $\kappa(i, j)$ of $t_{\mathbf{a}'}$ (or $t_{\mathbf{b}'}$) into an already modified column, with the aim of constructing $t_{\mathbf{a}}$ (or $t_{\mathbf{b}}$).

Each such operation increases $\Gamma(f_{\mathbf{a}'}) + \Gamma(f_{\mathbf{b}'})$ by one. On the other hand the operation has the following effect on the $m_S^{\mathbf{a}'}$ (or $m_S^{\mathbf{b}'}$): the multiplicities $m_{\{(i,-),(j,+)\}}$ and m_{S_0} decrease by one, and $m_{S_0+\{(i,-),(j,+)\}}$ increases by one. Here S_0 is the signature of the column to which the box (i, j) is moved. Hence the distance $\sum_S |m_S^{\mathbf{a}'} - m_S^{\mathbf{b}'}|$ decreases by at most three. Since initially this quantity was equal to $\sum_{i<j} |l_{i,j} - m_{i,j}|$, we need at least $\sum_{i<j} |l_{i,j} - m_{i,j}|/3$ such operations before reaching our goal $m_S^{\mathbf{a}} = m_S^{\mathbf{b}}$. This means that $\Gamma(f_{\mathbf{a}}) + \Gamma(f_{\mathbf{b}}) \geq |\mathbf{l} - \mathbf{m}|/3$.

Together with $\Gamma^b - \Gamma^a = |\mathbf{m}| - |\mathbf{l}|$, this result yields $\Gamma^a \geq (|\mathbf{l} - \mathbf{m}| + 3|\mathbf{l}| - 3|\mathbf{m}|)/6$. Moreover Γ^a is non-negative.

Replacing in the above equation yields (7.16). □

Proof of (7.18). Since $\mathbf{l} = \mathbf{m}$, Eqs. (7.8) and (7.12) prove that the bound (7.22) is saturated when $\Gamma^a = 0$, up to the error factor $\left(1 + O(n^{-1+2\eta}/\delta)\right)$. Hence the remainder term due to the other Γ consist in a geometric series with factor $\left(\frac{C|\mathbf{m}|^3}{n\delta^2}\right) = O(n^{1-3\eta}/\delta^2)$. □

Above we used the following lemma whose proof can be found in [29]:

Lemma 7.2. *If $x_n = O(n^{1/2-\epsilon})$, then*

$$\left(1 + \frac{x_n}{n}\right)^n = \exp(x_n)(1 + O(n^{-\epsilon})).$$

7.2. Proof of Lemma 5.4 and non-orthogonality issues.

Lemma 7.3. *Let $(\mathbf{m}, \lambda)$ and $(\mathbf{l}, \lambda)$ be semistandard Young tableaux with diagram λ and define $|\mathbf{m}| := \sum_{i<j} m_{ij}$ and $|\mathbf{l} - \mathbf{m}| := \sum_{i<j} |l_{i,j} - m_{ij}|$. If*

$$\sum_{j>i} m_{i,j} - \sum_{j<i} m_{j,i} \neq \sum_{j>i} l_{i,j} - \sum_{j<i} l_{j,i}$$

for some $1 \leq i \leq d$, then

$$\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda \rangle = 0.$$

Otherwise, we derive an upper bound under the following conditions. We assume that $\lambda_i - \lambda_{i+1} > \delta n$ for all $1 \leq i \leq d-1$ and $\lambda_d > \delta n$, for some $\delta > 0$. Furthermore we assume $|\mathbf{l}| \leq |\mathbf{m}| \leq n^\eta$ for some $\eta < 1/3$ and that $Cn^{3\eta-1}/\delta^2 < 1$, where $C = C(d)$ is a constant. Then:

$$\begin{aligned} |\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda \rangle| \leq\ & (C'n)^{-\eta(|\mathbf{m}|-|\mathbf{l}|)/4}\, (C'n)^{(9\eta-2)|\mathbf{m}-\mathbf{l}|/12}\, \delta^{(|\mathbf{m}|-|\mathbf{l}|)/2-|\mathbf{m}-\mathbf{l}|/3} \\ & \times (1 + O(n^{-1+3\eta}/\delta)), \end{aligned} \tag{7.23}$$

where $C' = C'(d, \eta)$ and the constant in the remainder term depends only on d. The right side is of order less than $n^{(9\eta-2)|\mathbf{m}-\mathbf{l}|/12}$ and converges to zero for $\eta < 2/9$ when $n \to \infty$.

Proof. We know that $|\mathbf{m}, \lambda\rangle$ is a linear combination of n-tensor product vectors in which the basis vector f_i appears exactly $\lambda_i - \sum_{j>i} m_{i,j} + \sum_{j<i} m_{j,i}$ times. As two tensor basis vectors are orthogonal if they do not have the same number of f_i in the decomposition, we get that $\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda\rangle = 0$ if $\sum_{j>i} m_{i,j} + \sum_{j<i} m_{j,i} \neq \sum_{j>i} l_{i,j} + \sum_{j<i} l_{j,i}$ for any $1 \leq i \leq d$.

In the general case,

$$\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda\rangle = \frac{\langle q_\lambda p_\lambda f_{\mathbf{m}} | q_\lambda p_\lambda f_{\mathbf{l}}\rangle}{\sqrt{\langle q_\lambda p_\lambda f_{\mathbf{m}} | q_\lambda p_\lambda f_{\mathbf{m}}\rangle \langle q_\lambda p_\lambda f_{\mathbf{l}} | q_\lambda p_\lambda f_{\mathbf{l}}\rangle}}. \tag{7.24}$$

We use the fact that q_λ is a projection, up to a constant factor (cf. (5.1),(5.3)), and erase the q_λ at the left of each scalar product, and we decompose $p_\lambda f_{\mathbf{m}}$ and $p_\lambda f_{\mathbf{l}}$ on orbits as in (7.3). Since the multiplicity of the elements in the orbits are the same in numerator and denominator, we end up with:

$$\begin{aligned}&\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda\rangle \\ &= \frac{\langle \sum_{f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}} | q_\lambda \sum_{f_{\mathbf{b}} \in \mathcal{O}_\lambda(\mathbf{l})} f_{\mathbf{b}}\rangle}{\langle \sum_{f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}} | q_\lambda \sum_{f_{\mathbf{a}'} \in \mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}'}\rangle \cdot \langle \sum_{f_{\mathbf{b}} \in \mathcal{O}_\lambda(\mathbf{l})} f_{\mathbf{b}} | q_\lambda \sum_{f_{\mathbf{b}'} \in \mathcal{O}_\lambda(\mathbf{l})} f_{\mathbf{b}'}\rangle}.\end{aligned} \tag{7.25}$$

We use (7.18) for the denominator:

$$\begin{aligned}&\langle \sum_{f_{\mathbf{a}} \in \mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}} | q_\lambda \sum_{f_{\mathbf{a}'} \in \mathcal{O}_\lambda(\mathbf{m})} f_{\mathbf{a}'}\rangle \cdot \langle \sum_{f_{\mathbf{b}} \in \mathcal{O}_\lambda(\mathbf{l})} f_{\mathbf{b}} | q_\lambda \sum_{f_{\mathbf{b}'} \in \mathcal{O}_\lambda(\mathbf{l})} f_{\mathbf{b}'}\rangle \\ &= \prod_{1 \leq i < j \leq d} \frac{(\lambda_i - \lambda_j)^{(m_{i,j}+l_{i,j})/2}}{\sqrt{m_{i,j}!\, l_{i,j}!}} (1 + O(n^{3\eta-1}/\delta))),\end{aligned}$$

and the numerator is bounded as in (7.16). Then, under the assumption $|\mathbf{m}| \geq |\mathbf{l}|$ we have

$$\begin{aligned}|\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda\rangle| \leq (C|\mathbf{m}|)^{|\mathbf{m}|-|\mathbf{l}|} \left(\frac{C|\mathbf{m}|^3}{\delta^2 n}\right)^{\Gamma_{min}} \cdot \prod_{i<j} (\lambda_i - \lambda_j)^{(l_{i,j}-m_{i,j})/2} \sqrt{\frac{m_{i,j}!}{l_{i,j}!}} \cdot \\ \times \left(1 + \left(O(n^{3\eta-1}/\delta)\right)\right),\end{aligned}$$

where $\Gamma_{min} = ((|\mathbf{l} - \mathbf{m}| + 3|\mathbf{l}| - 3|\mathbf{m}|)/6) \wedge 0$. The factorials can be bounded as

$$\prod_{i<j} \sqrt{\frac{m_{i,j}!}{l_{i,j}!}} \leq |\mathbf{m}|^{\sum (m_{i,j}-l_{i,j})_+/2} = |\mathbf{m}|^{(|\mathbf{m}-\mathbf{l}|+|\mathbf{m}|-|\mathbf{l}|)/4}.$$

Since $|\mathbf{m}| \leq n^\eta$ and $Cn^{3\eta-1}/\delta^2 < 1$, we have

$$\left(\frac{C|\mathbf{m}|^3}{\delta^2 n}\right)^{\Gamma_{min}} \leq \left(\frac{Cn^{3\eta-1}}{\delta^2}\right)^{(|\mathbf{l}-\mathbf{m}|+3|\mathbf{l}|-3|\mathbf{m}|)/6}.$$

Since $\lambda_i - \lambda_j > n\delta$ we have

$$\prod_{1 \leq i < j \leq d} (\lambda_i - \lambda_j)^{(l_{i,j}-m_{i,j})/2} \leq (n\delta)^{(|\mathbf{l}|-|\mathbf{m}|)/2}.$$

The constant $C = C(d)$ can be replaced by another constant $C' = C'(d, \eta)$ such that all powers of n appear in the form $(C'n)^{\gamma}$. Putting the bounds together we get

$$|\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda \rangle| \leq \delta^{(|\mathbf{m}|-|\mathbf{l}|)/2-|\mathbf{m}-\mathbf{l}|/3}(C'n)^{-\eta(|\mathbf{m}|-|\mathbf{l}|)/4}(C'n)^{(9\eta-2)|\mathbf{m}-\mathbf{l}|/12} \times(1 + O(n^{-1+3\eta}/\delta)).$$

□

Corollary 7.4. *Let $\eta < 2/9$ and let $(\mathbf{m}, \lambda)$ be such that $|\mathbf{m}| \leq n^{\eta}$. Assume as in Lemma 7.3 that $\lambda_i - \lambda_{i+1} > \delta n$ for all $1 \leq i \leq d-1$ and $\lambda_d > \delta n$, for some $\delta > 0$, and that $Cn^{3\eta-1}/\delta^2 < 1$, where $C = C(d)$ is a constant.*
Then there exists a constant $C'' = C''(d, \eta)$ such that

$$\sum_{\substack{|\mathbf{l}| \leq n^{\eta} \\ \mathbf{l} \neq \mathbf{m}}} |\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda \rangle| \leq (C''n)^{(9\eta-2)/12}\delta^{-1/3}. \tag{7.26}$$

Proof. Recall that the bound (7.23) is given for $|\mathbf{m}| \geq |\mathbf{l}|$. If on the contrary $|\mathbf{l}| > |\mathbf{m}|$, we must change all the $|\mathbf{m}-\mathbf{l}|$ into $|\mathbf{l}-\mathbf{m}|$, so that these terms are always positive. Now, they are always in exponents of values less than one. We shall therefore neglect all those terms.

Hence the expression on the left side of (7.26) is bounded from above by

$$2\sum_{k \geq 1} N(k) \left[(C'n)^{(9\eta-2)/12}\delta^{-1/3}\right]^k,$$

where $N(k)$ is the number of $\mathbf{l}$'s for which $|\mathbf{m} - \mathbf{l}| = k$.

Since there are $d(d-1)/2$ pairs $1 \leq i < j \leq d$, there are at most $(k+1)^{d(d-1)/2}$ different choices for the values $\{|l_{i,j} - m_{i,j}| : i < j\}$ satisfying $\sum |l_{i,j} - m_{i,j}| = k$. Moreover, there are $2^{d(d-1)/2}$ sign choices which fix $\mathbf{l} = \{l_{i,j}\}$ completely. Thus $N(k) \leq (2(k+1))^{d(d-1)/2} \leq c^k$ for some constant c which can be incorporated in the geometric series starting at $k = 1$, hence the desired estimate. □

We use this quasi-orthogonality to build an isometry $V_{\lambda} : \mathcal{H}_{\lambda} \to \mathcal{F}$ which maps the relevant finite-dimensional vectors $|\mathbf{m}, \lambda\rangle$ 'close' to their Fock counterparts $|\mathbf{m}\rangle$. This is the aim of Lemma 5.4.

Lemma 7.5. *Let A be a contraction (i.e. $A^*A \leq \mathbf{1}$) from a finite space $\mathcal{H}$ to an infinite space $\mathcal{K}$. Then there is an $R : \mathcal{H} \to \mathcal{K}$ such that $A + R$ is an isometry and* Range$(A) \perp$ Range(R). *As a consequence, for any unit vector f, we have $\|Rf\|^2 = 1 - \|Af\|^2$.*

Proof. As $\mathcal{K}$ is infinite-dimensional, we may consider a subspace $\mathcal{H}'$ of $\mathcal{K}$, orthogonal to Range(A), and the same dimension as $\mathcal{H}$, so that we can find an isomorphism I from $\mathcal{H}$ to $\mathcal{H}'$. We then take $R = I\sqrt{\mathbf{1} - A^*A}$. □

Proof of Lemma 5.4. Let $A_{\lambda} : \mathcal{H}_{\lambda} \to \mathcal{F}$ be defined by

$$A_{\lambda} := \frac{1}{\sqrt{1 + (Cn)^{(9\eta-2)/12}/\delta^{1/3}}} \sum_{|\mathbf{l}| \leq n^{\eta}} |\mathbf{l}\rangle \langle \mathbf{l}, \lambda| .$$

Then,

$$A_\lambda^* A_\lambda = \frac{1}{1+(Cn)^{(9\eta-2)/12}/\delta^{1/3}} \sum_{|\mathbf{l}|\le n^\eta} |\mathbf{l},\lambda\rangle\langle\mathbf{l},\lambda| \le \mathbf{1}_{\mathcal{H}_\lambda}.$$

where the last inequality follows from Corollary 7.4 and the following argument. It is enough to show that all eigenvalues of $A_\lambda^* A_\lambda$ are smaller than 1. Let $\sum_{\mathbf{m}} c_{\mathbf{m}} |\mathbf{m},\lambda\rangle$ be an eigenvector of $A_\lambda^* A_\lambda$, and a the corresponding eigenvalue. Then by the linear independence of $|\mathbf{m},\lambda\rangle$ we get that for each $\mathbf{l}$,

$$\frac{1}{1+(Cn)^{(9\eta-2)/12}/\delta^{1/3}} \sum_{|\mathbf{m}|\le n^\eta} \langle\mathbf{l},\lambda|\mathbf{m},\lambda\rangle\, c_{\mathbf{m}} = a c_{\mathbf{l}}.$$

If $\mathbf{l}_0$ is an index for which $|c_{\mathbf{l}}|$ is maximum, then by taking absolute values on both sides we obtain

$$a \le \frac{1}{1+(Cn)^{(9\eta-2)/12}/\delta^{1/3}} \sum_{|\mathbf{m}|\le n^\eta} |\langle\mathbf{l},\lambda|\mathbf{m},\lambda\rangle| \le 1.$$

Now we may apply Lemma 7.5, and find an R_λ such that $A_\lambda + R_\lambda$ is an isometry, and $\mathrm{Range}(R_\lambda) \perp \mathrm{Range}(A)$, so that $\langle\mathbf{m}|\, R_\lambda = 0$. We define $V_\lambda := A_\lambda + R_\lambda$. Then

$$\begin{aligned}\langle\mathbf{m}|\, V_\lambda &= \langle\mathbf{m}|\,(A_\lambda + R_\lambda) = \langle\mathbf{m}|\, A_\lambda \\ &= \frac{1}{\sqrt{1+(Cn)^{(9\eta-2)/12}/\delta^{1/3}}} \langle\mathbf{m}| \sum_{|\mathbf{l}|\le n^\eta} |\mathbf{l}\rangle\langle\mathbf{l},\lambda| \\ &= \frac{1}{\sqrt{1+(Cn)^{(9\eta-2)/12}/\delta^{1/3}}} \langle\mathbf{m},\lambda|\,.\end{aligned}$$

□

Recall By Lemma 7.3 we have $\langle\mathbf{m},\lambda|\mathbf{l},\lambda\rangle = 0$ if $m_i \ne l_i$ for some i, where m_i in the total number of i in $\mathbf{m}$ (cf. (7.38)). In particular, $|\mathbf{0},\lambda\rangle$ is orthogonal on all other basis vectors. This means that we can choose the isometry V_λ to satisfy $V_\lambda|\mathbf{0},\lambda\rangle = |\mathbf{0}\rangle$, and such that the relation above holds for all $0 < |\mathbf{m}| \le n^\eta$.

7.3. Proof of Lemma 6.4 on mapping rotations into displacements. We first recall a few definitions and notations. We denote by $D^{\vec{z}}$ the displacement operation (super-operator) acting on observables in the multimode Fock space $\mathcal{F}$ as

$$D^{\vec{z}}(W(\vec{y})) := \mathrm{Ad}[W(\vec{z})]\,(W(\vec{y})) = e^{2i\sigma(\vec{y},\vec{z})}\, W(\vec{z}+\vec{y}), \quad \vec{y},\vec{z} \in \mathbb{C}^{d(d-1)/2}.$$

The operation acts as displacement on coherent states, in particular

$$D^{\vec{\zeta}+\vec{z}}(|\mathbf{0}\rangle\langle\mathbf{0}|) = |\vec{\zeta}+\vec{z}\rangle\langle\vec{\zeta}+\vec{z}|.$$

Similarly, on the finite dimensional space $\left(\mathbb{C}^d\right)^{\otimes n}$ we have the action (cf. (7.1))

$$\Delta^{\vec{\zeta},\vec{\xi},n}(A) = \mathrm{Ad}[U(\vec{\zeta},\vec{\xi},n)](A) := U(\vec{\zeta}/\sqrt{n},\vec{\xi}/\sqrt{n})^{\otimes n}\, A\, U^*(\vec{\zeta}/\sqrt{n},\vec{\xi}/\sqrt{n})^{\otimes n},$$

whose restriction to the block λ is $\Delta_\lambda^{\vec{\zeta},\vec{\xi},n} = \mathrm{Ad}[U_\lambda(\vec{\zeta},\vec{\xi},n)]$.

The isometric embedding $T_\lambda(\cdot) := V_\lambda \cdot V_\lambda^*$ and its 'adjoint' $T_\lambda^*(\cdot) := V_\lambda^* \cdot V_\lambda$ satisfy

$$T_\lambda \Delta_\lambda^{\vec{\zeta}+\vec{z},\vec{\xi},n} T_\lambda^*(|\mathbf{0}\rangle\langle\mathbf{0}|) = V_\lambda|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle\langle\vec{\zeta}+\vec{z},\vec{\xi},\lambda|V_\lambda^*,$$

where $|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle := U_\lambda(\vec{\zeta}+\vec{z},\vec{\xi},n)\,|\mathbf{0},\lambda\rangle$ are the 'finite dimensional coherent states'.

According to Lemma 5.4, the coordinates of $V_\lambda|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle$ in the Fock basis are described by:

$$\langle\mathbf{m}|V_\lambda|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle = \begin{cases} \langle\mathbf{m},\lambda|U_\lambda(\vec{\zeta}+\vec{z},\vec{\xi},n)|\mathbf{0},\lambda\rangle(1+O(n^{(9\eta-2)/12}\delta^{-1/3})) & \text{if } |\mathbf{m}|\leq n^\eta; \\ \text{something not important if } |\mathbf{m}| > n^\eta. \end{cases} \tag{7.27}$$

Using the relation $\||f\rangle\langle f| - |f'\rangle\langle f'|\|_1 = 2\sqrt{1-|\langle f|f'\rangle|^2}$, which holds for unital vectors f, f', the statement of the lemma is equivalent to

$$\sup_{\|\vec{z}\|\leq n^\beta} \sup_{\vec{\zeta}\in\Theta_{n,\beta}} \sup_{\|\vec{\xi}\|\leq n^{-1/2+2\beta/\delta}} \sup_{\lambda\in\Lambda_{n,\alpha}} 1 - \left|\langle\vec{z}+\vec{\zeta}|V_\lambda|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle\right| = R(n)^2, \tag{7.28}$$

with $R(n)$ the original remainder term.

We shall prove formula (7.28) by decomposing these vectors in the Fock basis, that is

$$\langle\vec{\zeta}+\vec{z}|V_\lambda|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle = \sum_{\mathbf{m}} \langle\vec{\zeta}+\vec{z}|\mathbf{m}\rangle\langle\mathbf{m}|V_\lambda|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle. \tag{7.29}$$

The estimates are based on the following observations.

1) The coherent states have significant coefficients $\langle\vec{\zeta}+\vec{z}|\mathbf{m}\rangle$ only for 'small' $\mathbf{m}$'s, i.e. those in the set

$$\mathcal{M} := \{\mathbf{m} : m_{i,j} \leq |(\vec{\zeta}+\vec{z})_{i,j}|^2 n^\epsilon, \quad i<j\}. \tag{7.30}$$

In particular, since $2\beta+\epsilon<\eta$ we have $\mathcal{M}\subset\{\mathbf{m}:|\mathbf{m}|\leq n^\eta\}$.

2) The coefficients $\langle\mathbf{m}|V_\lambda|\vec{\zeta}+\vec{z},\vec{\xi},\lambda\rangle$ are uniformly close to $\exp(i\phi)\langle\vec{\zeta}+\vec{z}|\mathbf{m}\rangle$, where ϕ is a fixed real phase, in particular uniformly over $\mathbf{m}\in\mathcal{M}$.

3) If $a_\mathbf{m}$ and $b_\mathbf{m}$ are the two sets of coefficients, such that $\sum_\mathbf{m}|a_\mathbf{m}|^2 = \sum_\mathbf{m}|b_\mathbf{m}|^2 = 1$, then

$$\begin{aligned} 1 - \left|\sum_\mathbf{m} a_\mathbf{m} b_\mathbf{m}\right| &\leq 1 - \left|\sum_{\mathbf{m}\in\mathcal{M}} a_\mathbf{m} b_\mathbf{m}\right| + \left|\sum_{\mathbf{m}\notin\mathcal{M}} a_\mathbf{m} b_\mathbf{m}\right| \\ &\leq 2\left(1 - \left|\sum_{\mathbf{m}\in\mathcal{M}} a_\mathbf{m} b_\mathbf{m}\right|\right). \end{aligned} \tag{7.31}$$

The precise statement in point 1) is

$$\sum_{\mathbf{m}\notin\mathcal{M}} |\langle\vec{\zeta}+\vec{z}|\mathbf{m}\rangle|^2 \leq d^2 n^{-\beta}. \tag{7.32}$$

Indeed, the inner products can be written as a product over the (i, j) oscillators and we have the bound

$$\sum_{\mathbf{m}\notin\mathcal{M}} |\langle\vec{\zeta}+\vec{z}|\mathbf{m}\rangle|^2 \leq \sum_{i<j}\exp(-x_{i,j})\sum_{k>x_{i,j}n^{\epsilon}}\frac{x_{ij}^k}{k!}, \quad x_{i,j}=|(\vec{\zeta}+\vec{z})_{i,j}|^2.$$

Each of the terms in the sum is a tail of Poisson distribution and is bounded by $n^{-\epsilon n^{\beta}}$ if $x_{i,j}\geq 1$ and by $n^{-\beta}$ if $x_{i,j}<1$.

We turn now to point 2). From the third line of (7.27) we get

$$\begin{aligned}\langle\mathbf{m}|V_\lambda U_\lambda(\vec{\zeta}+\vec{z},\vec{\xi},n)|\mathbf{0},\lambda\rangle &= \frac{\langle y_\lambda f_{\mathbf{m}}|y_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n)|f_{\mathbf{0}}\rangle}{\sqrt{\langle y_\lambda f_{\mathbf{0}}|y_\lambda f_{\mathbf{0}}\rangle}\sqrt{\langle y_\lambda f_{\mathbf{m}}|y_\lambda f_{\mathbf{m}}\rangle}}(1+O(n^{(9\eta-2)/12}\delta^{-1/3}))\\ &= \frac{\langle p_\lambda f_{\mathbf{m}}|q_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n)f_{\mathbf{0}}\rangle}{\sqrt{\langle p_\lambda f_{\mathbf{m}}|q_\lambda p_\lambda f_{\mathbf{m}}\rangle}}(1+O(n^{(9\eta-2)/12}\delta^{-1/3})),\end{aligned}$$

where we have used (5.3) and (5.6).

We recall that $\mathcal{O}_\lambda(\mathbf{m})$ is the orbit in $(\mathbb{C}^d)^{\otimes n}$ of $f_{\mathbf{m}}$ under $\mathcal{R}_\lambda$ and that we have the decomposition

$$p_\lambda f_{\mathbf{m}} = \sum_{f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{m})}\frac{\#\mathcal{R}_\lambda}{\#\mathcal{O}_\lambda(\mathbf{m})}f_{\mathbf{a}}.$$

Then, by employing formulas (7.15) and (7.18), we can write

$$\begin{aligned}&\langle\mathbf{m}|V_\lambda U_\lambda(\vec{\zeta}+\vec{z},\vec{\xi},n)|\mathbf{0},\lambda\rangle\\ &= \frac{\sum_{f_{\mathbf{a}}\in\mathcal{O}_\lambda(\mathbf{m})}\langle f_{\mathbf{a}}|q_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n)f_{\mathbf{0}}\rangle}{\sqrt{\sum_{f_{\mathbf{a}},f_{\mathbf{b}}\in\mathcal{O}_\lambda(\mathbf{m})}\langle f_{\mathbf{a}}|q_\lambda f_{\mathbf{b}}\rangle}}(1+O(n^{(9\eta-2)/12}\delta^{-1/3})\\ &= e^{i\phi-\|\vec{\zeta}+\vec{z}\|_2^2/2}\prod_{i\leq j}\frac{(\vec{\zeta}+\vec{z})_{i,j}^{m_{i,j}}}{\sqrt{m_{i,j}!}}\left(\frac{n(\mu_i-\mu_j)}{\lambda_i-\lambda_j}\right)^{m_{i,j}/2}r(n).\end{aligned} \tag{7.33}$$

The corresponding remainder term is

$$\begin{aligned}r(n) = 1+O\Big(&n^{(9\eta-2)/12}\delta^{-1/3}, n^{-1+2\beta+\eta}\delta^{-1}, n^{-1/2+3\beta+2\epsilon}\delta^{-3/2}, n^{-1+\alpha+2\beta}\delta^{-1},\\ &n^{-1+\alpha+\eta}\delta^{-1}, n^{-1+3\eta}\delta^{-1}\Big)\end{aligned}$$

and the phase is:

$$\phi = \sqrt{n}\sum_{i=1}^{d-1}(\mu_i-\mu_{i+1})\xi_i.$$

Since $\lambda\in\Lambda_{n,\alpha}$ and the eigenvalues are separated by δ we have $\left(\frac{n(\mu_i-\mu_j)}{\lambda_i-\lambda_j}\right)^{m_{i,j}/2} = 1+O(n^{\alpha-1+\eta}/\delta)$ and the error can be absorbed in $r(n)$.

In conclusion, for $\mathbf{m}$ satisfying (7.30), we have:

$$\langle\mathbf{m}|V_\lambda U(\vec{\zeta}+\vec{z},\vec{\xi},n)|\mathbf{0},\lambda\rangle = \exp(i\phi)\langle\mathbf{m}|\vec{\zeta}+\vec{z}\rangle r(n).$$

Inserting this result into (7.29), and using (7.31) and (7.32), we get

$$1 - \left|\langle \vec{z} + \vec{\zeta}\,|V_\lambda U(\vec{\zeta} + \vec{z}, \vec{\xi}, n)|\mathbf{0}, \lambda\rangle\right| = O\left(1 - r(n), \sum_{\mathbf{m}\notin\mathcal{M}} |\langle \mathbf{m}|\vec{\zeta} + \vec{z}\rangle|^2\right) = R_2(n),$$

with

$$R_2(n) = O\left(n^{(9\eta-2)/12}\delta^{-1/3}, n^{-1+2\beta+\eta}\delta^{-1}, n^{-1/2+3\beta+2\epsilon}\delta^{-3/2}, n^{-1+\alpha+2\beta}\delta^{-1}, n^{-1+\alpha+\eta}\delta^{-1}, n^{-1+3\eta}\delta^{-1}, n^{-\beta}\right).$$

Through expression (7.28), noticing that $R_2(n) = R(n)^2$, we see that we have proved the lemma. □

7.4. Proof of Lemma 6.2 on typical Young diagrams. Recall that the state $\rho^{\theta,n} := \rho_{\theta/\sqrt{n}}^{\otimes n}$ has the decomposition over 'blocks' λ given by (4.8). The probability distribution over Young diagrams $p_\lambda^{\vec{\zeta},\vec{u},n}$ depends only on the diagonal parameters $\vec{u}$ and is given by

$$p_\lambda^{\vec{\zeta},\vec{u},n} = c_n^\lambda \sum_{\mathbf{m}\in\lambda} \prod_{i=1}^d (\mu_i^{\vec{u},n})^{\lambda_i} \prod_{j=i+1}^d \left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}},$$

with

$$c_n^\lambda = \binom{n}{\lambda_1, \lambda_2, \ldots, \lambda_d} \prod_{l=1}^d \frac{\lambda_l! \prod_{k=l+1}^d (\lambda_l - \lambda_k + k - l)}{(\lambda_l + d - l)!}.$$

The above formula can be understood as follows. By invariance under rotations we can take $\vec{\zeta} = 0$ and the state is diagonal in the standard basis $\left(\mathbb{C}^d\right)^{\otimes n}$ formed by the vector $f_{\mathbf{a}}$. Each eigenprojector carries a weight $\prod_{i=1}^d (\mu^{\vec{u},n})^{m_i}$, where m_i is the multiplicity of the vector f_i in the tensor product $f_{\mathbf{a}}$. Thus, we only need to add all multiplicities over vectors that are 'inside' the block λ. Since the irreducible representation has basis $f_{\mathbf{m}}$ labelled by semistandard Young tableaux, we get a factor

$$\prod_{i=1}^d (\mu_i^{\vec{u},n})^{m_i} = \prod_{i=1}^d (\mu_i^{\vec{u},n})^{\lambda_i} \prod_{j=i+1}^d \left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}}.$$

The additional factor c_n^λ is the dimension of $\mathcal{K}_\lambda$, on which the state is proportional to the identity.

Recall that $\mu_i^{\vec{u},n} = \mu_i + u_i/\sqrt{n}$ for $1 \le i \le (d-1)$ and $\mu_d^{\vec{u},n} = \mu_d - (\sum_i u_i)/\sqrt{n}$. If $\delta \ge 2dn^{\alpha-1} \ge 2dn^{\gamma-1/2}$, then $\mu_j^{\vec{u},n}/\mu_i^{\vec{u},n} \le 1$ for all $\|\vec{u}\| \le n^\gamma$. Moreover $m_{i,j} \le n$ for all (i, j), so the total number of $\mathbf{m}$'s is smaller than n^{d^2}. Thus

$$\sum_{\mathbf{m}} \prod_{i<j} (\mu^{\vec{u},n})^{\lambda_i} \left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}} \le n^{d^2}.$$

On the other hand $\mathbf{m} = \mathbf{0}$ is always in the set of possible $\mathbf{m}$, so that

$$\sum_{\mathbf{m}} \prod_{i<j} \left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}} \right)^{m_{i,j}} \geq 1.$$

One can easily verify that

$$1 \geq \prod_{l=1}^{d} \frac{\lambda_l! \prod_{k=l+1}^{d} (\lambda_l - \lambda_k + k - l)}{(\lambda_l + d - l)!} \geq \frac{1}{(n+d)^{d^2}}.$$

The remaining factor is the multinomial law. We now show that this is the dominating part. Let us write $(Y_1, \ldots, Y_d)$ for the multinomial random variable. Then we have

$$\mathbb{P}[|Y_i - n\mu_i^{\vec{u},n}| \geq x] \leq 2 \exp\left(-\frac{2x^2}{n}\right). \tag{7.34}$$

Indeed each Y_i is a sum of independent Bernoulli variables $X_1, \ldots, X_n$ with $\mathbb{P}(X_k = 1) = \mu_i^{\vec{u},n}$ and $\mathbb{P}(X_k = 0) = 1 - \mu_i^{\vec{u},n}$, and by Hoeffding's inequality [44],

$$\mathbb{P}[|\sum_{k=1}^{n} X_k - \mathbb{E}[X_k]| \geq x] \leq 2 \exp\left(-\frac{2x^2}{n}\right).$$

By definition, for any $\lambda \notin \Lambda_{n,\alpha}$ there exists an i such that $|\lambda_i - n\mu_i| \geq n^\alpha$, which implies $|\lambda_i - n\mu_i^{\vec{u},n}| \geq n^\alpha - dn^{\gamma+1/2}$. With $n^{\alpha-\gamma-1/2} > 2d$, the upper bound is simply $n^\alpha/2$ and we have

$$\begin{aligned} \sum_{\lambda \notin \Lambda n,\alpha} \|b_\lambda^{\theta,n}\|_1 = \mathbb{P}[\lambda \notin \Lambda_{n,\alpha}] &\leq n^{d^2} \sum_{i=1}^{d} \mathbb{P}[|Y_i - n\mu_i^{\vec{u},n}| \geq n^\alpha/2] \\ &\leq 2dn^{d^2} \exp(-n^{2\alpha-1}/2). \end{aligned}$$

□

7.5. Proof of Lemma 6.1 and Lemma 6.8 on classical LAN. We shall use multinomials as an intermediate step. Recalling that $b_\lambda^{\theta,n} = p_\lambda^{\theta,n} \tau_\lambda^n$, we can write:

$$\begin{aligned} \|\mathcal{N}(\vec{u}, V_\mu) - \sum_\lambda b_\lambda^{\theta,n}\|_1 \leq \|p^{\theta,n} - M^n_{\mu_1^{\vec{u},n},\ldots,\mu_d^{\vec{u},n}}\|_1 + \|\mathcal{N}(\vec{u}, V_\mu) \\ - \sum_\lambda M^n_{\mu_1^{\vec{u},n},\ldots,\mu_d^{\vec{u},n}}(\lambda) \tau_\lambda^n\|_1, \end{aligned} \tag{7.35}$$

where $M^n_{\mu_1^{\vec{u},n},\ldots,\mu_d^{\vec{u},n}}$ is the d-multinomial with coefficients $\mu_i^{\vec{u},n}$.

Concisely, what we really prove in this lemma is the asymptotic equivalence of the following classical experiments, together with an explicit rate:

$$\mathcal{P}_n = \left\{ p^{\vec{u},n}, \|\vec{u}\| \leq n^\gamma \right\},$$

$$\mathcal{M}_n = \left\{ M^n_{\mu_1^{\vec{u},n},\ldots,\mu_d^{\vec{u},n}}, \|\vec{u}\| \leq n^\gamma \right\},$$

$$\mathcal{G}_n = \left\{ \mathcal{N}(\vec{u}, V_\mu), \|\vec{u}\| \leq n^\gamma \right\}.$$

The equivalence of $\mathcal{M}_n$ and $\mathcal{G}_n$ is well known [31] and that of $\mathcal{P}_n$ and $\mathcal{M}_n$ can be treated similarly to the $d = 2$ case [16]. Complete details of the calculations leading to the desired rate of convergence for Lemma 6.1 can be found in [29].

From here, proving Lemma 6.8 (that is the inverse direction) is easy enough. Indeed, recall that $\sigma^n \tau^n p^{\theta,n} = p^{\theta,n}$ and that σ^n is a contraction. Then

$$\begin{aligned}\|\sigma^n \mathcal{N}(\vec{u}, V_\mu) - p^{\vec{\zeta},\vec{u},n}\|_1 &= \|\sigma^n \mathcal{N}(\vec{u}, V_\mu) - \sigma^n \tau^n p^{\vec{\zeta},\vec{u},n}\|_1 \\ &\leq \|\mathcal{N}(\vec{u}, V_\mu) - \tau^n p^{\vec{\zeta},\vec{u},n}\|_1.\end{aligned}$$

Thus we have the same speed and conditions as those of Lemma 6.1. □

7.6. Proof of Lemma 6.3 on convergence to the thermal equilibrium state. We recall that the state ϕ on $CCR(L^2(\rho), \sigma)$ was defined in (4.28) and is the product of a classical Gaussian distribution and $d(d-1)/2$ Gaussian states $\phi_{i,j}$ of quantum harmonic oscillators, one for each pair $i < j$. $\phi_{i,j}$ are thermal equilibrium states with inverse temperature $\beta = \ln(\mu_i/\mu_j)$ (cf. (4.14)). The joint state $\phi^{\vec{0}} := \bigotimes_{i<j} \phi_{i,j}$ is then displaced to obtain $\phi^{\vec{\zeta}}$ but Lemma 6.3 is only concerned with $\phi^{\vec{0}}$.

It is well known that thermal equilibrium states are diagonal in the number basis and in our case

$$\phi^{\vec{0}} = \sum_{\mathbf{m}\in\mathbb{N}^{d(d-1)/2}} \prod_{i<j} \frac{\mu_i - \mu_j}{\mu_i} \left(\frac{\mu_j}{\mu_i}\right)^{m_{i,j}} |\mathbf{m}\rangle\langle\mathbf{m}|. \tag{7.36}$$

As shown in (5.8), a similar formula holds for the finite dimensional block states $\rho_\lambda^{\vec{0},\vec{u},n}$:

$$\langle \mathbf{m}, \lambda | \rho_\lambda^{\vec{0},\vec{u},n} | \mathbf{m}, \lambda \rangle = C_\lambda^{\vec{u}} \prod_{i<j}^{d} \left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}}, \tag{7.37}$$

where $C_\lambda^{\vec{u}}$ is a normalisation constant, $\mu_i^{\vec{u},n} = \mu_i + u_i/\sqrt{n}$ for $1 \leq i \leq (d-1)$ and $\mu_d^{\vec{u},n} = \mu_d - (\sum_i u_i)/\sqrt{n}$.

However there is a caveat: although $|\mathbf{m}, \lambda\rangle$ are eigenvectors of $\rho_\lambda^{\vec{0},\vec{u},n}$, they are not orthogonal to each other so we cannot directly use $|\mathbf{m}, \lambda\rangle\langle\mathbf{m}, \lambda|$ as eigenprojectors in the spectral decomposition. However, Lemma 5.4 gives us an estimate of the error that we incur by doing just that.

Note first that the eigenvalues of $\rho_\lambda^{\vec{0},\vec{u},n}$ are labelled by the *total* multiplicities m_i of the index i in the semistandard Young tableaux:

$$m_i := \lambda_i - \sum_{j>i} m_{i,j} + \sum_{j<i} m_{j,i}. \tag{7.38}$$

By Lemma 7.3 we have that $\langle \mathbf{m}, \lambda | \mathbf{l}, \lambda \rangle = 0$ if $|\mathbf{m}| \neq |\mathbf{l}|$. This allows us to split $\mathcal{H}_\lambda$ into a direct sum of orthogonal subspaces

$$\mathcal{H}_{\lambda,\eta} := \mathrm{Lin}\{|\mathbf{m}, \lambda\rangle : |\mathbf{m}| \leq n^\eta\}, \quad \text{and} \quad \mathcal{H}_{\lambda,\eta}^{\perp} := \mathrm{Lin}\{|\mathbf{l}, \lambda\rangle : |\mathbf{l}| > n^\eta\},$$

and similarly for the Fock space $\mathcal{F} = \mathcal{F}_\eta \oplus \mathcal{F}_\eta^\perp$. Note the hidden dependence on n in the definition of the subspaces. Asymptotically, the state $\rho_\lambda^{\vec{0},\vec{u},n}$ and $\phi^{\vec{0}}$ concentrate on the 'low excitations' spaces $\mathcal{H}_{\lambda,\eta}$ and $\mathcal{F}_\eta$ with corresponding orthogonal projections $P_{\lambda,\eta}$ and P_η, respectively. More precisely,

$$\begin{aligned}\|T_\lambda(\rho_\lambda^{\vec{0},\vec{u},n}) - \phi^{\vec{0}}\|_1 &= \|T_\lambda(P_{\lambda,\eta}\rho_\lambda^{\vec{0},\vec{u},n}P_{\lambda,\eta}) - P_\eta\phi^{\vec{0}}P_\eta\|_1 + \|T_\lambda(P_{\lambda,\eta}^\perp\rho_\lambda^{\vec{0},\vec{u},n}P_{\lambda,\eta}^\perp) \\ &\quad - P_\eta^\perp\phi^{\vec{0}}P_\eta^\perp\|_1 \\ &\le 2\|T_\lambda(P_{\lambda,\eta}\rho_\lambda^{\vec{0},\vec{u},n}P_{\lambda,\eta}) - P_\eta\phi^{\vec{0}}P_\eta\|_1 + 2\|P_\eta^\perp\phi^{\vec{0}}P_\eta^\perp\|_1. \end{aligned} \tag{7.39}$$

From Definition 4.17 of $\phi^{\vec{0}}$ and that of thermal states (4.13) we see that the second term on the right side of order $\max_{j<k}(\mu_k/\mu_j)^{n^\eta} = O(\exp(-\delta n^\eta))$. For the rest of the proof we shall deal with the first term on the right side.

Let us denote by $\mathcal{H}(\{m_i\}) = \mathrm{Lin}\{|\mathbf{l} : \lambda\rangle, l_i = m_i\}$ the eigenspace of $\rho_\lambda^{\vec{0},\vec{u},n}$ and $P(\{m_i\})$ the corresponding eigenprojection. Then

$$\rho_\lambda^{\vec{0},\vec{u},n} = C_\lambda^{\vec{u}} \sum_{\{m_i\}} \prod_{i=1}^d (\mu_i^{\vec{u},n})^{m_i-\lambda_i} P(\{m_i\}).$$

As in Lemma 5.4 we have that for $|\mathbf{m}| \le n^\eta$,

$$P(\{m_i\}) = \frac{1}{1 + Cn^{(9\eta-2)/12}\delta^{-1/3}} \sum_{\mathbf{m}:\{m_i\}} |\mathbf{m}, \lambda\rangle\langle\mathbf{m}, \lambda| + E(\{m_i\}),$$

where the sum runs over those $\mathbf{m}$ with total multiplicities $\{m_i\}$. The (positive) remainder has trace norm

$$\mathrm{Tr}(E(\{m_i\})) = O(n^{(9\eta-2)/12}\delta^{-1/3}) \cdot \dim(\mathcal{H}(\{m_i\})).$$

By summing over all $\{m_i\}$ satisfying $|\mathbf{m}| \le n^\eta$ we get

$$P_{\lambda,\eta}\rho_\lambda^{\vec{0},\vec{u},n}P_{\lambda,\eta} = \frac{1}{1 + Cn^{(9\eta-2)/12}\delta^{-1/3}} \tilde{\rho}_\lambda^{\vec{0},\vec{u},n} + C_\lambda^{\vec{u}} \sum_{\{m_i\}} \prod_{i=1}^d (\mu_i^{\vec{u},n})^{m_i-\lambda_i} E(\{m_i\}),$$

where $\tilde{\rho}_\lambda^{\vec{0},\vec{u},n}$ is the approximate state

$$\tilde{\rho}_\lambda^{\vec{0},\vec{u},n} := C_\lambda^{\vec{u}} \sum_{\{m_i\}} \prod_{i=1}^d (\mu_i^{\vec{u},n})^{m_i-\lambda_i} \sum_{\mathbf{m}:\{m_i\}} |\mathbf{m}, \lambda\rangle\langle\mathbf{m}, \lambda|.$$

The error term has trace norm of the order

$$O(n^{(9\eta-2)/12}\delta^{-1/3}) \cdot C_\lambda^{\vec{u}} \sum_{\{m_i\}} \prod_{i=1}^d (\mu_i^{\vec{u},n})^{m_i-\lambda_i} \dim(\mathcal{H}(\{m_i\})) = O(n^{(9\eta-2)/12}\delta^{-1/3}),$$

where we have used the normalisation of the block state $\rho_\lambda^{\vec{0},\vec{u},n}$.

In conclusion

$$\|P_{\lambda,\eta}\rho_\lambda^{\vec{0},\vec{u},n}P_{\lambda,\eta}-\tilde{\rho}_\lambda^{\vec{0},\vec{u},n}\|_1=O(n^{(9\eta-2)/12}\delta^{-1/3}). \tag{7.40}$$

The next step is to show that the block states $\tilde{\rho}_\lambda^{\vec{0},\vec{u},n}$ are mapped by T_λ close to $P_\eta\phi^{\vec{0}}P_\eta$. Using (5.8), we can write

$$T_\lambda(\tilde{\rho}_\lambda^{\vec{0},\vec{u},n})=C_\lambda^{\vec{u}}\sum_{\mathbf{m}\in\lambda}\prod_{i<j}\left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}}T_\lambda(|\mathbf{m},\lambda\rangle\langle\mathbf{m},\lambda|). \tag{7.41}$$

If $n^{\alpha-1}\leq\delta/2$ and $\alpha>1/2>\eta$, we know that all $\mathbf{m}$ such that $|\mathbf{m}|\leq n^\eta$ 'fit into' λ.

Since $\mu_i^{\vec{u},n}=\mu_i+O(n^{-1/2+\gamma})$, when $|\mathbf{m}|\leq n^\eta$,

$$\left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}}=\left(\frac{\mu_j}{\mu_i}\right)^{m_{i,j}}(1+O(n^{-1/2+\gamma+\eta}/\delta)). \tag{7.42}$$

For the normalisation constant we can write:

$$(C_\lambda^{\vec{u}})^{-1}=\sum_{|\mathbf{m}|\leq n^\eta}\prod_{i<j}\left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}}+\sum_{\mathbf{m}\in\lambda:|\mathbf{m}|\geq n^\eta}\prod_{i<j}\left(\frac{\mu_j^{\vec{u},n}}{\mu_i^{\vec{u},n}}\right)^{m_{i,j}}.$$

If $2dn^{\gamma-1/2}<\delta/2$ then the second part is less than $n^{d^2}(1-\delta/2)^{n^\eta}$ which is negligible compared to the other error terms. Hence:

$$\begin{aligned}(C_\lambda^{\vec{u}})^{-1}&=\sum_{|\mathbf{m}|\leq n^\eta}\prod_{i<j}\left(\frac{\mu_j}{\mu_i}\right)^{m_{i,j}}(1+O(n^{-1/2+\gamma+\eta}/\delta))\\&=\sum_{\mathbf{m}\in\mathbb{N}^{d(d-1)/2}}\prod_{i<j}\left(\frac{\mu_j}{\mu_i}\right)^{m_{i,j}}(1+O(n^{-1/2+\gamma+\eta}/\delta)\\&=\prod_{i<j}\frac{\mu_i}{\mu_i-\mu_j}(1+O(n^{-1/2+\gamma+\eta}/\delta)).\end{aligned} \tag{7.43}$$

We then recall that for unit vectors, we have $\||f\rangle\langle f|-|f'\rangle\langle f'|\|_1=2\sqrt{1-|\langle f|f'\rangle|^2}$. So that, using Lemma 5.4, we get that for $|\mathbf{m}|\leq n^\eta$,

$$\begin{aligned}\|T_\lambda(|\mathbf{m},\lambda\rangle\langle\mathbf{m},\lambda|)-|\mathbf{m}\rangle\langle\mathbf{m}|\|_1&=\|V_\lambda|\mathbf{m},\lambda\rangle\langle\mathbf{m},\lambda|V_\lambda^*-|\mathbf{m}\rangle\langle\mathbf{m}|\|_1\\&=O(n^{(9\eta-2)/24}/\delta^{1/6}).\end{aligned} \tag{7.44}$$

Putting the estimates (7.42), (7.43), (7.44) back into formula (7.41), we obtain $T_\lambda(\tilde{\rho}_\lambda^{\vec{0},\vec{u},n})$, so that

$$\begin{aligned}T_\lambda(\tilde{\rho}_\lambda^{\vec{0},\vec{u},n})=&\sum_{|\mathbf{m}|\leq n^\eta}\prod_{i<j}\frac{\mu_i-\mu_j}{\mu_i}\left(\frac{\mu_j}{\mu_i}\right)^{m_{i,j}}|\mathbf{m}\rangle\langle\mathbf{m}|\\&+O(n^{-1/2+\gamma+\eta}/\delta,n^{(9\eta-2)/24}/\delta^{1/6}).\end{aligned} \tag{7.45}$$

Comparing with (7.36), and using (7.39) and (7.40) we get the desired result. □

7.7. Proof of Lemma 6.5 on local linearity of $SU(d)$. The key is to notice that, as we are dealing with a group, there is an r such that

$$U^{-1}(\vec{\zeta}+\vec{z},\vec{0},n)U(\vec{\zeta},\vec{0},n)U(\vec{z},\vec{0},n) = U(-\vec{\zeta}-\vec{z},\vec{0},n)U(\vec{\zeta},\vec{0},n)U(\vec{z},\vec{0},n)$$
$$= U(\vec{r},\vec{s},n),$$

and similarly for the operation Δ. We shall prove below that under the condition that both $\vec{\zeta}$ and $\vec{z}$ are smaller than n^{β}, then $\|\vec{r}\|+\|\vec{s}\| = O(n^{-1/2+2\beta}/\delta)$. Let us call this the *domination hypothesis* for further reference.

Now, as the actions are unitary, we may rewrite the norm in Lemma as 6.5:

$$\begin{aligned} A &= \|[\Delta_{\lambda}^{\vec{\zeta}+\vec{z},n} - \Delta_{\lambda}^{\vec{\zeta},n}\Delta_{\lambda}^{\vec{z},n}](|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|)\|_1 \\ &= \|\Delta_{\lambda}^{-(\vec{\zeta}+\vec{z}),n}[\Delta_{\lambda}^{\vec{\zeta}+\vec{z},n} - \Delta_{\lambda}^{\vec{\zeta},n}\Delta_{\lambda}^{\vec{z},n}](|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|)\|_1 \\ &= \|[\mathrm{Id} - \Delta_{\lambda}^{\vec{r},\vec{s},n}](|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|)\|_1. \end{aligned}$$

As T_λ is an isometry, we may also let it act on the left and T_λ^* on the right and get:

$$\begin{aligned} A &= \|T_\lambda(|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|) - T_\lambda\Delta_{\lambda}^{\vec{r},\vec{s},n}T_\lambda^*(|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1 \\ &\le \|\,|\mathbf{0}\rangle\langle\mathbf{0}| - |\vec{r}\rangle\langle\vec{r}|\,\|_1 \\ &\quad + \|\,|\vec{r}\rangle\langle\vec{r}| - T_\lambda\Delta_{\lambda}^{\vec{r},\vec{s},n}T_\lambda^*(|\mathbf{0}\rangle\langle\mathbf{0}|)\|_1 + \|T_\lambda(|\mathbf{0},\lambda\rangle\langle\mathbf{0},\lambda|) - |\mathbf{0}\rangle\langle\mathbf{0}|\,\|_1. \end{aligned}$$

By the domination hypothesis, the norm of $\vec{r}$ is smaller than $n^{-1/2+2\beta}/\delta$, hence $\langle\vec{r}|\mathbf{0}\rangle = 1 - O(n^{-1+4\beta}/\delta^2)$. Using $\||f\rangle\langle f| - |f'\rangle\langle f'|\|_1 = 2\sqrt{1-|\langle f|f'\rangle|^2}$ we get that the first term on the right side of the inequality is $O(n^{-1/2+2\beta}/\delta)$. Notice that this is dominated by $R(n)$ given in Eq. (6.11) since $\eta > 2\beta$.

For the second term, we apply Lemma 6.4, with $\vec{z}=0$. By the domination hypothesis, $\|\vec{s}\| \le n^{-1/2+2\beta}/\delta$, so we may apply Lemma 6.4, and the remainder is given by $R(n)$ in Eq. (6.11).

The last term is $O(n^{(9\eta-2)/24}/\delta^{1/6})$ as shown in (7.44) which is dominated by $R(n)$.

We finish the proof of the lemma, and simultaneously that of Theorem 4.3, by proving the domination hypothesis. Recall that an arbitrary element in $SU(d)$ can be written in the exponential form

$$U(\vec{r},\vec{s}) := \exp\left[\, i\left(\sum_{i=1}^{d-1} s_i H_i + \sum_{1\le j<k\le d} \frac{\mathrm{Re}(r_{j,k})T_{j,k} + \mathrm{Im}(r_{j,k})T_{k,j}}{\sqrt{\mu_j-\mu_k}}\right)\right],$$

where $(\vec{r},\vec{s}) \in \mathbb{C}^{d(d-1)/2}\times\mathbb{R}^{d-1}$, and $T_{i,j}$, H_i are the generators of $SU(d)$ defined in (7.2). A special case of this is $U(\vec{r}) := U(\vec{r},\vec{0})$. In general, the map $(\vec{r},\vec{s}) \mapsto U(\vec{r},\vec{s})$ is not injective but becomes so if we restrict to a small enough neighbourhood $\mathcal{C}$ of the origin $(0,0) \in \mathbb{C}^{d(d-1)/2}\times\mathbb{R}^{d-1}$. On this neighbourhood it makes sense to define the inverse as a sort of 'logarithm' $\log U(\vec{r},\vec{s}) := (\vec{r},\vec{s})$, which is a C^∞ function.

By continuity of the product, if $\vec{x},\vec{y} \in \mathbb{C}^{d(d-1)/2}$ are small enough, then $U(-\vec{x}-\vec{y})U(\vec{x})U(\vec{y}) \in \mathcal{C}$. Since $\|\vec{\zeta}\|+\|\vec{z}\|/\sqrt{n} \le n^{\beta-1/2}/\delta$, we can apply this to $\vec{x} = \vec{\zeta}/\sqrt{n}$, $\vec{y} = \vec{z}/\sqrt{n}$ for $n > (C/\delta)^{\frac{1}{1/2-\beta}}$ with the constant C depending only on the dimension, and get

$$(\vec{r}/\sqrt{n},\vec{s}/\sqrt{n}) = f(\vec{\zeta}/\sqrt{n},\vec{z}/\sqrt{n}) := \log\Big[U(-(\vec{\zeta}+\vec{z})/\sqrt{n})U(\vec{\zeta}/\sqrt{n})U(\vec{z}/\sqrt{n})\Big].$$

Since f is a C^∞ function we can expand in Taylor series and it is easy to show that $f(\vec{0},\vec{0}) = (\vec{0},\vec{0})$, the first order partial derivatives are zero as well, and the second order derivatives are uniformly bounded in a neighbourhood of the origin. Thus we get

$$\vec{r} = \sqrt{n}\, O\left(\frac{\|z_{i,j}\|^2}{n(\mu_i - \mu_j)}, \frac{\|\zeta_{i,j}\|^2}{n(\mu_i - \mu_j)}\right) = O(n^{-1/2+2\beta}/\delta).$$

□

Acknowledgements. We thank Richard Gill for suggesting the research topic and for many useful discussions. J. Kahn acknowledges the support of the French Agence Nationale de la Recherche (ANR), under grant StatQuant (JC07 205763) "Quantum Statistics". M. Guţă's research is supported by the Engineering and Physical Sciences Research Council (EPSRC).

References

1. Armen, M.A., Au, J.K., Stockton, J.K., Doherty, A.C., Mabuchi, H.: Adaptive Homodyne Measurement of Optical Phase. Phys. Rev. Lett. **89**, 133602 (2002)
2. Artiles, L., Gill, R., Guţă, M.: An invitation to quantum tomography. J. Royal Statist. Soc. B (Methodological) **67**, 109–134 (2005)
3. Audenaert, K.M.R., Nussbaum, M., Szkola, A., Verstraete, F.: Asymptotic Error Rates in Quantum Hypothesis Testing. Commun. Math. Phys. **279**, 251–283 (2008)
4. Bagan, E., Baig, M., Munoz-Tapia, R.: Optimal Scheme for Estimating a Pure Qubit State via Local Measurements. Phys. Rev. Lett. **89**, 277904 (2002)
5. Bagan, E., Ballester, M.A., Gill, R.D., Monras, A., Munõz-Tapia, R.: Optimal full estimation of qubit mixed states. Phys. Rev. A **73**, 032301 (2006)
6. Barndorff-Nielsen, O.E., Gill, R., Jupp, P.E.: On quantum statistical inference (with discussion). J. R. Statist. Soc. B **65**, 775–816 (2003)
7. Belavkin, V.P.: Optimal multiple quantum statistical hypothesis testing. Stochastics **1**, 315–345 (1975)
8. Belavkin, V.P.: Generalized Heisenberg uncertainty relations, and efficient measurements in quantum systems. Theor. Math. Phys. **26**, 213–222 (1976)
9. Cirac, J.I., Ekert, A.K., Macchiavello, C.: Optimal Purification of Single Qubits. Phys. Rev. Lett. **82**, 4344 (1999)
10. Erdélyi, A., Magnus, W., Oberhettinger, F., Tricomi, F.G.: *Higher Transcendental Functions*. vol. **II**. New York-Toronto-London: McGraw-Hill, 1953
11. Fulton, W.: *Young Tableaux, with Applications to Representation Theory and Geometry*. Cambridge: Cambridge University Press, 1997
12. Fulton, W., Harris, J.: *Representation Theory, A First Course*. New York: Springer Verlag, 1991
13. Gill, R.D.: Asymptotic information bounds in quantum statistics. In: *Quantum Stochatics and Information: Statstics Filtering and Control*, Belavin, V.P., Guta, M. (eds.), River Edge, NJ: World Scientific, 2008
14. Gill, R.D., Massar, S.: State estimation for large ensembles. Phys. Rev. A **61**, 042312 (2000)
15. Goodman, R., Wallach, N.R.: *Representations and Invariants of the Classical Groups*. Cambridge: Cambridge University Press, 1998
16. Guţă, M., Janssens, B., Kahn, J.: Optimal estimation of qubit states with continuous time measurements. Commun. Math. Phys. **277**, 127–160 (2008)
17. Guţă, M., Jençová, A.: Local asymptotic normality in quantum statistics. Commun. Math. Phys. **276**, 341–379 (2007)
18. Guţă, M., Kahn, J.: Optimal estimation of d-dimensional quantum states. In preparation
19. Guţă, M., Kahn, J.: Local asymptotic normality for qubit states. Phys. Rev. A **73**, 052108 (2006)
20. Hannemann, T., Reiss, D., Balzer, C., Neuhauser, W., Toschek, P.E., Wunderlich, C.: Self-learning estimation of quantum states. Phys. Rev. A **65**, 050303 (2002)
21. Hayashi, M.: *Presentations at MaPhySto and QUANTOP Workshop on Quantum Measurements and Quantum Stochastics*, Aarhus, 2003, and Special Week on Quantum Statistics, Isaac Newton Institute for Mathematical Sciences, Cambridge, 2004
22. Hayashi, M.: Quantum estimation and the quantum central limit theorem. *Bulletin of the Mathematical Society of Japan* **55**, 368–391 (2003) (in Japanese; Translated into English in quant-ph/0608198)

23. Hayashi, M., (ed.): *Asymptotic Theory of Quantum Statistical Inference: Selected Papers*, River Edge, NJ: World Scientific, 2005
24. Hayashi, M.: *Quantum Information*. Berlin-Heidelberg: Springer-Verlag, 2006
25. Hayashi, M., Matsumoto, K.: Statistical Model with Measurement Degree of Freedom and Quantum Physics. In: Hayashi, M. ed., *Asymptotic theory of quantum statistical inference: selected papers*, River Edge, NJ: World Scientific, 2005, pp. 162–170 (English translation of a paper in Japanese published in Surikaiseki Kenkyusho Kokyuroku, vol. **35**, pp. 7689–7727, 2002)
26. Hayashi, M., Matsumoto, K.: *Asymptotic performance of optimal state estimation in quantum two level system*. http://arxiv.org/abs/quant-ph/0411073v2, 2006
27. Helstrom, C.W.: Quantum detection and estimation theory. J. Stat. Phys. **1**, 231–252 (1969)
28. Holevo, A.S.: *Probabilistic and Statistical Aspects of Quantum Theory*. Amsterdam: North-Holland, 1982
29. Kahn, J.: *Quantum Local Asymptotic Normality and other Questions in Quantum Statistics*. Ph.D. thesis, University of Leiden, 2008
30. Keyl, M., Werner, R.F.: Estimating the spectrum of a density operator. Phys. Rev. A **64**, 052311 (2001)
31. Le Cam, L.: *Asymptotic Methods in Statistical Decision Theory*. New York: Springer Verlag, 1986
32. Massar, S., Popescu, S.: Optimal extraction of information from finite quantum ensembles. Phys. Rev. Lett. **74**, 1259–1263 (1995)
33. Matsumoto, K.: Unpublished manuscript
34. Ohya, M., Petz, D.: *Quantum Entropy and its Use*. Berlin-Heidelberg: Springer Verlag, 2004
35. Paris, M.G.A., Řeháček, J. (eds.): *Quantum State Estimation*, Lecture Notes in Phys. vol. **649**, New York: Springer, 2004
36. Petz, D.: Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. Commun. Math. Phys. **105**, 123–131 (1986)
37. Petz, D.: *An Invitation to the Algebra of Canonical Commutation Relations*. Leuven: Leuven University Press, 1990
38. Petz, D., Jencova, A.: Sufficiency in quantum statistical inference. Commun. Math. Phys. **263**, 259–276 (2006)
39. Schiller, S., Breitenbach, G., Pereira, S.F., Müller, T., Mlynek, J.: Quantum statistics of the squeezed vacuum by measurement of the density matrix in the number state representation. Phys. Rev. Lett. **77**, 2933–2936 (1996)
40. Smith, G.A., Silberfarb, A., Deutsch, I.H., Jessen, P.S.: Efficient Quantum-State Estimation by Continuous Weak Measurement and Dynamical Control. Phys. Rev. Lett. **97**, 180403 (2006)
41. Strasser, H.: *Mathematical Theory of Statistics*. Berlin-New York: De Gruyter, 1985
42. Torgersen, E.: *Comparison of Statistical Experiments*. Cambridge: Cambridge University Press, 1991
43. van der Vaart, A.: *Asymptotic Statistics*. Cambridge: Cambridge University Press, 1998
44. van der Vaart, A., Wellner, J.A.: *Weak Convergence and Empirical Processes*. New York: Springer, 1996
45. Vidal, G., Latorre, J.I., Pascual, P., Tarrach, R.: Optimal minimal measurements of mixed states. Phys. Rev. A **60**, 126 (1999)
46. Wald, A.: *Statistical Decision Functions*. New York: John Wiley & Sons, 1950
47. Yuen, H., Kennedy, R., Lax, M.: Optimum testing of multiple hypotheses in quantum detection theory. IEEE Trans. Inform. Theory **21**, 125–134 (1975)
48. Yuen, H.P., Lax, M.: Multiple-parameter quantum estimation and measurement of non-selfadjoint observables. IEEE Trans. Inform. Theory **19**, 740 (1973)

Communicated by M. B. Ruskai