Bernoulli **30**(1), 2024, 610–635
https://doi.org/10.3150/23-BEJ1610

Minimax estimation of low-rank quantum states and their linear functionals

SAMRIDDHA LAHIRY[1,a] and MICHAEL NUSSBAUM[2,b]

[1]*Department of Statistics, Science Center, Harvard University, Cambridge, MA 02138, USA,*
[a]*slahiry@fas.harvard.edu*
[2]*Department of Mathematics, Malott Hall, Cornell University, Ithaca, NY 14853, USA,*
[b]*nussbaum@math.cornell.edu*

In classical statistics, a well known paradigm consists in establishing asymptotic equivalence between an experiment of i.i.d. observations and a Gaussian shift experiment, with the aim of obtaining optimal estimators in the former complicated model from the latter simpler model. In particular, a statistical experiment consisting of n i.i.d. observations from d-dimensional multinomial distributions can be well approximated by an experiment consisting of $d-1$ dimensional Gaussian distributions. In a quantum version of the result, it has been shown that a collection of n qudits (d-dimensional quantum states) of full rank can be well approximated by a quantum system containing a classical part, which is a $d-1$ dimensional Gaussian distribution, and a quantum part containing an ensemble of $d(d-1)/2$ shifted thermal states. In this paper, we obtain a generalization of this result when the qudits are not of full rank. We show that when the rank of the qudits is r, then the limiting experiment consists of an $r-1$ dimensional Gaussian distribution and an ensemble of both shifted pure and shifted thermal states. For estimation purposes, we establish an asymptotic minimax result in the limiting Gaussian model. Analogous results are then obtained for estimation of a low rank qudit from an ensemble of identically prepared, independent quantum systems, using the local asymptotic equivalence result. We also consider the problem of estimation of a linear functional of the quantum state. We construct an estimator for the functional, analyze the risk and use quantum local asymptotic equivalence to show that our estimator is also optimal in the minimax sense.

Keywords: Functional estimation; low rank states; quantum local asymptotic normality; quantum minimax estimation

1. Introduction

Recent breakthroughs in quantum technology, such as quantum computing, communication, and metrology (cf. [52]) have created renewed interest in the probabilistic and statistical problems that arise from quantum information theory. In particular, the age old tools of mathematical statistics have found their way into the toolbox of scientists working with quantum information.

One of the fundamental problems in quantum statistics is quantum estimation, i.e. estimation of a quantum state or functionals of a state. In the case of state estimation one deals with a statistical inverse problem of inferring unknown state parameters from the measurement data obtained by probing a large number of individual quantum systems. On the other hand, in the case of functional estimation, one is interested in measuring some particular property of the state instead of the whole state. In analogy to the classical decision theoretic approach, one can develop a *quantum decision theoretic framework* for such problems where the "best estimate" also involves optimizing over the measurements. In this direction, a *quantum Cramer-Rao bound* has been obtained in [9,27,28] for the covariance matrix of an unbiased estimator, as a first step in uncovering this framework. Another closely related problem pertains to the discrimination between quantum states, in which case one can devise optimal testing procedures. In the asymptotic setup, the *quantum Stein's lemma* (cf. [44]) and the *quantum Chernoff bound* have been established (cf. [6,7,40,43]). For an overview of the evolving literature on quantum

statistics cf. [8,12,19,21,23] and references therein, and the monographs [26,28,46]. Statistical aspects of quantum algorithms and simulation are discussed in [50,51,53].

In classical statistics, the first step for obtaining optimal decision theoretic procedures often consists in approximating complicated experiments (families of laws, or models) by simpler ones (cf. [39] for details). An important example under this paradigm is called local asymptotic normality (LAN), where one establishes asymptotic equivalence between i.i.d. models indexed by a local parameter and a Gaussian shift model (with the shift given by the same local parameter). To construct an optimal estimator, one constructs a preliminary estimate of the parameter first and then uses LAN in the neighborhood of the estimated value. The optimal procedure in the Gaussian model then ensures an (asymptotically) optimal procedure in the i.i.d. model. Global equivalence has also been shown between several nonparametric estimation problems like nonparametric regression (cf. [10]), density estimation (cf. [42]) and the Gaussian white noise model. The theory of LAN can also be used to show optimality in the estimation of functionals by restating the functional estimation problem as a smooth parametric estimation problem (cf. [32]). In the quantum setup, *quantum LAN theory* for parametric models, established in [23,24,34,54], shows that a model given by a large collection of identically prepared finite dimensional states can be approximated by a quantum Gaussian shift model in a local neighborhood. An extension of quantum LAN theory towards a quantum version of local asymptotic equivalence in nonparametric models can be found in [12] where it is shown that an ensemble of pure states in infinite dimensional Hilbert space can be approximated by coherent states, which constitute a quantum counterpart of the Gaussian white noise model.

Our contribution in this paper is threefold; as a first result we establish LAN for low-rank quantum states. We restrict ourselves to finite dimensional states, in particular to a setup similar to [22,24,34]. In [24] the authors show that a quantum statistical experiment consisting of a large number of qubits (two dimensional quantum states) with rotational shift can be approximated by a quantum Gaussian state, specifically a shifted thermal state, while in [22] the authors improved the result to include diagonal perturbations of the qubits. They showed that the limiting experiment in this case contains a classical part which is Gaussian and a quantum part which is a shifted thermal state. Later in [34] the authors extended the results for qubits to qudits (d-dimensional quantum systems) and showed that a large number of full rank qudits can be well approximated by a quantum experiment consisting of a $d-1$ dimensional "classical" Gaussian part (corresponding to the diagonal elements of the original qudit model) and a quantum part which is a tensor product of $d(d-1)/2$ shifted thermal states (corresponding to the off-diagonal elements of the original qudit model). The latter paper used technical results on symmetric and general linear groups but was crucially dependent on the assumption that the original qudit was of full rank. We extend this result to the case of rank-r qudits and obtain an interesting limiting experiment. While the classical part contains an $r-1$ dimensional classical Gaussian as expected from the degeneracy of the eigenvalues, the quantum part contains a tensor product of both shifted thermal states ($r(r-1)/2$ many) and shifted pure states ($r(d-r)$ many). We observe that for the full rank case there are only shifted thermal states in the quantum part of the limit as in [22] and [34], while for the rank $r=1$ case we only have shifted pure states (without a classical part). On the other hand, one can easily modify the proof of local asymptotic equivalence in [12] to show that the limiting experiment of an ensemble of finite dimensional pure states gives rise to a tensor product of $d-1$ shifted pure states (without a classical part) in the limit. Our result for rank 1 qudits matches this result. Thus we obtain a unifying picture that suitably generalizes previous results on finite dimensional quantum LAN.

Secondly, we address the question of optimal estimation of low-rank qudits using the LAN result. Recently there has been considerable interest in the study of low-rank quantum models, although in most cases the authors study them in the high dimensional low-rank setup and obtain finite sample bounds in contrast to our asymptotic setup. In the framework of low-rank quantum state tomography, the problem is treated in [15,35,36] where the authors consider a trace regression model with the

observables being random and then establish minimax bounds in the estimation of the low-rank state. A similar problem has been studied in [14] but with a sparsity assumption different from the low-rank structure. The noiseless case (also called quantum compressed sensing) has been studied in [17,20,41] and similar methods of quantum tomography (using a small number of randomized measurements to reconstruct a low rank state) have been considered in [1,2]. A spectral thresholding and a rank penalized method are used for the low-rank quantum state tomography problem in [11] and [5] respectively, while [25] uses a projected least square approach. A comparative study of these approaches can be found in [3]. However, most of these methods are only optimal in the rate sense and fail to obtain sharp constants. We employ the methods of [11] and [25] to obtain a preliminary estimator which is both rank consistent and lies close to the original state and then use our LAN result to obtain the minimax risk up to sharp constants. The optimal minimax risk for the full rank case is obtained in [33]; our result for the rank r case matches with former for $r = d$.

Finally, in the spirit of [32,37], we use LAN to address the question of optimal estimation of linear functionals of low-rank quantum states. While property estimation is a well-established concept in quantum information theory (for example see [4] for estimation of von-Neumann entropy), the optimal procedures for estimation of a linear functional of a quantum state are less well known. In this direction, the quantum analog of the concept of a least favorable sub-family is treated in [49]. We construct an estimator of the linear functional and use the latter concept and LAN to show optimality (in the sense of sharp constants) of our estimator. It should be noted that while [49] gives a pointwise optimality result, we establish a minimax result; the classical analog of the latter can be found in [37].

1.1. Outline of results

The paper is organized as follows. In Section 2, we review the basic quantum mechanical concepts of states, measurements, observables, and quantum channels. Section 3 reviews classical parametric LAN and quantum LAN for the full rank case as obtained in [34]. It also contains our main theorem (Theorem 3.3) where we establish LAN for the low-rank model.

Section 4 describes the Bayes estimator in the Gaussian case from [29,30], which will be useful in our construction of the optimal estimator and also in establishing the minimax lower bound. Analogously to the classical case we also observe a shrinkage phenomenon for this Bayes estimator.

In Section 5 we construct the optimal estimator for rank r qudits. First, in Theorem 5.1, using a part of the sample, we obtain an estimator which lies close to the original state and has rank r with high probability. We then use the quantum channel on the remaining sample to transfer the estimation problem of qudits to the estimation of parameters in a limiting Gaussian model which contains a classical Gaussian law and a tensor product of both shifted and pure thermal states in its quantum part. An asymptotic minimax result for the limiting Gaussian model is given in Theorem 5.2. To prove this theorem we use the result from Section 4 to get a lower bound of the minimax risk by the Bayes risk and then use a covariant measurement to give a matching upper bound. Finally, we observe that the Hilbert-Schmidt norm between the two qudits is locally quadratic approximately and then transfer the risk from the Gaussian model to obtain the optimal risk in the low-rank qudit model in Theorem 5.4.

Estimation of a linear functional is treated in Section 6. We construct an estimate of the functional in question using the appropriate observable to give an upper bound. The lower bound is given by constructing a least favorable parametric subfamily and then using LAN to obtain a lower bound in the limiting Gaussian model. The latter is indexed by a one-dimensional parameter and we use a Bayesian result from [30] to give the lower bound.

Some representation theoretic tools are needed for proving Theorem 3.3; they are reviewed in Appendix A. The Bayesian result for a one dimensional parameter is discussed in Appendix B. Proofs of

the main theorems are given in Appendix C while the proofs of the more technical lemmas are deferred to Appendix D. The appendices are included in [38].

1.2. Notation

In physics, the vectors of a Hilbert space $\mathcal{H}$ (assumed separable) are written as "ket" $|v\rangle$, v^* (a vector in the dual space $\mathcal{H}^*$) as "bra" $\langle v|$ and the inner product of two vectors as the "bra-ket" $\langle u|v\rangle \in \mathbb{C}$ which is linear with respect to the right entry and anti-linear with respect to the left entry. Similarly, $M := |u\rangle\langle v|$ is the rank one operator acting as $M : |w\rangle \mapsto M|w\rangle = \langle v|w\rangle|u\rangle$. For an operator A the expression $\langle u|Av\rangle$ will sometimes be denoted as $\langle u|A|v\rangle$. The space of bounded linear operators on $\mathcal{H}$ is denoted by $\mathcal{L}(\mathcal{H})$. Of particular interest are the following two subspaces of $\mathcal{L}(\mathcal{H})$.

1. $\mathcal{T}_1(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$ – the trace class defined by $\mathcal{T}_1(\mathcal{H}) = \{A : \mathcal{H} \to \mathcal{H} : \mathrm{Tr}(A^*A)^{1/2} < \infty\}$. Operators in $\mathcal{T}_1(\mathcal{H})$ are equipped with the norm $\mathrm{Tr}(A^*A)^{1/2}$.
2. $\mathcal{T}_2(\mathcal{H}) \subset \mathcal{L}(\mathcal{H})$ – the Hilbert Schmidt operators defined by $\mathcal{T}_2(\mathcal{H}) = \{A : \mathcal{H} \to \mathcal{H} : \mathrm{Tr}(A^*A) < \infty\}$. Operators in $\mathcal{T}_2(\mathcal{H})$ are equipped with the norm $(\mathrm{Tr}(A^*A))^{1/2}$. The class $\mathcal{T}_2(\mathcal{H})$ is a Hilbert space with respect to the inner product $(A, B) := \mathrm{Tr}(A^*B)$.

It is well known that $\mathcal{T}_1(\mathcal{H}) \subset \mathcal{T}_2(\mathcal{H})$. For any Hilbert space, the usual norm will be denoted by $||.||$ and the identity operator on that space by $\mathbf{1}$ where the particular space will be understood from the context. We will denote by $||\mu - \nu||_{\mathrm{TV}}$ the total variation norm between two measures μ and ν. By $a \vee b$ and $a \wedge b$ we will denote $\max(a, b)$ and $\min(a, b)$ respectively and a_+ will be used to denote $a \vee 0$. By $\lfloor a \rfloor$ and $\lceil a \rceil$, we will denote the largest integer less than or equal to a and the smallest integer greater than or equal to a respectively. We will use the notation $a_n \asymp b_n$ whenever $c < \liminf_n(a_n/b_n) \leq \limsup_n(a_n/b_n) < C$ for some constants $c, C > 0$. Throughout the paper, c and C will denote arbitrary constants.

2. Quantum mechanics preliminaries

The outline for this section is as follows. In Subsection 2.1 we describe the concepts of quantum states, measurements and observables. We also discuss quantum channels which are essential for exchanging information between two quantum systems. In Subsection 2.2 we consider qudits or d-dimensional quantum states; the problem of finding optimal measurements for these states and their linear functionals is the main objective of this article. For optimal estimation of qudits (which will also be called an i.i.d. model) it is convenient to analyze a quantum Gaussian model first and then relate it to the i.i.d. model. Following this path, we describe quantum Gaussian states in Subsection 2.3, in particular we describe a limiting Gaussian model. In Section 3 we state the theorem (Theorem 3.3) which guarantees that the i.i.d. model can be approximated by the Gaussian model described in Subsection 2.3. In subsection 2.4, we describe the general problem of quantum statistical inference and the concepts of Bayes and minimax risk in quantum models, which are used as benchmarks for optimality. We also describe the concept of quantum asymptotic equivalence and how risks can be transferred between two asymptotically equivalent models using quantum channels. In Section 5 we compute the risk in the limiting Gaussian model (Theorem 5.2) and then use the concept of risk transfer together with Theorem 3.3 to obtain the optimal risk in the i.i.d. model. A similar method of risk transfer will also be employed to establish optimality of an estimator of a linear functional in Section 6.

2.1. States, measurements, and observables

A *state of a quantum system* is described by a self-adjoint trace class operator ρ on a complex Hilbert space $\mathcal{H}$, which is positive ($\rho \geq 0$) and normalized to $\mathrm{Tr}(\rho) = 1$ (a density operator). A state is called *pure* if it is of the form $\rho = |\psi\rangle\langle\psi|$, otherwise it is called a *mixed state*. We denote the set of states by $\mathcal{S}(\mathcal{H})$.

Data on a quantum system are obtained from *observables* which are self-adjoint operators S in the Hilbert space $\mathcal{H}$. If S has spectral decomposition $S = \sum_j \lambda_j \Pi_j$ where Π_js are projectors, then a measurement generates a *discrete random variable* X_S taking values in the set of eigenvalues $\{\lambda_1, \lambda_2, \ldots\}$ with probabilities $p_j = \mathrm{Tr}(\rho \cdot \Pi_j)$. The expectation of X_S under the state ρ is then given by the *Born-von Neumann postulate:*

$$E_\rho X_S = \sum_j \lambda_j \,\mathrm{Tr}\left(\rho \Pi_j\right) = \mathrm{Tr}\left(\rho S\right).$$

In quantum mechanics, one needs generalized versions of the above definitions of observables and measurements because the spectral decomposition of self-adjoint operators in the form of a weighted sum of projectors may fail to hold when the Hilbert space is infinite dimensional. If a measurement has outcomes in a measurable space $(\Omega, \mathfrak{B})$, it is determined by a positive operator-valued measure.

Definition 1. A positive operator valued measure (POVM) is a map $M : \mathfrak{B} \to \mathcal{L}(\mathcal{H})$ having the following properties

1) positivity: $M(B) \geq 0$ for all events $B \in \mathfrak{B}$ (hence M(B) is self-adjoint)
2) σ-additivity: $M(\cup_i B_i) = \sum_i M(B_i)$ for any countable set of mutually disjoint events B_i (here the convergence is in the weak operator topology of $\mathcal{L}(\mathcal{H})$)
3) normalization: $M(\Omega) = \mathbf{1}$.

If the operators $M(B)$ are also orthogonal projections, i.e. $M(A)^2 = M(A)$ and $M(B)M(A) = 0$ when $A \cap B = \varnothing$, then it is called a *simple measurement*. The collection of projectors $\{\Pi_j\}$ in the spectral decomposition $S = \sum_j \lambda_j \Pi_j$ is an example of a simple measurement. The outcome of the measurement has probability distribution

$$P_\rho(B) = \mathrm{Tr}(\rho M(B)), \qquad B \in \mathfrak{B}. \tag{1}$$

The spectral theorem shows that any self-adjoint operator $S : \mathcal{H} \to \mathcal{H}$ can be diagonalized as follows:

$$S = \int_{\sigma(S)} x M(dx),$$

where $\sigma(S)$ is the spectrum of S and M is a POVM, also called spectral measure associated with the operator S. When S is an observable with a continuous spectrum, it generates a *continuous random variable* X_S with probabilities given by (1). Also, it easily follows that

$$E[X_S] = \mathrm{Tr}(S\rho).$$

The expected value of an observable S is often denoted as $\langle S \rangle$, when the state dependence is not explicitly shown. There are POVMs (called *generalized measurements*) where the orthogonality does not hold, but these can be extended to a POVM in a larger Hilbert space where the extended version is orthogonal. Let $POVM(\Omega, \mathcal{H})$ be the set of POVMs with values in $\mathcal{L}(\mathcal{H})$ and outcome space Ω

and let $\mathcal{H}_0$ be another Hilbert space with a density operator ρ_0. Then any simple measurement M' in $POVM(\Omega, \mathcal{H} \otimes \mathcal{H}_0)$ induces a measurement M in $POVM(\Omega, \mathcal{H})$ which is determined by

$$\mathrm{Tr}(\rho M(B)) = \mathrm{Tr}((\rho \otimes \rho_0)M'(B)), B \in \mathfrak{B},$$

for all states ρ on $\mathcal{H}$. The pair $(\mathcal{H}_0, \rho_0)$ is called an *ancilla* and it is known that (cf. [28], Section 2.5) given any measurement M in $\mathcal{H}$, there exists an ancilla $(\mathcal{H}_0, \rho_0)$ and a simple measurement M' such that the above equation holds. The triple $(\mathcal{H}_0, \rho_0, M')$ is called a *realization* of the measurement M and the notion of adding an ancilla before taking simple measurements is called *quantum randomization* in [8].

In many cases, it is convenient to perform a measurement after "changing" the state of the original system by interacting with other systems. The maps describing such transformations are called quantum channels.

Definition 2. A quantum channel between systems with Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$ is a mapping T which assigns to every state ρ on $\mathcal{H}_1$ the state $T(\rho)$ on $\mathcal{H}_2$ given by

$$T(\rho) = \sum_{i=1}^{\infty} K_i \rho K_i^*,$$

where $\{K_i\}$ are bounded operators $K_i : \mathcal{H}_1 \to \mathcal{H}_2$ such that $\sum_{i=1}^{\infty} K_i^* K_i = \mathbf{1}$ (the series converging in the strong operator topology of $\mathcal{L}(\mathcal{H})$).

It can be shown that the map T is trace preserving and *completely positive*, i.e. $\mathrm{Id}_m \otimes T$ is positive for all $m \geq 1$, where Id_m is the identity map on the space of m dimensional matrices. The simplest example of a quantum channel is a transformation $\rho \mapsto U\rho U^*$, where U is a unitary operator on $\mathcal{H}$. More generally, if $|\varphi\rangle \in \mathcal{K}$ is a pure state of an ancillary system, and V is a unitary on $\mathcal{H} \otimes \mathcal{K}$, then

$$\rho \mapsto T(\rho) := \mathrm{Tr}_{\mathcal{K}}(V(\rho \otimes |\varphi\rangle\langle\varphi|)V^*)$$

is a quantum channel where $\mathrm{Tr}_{\mathcal{K}}$ is the partial trace over $\mathcal{K}$ (with respect to an orthonormal basis $\{|f_i\rangle\}_{i=1}^{\dim \mathcal{K}}$). If we define operators K_i on $\mathcal{H}$ such that $\langle\psi|K_i|\psi'\rangle := \langle\psi \otimes f_i|V|\psi' \otimes \varphi\rangle$, then it can be seen that $T(\rho)$ can be written as in the form given in Definition 2. We define a dual map T^* of a quantum channel as follows

$$T^* : POVM(\Omega, \mathcal{H}_2) \to POVM(\Omega, \mathcal{H}_1)$$

$$T^*(M)(B) = \sum_{i=1}^{\infty} K_i^* M(B) K_i,$$

where the $\sum_{i=1}^{\infty} K_i^* M(B) K_i$ is a strongly convergent sum. From the definition, it can be easily verified that $T^*(M)$ is indeed an element of $POVM(\Omega, \mathcal{H}_1)$ (i.e. a POVM satisfying properties 1,2 and 3 of Definition 1) and that it satisfies the following duality relation

$$\mathrm{Tr}(\rho T^*(M)(B)) = \mathrm{Tr}(T(\rho)M(B)), \quad \forall B \in \mathfrak{B}$$

and all states ρ on $\mathcal{H}_1$ (cf. 29.9 of [45]).

For estimation purposes, we will need to define the following distances between two quantum states. The *trace-norm* distance between two states $\rho_0, \rho_1 \in \mathcal{S}(\mathcal{H})$ is given by

$$\|\rho_0 - \rho_1\|_1 := \mathrm{Tr}(|\rho_0 - \rho_1|),$$

where $|\tau| := \sqrt{\tau^*\tau}$ denotes the absolute value of τ. An interpretation of this metric in terms of quantum testing can be found in [26]. In the special case of pure states, the trace-norm distance is given by

$$\||\psi_0\rangle\langle\psi_0| - |\psi_1\rangle\langle\psi_1|\|_1 = 2\sqrt{1 - |\langle\psi_0|\psi_1\rangle|^2}. \tag{2}$$

Similarly one can define the L^2 distance between two states induced by the Hilbert-Schmidt norm as

$$\|\rho_0 - \rho_1\|_2 := [\mathrm{Tr}((\rho_0 - \rho_1)^*(\rho_0 - \rho_1))]^{1/2}.$$

2.2. Qudits under local parametrization

Consider a qudit or a d-dimensional density matrix, i.e. $\rho \in M_d(\mathbb{C})$ (the space of $d \times d$ complex matrices), with $\rho \geq 0$ and $\mathrm{Tr}(\rho) = 1$. A natural way to parametrize qudits is to write it in the form $\mathscr{U}(\zeta)\rho_0\mathscr{U}^*(\zeta)$ with $\rho_0 = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_d)$ and $\sum_{i=1}^d \mu_i = 1$, where $\mathscr{U}(\zeta)$ are unitary matrices, i.e. elements of the Lie group $SU(d)$ parametrized by $\zeta \in \mathbb{C}^{d(d-1)/2}$. We describe the unitaries in more detail. Consider the generators of the Lie algebra $\mathfrak{su}(d)$ (cf. [18]):

$$\begin{aligned} H_j &= E_{jj} - E_{j+1,j+1} \text{ for } 1 \leq j \leq d-1 \\ T_{j,k} &= iE_{j,k} - iE_{k,j} \text{ for } 1 \leq j < k \leq d \\ T_{k,j} &= E_{j,k} + E_{kj} \text{ for } 1 \leq j < k \leq d, \end{aligned}$$

where $E_{i,j}$ is the matrix with $(i,j)^{th}$ entry equal to 1, and all other entries equal to 0.

Assume $\mu_1 > \mu_2 > \ldots > \mu_d > 0$ and define

$$\mathscr{U}(\zeta) = \exp\left[i\left(\sum_{1 \leq j < k \leq d} \frac{Re(\zeta_{j,k})T_{j,k} + Im(\zeta_{j,k})T_{k,j}}{\mu_j - \mu_k}\right)\right].$$

Next, we consider local models by first perturbing the eigenvalues only:

$$\rho_{0,u} = \mathrm{diag}(\mu_1 + u_1, \mu_2 + u_2, \ldots, \mu_d + u_d),$$

where $\sum u_i = 0$; that gives a state in a local neighborhood of ρ_0. To describe all possible states in the local neighborhood of ρ_0, we should also consider rotations using unitaries, i.e.

$$\rho_\vartheta = \mathscr{U}(\zeta)\rho_{0,u}\mathscr{U}^*(\zeta), \tag{3}$$

where $\vartheta = (u, \zeta)$. Similarly in the low rank case we assume $\mu_1 > \mu_2 > \ldots > \mu_r > \mu_{r+1} = \ldots = \mu_d = 0$ and define

$$\mathscr{U}^r(\zeta) = \exp\left[i\left(\sum_{\substack{1 \leq j \leq r \\ j < k \leq d}} \frac{Re(\zeta_{j,k})T_{j,k} + Im(\zeta_{j,k})T_{k,j}}{\mu_j - \mu_k}\right)\right].$$

Next we consider a low-rank state in the local neighborhood of $\rho_{0,r} = \mathrm{diag}(\mu_1, \mu_2, \ldots, \mu_r, 0, \ldots, 0)$:

$$\rho_{0,u,r} = \mathrm{diag}(\mu_1 + u_1, \mu_2 + u_2, \ldots, \mu_r + u_r, 0, \ldots, 0) \tag{4}$$

$$\rho_{\vartheta,r} = \mathscr{U}^r(\zeta)\rho_{0,u,r}\mathscr{U}^{r*}(\zeta). \tag{5}$$

We note that it is enough to parametrize low rank qudits using $\mathscr{U}^r(\zeta)$ instead of $\mathscr{U}(\zeta)$ since the first order terms in the Taylor expansion of $\mathscr{U}^r(\zeta)\rho_{0,u,r}\mathscr{U}^{r*}(\zeta)$ and $\mathscr{U}(\zeta)\rho_{0,u,r}\mathscr{U}^*(\zeta)$ are identical and it is this term that determines the limiting model (cf. [2]).

Let $\mathfrak{z}_{ij} = \frac{\zeta_{ij}}{\sqrt{\mu_i - \mu_j}}$. For notational convenience we will denote $\theta = (u, \mathfrak{z})$ and redefine the unitaries $\mathscr{U}(\zeta)$ and $\mathscr{U}^r(\zeta)$ as $U(\mathfrak{z})$ and $U^r(\mathfrak{z})$ respectively. We will denote the state by ρ_ϑ ($\rho_{\vartheta,r}$ respectively) or ρ_θ ($\rho_{\theta,r}$ respectively) depending on whether it is indexed by ϑ or θ.

In classical LAN we are interested in the limit for n i.i.d. copies of the experiment with the local parameter lying in an $n^{-1/2}$ neighborhood of 0 (see Section 3 for more details). In the quantum setup, this amounts to studying the following operators

$$\rho^{\theta,n} = \rho^{\otimes n}_{\theta/\sqrt{n}}, \quad \rho^{\theta,r,n} = \rho^{\otimes n}_{\theta/\sqrt{n},r}.$$

2.3. Gaussian states and Fock spaces

To obtain Gaussian random variables, in the space $\mathcal{H} = L^2(\mathbb{R})$ one considers two special observables Q, P with continuous spectrum:

$$(Qf)(x) = xf(x),\ (Pf)(x) = -i\frac{df}{dx}(x), \quad f \in D \subset L^2(\mathbb{R})$$

(defined on an appropriate domain D) often associated to position (Q) and momentum (P) of a particle. These operators satisfy the Heisenberg commutation relations

$$[Q, P] = i\mathbf{1}.$$

It can be shown that $Z_u := u_1 Q + u_2 P$, $u \in \mathbb{R}^2$ are observables (called the *canonical observables*). In this context we define the *quantum characteristic function* as $\tilde{W}_\rho(u_1, u_2) = \mathrm{Tr}(\rho \exp(iZ_u))$. If the following relation holds

$$E_\rho \exp(iZ_u) = \mathrm{Tr}(\rho \exp(iZ_u)) = \exp\left(iu^T\mu - \frac{1}{2}u^T\Sigma u\right),\ u \in \mathbb{R}^2,$$

then ρ is called a Gaussian state with mean μ and covariance matrix Σ. For such quantum Gaussian states in $L^2(\mathbb{R})$ we adopt a compact notation, resembling the one for the 2-variate normal law:

$$\rho = \mathbb{N}_2(\mu, \Sigma). \tag{6}$$

Here Σ is a 2×2 real matrix such that

$$\Sigma \geq \pm\frac{i}{2}\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

To define the simplest Gaussian state, let $\psi_0 = \sqrt{\varphi_{1/2}}$ be the square root of the density function of the normal $\mathcal{N}(0, 1/2)$ distribution and consider the operator ρ_0 acting by $\rho_0 f = \psi_0 \langle \psi_0, f \rangle$, $f \in L^2(\mathbb{R})$. Since ψ_0 is a unit vector in $L^2(\mathbb{R})$, the operator ρ_0 (henceforth called the vacuum state) is a projection (written $\rho_0 = |\psi_0\rangle\langle\psi_0|$ in Dirac notation) and it can be shown that $\rho_0 = \mathbb{N}_2(0, I_2/2)$ in the notation described above.

An important class is the collection of *coherent states* $\mathbb{N}_2(\mu, I_2/2)$; these are pure states which can be interpreted as a vacuum shifted by $\mu \in \mathbb{R}^2$ (similar to the Gaussian shift model in classical statistics). Consider the operators $a^* = (Q - iP)/\sqrt{2}$ (the *creation operator*), $a = (Q + iP)/\sqrt{2}$ (the *annihilation*

operator) and $N = a^*a$ (the number operator). It is well known that the *Hermite basis* $\{|0\rangle, |1\rangle \ldots\}$ forms an eigenbasis of the number operator, i.e. $N|k\rangle = k|k\rangle$. For any $z \in \mathbb{C}$ define the displacement operator as

$$D(z) = \exp(za^* - \bar{z}a)$$

and the *coherent state* as

$$|G(z)\rangle = D(z)|0\rangle = \exp(-|z|^2/2) \sum_{k=0}^{\infty} \frac{z^k}{\sqrt{k!}} |k\rangle. \tag{7}$$

In the density operator notation this pure Gaussian state is $|G(z)\rangle\langle G(z)|$. The expectations of the canonical observables Q and P under the state $|G(z)\rangle\langle G(z)|$ are

$$\langle Q\rangle = \sqrt{2}\,\mathrm{Re}\, z,\ \langle P\rangle = \sqrt{2}\,\mathrm{Im}\, z$$

and the characteristic function of $|G(z)\rangle\langle G(z)|$ is

$$\varphi(t) = \exp\left(i\left(t_1\sqrt{2}\,\mathrm{Re}\, z + t_2\sqrt{2}\,\mathrm{Im}\, z\right) - \frac{1}{4}\left(t_1^2 + t_2^2\right)\right),\ t \in \mathbb{R}^2.$$

The presence of the factor $\sqrt{2}$ motivates us to adopt a modified notation for the coherent vector: setting $\mu = \left(\sqrt{2}\,\mathrm{Re}\, z, \sqrt{2}\,\mathrm{Im}\, z\right)$, we will write $|G(z)\rangle = |\psi_\mu\rangle$ so that now the expectations are $(\langle Q\rangle, \langle P\rangle) = \mu$. The characteristic function $\varphi(t)$ is that of $\mathcal{N}_2(\mu, I/2)$ and hence in the notation of (6)

$$|G(z)\rangle\langle G(z)| = |\psi_\mu\rangle\langle\psi_\mu| = \mathbb{N}_2(\mu, I/2). \tag{8}$$

Other important classes of Gaussian states are thermal states and shifted thermal states; these are mixed states unlike the vacuum and coherent states. We define below the thermal state with temperature β as

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

Shifted thermal states are defined using the shift operator $D(z)$ as follows:

$$\phi_\beta^z = D(z)\phi_\beta D^*(z).$$

One can show that the quantum characteristic function $\mathrm{Tr}(\phi_\beta^z \exp(iu_1Q + iu_2P))$ of the shifted thermal state is given by

$$\mathrm{Tr}(\phi_\beta^z \exp(iu_1Q + iu_2P)) = \exp(i(u_1\sqrt{2}Re(z) + u_2\sqrt{2}Im(z)) - \frac{\sigma_\beta^2}{2}(u_1^2 + u_2^2)),$$

where $\sigma_\beta^2 = \frac{\coth(\beta/2)}{2}$. Defining $\mu = (\sqrt{2}Re(z), \sqrt{2}Im(z))$, we write that

$$\phi_\beta^z = \mathbb{N}_2(\mu, \sigma_\beta^2 I_2). \tag{9}$$

To define a k-mode Gaussian state one considers the space $\bigotimes_{i=1}^k L^2(\mathbb{R})$ and identifies the number basis as follows

$$|\mathbf{m}\rangle = \otimes_{1\le i\le k} |m_i\rangle, \quad \mathbf{m} = \{m_i \in \mathbb{N} : 1 \le i \le k\}. \tag{10}$$

We consider the collection of operators $\{Q_1, P_1, \ldots, Q_k, P_k\}$ where each Q_i and P_i are position and momentum operators of a particular mode (i.e. acting on a particular $L^2(\mathbb{R})$). These operators satisfy joint commutation relations as follows:

$$[Q_i, P_j] = i\delta_{i,j}\mathbf{1}, \quad [Q_i, Q_j] = 0, \quad [P_i, P_j] = 0. \tag{11}$$

One can then define the creation and annihilation operators a_i^* and a_i for each mode and proceed to define the displacement operator as

$$D(\mathbf{z}) = \exp(\mathbf{z}.\mathbf{a}^* - \bar{\mathbf{z}}.\mathbf{a}),$$

where $\mathbf{z}.\mathbf{a}^* = \sum_{i=1}^k z_i a_i^*$ and similarly for $\bar{\mathbf{z}}.\mathbf{a}$. Then we have

$$|G(\mathbf{z})\rangle = D(\mathbf{z})|\mathbf{0}\rangle = \exp(-||\mathbf{z}||^2/2) \sum_{m_1,\ldots,m_k=0}^{\infty} \prod_{i=1}^{k} \frac{z_i^{m_i}}{\sqrt{m_i!}} |\mathbf{m}\rangle. \tag{12}$$

To describe our limiting model we need the following multimode Fock spaces:

$$\mathcal{F} := \bigotimes_{1\le i<j\le d} L^2(\mathbb{R}) \tag{13}$$

$$\mathcal{F}^r := \bigotimes_{1\le i\le r, i<j\le d} L^2(\mathbb{R}). \tag{14}$$

Recall the shifted thermal state $\mathbb{N}_2(\mu, \sigma_\beta^2 I_2)$; in the complex notations it is denoted by ϕ_β^z (see equation (9)). Similarly the shifted pure state in (8) can be written as ϕ_∞^z (noting that the case $\beta = \infty$ corresponds to the pure case). The limiting model considered in [34] is as follows

$$\Phi^\theta = \mathcal{N}_{d-1}(u, V_\mu) \otimes \bigotimes_{1\le i<j\le d} \phi_{\beta_{ij}}^{\eth_{ij}}, \tag{15}$$

where $\beta_{ij} = \ln(\mu_i/\mu_j)$ and $\eth_{ij} = \frac{\zeta_{ij}}{\sqrt{\mu_i - \mu_j}}$ i.e. the diagonal perturbation only appears in the classical part which is a $d-1$ multivariate normal experiment with mean u and covariance matrix V_μ. The latter is the covariance matrix of a multinomial random variable with probabilities μ_i, while the rotation perturbations determine the $d(d-1)/2$ shifted thermal states. Since β_{ij} are constants we will use the abbreviated notation

$$\phi^\eth = \bigotimes_{1\le i<j\le d} \phi_{\beta_{ij}}^{\eth_{ij}} \quad \in \mathcal{T}_1(\mathcal{F}), \tag{16}$$

where $\eth = (\eth_{ij})_{1\le i<j\le d}$ is a vector in $\mathbb{C}^{d(d-1)/2}$.

We also consider the following model which is a tensor product of both thermal and pure quantum states along with a classical part which is given by a multivariate normal random variable.

$$\Phi^{\theta,r} = \mathcal{N}_{r-1}(u, V_\mu) \otimes \bigotimes_{1\le i<j\le r} \phi_{\beta_{ij}}^{\eth_{ij}} \otimes \bigotimes_{\substack{1\le i\le r \\ r+1\le j\le d}} \phi_\infty^{\eth_{ij}} = \mathcal{N}_{r-1}(u, V_\mu) \otimes \phi^{\eth,r}, \tag{17}$$

where

$$\phi^{\eth,\mathbf{r}} = \bigotimes_{1\le i<j\le r} \phi_{\beta_{ij}}^{\eth_{ij}} \otimes \bigotimes_{\substack{1\le i\le r \\ r+1\le j\le d}} \phi_\infty^{\eth_{ij}} \quad \in \mathcal{T}_1(\mathcal{F}^r).$$

In Section 3 we show that the states $\Phi^{\theta,r}$ arise as limiting models for low-rank qudits.

The classical-quantum limiting state also motivates us to adopt a notation (similar to [19]) that allows us to describe the commutation relations in the hybrid system in a compact fashion. Note that for the k-mode system we have defined operators $\{Q_i, P_i\}_{i=1}^k$ given by (11). In addition to this, we allow l classical random variables $C_1, \ldots, C_l$ that commute with each other and with all (Q_i, P_i). We can denote the $m = 2k + l$ variables as

$$(X_1, \ldots, X_m) \equiv (C_1, \ldots, C_l, Q_1, P_1, \ldots, Q_k, P_k) \tag{18}$$

and the commutation relations as

$$[X_i, X_j] = iS_{ij}\mathbf{1},$$

where $S = 0_{l\times l} \oplus \bigoplus_{i=1}^{2k} \Omega$ and

$$\Omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

We can define a state $\varrho = f \otimes \rho$ in the space $L^1(\mathbb{R}^l) \otimes \mathcal{T}_1(\mathcal{F}^k)$, with ρ a quantum state and f a probability density. Now define the hybrid characteristic function of the state as

$$E_\varrho(e^{i\sum_{j=1}^m u_j X_j}) := \int \mathrm{Tr}(\rho e^{i\sum_{j=l+1}^{2k+l} u_j X_j}) f(y) e^{i\sum_{j=1}^l u_j y_j}\, dy_1 \ldots dy_l.$$

Definition 3. A hybrid state ϱ is called classical-quantum Gaussian if the characteristic function has the following form:

$$E_\varrho(e^{i\sum_{i=1}^m u_j X_j}) = e^{iu^T\tau - u^T\Sigma u/2},$$

where $\tau \in \mathbb{R}^m$ and the covariance matrix Σ is a $m \times m$ real matrix such that $\Sigma \geq \pm\frac{i}{2}S$.

Note that each shifted thermal state can be denoted by $\phi_{\beta_{ij}}^{\eth_{ij}} = \mathbb{N}_2(\nu_{ij}, \sigma^2_{\beta_{ij}} I_2)$ where

$$\nu_{ij} = (\sqrt{2}Re(\eth_{ij}), \sqrt{2}Im(\eth_{ij})).$$

Since the classical part in the limiting model is also a Gaussian (given by $\mathcal{N}_{r-1}(u, V_\mu)$), we adopt the following alternate notation for the classical-quantum Gaussian:

$$\Phi^{\theta,r} = \mathfrak{N}(\tau, \mathscr{S}), \tag{19}$$

where

$$\Sigma = V_\mu \oplus \bigoplus_{1\leq i\leq r, i<j\leq d} \sigma^2_{\beta_{ij}} I_2$$

$$S = 0_{r-1\times r-1} \oplus \bigoplus_{1\leq i\leq r, i<j\leq d} \Omega$$

$$\mathscr{S} = \Sigma + \frac{i}{2}S, \quad \tau = u \oplus \bigoplus_{1\leq i\leq r, i<j\leq d} \nu_{ij}.$$

We use $\mathscr{S}$ instead of Σ in $\mathfrak{N}(\tau, \mathscr{S})$ that captures the underlying non-commutative structure via S. We note that a similar notation incorporating the commutation relations into a complex covariance matrix

has been used in [54]. This notation will be useful in computing a Bayes risk for a one dimensional parameter (see Appendix B) which will subsequently be used in establishing a minimax lower bound for the estimation of a linear functional of the state.

2.4. Quantum statistical inference

In this section we formalize the quantum counterparts of the basic notions of optimality in classical statistical inference. In classical statistics, an experiment is defined to be a family of probability measures on a sample space and denoted by $\mathcal{E} = \{P_\theta, \theta \in \Theta\}$ where Θ is the parameter space.

Definition 4. A quantum statistical model over a parameter space Θ consists of a family of quantum states $\mathcal{Q} = \{\rho_\theta : \theta \in \Theta\}$ on a Hilbert space $\mathcal{H}$, indexed by an unknown parameter $\theta \in \Theta$.

Inference in quantum models generally involves two steps. In the first step one performs a measurement on the state ρ_θ and generates data, while in the second step, one uses standard statistical tools to solve the specific decision problem using data from the first step. If one performs a measurement M on the system in state ρ_θ, a random outcome is obtained with distribution $P_\theta^M(E) := \mathrm{Tr}(\rho_\theta M(E))$ (cf. Subsection 2.1). The measurement data is therefore described by the classical model $\mathcal{P}^M := \{P_\theta^M : \theta \in \Theta\}$ and the estimation problem can be treated using "classical" statistical methods. However, in many scenarios, the optimal estimators for individual components of a parameter are incompatible with each other and the optimal joint estimator for the two components can be entirely different from the optimal estimators of the individual components.

In the classical setup, a randomized decision function is given by a Markov kernel ν. If $L(\theta,u)$ is the loss function then the risk is given by

$$R(\theta,\nu) = \int\int L(\theta,u)\nu_x(du)\mu_\theta(dx) = \int L(\theta,u)\int \nu_x(du)\mu_\theta(dx) = \int L(\theta,u)\tilde{\nu}_\theta(du), \tag{20}$$

where $\tilde{\nu}_\theta(A) = \int \nu_x(A)\mu_\theta(dx)$.

Section 2.2.4 of [31] discusses the quantum counterpart of this classical formulation. Let ρ_θ be the quantum state and $\mu_\theta^M(B) = Tr(\rho_\theta M(B))$ be the probability measure generated by the POVM M. Then the risk is given by

$$R(\theta,M) = \int L(\theta,u)\mu_\theta^M(du).$$

By using the fact that every affine map $\rho_\theta \to \mu_\theta()$ can be associated with a POVM, we see that M is an analog of the classical randomized decision function ν given in (20). We can easily define the Bayes and minimax problems for quantum estimation.

Minimax problem

$$\inf_M \sup_{\theta\in\Theta} R(\theta,M) = \inf_M \sup_{\theta\in\Theta} \int L(\theta,u)\mu_\theta^M(du) = \inf_{\hat{m}} \sup_{\theta\in\Theta} E_\theta[L(\theta,\hat{m})]$$

Bayes problem

$$\inf_M \int_\Theta R(\theta,M)\pi(d\theta) = \inf_M \int_\Theta \int L(\theta,u)\mu_\theta^M(du)\pi(d\theta) = \inf_{\hat{m}} \int E_\theta[L(\theta,\hat{m})]\pi(d\theta).$$

The notations $\inf_{\hat{m}} \sup_{\theta\in\Theta} E_\theta[L(\theta,\hat{m})]$ and $\inf_{\hat{m}} \int E_\theta[L(\theta,\hat{m})]\pi(d\theta)$ will be called *condensed notations* and will be used henceforth. Note that the infimum is over all POVM and the notation $\hat{m}$ should not be

confused with a deterministic estimator seen in the classical setup. We will also denote the Bayes risk as $\inf_{\hat{m}} E[L(\theta,\hat{m})]$ where the expectation is also taken over the parameter θ.

In classical statistics, a well-known paradigm is using asymptotic equivalence of experiments to transfer risk bounds from one experiment to another. Suppose we have two experiments $\mathcal{E} = \{P_\theta, \theta \in \Theta\}$ on a sample space $(\Omega_1, \mathcal{A}_1)$ and $\mathcal{F} = \{Q_\theta, \theta \in \Theta\}$ on a sample space $(\Omega_2, \mathcal{A}_2)$. Also let the loss function satisfy the condition $0 \leq L(\theta,u) \leq 1$. If there exists a Markov kernel K such that

$$\sup_{\theta\in\Theta} ||KP_\theta - Q_\theta||_{TV} \leq \epsilon,$$

then for any randomized decision function μ of θ in the model Q_θ, the randomized decision function $\nu = \mu \circ K$ (composition of two Markov kernels) satisfies

$$R^1(\theta,\nu) \leq R^2(\theta,\mu) + \epsilon,$$

where $R^1(\theta,\nu)$ and $R^2(\theta,\mu)$ are the risks in the models $\mathcal{E}$ and $\mathcal{F}$ respectively. We discuss the generalization of this paradigm to the quantum setup and also generalize it to the case of unbounded loss.

The quantum equivalent of a Markov kernel is the transformation by quantum channels. The quantum model $\mathcal{Q}$ can be transformed into another quantum model $\mathcal{Q}' := \{\rho'_\theta : \theta \in \Theta\}$ on a Hilbert space $\mathcal{H}'$ by applying a quantum channel

$$T : \mathcal{T}_1(\mathcal{H}) \to \mathcal{T}_1(\mathcal{H}')$$
$$T : \rho_\theta \mapsto \rho'_\theta.$$

In this context, we define the *quantum Le Cam distance* between two models from [12].

Definition 5. Let $\mathcal{Q}$ and $\mathcal{Q}'$ be two quantum models over Θ. The deficiency of $\mathcal{Q}$ with respect to $\mathcal{Q}'$ is defined by

$$\delta(\mathcal{Q},\mathcal{Q}') := \inf_T \sup_{\theta\in\Theta} \|T(\rho_\theta) - \rho'_\theta\|_1,$$

where the infimum is taken over all channels T. The *Le Cam* distance between $\mathcal{Q}$ and $\mathcal{Q}'$ is defined as

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

Its interpretation is that models which are "close" in the Le Cam distance have similar risk bounds. Suppose we have two sequences of quantum models (or experiments) $\mathcal{E}^{(n)} = \{\rho_\theta^{(1,n)} : \theta \in \Theta\}$ and $\mathcal{F}^{(n)} = \{\rho_\theta^{(2,n)} : \theta \in \Theta\}$ with associated sequences of Hilbert spaces $\mathcal{H}^{1,n}$ and $\mathcal{H}^{2,n}$. Assume that $\Delta\left(\mathcal{E}^{(n)},\mathcal{F}^{(n)}\right) \to 0$; this implies $\delta\left(\mathcal{E}^{(n)},\mathcal{F}^{(n)}\right) \to 0$ and in particular there exists a sequence of quantum channels T_n, such that

$$||T_n(\rho_\theta^{(1,n)}) - \rho_\theta^{(2,n)}||_1 = o(1).$$

Let the loss function also change with n and satisfy the relation $0 \leq L_n(\theta,u) \leq c_n$. Also, assume that the sequence of quantum channels T_n is such that

$$c_n \sup_{\theta\in\Theta} ||T_n(\rho_\theta^{(1,n)}) - \rho_\theta^{(2,n)}||_1 = o(1). \tag{21}$$

Recall the dual map T^* of a quantum channel T. It follows that for any $M \in POVM(\Omega, \mathcal{H}^{2,n})$

$$\begin{aligned}
R_n^1(\theta, T_n^*(M)) &= \int L_n(\theta,u)\,\mathrm{Tr}(\rho_\theta^{(1,n)} T_n^*(M(du))) \\
&= \int L_n(\theta,u)\,\mathrm{Tr}(\rho_\theta^{(2,n)} M(du)) \\
&\quad + \int L_n(\theta,u)[\mathrm{Tr}(\rho_\theta^{(1,n)} T_n^*(M(du))) - \mathrm{Tr}(\rho_\theta^{(2,n)} M(du))] \\
&= R_n^2(\theta,M) + \int L_n(\theta,u)[\mathrm{Tr}(T_n(\rho_\theta^{(1,n)})(M(du))) - \mathrm{Tr}(\rho_\theta^{(2,n)} M(du))] \\
&\le R_n^2(\theta,M) + c_n ||T_n(\rho_\theta^{(1,n)}) - \rho_\theta^{(2,n)}||_1 \\
&\le R_n^2(\theta,M) + o(1), \qquad (22)
\end{aligned}$$

the term $o(1)$ tending to 0 uniformly over all θ. Thus we can compare the risks of the two models $\mathcal{E}^{(n)}$ and $\mathcal{F}^{(n)}$ if (21) holds. Note that we have similar relations for minimax risks and Bayes risks, by taking a supremum over Θ or integrating with respect to a prior, respectively, and then taking an infimum over all estimators:

$$\inf_M \sup_{\theta\in\Theta} R_n^1(\theta,M) \le \inf_M \sup_{\theta\in\Theta} R_n^2(\theta,M) + o(1). \qquad (23)$$

$$\inf_M \int_\Theta R_n^1(\theta,M)\pi(d\theta) \le \inf_M \int_\Theta R_n^2(\theta,M)\pi(d\theta) + o(1).$$

3. Local asymptotic normality in low-rank systems

3.1. Classical LAN

Consider a collection of i.i.d. random variables $\{X_1,\dots,X_n\}$ taking values in a measurable space $(\mathcal{X},\Sigma_\mathcal{X})$ with $X_i \sim P_\theta$ where θ belongs to Θ, which is an open subset of $\mathbb{R}^d$. We can consider a local perturbation around a fixed point θ_0 and if we denote $\theta = \theta_0 + u/\sqrt{n}$ (with u bounded), then we can represent the aforementioned collection of random variables by a statistical experiment $\mathcal{E}_n = \{P^n_{\theta_0+u/\sqrt{n}}, ||u|| \le C\}$ on a sample space $(\mathcal{X}^n,\Sigma_\mathcal{X}^n)$ where $P^n_{\theta_0+u/\sqrt{n}}$ is an n-fold product of $P_{\theta_0+u/\sqrt{n}}$. Under some regularity assumptions (see theorem below) $\mathcal{E}_n$ can be approximated by a Gaussian shift experiment $\mathcal{F} = \{N(u, I_{\theta_0}^{-1}), ||u|| \le C\}$ where I_{θ_0} is the Fisher information matrix at θ_0. The following result is well known (cf. [47], Theorem 2.9 along with [48], 79.3).

Theorem 3.1. *Assume $(\mathcal{X},\Sigma_\mathcal{X})$ is a Polish (complete separable metric) space with its Borel σ-algebra. Assume further that*

(i) the experiment $\mathcal{E}_n$ is dominated: $P_\theta \ll \mu$, $\theta \in \Theta$ where μ is a σ-finite measure on $\Sigma_\mathcal{X}$,

(ii) the densities $p_\theta(x) = (dP_\theta/d\mu)(x)$ are jointly measurable in (x,θ) and differentiable in quadratic mean at $\theta = \theta_0$, i.e. for some measurable function $\ell_\theta : \mathcal{X} \to \mathbb{R}^d$

$$\int \left[p_{\theta+u}^{1/2} - p_\theta^{1/2} - u^T \ell_\theta p_\theta^{1/2}\right]^2 d\mu = o\left(\|u\|^2\right) \text{ as } u \to 0,$$

(iii) the Fisher information matrix $I_\theta = 4E_\theta[\ell_\theta \ell_\theta^T]$ is nonsingular at $\theta = \theta_0$.

Then the experiments $\mathcal{E}_n$ and $\mathcal{F}$ are asymptotically equivalent.

In other words, there exist sequences of Markov kernels T_n and S_n, such that:

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

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

3.2. Quantum LAN

Consider the following domain of the local parameters

$$\Theta_{n,\beta,\gamma} = \{(u,\mathfrak{z}) : |u_k| \leq n^{\gamma}, |\mathfrak{z}_{ij}| \leq n^{\beta}, \forall 1 \leq k \leq d-1,\ 1 \leq i < j \leq d\},$$

for some $\beta > 0, \gamma > 0$.

Recall the d-dimensional state ρ_θ given in equation (3) and the corresponding Gaussian state given in (15) indexed by the same local parameter θ. Note that we have used the alternate notation using θ instead of ϑ (see the discussion after (5)). We consider the following two experiments:

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

We state the LAN theorem for full rank states (proved in [34]) which shows that these local models are asymptotically equivalent.

Theorem 3.2. *Recall the Fock space described in (13). Then for $0 < \gamma < 1/4$ and $0 < \beta < 1/9$ there exist quantum channels T_n and S_n*

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

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

such that

$$\sup_{\theta\in\Theta_{n,\beta,\gamma}} ||\Phi^{\theta} - T_n(\rho^{\theta,n})||_1 = O(n^{-\kappa})$$

$$\sup_{\theta\in\Theta_{n,\beta,\gamma}} ||S_n(\Phi^{\theta}) - \rho^{\theta,n}||_1 = O(n^{-\kappa}),$$

for some $\kappa > 0$ which depends on β and γ.

Now we are ready to state the low-rank version of the above theorem. Recall that a qudit of rank r can be parametrized as $\rho_{\theta,r}$ (see equation (5) and use the alternate notation, i.e. θ instead of ϑ) and consider the corresponding Gaussian state given in (17) indexed by a local parameter θ. Define

$$\Theta_{n,r,\beta,\gamma} = \{(u,\mathfrak{z}) : |u_k| \leq n^{\gamma}, |\mathfrak{z}_{ij}| \leq n^{\beta}\ \forall\ 1 \leq k \leq r-1, 1 \leq i \leq r, i < j \leq d\},$$

for some $\beta > 0, \gamma > 0$. We now have the low-rank versions of the earlier models:

$$\mathcal{Q}^r_n = \{\rho^{\theta,r,n} : \theta \in \Theta_{n,r,\beta,\gamma}\} \quad \mathcal{R}^r_n = \{\Phi^{\theta,r} : \theta \in \Theta_{n,r,\beta,\gamma}\}.$$

The following theorem shows that the above models are asymptotically equivalent.

Theorem 3.3. *Recall the Fock space described in (14). Then for* $0 < \gamma < 1/4$ *and* $0 < \beta < 1/9$ *there exist quantum channels* T_n^r *and* S_n^r

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

such that

$$\sup_{\theta \in \Theta_{n,r,\beta,\gamma}} ||\Phi^{\theta,r} - T_n^r(\rho^{\theta,r,n})||_1 = O(n^{-\kappa}) \tag{24}$$

$$\sup_{\theta \in \Theta_{n,r,\beta,\gamma}} ||S_n^r(\Phi^{\theta,r}) - \rho^{\theta,r,n}||_1 = O(n^{-\kappa}), \tag{25}$$

for some $\kappa > 0$ *which depends on* β *and* γ.

We observe that the classical part corresponds to a low dimensional ($r-1$- variate) normal experiment while the quantum part contains two subparts. When μ_j and μ_k are both positive (i.e. $1 \le j < k \le r$), we get shifted thermal states with temperatures given as before ($\beta_{jk} = \ln(\mu_j/\mu_k)$). When $\mu_j > 0$ and $\mu_k = 0$, we get shifted pure states. We can compare our result with other available results in quantum LAN.

Comparison with other LAN results

1. The diagonal case: In the absence of the rotation by unitaries the diagonal state represents a local multinomial model and the limiting model contains only the classical Gaussian part, i.e. $\mathcal{N}_{r-1}(u, V_\mu)$, and we recover the classical LAN result of Theorem 3.1. A global version of this approximation (with possibly increasing dimension of the multinomial) was obtained in [16].
2. $d = 2, r = 1$: In this case, we observe that the limiting model consists of a single shifted pure Gaussian state and no classical component. This case was discussed in [24] using a heuristic argument.
3. $d = r > 2$: The limiting model consists of a $d-1$ dimensional normal distribution in the classical part and only shifted thermal states in the quantum part. We recover the full rank case of [34], i.e. Theorem 3.2.
4. $d > r = 1$: The limiting model consists of $d-1$ shifted pure states and no classical component. The case when $d = \infty$ and $r = 1$, i.e. the case of infinite dimensional pure states was treated in [12] under a slightly different parametrization. The limiting model in [12] is quantum white noise, or equivalently a tensor product of infinitely many shifted pure states. A slight modification of the proof of Theorem 4.1 in [13] (proving it for finite dimensional pure states) shows that the limiting model is indeed a tensor product of $d-1$ shifted pure states and agrees with our current result.

4. Measurement of the shift parameter for a Gaussian state

In this section, we consider measurement of the shift parameter μ of the model $\rho = \mathbb{N}_2(\mu, \sigma^2 I_2)$. We describe the particular generalized measurement (the *covariant measurement*) that is used to measure the shift parameter. It can be shown that a Bayes estimator can be constructed by appropriately "shrinking" the covariant measurement.

An essential fact is that the coherent vectors $\left\{|\psi_m\rangle, m \in \mathbb{R}^2\right\}$ form an "overcomplete system" (if multiplied by a factor $1/\sqrt{2\pi}$), i.e. fulfill

$$\frac{1}{2\pi}\int_{\mathbb{R}^2} |\psi_m\rangle \langle\psi_m| \, dm = \mathbf{1}, \tag{26}$$

where $\mathbf{1}$ is the identity operator in the Hilbert space $\mathcal{H} = L^2(\mathbb{R})$ (see equation 3.5.45, p. 101 of [28], with proof after Proposition 3.5.1). A complete orthonormal system $\{|\psi_m\rangle\}$ is an example of an overcomplete system (with integration replaced by summation); however, in the general case the vectors ψ_m can be non-orthogonal and linearly dependent. The system of coherent vectors $\left\{|\psi_m\rangle, m \in \mathbb{R}^2\right\}$ generates a *resolution of the identity*, i.e. a normalized POVM M on the measurable space $\left(\mathbb{R}^2, \mathfrak{B}\right)$ (with $\mathfrak{B}$ being the Borel sigma-algebra) according to

$$M(B) = \frac{1}{2\pi}\int_B |\psi_m\rangle \langle\psi_m| \, dm, \, B \in \mathfrak{B}.$$

The POVM M then generates a (generalized) observable X_M with values in $\mathbb{R}^2$ which under the state ρ has probability distribution

$$P(X_M \in B) = \operatorname{Tr} \rho M(B), \, B \in \mathfrak{B}.$$

This is called the *canonical covariant measurement* in Section 3.6 of [28], covariance referring to the action of the Weyl unitaries (or the displacement operators). There also optimality properties are proved, as well as equivalence to simple measurements on an extended system (corresponding to an *orthogonal* resolution of the identity there). When the covariant measurement is clear from the context we will often write X_M as X. It can be shown that when ρ is a shifted pure state, i.e. $\rho = \mathbb{N}_2(\mu, I_2/2)$ then $X \sim \mathcal{N}_2(\mu, I_2)$ and if ρ is a shifted thermal state, i.e. $\rho = \mathbb{N}_2\left(\mu, \sigma_\beta^2 I_2\right)$ then $X \sim \mathcal{N}_2\left(\mu, \frac{(2\sigma_\beta^2+1)}{2} I_2\right)$.

An equivalent description can be given as follows; cf. Proposition 3.6.1 of [28] and also relation (3.18) in [30]. Let $\mathcal{H} = L^2(\mathbb{R})$ and let $\mathcal{H}_0$ be an identical Hilbert space with canonical observables Q_0 and P_0. In the tensor product $\mathcal{H} \otimes \mathcal{H}_0$, consider the operators

$$\tilde{Q} = Q \otimes \mathbf{1}_0 + \mathbf{1} \otimes Q_0, \; \tilde{P} = P \otimes \mathbf{1}_0 - \mathbf{1} \otimes P_0,$$

where $\mathbf{1}_0$ is the identity operator in $\mathcal{H}_0$. Let ρ be the state in $\mathcal{H}$ to be measured and ρ_0 be an auxiliary state in $\mathcal{H}_0$ to be chosen; then a simple measurement of $\rho \otimes \rho_0$ can be understood as a "randomized" measurement of ρ. These randomized measurements correspond to nonorthogonal resolutions like (26).

It can be shown that $\tilde{Q}$ and $\tilde{P}$ commute, which means that the observables $\tilde{Q}, \tilde{P}$ are jointly measurable in the system given by $\mathcal{H} \otimes \mathcal{H}_0$. The operators $\tilde{Q}, \tilde{P}$ are self-adjoint and thus generate jointly distributed real valued random variables $X_{\tilde{Q}}, X_{\tilde{P}}$. We define $\tilde{X}$ as follows

$$\tilde{X} = \left(X_{\tilde{Q}}, X_{\tilde{P}}\right) \tag{27}$$

and it can be checked that if $\rho = \left|\psi_\mu\right\rangle \left\langle\psi_\mu\right|$ and the auxiliary state ρ_0 is the vacuum $\rho_0 = |\psi_0\rangle \langle\psi_0|$ then the distribution of $\tilde{X}$ coincides with the distribution of X (obtained with the covariant measurement), i.e. with $\mathcal{N}_2(\mu, I_2)$. Similarly when $\rho = \mathbb{N}_2(\mu, \sigma_\beta^2 I_2)$ then the distribution of $\tilde{X}$ coincides with $\mathcal{N}_2\left(\mu, \frac{(2\sigma_\beta^2+1)}{2} I_2\right)$.

Next, for the state $\mathbb{N}_2(\mu, \sigma^2 I_2)$, we consider the problem of Bayes estimation with quadratic loss of the parameter $\mu \in \mathbb{R}^2$ under a normal prior $\mu \sim \mathcal{N}_2\left(0, \sigma_0^2 I_2\right)$. The solution for a quadratic risk is given in [31], p. 55, with details and proofs in [30]. Consider the loss function

$$L(\hat{\mu}, \mu) = (\hat{\mu}_1 - \mu_1)^2 + (\hat{\mu}_2 - \mu_2)^2 .$$

The observable which is optimal in the Bayes sense (along with the equivalent POVM) is discussed in [29], Section 3. By Proposition 3 there, the optimal POVM is given by (denoting $m = (x, y)$)

$$M_c(B) = \frac{1}{2\pi c^2} \int_B \left|\psi_{m/c}\right\rangle \left\langle\psi_{m/c}\right| dm,\ B \in \mathfrak{B}.$$

From (26) and a change of variables, it can be easily verified that

$$\frac{1}{2\pi c^2} \int_{\mathbb{R}^2} \left|\psi_{m/c}\right\rangle \left\langle\psi_{m/c}\right| dm = \mathbf{1}$$

and hence M_c is a resolution of identity (the other properties of a POVM are also easy to check). Here $c = 2\sigma_0^2/(2\sigma_0^2 + 2\sigma^2 + 1)$ and the Bayes risk is given by (writing σ^2 as σ_β^2)

$$\inf_M R(M, \pi) = \frac{2\sigma_0^2(2\sigma_\beta^2 + 1)}{2(\sigma_0^2 + \sigma_\beta^2) + 1}. \tag{28}$$

When $\sigma^2 = 1/2$ (i.e. the state is pure), we have $c = \frac{\sigma_0^2}{\sigma_0^2+1}$ and the Bayes risk according to (17) in [29] thus becomes

$$\inf_M R(M, \pi) = \frac{2\sigma_0^2}{\sigma_0^2 + 1}. \tag{29}$$

The randomized measurements are then given by

$$\tilde{Q}_c = c(Q \otimes \mathbf{1}_0 + \mathbf{1} \otimes Q_0), \quad \tilde{P}_c = c(P \otimes \mathbf{1}_0 - \mathbf{1} \otimes P_0), \tag{30}$$

cf. equation (12) of [29]; it can easily be shown that these randomized measurements and M_c generate the same random variables.

From (30) it is clear that $\tilde{X}_c = (X_{\tilde{Q}_c}, X_{\tilde{P}_c})$ satisfies $\tilde{X}_c = c\tilde{X}$ where $\tilde{X}$ is given in (27). As $c < 1$, we note that Bayes estimation in the quantum case exhibits the same shrinkage phenomenon as witnessed in the classical counterpart.

However we note that the risk in (29) is greater than the Bayes risk in the classical model $\mathcal{N}_2(\mu, I_2/2)$ with the same prior, i.e. we obtain $2\sigma_0^2/(\sigma_0^2 + 1)$ in the quantum case compared to $2\sigma_0^2/(2\sigma_0^2 + 1)$ in the classical case. A similar statement holds for the thermal state as well. This inflation of risk is essentially due to the quantum nature of the data, in particular to the non-commutativity of the observables Q and P (as a consequence of which the random variables generated by Q and P separately do not admit a joint distribution).

5. Optimal estimation of the low-rank state

In this section, we outline a two-stage procedure to construct an estimator which is asymptotically minimax optimal. To construct the estimator, we split the sample into two parts. In Subsection 5.1 we

use the first part (which has sample size $\lfloor n^\delta \rfloor$) to obtain a preliminary estimator of ρ which lies within a ball of shrinking radius, centered at ρ, with high probability. Then we treat the initial estimator $\tilde{\rho}_n$ as the central state and parametrize the original state ρ as ρ_θ, so that it suffices to estimate θ. Now we use Theorem 3.3 to approximate the i.i.d. model by a limiting Gaussian model using the remaining part of the sample of size $n - \lfloor n^\delta \rfloor$. The risk in the limiting Gaussian model is computed in Section 5.2. In Subsection 5.3, we use the channel T_n^r and the risk transfer mechanism (22) to give an upper bound to the risk in the original i.i.d model. To show that the constructed estimator is asymptotically minimax optimal we choose an arbitrary state σ (which lies in ball of shrinking radius around ρ) as the central state and parametrize the original state ρ with a local parameter θ. Again, the i.i.d model is approximated by the limiting Gaussian model and using the channel S_n^r and equation (22), we obtain a lower bound for the minimax risk in the i.i.d model. We show that the upper and lower bounds match, ensuring that the constructed estimator is asymptotically minimax optimal.

For notational convenience, we work with n instead of $\lfloor n^\delta \rfloor$ while showing the properties of the preliminary estimator and replace n with $\lfloor n^\delta \rfloor$ in the final computations.

5.1. Preliminary estimator

Rank penalized and "physical" estimators have been constructed in [11] in the context of estimation of the joint quantum state of k two-dimensional systems (qubits), i.e. in a special case of our setting where $d = 2^k$. The authors obtain a least square estimator constructed using Pauli observables and then use rank penalization and spectral thresholding to obtain rank penalized and physical estimators, respectively. Not only do these estimators lie within a ball of radius $n^{-1/2+\epsilon}$ (for $\epsilon < 1/2$) centered at the original state with high probability; they are also rank consistent in the sense that the ranks of estimated states match the rank of the original state with high probability.

The general d dimensional case has been considered in [25], where the authors construct a least squares estimator using a so-called uniform POVM and project it into the space of density matrices to obtain an estimator which lies within a ball of radius $n^{-1/2+\epsilon}$ (for arbitrary small ϵ) centered at the original state with high probability. Since the authors are interested in the error of estimation in the trace norm, the rank consistency is not discussed in the paper. We show the least squares estimator of [25], when processed suitably like the physical estimator in [11], enjoys the rank consistency property. We note that in order to use LAN we need a local parametrization around a preliminary estimate of the state; it is essential that its rank matches that of the original state and hence rank consistency is crucial. The following theorem (proved in Appendix C) summarizes the above discussion.

Theorem 5.1. *Let ρ be a rank r state in $\mathbb{C}^d$ with the minimum eigenvalue $\lambda_r > 6\epsilon$. Then there exists an estimator $\tilde{\rho}_n$ of ρ such that the following concentration inequality holds:*

$$P[||\rho - \tilde{\rho}_n||_2^2 \geq 25r\epsilon^2] \leq de^{-3n\epsilon^2/16d}.$$

Further, if $\hat{r}$ is the rank of the state $\tilde{\rho}_n$, then

$$P[r = \hat{r}] \geq 1 - de^{-3n\epsilon^2/16d}.$$

5.2. Optimal risk in the Gaussian model

Here we consider minimax estimation in the limiting Gaussian model given in (17). Recall that the local parameter θ has two components: u corresponds to the classical part and $\mathfrak{z}$ corresponds to the

quantum part. First, we define the loss function with respect to which the risk will be computed:

$$\mathcal{L}(\theta,\hat{\theta}) = \sum_{i=1}^{r-1}(u_i - \hat{u}_i)^2 + (\sum_{i=1}^{r-1}(u_i - \hat{u}_i))^2 + 2\sum_{\substack{1\le i\le r\\ i<j\le d}}(\mu_i - \mu_j)|\eth_{ij} - \hat{\eth}_{ij}|^2. \tag{31}$$

Although the particular form of loss may seem arbitrary, we will see that the L^2 distance between two density matrices is approximately locally quadratic and then given by this loss. The following theorem gives the asymptotic minimax bound for estimation of θ:

Theorem 5.2. *Let $\Phi^{\theta,r}$ be the limiting Gaussian model given in (17), then*

$$\lim_{n\to\infty}\inf_{\hat{\theta}}\sup_{\theta\in\Theta_{n,r,\beta,\gamma}} E_\theta[\mathcal{L}(\theta,\hat{\theta})] = \lim_{n\to\infty}\inf_{M}\sup_{\theta\in\Theta_{n,r,\beta,\gamma}}\int \mathcal{L}(\theta,\hat{\theta})\,\mathrm{Tr}(\Phi^{\theta,r}M(d\hat{\theta}))$$

$$= \sum_{i=1}^{r}\mu_i(1-\mu_i) + \sum_{\substack{1\le i\le r\\ i<j\le d}} 2\mu_i.$$

Note that the limiting model is a tensor product of classical and quantum shifted Gaussian states. To show the upper bound we use a classical estimator and a POVM for the classical and quantum parts respectively. For the classical part, we use $Z \sim \mathcal{N}_{r-1}(u,V_\mu)$, as the estimator $\hat{u}$. For the quantum part, we note that each component of the tensored quantum Gaussian state is either a shifted thermal state or a shifted pure state indexed by (i,j). We will use the covariant measurement M (see Section 4) on each component and hence the joint POVM will be given by $\bar{M} = \bigotimes_{\substack{1\le i\le r\\ i<j\le d}} M$.

We use a Bayes risk to give a lower bound for the minimax risk, by setting up Gaussian priors for the parameters in both the classical and quantum parts. The Bayes estimator for the classical part is well known and is given by a shrinkage of the Gaussian random variable. For the quantum part, we directly compute the Bayes risk using equations (28) or (29) depending on whether the component is a thermal or pure Gaussian state. We again note that the optimal POVMs for each component are obtained by appropriately shrinking covariant measurements. We defer the proof to Appendix C.

5.3. Localization and the optimal risk

Finally, we are ready to compute the exact asymptotics of the minimax risk in the low-rank qudit model. First, we state a lemma (proved in Appendix D) that shows that the L^2 distance between two states is approximately a quadratic loss in the local parameters.

Lemma 5.3. *Let $\rho_{\theta^{(1)}/\sqrt{n},r}$ and $\rho_{\theta^{(2)}/\sqrt{n},r}$ be two states indexed by the local parameters $\theta^{(1)}$ and $\theta^{(2)}$, where $\theta^{(1)},\theta^{(2)} \in \Theta_{n,r,\beta,\gamma}$, then*

$$||\rho_{\theta^{(1)}/\sqrt{n},r} - \rho_{\theta^{(2)}/\sqrt{n},r}||_2^2 = \frac{1}{n}\sum_{i=1}^{r}(u_i^{(1)} - u_i^{(2)})^2 + \frac{2}{n}\sum_{\substack{1\le i\le r\\ i<j\le d}}(\mu_i-\mu_j)|\eth_{ij}^{(1)} - \eth_{ij}^{(2)}|^2$$

$$+ O\left(\frac{||\theta^{(1)}||^3, ||\theta^{(2)}||^3}{n^{3/2}}\right),$$

where $u_r^{(i)} = -\sum_{k=1}^{r-1} u_k^{(i)}$ for $i = 1,2$.

In Subsection 5.1 we have used the first part of the sample to construct a preliminary estimator $\tilde{\rho}_n$ which lies within an $n^{-1/2+\epsilon}$ ball of the original state ρ.

Note that $\tilde{\rho}_n$ is of standard form $\tilde{\mathcal{V}}\tilde{\rho}_{0,n}\tilde{\mathcal{V}}^*$ with $\tilde{\rho}_{0,n}$ diagonal and $\tilde{\mathcal{V}}$ unitary. Thus it can be brought into diagonal form by the channel $\tilde{L} : \rho \mapsto \tilde{\mathcal{V}}^*\rho\tilde{\mathcal{V}}$. All states parametrized by θ around the central state $\tilde{\rho}_{0,n}$ can be reparametrized around the central state $\tilde{\rho}_n$ by using the inverse transformation $\tilde{L}^* : \rho_\theta \mapsto \tilde{\mathcal{V}}\rho_\theta\tilde{\mathcal{V}}^*$. Since the norm $||.||_2$ is invariant under unitary transformation of states we will have

$$||\tilde{\rho}_{0,n} - \rho_\theta||_2 = ||\tilde{\rho}_n - L^*(\rho_\theta)||_2.$$

Henceforth we will assume that $\tilde{\rho}_n$ is diagonal i.e. we set it to be $\rho_{0,0,r}$ (see (4)). Thus the transformations L and L^* are understood to be part of the two stage estimation process, but in the notation, following [24,34], we will suppress them for the sake of brevity.

Now ρ, which lies within a $n^{-1/2+\epsilon}$ ball around $\tilde{\rho}_n$, can be parametrized with the local parameter θ, whereupon it is sufficient to estimate θ using $\hat{\theta}_n$. Thus we let $\rho = \rho_\theta$ and denote its estimator $\hat{\rho}_n$ by $\rho_{\hat{\theta}}$. Let $\mu_1 > \mu_2 > \ldots > \mu_r > 0$ be the eigenvalues of a density matrix σ of rank r and $||\rho - \sigma||_2 < n^{-1/2+\varepsilon}$. Since (24) holds one can easily transfer the risk between the i.i.d. and Gaussian models invoking (21) and (22), i.e. we have

$$\limsup_{n\to\infty} \sup_{||\rho-\sigma||_2<n^{-1/2+\varepsilon}} E_\rho n||\rho - \hat{\rho}_n||_2^2 \le \sum_{i=1}^{r} \mu_i(1-\mu_i) + \sum_{\substack{1\le i\le r\\ i<j\le d}} 2\mu_i.$$

The last part uses the upper bound of Theorem 5.2 and the continuity of eigenvalues, i.e. the eigenvalues of σ and the central state $\tilde{\rho}_n$ are asymptotically the same (see Appendix C for details). For the lower bound we choose an arbitrary σ and again parametrize the state and its estimate by local parameters θ and $\hat{\theta}$, i.e we let $\rho = \rho_\theta$ and $\hat{\rho}_n = \rho_{\hat{\theta}}$. Since (25) holds, we can transfer the risk between the i.i.d and Gaussian models invoking (21) and (22) and then using the lower bound of Theorem 5.2 we have

$$\liminf_{n\to\infty} \inf_{\hat{\rho}_n} \sup_{||\rho-\sigma||_2<n^{-1/2+\varepsilon}} E_\rho n||\rho - \hat{\rho}_n||_2^2 \ge \sum_{i=1}^{r} \mu_i(1-\mu_i) + \sum_{\substack{1\le i\le r\\ i<j\le d}} 2\mu_i.$$

We note that since the dimension d is kept constant, the $||.||_1$ and $||.||_2$ norms are equivalent and hence we can easily replace the L^2 neighborhood with a trace norm neighborhood. We define the trace norm neighborhood as follows:

$$\Sigma_{n,\varepsilon}(\sigma) := \left\{\rho \text{ is a rank } r \text{ state in } \mathcal{H} : \|\rho - \sigma\|_1 \le n^{-1/2+\varepsilon}\right\}. \tag{32}$$

The upper and lower bounds can now be summarized in the following theorem.

Theorem 5.4. *Let $\mu_1 > \mu_2 > \ldots > \mu_r > 0$ be the eigenvalues of a density matrix σ of rank r. Then there exists an $\varepsilon < 1/2$ such that*

$$\lim_{n\to\infty} \inf_{\hat{\rho}_n} \sup_{\rho\in\Sigma_{n,\varepsilon}(\sigma)} E_\rho n||\rho - \hat{\rho}_n||_2^2 = \sum_{i=1}^{r} \mu_i(1-\mu_i) + \sum_{\substack{1\le i\le r\\ i<j\le d}} 2\mu_i.$$

Note that in the above display we have used the first of the following equivalent expressions:

$$E_\rho n||\rho - \hat{\rho}_n||_2^2 = \int n||\rho - \hat{\rho}_n||_2^2 \operatorname{Tr}(\rho^{\otimes n} M(d\hat{\rho}_n))$$

and the infimum over $\hat{\rho}_n$ is an abbreviated notation for an infimum over all POVMs M.

6. Optimal estimation of linear functional of a low-rank state

We first discuss the classical problem of estimation of a linear functional of a density. Consider i.i.d observations $X_1, \ldots, X_n$ having density f on $[0,1]$, and consider a linear functional

$$\Psi(f) = \int \varphi(x) f(x) dx,$$

where φ is a bounded function on [0,1]. The estimator

$$\hat{\Psi}_n = n^{-1} \sum_{i=1}^{n} \varphi(X_i)$$

satisfies the CLT

$$n^{1/2} \left(\hat{\Psi}_n - \Psi(f) \right) \rightsquigarrow N\left(0, V_f^2\right),$$

where

$$V_f^2 = \operatorname{Var}_f(\varphi(X)).$$

One can show by LAN type results that $\hat{\Psi}_n$ is asymptotically optimal (see [37]), i.e. V_f^2 is the best possible variance in a local asymptotic minimax sense. In this section, we discuss an analogous result for the quantum case.

Suppose ρ is a state in a d-dimensional Hilbert space $\mathcal{H}$ of rank r and consider the functional

$$\Psi(\rho) = \operatorname{Tr}(A\rho), \tag{33}$$

where A is a self-adjoint matrix. We will denote the operator norm of A by $||A||$. Let X_A be the r.v. generated by A; then recall that the probability distribution of X_A is given by (1) and we denote the expectation by $E_\rho X_A = \langle A \rangle_\rho$. Also the variance of X_A is given by

$$\operatorname{Var}_\rho(X_A) = \left\langle A^2 \right\rangle_\rho - \langle A \rangle_\rho^2 = \left\langle \tilde{A}^2 \right\rangle_\rho =: V_\rho^2, \text{ where}$$
$$\tilde{A} := A - \langle A \rangle_\rho \mathbf{1}.$$

Let $\Sigma_{n,\varepsilon}(\rho_0)$ be a shrinking neighborhood in trace norm around a certain rank r state ρ_0 defined as in (32). We aim to show that with a sample $X_1, \ldots, X_n$, the estimator $\bar{X}_n$ attains the asymptotic variance $V_{\rho_0}^2$ uniformly over $\rho \in \Sigma_{n,\varepsilon}(\rho_0)$. For the purpose of proving an asymptotic minimax theorem, consider the expression

$$\sup_{\rho \in \Sigma_{n,\varepsilon}(\rho_0)} E_\rho \left(n^{1/2} \left(\bar{X}_n - \Psi(\rho) \right) \right)^2.$$

We will show that this expression converges to $V_{\rho_0}^2$.

To give a lower bound to the risk, we will have to find, for a d-dimensional center state ρ_0 and a parametrization of all d-dimensional nearby states, a "least favorable parametric subfamily" so that the information bound for estimating $\Psi(\rho)$ along this family coincides with $V_{\rho_0}^2$. We consider the following local model indexed by the single parameter t:

$$\rho_t = \rho_0 + n^{-1/2} tH, \quad t \in (-n^{\varepsilon}, n^{\varepsilon}),$$

where the matrix H is suitably chosen. The problem of estimation of the functional Ψ can be shown to be equivalent to the estimation of the parameter t in the limiting model of the form $\mathfrak{N}(t\tau, \mathscr{S})$ (see (19)) for appropriately chosen τ and $\mathscr{S}$. We then use a Bayesian result from [30] to establish a minimax bound for the estimation of t in the limiting model. We have the following theorem, the proof of which is deferred to Appendix C.

Theorem 6.1. *Let Ψ be the functional given in (33); then there exists an $\varepsilon < 1/2$ such that*

$$\lim_{n\to\infty} \inf_{\hat{\Psi}} \sup_{\rho \in \Sigma_{n,\varepsilon}(\rho_0)} E_\rho \left(n^{1/2} \left(\hat{\Psi} - \Psi(\rho) \right) \right)^2 = V_{\rho_0}^2.$$

Again note that in the above display we have used the first of the following equivalent expressions:

$$E_\rho \left(n^{1/2} \left(\hat{\Psi} - \Psi(\rho) \right) \right)^2 = \int \left(n^{1/2} \left(\hat{\Psi} - \Psi(\rho) \right) \right)^2 \mathrm{Tr}(\rho^{\otimes n} M(d\hat{\Psi}))$$

and the infimum over $\hat{\Psi}$ stands for an infimum taken over all POVMs M.

7. Discussion

We have obtained a generalization of the LAN result of [34] in this paper. Since a local asymptotic equivalence result was proved in [12] for rank 1 states in the infinite dimensional case, a natural question is about the limiting model in the case of an ensemble of rank r states if the Hilbert space is infinite dimensional. We conjecture that the limiting state will consist of three parts: a classical multivariate normal law, a tensor product of $r(r-1)/2$ shifted thermal states and a tensor product of infinitely many shifted pure states. Although this result is outside the scope of this paper, we observe that a restriction of the conjectured model to finite dimension will correspond to our result while a restriction to the rank 1 case (in infinite dimension) will correspond to the result obtained in [12].

Another interesting phenomenon occurs in the estimation of a functional where the limiting model is a tensor product of classical and quantum states indexed by a one-dimensional parameter. It is well known that if the loss function $L(\hat{\mu}, \mu)$ in Section 4 is replaced by a weighted loss function $L_g(\hat{\mu}, \mu)$ given by

$$L_g(\hat{\mu}, \mu) = g_1 (\hat{\mu}_1 - \mu_1)^2 + g_2 (\hat{\mu}_2 - \mu_2)^2,$$

then for $g_1 = 1$ and $g_2 = 0$, the POVM induced by Q suffices for optimal estimation, i.e. the Bayes risk is obtained by a simple measurement (the spectral measure of Q) and the associated Bayes risk matches the one in the the classical bivariate normal distribution $\mathcal{N}_2(\mu, \sigma^2 I_2)$. Analogously the risk in the limiting model for the functional estimation matches the same for a related classical model (see the remark after the proof of Theorem 6.1 in Appendix C). This is in sharp contrast with the risk observed in the measurement of the shift parameter of the coherent or thermal state ($g_1 = g_2 = 1$ in the expression of $L_g(\hat{\mu}, \mu)$), where we observe an inflation of the risk compared to the classical case

(see the last paragraph of Section 4). However, we emphasize that the classical-quantum Gaussian shift model indexed by a one-dimensional parameter is *not* equivalent to a classical model (in terms of Le Cam equivalence). A discussion of this phenomenon in a similar model can be found in [24] where the authors argue that although measuring the position observable results in a classical model whose "classical Fisher information" matches that of the quantum Fisher information of the original model, optimal testing procedures in the classical model are suboptimal in the sense that the associated risk is higher than that of the optimal testing procedure of the original quantum model. A more in-depth discussion in terms of sufficient statistics and Le Cam equivalence can be found in [23].

Acknowledgements

We would like to thank the Associate Editor and two anonymous referees whose suggestions helped in improving the quality of the paper.

Funding

The second author was supported in part by NSF Grant DMS-1915884.

Supplementary Material

Supplement to "Minimax estimation of low-rank quantum states and their linear functionals" (DOI: 10.3150/23-BEJ1610SUPP; .pdf). Appendix A of the supplementary material contains a brief review of some representation theoretic tools needed for proving the LAN result. The Bayesian risk for a one dimensional paramter is given in Appendix B. The proofs of the main theorems and the supporting technical lemmas are given in Appendix C and D respectively.

References

[1] Acharya, A. and Guţă, M. (2017). Statistical analysis of compressive low rank tomography with random measurements. *J. Phys. A* **50** 195301, 21 pp. MR3635730 https://doi.org/10.1088/1751-8121/aa682e

[2] Acharya, A., Kypraios, T. and Guţă, M. (2016). Statistically efficient tomography of low rank states with incomplete measurements. *New J. Phys.* **18** 043018. https://doi.org/10.1088/1367-2630/18/4/043018

[3] Acharya, A., Kypraios, T. and Guţă, M. (2019). A comparative study of estimation methods in quantum tomography. *J. Phys. A* **52** 234001, 36 pp. MR3953819 https://doi.org/10.1088/1751-8121/ab1958

[4] Acharya, J., Issa, I., Shende, N.V. and Wagner, A.B. (2020). Estimating quantum entropy. *IEEE J. Sel. Areas Inf. Theory* **1** 454–468. https://doi.org/10.1109/JSAIT.2020.3015235

[5] Alquier, P., Butucea, C., Hebiri, M., Meziani, K. and Morimae, T. (2013). Rank-penalized estimation of a quantum system. *Phys. Rev. A* **88** 032113. https://doi.org/10.1103/PhysRevA.88.032113

[6] Audenaert, K.M.R., Calsamiglia, J., Muñoz Tapia, R., Bagan, E., Masanes, L., Acin, A. and Verstraete, F. (2007). Discriminating states: The quantum Chernoff bound. *Phys. Rev. Lett.* **98** 160501. https://doi.org/10.1103/PhysRevLett.98.160501

[7] Audenaert, K.M.R., Nussbaum, M., Szkoła, A. and Verstraete, F. (2008). Asymptotic error rates in quantum hypothesis testing. *Comm. Math. Phys.* **279** 251–283. MR2377635 https://doi.org/10.1007/s00220-008-0417-5

[8] Barndorff-Nielsen, O.E., Gill, R.D. and Jupp, P.E. (2003). On quantum statistical inference. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **65** 775–816. MR2017871 https://doi.org/10.1111/1467-9868.00415

[9] Braunstein, S.L. and Caves, C.M. (1994). Statistical distance and the geometry of quantum states. *Phys. Rev. Lett.* **72** 3439–3443. MR1274631 https://doi.org/10.1103/PhysRevLett.72.3439
[10] Brown, L.D. and Low, M.G. (1996). Asymptotic equivalence of nonparametric regression and white noise. *Ann. Statist.* **24** 2384–2398. MR1425958 https://doi.org/10.1214/aos/1032181159
[11] Butucea, C., Guţă, M. and Kypraios, T. (2015). Spectral thresholding quantum tomography for low rank states. *New J. Phys.* **17** 113050. https://doi.org/10.1088/1367-2630/17/11/113050
[12] Butucea, C., Guţă, M. and Nussbaum, M. (2018). Local asymptotic equivalence of pure states ensembles and quantum Gaussian white noise. *Ann. Statist.* **46** 3676–3706. MR3852665 https://doi.org/10.1214/17-AOS1672
[13] Butucea, C., Guţă, M. and Nussbaum, M. (2018). Supplementary material to "Local asymptotic equivalence of pure states ensembles and quantum Gaussian white noise." Ann. Statist. **46**. https://doi.org/10.1214/17-AOS1672SUPP
[14] Cai, T., Kim, D., Wang, Y., Yuan, M. and Zhou, H.H. (2016). Optimal large-scale quantum state tomography with Pauli measurements. *Ann. Statist.* **44** 682–712. MR3476614 https://doi.org/10.1214/15-AOS1382
[15] Carpentier, A., Eisert, J., Gross, D. and Nickl, R. (2019). Uncertainty quantification for matrix compressed sensing and quantum tomography problems. In *High Dimensional Probability VIII – The Oaxaca Volume*. *Progress in Probability* **74** 385–430. Cham: Birkhäuser/Springer. MR4181374 https://doi.org/10.1007/978-3-030-26391-1_18
[16] Carter, A.V. (2002). Deficiency distance between multinomial and multivariate normal experiments. *Ann. Statist.* **30** 708–730. MR1922539 https://doi.org/10.1214/aos/1028674839
[17] Flammia, S.T., Gross, D., Liu, Y.-K. and Eisert, J. (2012). Quantum tomography via compressed sensing: Error bounds, sample complexity, and efficient estimators. *New J. Phys.* **14** 095022. https://doi.org/10.1088/1367-2630/14/9/095022
[18] Fulton, W. and Harris, J. (1991). *Representation Theory: A First Course, Readings in Mathematics*. *Graduate Texts in Mathematics* **129**. New York: Springer. MR1153249 https://doi.org/10.1007/978-1-4612-0979-9
[19] Gill, R.D. and Guţă, M.I. (2013). On asymptotic quantum statistical inference. In *From Probability to Statistics and Back: High-Dimensional Models and Processes*. *Inst. Math. Stat. (IMS) Collect.* **9** 105–127. Beachwood, OH: IMS. MR3186752 https://doi.org/10.1214/12-IMSCOLL909
[20] Gross, D., Liu, Y.-K., Flammia, S.T., Becker, S. and Eisert, J. (2010). Quantum state tomography via compressed sensing. *Phys. Rev. Lett.* **105** 150401. https://doi.org/10.1103/PhysRevLett.105.150401
[21] Guţă, M. and Butucea, C. (2010). Quantum U-statistics. *J. Math. Phys.* **51** 102202, 24 pp. MR2761295 https://doi.org/10.1063/1.3476776
[22] Guţă, M., Janssens, B. and Kahn, J. (2008). Optimal estimation of qubit states with continuous time measurements. *Comm. Math. Phys.* **277** 127–160. MR2357428 https://doi.org/10.1007/s00220-007-0357-5
[23] Guţă, M. and Jenčová, A. (2007). Local asymptotic normality in quantum statistics. *Comm. Math. Phys.* **276** 341–379. MR2346393 https://doi.org/10.1007/s00220-007-0340-1
[24] Guţă, M. and Kahn, J. (2006). Local asymptotic normality for qubit states. *Phys. Rev. A (3)* **73** 052108, 15 pp. MR2229156 https://doi.org/10.1103/PhysRevA.73.052108
[25] Guţă, M., Kahn, J., Kueng, R. and Tropp, J.A. (2020). Fast state tomography with optimal error bounds. *J. Phys. A* **53** 204001, 28 pp. MR4093475 https://doi.org/10.1088/1751-8121/ab8111
[26] Hayashi, M. (2017). *Quantum Information Theory: Mathematical Foundation*, 2nd ed. *Graduate Texts in Physics*. Berlin: Springer. MR3558531 https://doi.org/10.1007/978-3-662-49725-8
[27] Helstrom, C.W. (1976). *Quantum Detection and Estimation Theory*. New York: Academic Press.
[28] Holevo, A. (2011). *Probabilistic and Statistical Aspects of Quantum Theory*, 2nd ed. *Quaderni/Monographs* **1**. Pisa: Edizioni della Normale. MR2797301 https://doi.org/10.1007/978-88-7642-378-9
[29] Holevo, A.S. (1973). Optimal quantum measurements. *Theoret. Math. Phys.* **17** 1172–1177. https://doi.org/10.1007/BF01037594
[30] Holevo, A.S. (1978). *Investigations in the General Theory of Statistical Decisions*. Providence, RI: Amer. Math. Soc. MR0652516
[31] Holevo, A.S. (2001). *Statistical Structure of Quantum Theory*. *Lecture Notes in Physics. Monographs* **67**. Berlin: Springer. MR1889193 https://doi.org/10.1007/3-540-44998-1
[32] Ibragimov, I.A. and Has'minskiĭ, R.Z. (1981). *Statistical Estimation: Asymptotic Theory*. *Applications of Mathematics* **16**. New York–Berlin: Springer. MR0620321

[33] Kahn, J. and Guţă, M. (2008). Local asymptotic normality and optimal estimation for d-dimensional quantum systems. In *Quantum Stochastics and Information: Statistics, Filtering and Control* 300–322. Hackensack, NJ: World Sci. Publ. MR2562438 https://doi.org/10.1142/9789812832962_0014

[34] Kahn, J. and Guţă, M. (2009). Local asymptotic normality for finite dimensional quantum systems. *Comm. Math. Phys.* **289** 597–652. MR2506764 https://doi.org/10.1007/s00220-009-0787-3

[35] Koltchinskii, V. (2011). Von Neumann entropy penalization and low-rank matrix estimation. *Ann. Statist.* **39** 2936–2973. MR3012397 https://doi.org/10.1214/11-AOS926

[36] Koltchinskii, V. and Xia, D. (2015). Optimal estimation of low rank density matrices. *J. Mach. Learn. Res.* **16** 1757–1792. MR3417798

[37] Koševnik, J.A. and Levit, B.J. (1976). On a nonparametric analogue of the information matrix. *Theory Probab. Appl.* **21** 738–753. https://doi.org/10.1137/1121087

[38] Lahiry, S. and Nussbaum, M. (2024). Supplement to "Minimax estimation of low-rank quantum states and their linear functionals." https://doi.org/10.3150/23-BEJ1610SUPP

[39] Le Cam, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. *Springer Series in Statistics*. New York: Springer. MR0856411 https://doi.org/10.1007/978-1-4612-4946-7

[40] Li, K. (2016). Discriminating quantum states: The multiple Chernoff distance. *Ann. Statist.* **44** 1661–1679. MR3519936 https://doi.org/10.1214/16-AOS1436

[41] Liu, Y.-K. (2011). Universal low-rank matrix recovery from Pauli measurements. *Adv. Neural Inf. Process. Syst.* **24**.

[42] Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24** 2399–2430. MR1425959 https://doi.org/10.1214/aos/1032181160

[43] Nussbaum, M. and Szkoła, A. (2009). The Chernoff lower bound for symmetric quantum hypothesis testing. *Ann. Statist.* **37** 1040–1057. MR2502660 https://doi.org/10.1214/08-AOS593

[44] Ogawa, T. and Nagaoka, H. (2000). Strong converse and Stein's lemma in quantum hypothesis testing. *IEEE Trans. Inf. Theory* **46** 2428–2433. MR1806811 https://doi.org/10.1109/18.887855

[45] Parthasarathy, K.R. (1992). *An Introduction to Quantum Stochastic Calculus*. *Modern Birkhäuser Classics*. Basel: Birkhäuser/Springer Basel AG. MR3012668

[46] Petz, D. (2008). *Quantum Information Theory and Quantum Statistics*. *Theoretical and Mathematical Physics*. Berlin: Springer. MR2363070

[47] Shiryaev, A.N. and Spokoiny, V.G. (2000). *Statistical Experiments and Decisions: Asymptotic Theory*. *Advanced Series on Statistical Science & Applied Probability* **8**. River Edge, NJ: World Scientific Co., Inc. MR1791434 https://doi.org/10.1142/9789812779243

[48] Strasser, H. (1985). *Mathematical Theory of Statistics: Statistical Experiments and Asymptotic Decision Theory*. *De Gruyter Studies in Mathematics* **7**. Berlin: de Gruyter. MR0812467 https://doi.org/10.1515/9783110850826

[49] Tsang, M., Albarelli, F. and Datta, A. (2020). Quantum semiparametric estimation. *Phys. Rev. X* **10** 031023. https://doi.org/10.1103/PhysRevX.10.031023

[50] Wang, Y. (2011). Quantum Monte Carlo simulation. *Ann. Appl. Stat.* **5** 669–683. MR2840170 https://doi.org/10.1214/10-AOAS406

[51] Wang, Y. (2012). Quantum computation and quantum information. *Statist. Sci.* **27** 373–394. MR3012432 https://doi.org/10.1214/11-STS378

[52] Wang, Y. and Song, X. (2020). Quantum science and quantum technology. *Statist. Sci.* **35** 51–74. MR4071358 https://doi.org/10.1214/19-STS745

[53] Wang, Y., Wu, S. and Zou, J. (2016). Quantum annealing with Markov chain Monte Carlo simulations and D-wave quantum computers. *Statist. Sci.* **31** 362–398. MR3552740 https://doi.org/10.1214/16-STS560

[54] Yamagata, K., Fujiwara, A. and Gill, R.D. (2013). Quantum local asymptotic normality based on a new quantum likelihood ratio. *Ann. Statist.* **41** 2197–2217. MR3127863 https://doi.org/10.1214/13-AOS1147

Received March 2022 and revised March 2023