The Annals of Statistics
2023, Vol. 51, No. 3, 1159–1182
https://doi.org/10.1214/23-AOS2285

EFFICIENCY OF ESTIMATORS FOR LOCALLY ASYMPTOTICALLY NORMAL QUANTUM STATISTICAL MODELS

BY AKIO FUJIWARA[1,a] AND KOICHI YAMAGATA[2,b]

[1]*Department of Mathematics, Osaka University,* [a]*fujiwara@math.sci.osaka-u.ac.jp*
[2]*Principles of Informatics Research Division, National Institute of Informatics,* [b]*yamagata@nii.ac.jp*

We herein establish an asymptotic representation theorem for locally asymptotically normal quantum statistical models. This theorem enables us to study the asymptotic efficiency of quantum estimators, such as quantum regular estimators and quantum minimax estimators, leading to a universal tight lower bound beyond the i.i.d. assumption. This formulation complements the theory of quantum contiguity developed in the previous paper [Fujiwara and Yamagata, *Bernoulli* **26** (2020) 2105–2141], providing a solid foundation of the theory of weak quantum local asymptotic normality.

1. Introduction. In classical statistics a sequence $\{P_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ of statistical models on measurable spaces $(\Omega^{(n)}, \mathcal{F}^{(n)})$ is called *locally asymptotically normal* (LAN) at $\theta_0 \in \Theta$ (in the "weak" sense) if the log-likelihood ratio $\log(dP_\theta^{(n)}/dP_{\theta_0}^{(n)})$ is expanded in the local parameter $h := \sqrt{n}(\theta - \theta_0)$ as

$$\log \frac{dP^{(n)}_{\theta_0+h/\sqrt{n}}}{dP^{(n)}_{\theta_0}} = h^i \Delta_i^{(n)} - \frac{1}{2} h^i h^j J_{ij} + o_{P_{\theta_0}}(1). \tag{1}$$

Here $\Delta^{(n)} = (\Delta_1^{(n)}, \ldots, \Delta_d^{(n)})$ is a list of d-dimensional random vectors on each $(\Omega^{(n)}, \mathcal{F}^{(n)})$ that exhibits

$$\Delta^{(n)} \overset{0}{\rightsquigarrow} N(0, J)$$

with J being a $d \times d$ real symmetric strictly positive matrix, the arrow $\overset{h}{\rightsquigarrow}$ stands for the convergence in distribution under $P^{(n)}_{\theta_0+h/\sqrt{n}}$, the remainder term $o_{P_{\theta_0}}(1)$ converges in probability to zero under $P^{(n)}_{\theta_0}$, and Einstein's summation convention is used.

There is an obvious similarity between (1) and the log-likelihood ratio of the Gaussian shift model,

$$\log \frac{dN(Jh, J)}{dN(0, J)}(X_1, \ldots, X_d) = h^i X_i - \frac{1}{2} h^i h^j J_{ij}.$$

In fact, this similarity is a manifestation of a profound connection between the local parameter model $\{P^{(n)}_{\theta_0+h/\sqrt{n}} : h \in \mathbb{R}^d\}$ and the Gaussian shift model $\{N(Jh, J) : h \in \mathbb{R}^d\}$, playing an important role in asymptotic statistics [21].

In general, a statistical theory comprises two parts: one is to prove the existence of a statistic that possesses a certain desired property (direct part), and the other is to prove the nonexistence of a statistic that exceeds that property (converse part). In the problem of asymptotic efficiency, the converse part, the impossibility to do asymptotically better than the best,

Received August 2022; revised February 2023.
MSC2020 subject classifications. Primary 62F12, 81P50; secondary 62A01.
Key words and phrases. Quantum statistics, local asymptotic normality, asymptotic representation theorem, asymptotic efficiency, regular estimator, minimax estimator.

which can be done in the limit situation, is ensured by the so-called asymptotic representation theorem [21], Theorem 7.10.

THEOREM 1.1 (Asymptotic representation theorem). *Assume that* $\{P_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ *is LAN at* $\theta_0 \in \Theta$. *Let* $T^{(n)}$ *be statistics on the local models* $P_{\theta_0+h/\sqrt{n}}^{(n)}$ *that are weakly convergent under every* $h \in \mathbb{R}^d$. *Then there exists a randomized statistic* T *on the Gaussian shift model* $N(Jh, J)$ *such that* $T^{(n)} \overset{h}{\rightsquigarrow} T$ *for every* h.

For an accessible proof, see Section A of the Supplementary Material [4]. Theorem 1.1 allows us to deduce in several precise mathematical senses that no estimator can asymptotically do better than what can be achieved in the limiting Gaussian shift model. For example, this theorem leads to the convolution theorem, which tells us that *regular* estimators (estimators whose asymptotic behavior in a small neighborhood of θ_0 is more or less stable as the parameter varies) have a limiting distribution which in a very strong sense is more disperse than the optimal limiting distribution, which we expect from the limiting statistical problem. Another option is to use the representation theorem to derive the asymptotic minimax theorem, telling us that the worst behavior of an estimator as θ varies in a shrinking neighborhood of θ_0 cannot improve on what we expect from the limiting problem. This theorem applies to *all* possible estimators but only discusses their *worst* behavior in a neighborhood of θ_0.

Extending the notion of local asymptotic normality to the quantum domain was pioneered by Guţă and Kahn [8, 17]. They proved that, given a quantum parametric model $\mathcal{S}(\mathbb{C}^D) = \{\rho_\theta > 0 : \theta \in \Theta \subset \mathbb{R}^{D^2-1}\}$ comprising the totality of faithful density operators on a D-dimensional Hilbert space and a point θ_0 on the parameter space Θ at which ρ_{θ_0} is nondegenerate (i.e., every eigenvalue of ρ_{θ_0} is simple), there exist quantum channels $\Gamma^{(n)}$ and $\Lambda^{(n)}$ as well as compact sets $K^{(n)} \subset \mathbb{R}^{D^2-1}$, satisfying $K^{(n)} \uparrow \mathbb{R}^{D^2-1}$, such that

$$\lim_{n\to\infty} \sup_{h\in K^{(n)}} \left\| \sigma_h - \Gamma^{(n)}\left(\rho_{\theta_0+h/\sqrt{n}}^{\otimes n}\right) \right\|_1 = 0 \quad \text{and} \quad \lim_{n\to\infty} \sup_{h\in K^{(n)}} \left\| \Lambda^{(n)}(\sigma_h) - \rho_{\theta_0+h/\sqrt{n}}^{\otimes n} \right\|_1 = 0,$$

where $\{\sigma_h : h \in \mathbb{R}^{D^2-1}\}$ is a family of classical/quantum-mixed Gaussian shift model. Later, Lahiry and Nussbaum [19] extended their formulation to models that comprise nonfaithful density operators but have the same rank. Note that these formulations are not a direct analogue of the weak LAN defined by (1); in particular, the convergence to a quantum Gaussian shift model is evaluated not by the convergence in distribution but by the convergence in trace norm. In this sense their formulation could be referred to as a "strong" q-LAN (cf. [5]). Meanwhile, Guţă and Jenčová [7] also tried to formulate a "weak" q-LAN, based on the Connes cocycle derivative, which was sometimes regarded as a proper quantum analogue of the likelihood ratio. However, they did not establish an asymptotic expansion formula, which would be directly analogous to (1).

A different approach to a "weak" q-LAN was put forward by the present authors [3, 23]. Given two quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$ on a finite dimensional Hilbert space $\mathcal{H}$, define the *square-root likelihood ratio* $\mathcal{R}(\sigma \mid \rho)$ of σ relative to ρ as the positive operator R, satisfying the quantum Lebesgue decomposition $\sigma = R\rho R + \sigma^\perp$, where the singular part $\sigma^\perp$ is the positive operator that satisfies $\operatorname{Tr} \rho\sigma^\perp = 0$. The notion of (weak) q-LAN is defined as follows. (See [3] for details.)

DEFINITION 1.2 (q-LAN). A sequence $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} \mid \theta \in \Theta \subset \mathbb{R}^d\}$ of quantum statistical models on Hilbert spaces $\mathcal{H}^{(n)}$ is called *quantum locally asymptotically normal* (q-LAN) at $\theta_0 \in \Theta$ if the square-root likelihood ratio $R_h^{(n)} = \mathcal{R}(\rho_{\theta_0+h/\sqrt{n}}^{(n)} \mid \rho_{\theta_0}^{(n)})$ is expanded in $h \in \mathbb{R}^d$ as

$$\log\bigl(R_h^{(n)} + o_{L^2}(\rho_{\theta_0}^{(n)})\bigr)^2 = h^i \Delta_i^{(n)} - \frac{1}{2}(h^i h^j J_{ij}) I^{(n)} + o_D(h^i \Delta_i^{(n)}, \rho_{\theta_0}^{(n)}). \tag{2}$$

Here $\Delta^{(n)} = (\Delta_1^{(n)}, \ldots, \Delta_d^{(n)})$ is a list of observables on each $\mathcal{H}^{(n)}$ that exhibits

$$\Delta^{(n)} \overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N(0, J)$$

with J being a $d \times d$ complex nonnegative matrix[1] satisfying $\operatorname{Re} J > 0$; the arrow $\overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow}$ stands for the quantum convergence in distribution under $\rho_{\theta_0}^{(n)}$, defined by the convergence of the quasi-characteristic function, and $o_{L^2}(\rho_{\theta_0}^{(n)})$ and $o_D(h^i \Delta_i^{(n)}, \rho_{\theta_0}^{(n)})$ are infinitesimal remainder terms in L^2 and in distribution, respectively.

One may recognize a clear parallelism between the classical definition (1) and the quantum one (2). In fact, the theory of weak q-LAN, based on (2), has been successfully applied to quantum statistical models satisfying mild regularity conditions, culminating in the derivation of (an abstract version of) the quantum Le Cam third lemma [3]. However, this theory is not yet fully satisfactory because it lacks tools to cope with the converse problems, that is, to prove the impossibility of doing asymptotically better than the best which can be done on the limiting model specified by the quantum Le Cam third lemma. For example, we do not know conditions to get rid of asymptotically super-efficient estimators that break the Holevo bound in an i.i.d. model.

In the context of these circumstances, we aim to establish a noncommutative counterpart of Theorem 1.1 that enables us to study the converse part in quantum asymptotic statistics. The paper is organized as follows. In Section 2 we summarize the main results, including the asymptotic quantum representation theorem for q-LAN models, and a universal tight bound for efficiency that generalizes the Holevo bound to generic (not necessarily i.i.d.) models. This section will also serve as an overview of the paper. In Section 3 we provide some mathematical tools and a number of lemmas that are used in the proof of the representation theorem, and the proof of the theorem itself is carried over to the succeeding Section 4. In Section 5 we apply the representation theorem to the analysis of efficiency for sequences of quantum estimators such as the quantum Hodges estimator, quantum regular estimators, quantum minimax estimators, and the quantum James–Stein estimator. Section 6 is devoted to concluding remarks.

Some additional materials are provided in the Supplementary Material [4], including a proof of Theorem 1.1 (Section A), a comprehensible account of degenerate canonical commutation relation (CCR) and hybrid classical/quantum Gaussian states (Section B), a detailed account of the notion of D-extendibility (Section C), and proofs of lemmas and theorems presented in Sections 3 and 5 (Section D and E, respectively).

2. Main results. Assume that a sequence $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ of quantum statistical models is q-LAN at $\theta_0 \in \Theta$, as in Definition 1.2. In view of the classical representation theorem (Theorem 1.1), one may envisage the following.

[1]For a complex covariance matrix J, the state $N(0, J)$ is regarded as a hybrid classical/quantum Gaussian state. Specifically, $N(0, J)$ is classical if and only if $\operatorname{Im} J = 0$ and is purely quantum if and only if $\operatorname{Im} J$ is invertible. For more information, see Section B of the Supplementary Material.

CONJECTURE. Let $M^{(n)} = \{M^{(n)}(B)\}_{B\in\mathcal{B}(\mathbb{R}^s)}$ be a sequence of POVMs over the Borel σ-algebra $\mathcal{B}(\mathbb{R}^s)$ of $\mathbb{R}^s$ such that the corresponding sequence of classical probability measures

$$\mathcal{L}_h^{(n)} := \operatorname{Tr} \rho^{(n)}_{\theta_0+h/\sqrt{n}} M^{(n)}$$

is weakly convergent to some probability measure $\mathcal{L}_h$ for every h. Then there would exist a POVM $M^{(\infty)} = \{M(B)\}_{B\in\mathcal{B}(\mathbb{R}^s)}$ on CCR(Im J) such that

$$\phi_h\bigl(M^{(\infty)}(B)\bigr) = \mathcal{L}_h(B)$$

for every h, where $\phi_h \sim N((\operatorname{Re} J)\,h, J)$.

However, such a naive guess fails, as the following example shows.

EXAMPLE 2.1. Let us consider the following one-dimensional pure state model:

$$\rho_\theta = \frac{2}{e^\theta + e^{-\theta}}\, e^{\frac{\theta}{2}\sigma_x}\,\rho_0\, e^{\frac{\theta}{2}\sigma_x}, \quad (\theta \in \mathbb{R}),$$

where

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \rho_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

This model has an SLD σ_x at $\theta = 0$. Let

$$\Delta^{(n)} := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} I^{\otimes(k-1)} \otimes \sigma_x \otimes I^{\otimes(n-k)}.$$

Then it is shown (cf. [23], Section 3.2, [3], Section 7.3) that $\rho_\theta^{\otimes n}$ is q-LAN at $\theta = 0$, and

$$\Delta^{(n)} \overset{\rho^{\otimes n}_{h/\sqrt{n}}}{\rightsquigarrow} N(h, 1).$$

However, there is a sequence of POVMs that does not have a limiting POVM on the (classical) Gaussian shift model $N(h, 1)$.

Let $M^{(n)}$ be a binary-valued POVM on $\rho^{\otimes n}_{h/\sqrt{n}}$, defined by

$$M^{(n)}(0) = \rho_0^{\otimes n}, \qquad M^{(n)}(1) = I^{(n)} - \rho_0^{\otimes n}.$$

Then

$$\lim_{n\to\infty} \operatorname{Tr} \rho^{\otimes n}_{h/\sqrt{n}} M^{(n)}(0) = \lim_{n\to\infty} (\operatorname{Tr} \rho_{h/\sqrt{n}}\rho_0)^n = e^{-\frac{1}{4}h^2},$$

and thus the sequence of POVMs has a limiting distribution

$$\mathcal{L}_h(0) = e^{-\frac{1}{4}h^2}, \qquad \mathcal{L}_h(1) = 1 - e^{-\frac{1}{4}h^2}$$

for each $h \in \mathbb{R}$.

Now suppose that this distribution is realized by a binary-valued POVM $M^{(\infty)}$ that is independent of h. Since the limiting Gaussian shift model $N(h, 1)$ is classical, $M^{(\infty)}$ is represented by a measurable function $m(x)$ on $\mathbb{R}$ such that

$$M^{(\infty)}(0) = m(x), \qquad M^{(\infty)}(1) = 1 - m(x).$$

Specifically, $0 \le m(x) \le 1$ for all $x \in \mathbb{R}$, and

$$e^{-\frac{1}{4}h^2} = \int_{-\infty}^{\infty} m(x)\, p_h(x)\, dx \tag{3}$$

for all $h \in \mathbb{R}$, where $p_h(x) = \frac{1}{\sqrt{2\pi}} e^{-(x-h)^2/2}$ is the density function of $N(h,1)$. However, (3) has the solution

$$m(x) = \sqrt{2}\, e^{-\frac{1}{2}x^2} \quad \text{(a.e.)},$$

which does not fulfill the requirement that $0 \le m(x) \le 1$. This is a contradiction.

Example 2.1 demonstrates that we need some additional condition to establish an asymptotic representation theorem in the quantum domain. In fact, the following condition will prove to be sufficient.

DEFINITION 2.2 (D-extendibility). Given a sequence $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ of quantum statistical models on $\mathcal{H}^{(n)}$, a sequence $X^{(n)} = (X_1^{(n)}, \dots, X_r^{(n)})$ of observables on $\mathcal{H}^{(n)}$ is called *asymptotically D-invariant* at $\theta_0 \in \Theta$ if it fulfills the following requirements:

$$X^{(n)} \overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N(0, \Sigma) \tag{4}$$

for some $r \times r$ nonnegative matrix Σ with $\operatorname{Re}\Sigma > 0$, and

$$\lim_{n\to\infty} \operatorname{Tr} \sqrt{\rho_{\theta_0}^{(n)}}\, e^{\sqrt{-1}\xi^i X_i^{(n)}} \sqrt{\rho_{\theta_0}^{(n)}}\, e^{\sqrt{-1}\eta^i X_i^{(n)}} = e^{-\frac{1}{2}\binom{\xi}{\eta}^\top \begin{pmatrix} \Sigma & \Sigma\#\Sigma^\top \\ \Sigma\#\Sigma^\top & \Sigma^\top \end{pmatrix}\binom{\xi}{\eta}} \tag{5}$$

for all $\xi, \eta \in \mathbb{R}^r$, where $\#$ stands for the operator geometric mean [1, 18].

A sequence $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ of quantum statistical models that is q-LAN at $\theta_0 \in \Theta$ is called *D-extendible* at θ_0 if there exists a sequence $X^{(n)} = (X_i^{(n)})_{1 \le i \le r}$ of observables as well as an $r \times d$ real matrix F such that

$$\Delta_k^{(n)} = F_k^i X_i^{(n)} \quad (1 \le k \le d, n \in \mathbb{N}) \tag{6}$$

and $X^{(n)}$ is asymptotically D-invariant at $\theta_0 \in \Theta$. Such a sequence $X^{(n)}$ is called a *D-extension* of $\Delta^{(n)}$.

REMARK 2.3. One may have the impression that the condition (5) is strange and intractable; but in reality it is not too restrictive in applications. For example, let $\mathcal{S} = \{\rho_\theta : \theta \in \Theta \subset \mathbb{R}^d\}$ be a quantum statistical model on a finite dimensional Hilbert space $\mathcal{H}$. Then, under some mild regularity conditions, the sequence $\mathcal{S}^{(n)} := \{\rho_\theta^{\otimes n} : \theta \in \Theta \subset \mathbb{R}^d\}$ of i.i.d. models on $\mathcal{H}^{\otimes n}$ is not only q-LAN at a given $\theta_0 \in \Theta$ [3], Theorem 7.6, but also D-extendible at θ_0. For a proof, see Section C of the Supplementary Material [4], where the idea behind the term "asymptotic D-invariance" is also clarified and a proper perspective on the model in Example 2.1 is demonstrated. There are, of course, models $\mathcal{S}^{(n)}$ that are non-i.i.d. but are, nevertheless, q-LAN and D-extendible; a simple example is provided in Section C of the Supplementary Material [4].

With this additional requirement of D-extendibility, we can prove the following.

THEOREM 2.4 (Asymptotic quantum representation theorem). *Assume that a sequence $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ of quantum statistical models is q-LAN and D-extendible at $\theta_0 \in \Theta$. Let $M^{(n)} = \{M^{(n)}(B)\}_{B \in \mathcal{B}(\mathbb{R}^s)}$ be a sequence of POVMs over $\mathbb{R}^s$ such that the corresponding sequence of classical probability measures*

$$\mathcal{L}_h^{(n)} := \operatorname{Tr} \rho_{\theta_0 + h/\sqrt{n}}^{(n)} M^{(n)}$$

is weakly convergent to some probability measure $\mathcal{L}_h$ for every h. Then there exists a POVM $M^{(\infty)} = \{M(B)\}_{B\in\mathcal{B}(\mathbb{R}^s)}$ on CCR(Im Σ) *such that*

$$\phi_h(M^{(\infty)}(B)) = \mathcal{L}_h(B)$$

for every h, where $\phi_h \sim N((\operatorname{Re}\tau)h, \Sigma)$ with $\tau = \Sigma F$.

Theorem 2.4 allows us to convert a statistical problem for the local parameter model $\{\rho^{(n)}_{\theta_0+h/\sqrt{n}} : h \in \mathbb{R}^d\}$ into another one for the limiting quantum Gaussian shift model $\{N((\operatorname{Re}\tau)h, \Sigma) : h \in \mathbb{R}^d\}$. It is thus natural to expect that the Holevo bound[2] for the limiting model $N((\operatorname{Re}\tau)h, \Sigma)$, given a weight matrix $G > 0$, that is,

$$c_G^{(\mathrm{rep})} := \min_K\{\operatorname{Tr} GZ + \operatorname{Tr}|\sqrt{G}\operatorname{Im} Z\sqrt{G}| : Z = K^\top \Sigma K, \quad K \text{ is an } r\times d \text{ real matrix satisfying } K^\top(\operatorname{Re}\tau) = I\} \tag{7}$$

will be of fundamental importance in quantum asymptotics. Note that the D-extension in Definition 2.2 is not unique; however, it can be shown that the bound $c_G^{(\mathrm{rep})}$ is independent of the choice of a D-extension (Corollary 5.4). In what follows, we shall call this universal bound the asymptotic representation bound.

Indeed, the bound $c_G^{(\mathrm{rep})}$ plays a crucial role in asymptotic quantum statistics. For example, it gives the ultimate limit of estimation for regular estimators (Theorems 5.2 and 5.3) and minimax estimators (Theorem 5.6). Moreover, the bound $c_G^{(\mathrm{rep})}$ for an i.i.d. model $\mathcal{S}^{(n)} = \{\rho_\theta^{\otimes n}\}$ is identical to the standard Holevo bound $c_G^{(H)}$ for the base model ρ_θ (Theorem 5.3 and Section C of the Supplementary Material [4]). Thus, the asymptotic representation bound $c_G^{(\mathrm{rep})}$ can be regarded as a fully generalized version of the Holevo bound that is also applicable to non-i.i.d. models.

Incidentally, as one can see from the proof, Theorem 2.4 is valid, even if the scaling factors $\sqrt{n}$ in Definition 1.2 and Theorem 2.4 are both replaced with an arbitrary monotone increasing positive sequence $r_n \uparrow \infty$. Also, one can replace the domain $\mathbb{R}^d$ of the local parameter h to an arbitrary subset of $\mathbb{R}^d$. Classical analogues of these generalizations are found, for example, in [21], Definition 7.14, Theorem 9.4.

3. Preliminaries. In this section we devise some mathematical tools and prepare a number of lemmas toward the proof of Theorem 2.4. First, we give a condition for a quantum Gaussian state to be pure. We then introduce a new way of representing bounded operators on a CCR(S), which is analogous to the Husimi representation [12]. We further extend quantum Lévy–Cramér continuity theorem [13] and quantum Le Cam third lemma [3] so that they are directly applicable to the proof of Theorem 2.4. All the proofs of the lemmas and corollaries presented in this section are deferred to Section D of the Supplementary Material [4]. For the definition of von Neumann algebra CCR(S) with possibly degenerate S and quantum Gaussian states on it, see Section B of the Supplementary Material [4].

---

[2]The Holevo bound $c_G^{(H)}$ for a generic quantum statistical model $\{\rho_\theta : \theta \in \Theta \subset \mathbb{R}^d\}$ on a Hilbert space $\mathcal{H}$ is given by the minimum of $\operatorname{Tr} GZ(B) + \operatorname{Tr}|\sqrt{G}\operatorname{Im} Z(B)\sqrt{G}|$ over all Hermitian operators $B = (B_1, \dots, B_d)$ on $\mathcal{H}$ satisfying the local unbiasedness condition $\operatorname{Re}\operatorname{Tr}\rho_\theta L_i B_j = \delta_{ij}$, where L_i is the ith SLD and $Z(B)$ is the $d \times d$ matrix whose (i, j)th entry is $Z_{ij}(B) := \operatorname{Tr}\rho_\theta B_j B_i$. The reduced expression (7) for the quantum Gaussian shift model $\{N((\operatorname{Re}\tau)h, \Sigma) : h \in \mathbb{R}^d\}$ is derived in [23], Appendix B.

3.1. *Condition for a quantum Gaussian state to be pure.*

LEMMA 3.1 (Minimum uncertainty). *Let $J = V + \sqrt{-1}S$ be a $d \times d$ nonnegative matrix in which $S = \operatorname{Im} J$ is invertible. Then the quantum Gaussian state $N(0, J)$ on the von Neumann algebra* $\mathrm{CCR}(S)$ *is pure if and only if* $\det V = \det S$.

PROOF. See Section D of the Supplementary Material [4]. $\square$

COROLLARY 3.2. *Let $J = V + \sqrt{-1}S$ be a $d \times d$ nonnegative matrix in which both $V = \operatorname{Re} J$ and $S = \operatorname{Im} J$ are invertible. Then the quantum Gaussian state*

$$N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} J & J\#J^\top \\ J\#J^\top & J^\top \end{pmatrix}\right)$$

is pure.

PROOF. See Section D of the Supplementary Material [4]. $\square$

3.2. *Sandwiched coherent state representation of operators on a CCR algebra.* Let $\mathcal{H}$ be a Hilbert space that represents the von Neumann algebra $\mathrm{CCR}(S)$, where S is a skew-symmetric real $d \times d$ matrix that is not necessarily invertible, and let $\{X_i\}_{1 \le i \le d}$ be the canonical observables of $\mathrm{CCR}(S)$. Fix a cyclic[3] unit vector $\psi \in \mathcal{H}$ for $\mathrm{CCR}(S)$, and let

$$\psi(\xi) := e^{\sqrt{-1}\xi^i X_i}\psi, \quad (\xi \in \mathbb{R}^d).$$

Associated with a bounded operator $A \in B(\mathcal{H})$ is a continuous function $\varphi_A : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ defined by

$$\varphi_A(\xi; \eta) := \langle \psi(\xi), A\psi(\eta) \rangle, \quad (\xi, \eta \in \mathbb{R}^d).$$

We shall call φ_A the *sandwiched coherent state representation* of a bounded operator A.

We are interested in the converse problem: when does a function $\varphi : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{C}$ uniquely determine an operator $A \in B(\mathcal{H})$ satisfying $\varphi(\xi; \eta) = \langle \psi(\xi), A\psi(\eta) \rangle$? Let D be a dense subset of $\mathbb{R}^d$. A function $\varphi : D \times D \to \mathbb{C}$ is called *positive semidefinite* if, for all $r \in \mathbb{N}$ and $\{\xi^{(i)}\}_{1 \le i \le r} \subset D$, the $r \times r$ matrix whose (i, j)th entry is $\varphi(\xi^{(i)}; \xi^{(j)})$ is positive semidefinite, that is,

$$[\varphi(\xi^{(i)}; \xi^{(j)})]_{1 \le i, j \le r} \ge 0.$$

In this case we denote $\varphi \succ 0$. Further, for two functions φ_1 and φ_2, we denote $\varphi_1 \succ \varphi_2$ if $\varphi_1 - \varphi_2 \succ 0$.

LEMMA 3.3. *Suppose that $\varphi : D \times D \to \mathbb{C}$ satisfies $0 \prec \varphi \prec \varphi_I$. Then there exists a unique operator A satisfying $0 \le A \le I$ and $\varphi = \varphi_A$. Consequently, φ is continuously extended to the totality of $\mathbb{R}^d \times \mathbb{R}^d$.*

PROOF. See Section D of the Supplementary Material [4]. $\square$

Lemma 3.3 establishes a one-to-one correspondence between bounded operators A, satisfying $0 \le A \le I$, and functions φ satisfying $0 \prec \varphi \prec \varphi_I$. In what follows, the operator A that is recovered from the function φ is denoted by $V(\varphi)$.

[3]A vector $\psi \in \mathcal{H}$ is called *cyclic* for a linear subspace $\mathcal{A}$ of $B(\mathcal{H})$ if the linear space $\mathcal{A}\psi := \{A\psi : A \in \mathcal{A}\}$ is norm-dense in $\mathcal{H}$.

Now let $S = O_c \oplus S_q \oplus S_a$, where O_c is a $d_c \times d_c$ zero matrix, S_q a $d_q \times d_q$ skew-symmetric real invertible matrix, and S_a a $d_a \times d_a$ skew-symmetric real invertible matrix.[4] Then $\mathrm{CCR}(S) = \mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes \mathrm{CCR}(S_a)$, and the canonical observables are

$$\{\hat{X}_{c,i} := X_{c,i} \otimes I_q \otimes I_a\}_i \cup \{\hat{X}_{q,j} := I_c \otimes X_{q,j} \otimes I_a\}_j \cup \{\hat{X}_{a,k} := I_c \otimes I_q \otimes X_{a,k}\}_k,$$

where $\{X_{c,i}\}_i$, $\{X_{q,j}\}_j$, and $\{X_{a,k}\}_k$ are the canonical observables of $\mathrm{CCR}(O_c)$, $\mathrm{CCR}(S_q)$, and $\mathrm{CCR}(S_a)$, respectively. In the Schrödinger representation, the algebra $\mathrm{CCR}(S)$ is represented on the Hilbert space $\mathcal{H} := \mathcal{H}_c \otimes \mathcal{H}_q \otimes \mathcal{H}_a$, where $\mathcal{H}_c := L^2(\mathbb{R}^{d_c})$, $\mathcal{H}_q := L^2(\mathbb{R}^{d_q/2})$, and $\mathcal{H}_a := L^2(\mathbb{R}^{d_a/2})$, and

$$\begin{aligned}
\mathrm{CCR}(O_c) &= \overline{\mathrm{Span}}^{\mathrm{SOT}}\{e^{\sqrt{-1}\xi_c^i X_{c,i}}\}_{\xi_c \in \mathbb{R}^{d_c}} = L^\infty(\mathbb{R}^{d_c}),\\
\mathrm{CCR}(S_q) &= \overline{\mathrm{Span}}^{\mathrm{SOT}}\{e^{\sqrt{-1}\xi_q^j X_{q,j}}\}_{\xi_q \in \mathbb{R}^{d_q}} = B(\mathcal{H}_q),\\
\mathrm{CCR}(S_a) &= \overline{\mathrm{Span}}^{\mathrm{SOT}}\{e^{\sqrt{-1}\xi_a^k X_{a,k}}\}_{\xi_a \in \mathbb{R}^{d_a}} = B(\mathcal{H}_a),
\end{aligned}$$

where $\overline{\mathrm{Span}}^{\mathrm{SOT}}$ denotes the closure of the linear span with respect to the strong operator topology (SOT). Since $L^\infty(\mathbb{R}^{d_c})$ is a maximal abelian subalgebra[5] of $B(\mathcal{H}_c)$, the celebrated commutant theorem [10, 16] yields

$$\big(\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes I_a\big)' = \mathrm{CCR}(O_c) \otimes I_q \otimes \mathrm{CCR}(S_a). \tag{8}$$

In this identity I_q and I_a symbolically represent the centers of $\mathrm{CCR}(S_q)$ and $\mathrm{CCR}(S_a)$, respectively.

Let $\psi \in \mathcal{H}$ be a cyclic unit vector for $\mathrm{CCR}(S)$. Then the sandwiched coherent state representation of $A \in \mathrm{CCR}(S)$ is given by

$$\varphi_A(\xi_c, \xi_q, \xi_a; \eta_c, \eta_q, \eta_a) = \langle \psi(\xi_c, \xi_q, \xi_a), A\psi(\eta_c, \eta_q, \eta_a)\rangle,$$

where $\xi_c, \eta_c \in \mathbb{R}^{d_c}$, $\xi_q, \eta_q \in \mathbb{R}^{d_q}$, $\xi_a, \eta_a \in \mathbb{R}^{d_a}$, and

$$\psi(\xi_c, \xi_q, \xi_a) = e^{\sqrt{-1}(\xi_c^i \hat{X}_{c,i} + \xi_q^j \hat{X}_{q,j} + \xi_a^k \hat{X}_{a,k})}\psi.$$

Conversely, due to Lemma 3.3, a bounded continuous function $\varphi(\xi_c, \xi_q, \xi_a; \eta_c, \eta_q, \eta_a)$, satisfying $0 \prec \varphi \prec \varphi_I$, uniquely determines an operator $A = V(\varphi)$ satisfying $0 \le A \le I$. Moreover, the following lemma gives a criterion for $V(\varphi)$ to be an element of $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes I_a$, which means that $V(\varphi)$ can be regarded as an operator acting on $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$.

LEMMA 3.4. *Suppose that a bounded continuous function $\varphi(\xi_c, \xi_q, \xi_a; \eta_c, \eta_q, \eta_a)$ that fulfills the condition $0 \prec \varphi \prec \varphi_I$ satisfies the identity*

$$\varphi(\xi_c, \xi_q, \xi_a; \eta_c, \eta_q, \eta_a) = e^{-\sqrt{-1}\xi_a^\top S_a \eta_a}\varphi(\xi_c - \eta_c, \xi_q, \xi_a - \eta_a; 0, \eta_q, 0)$$

for all $\xi_c, \eta_c \in \mathbb{R}^{d_c}$, $\xi_q, \eta_q \in \mathbb{R}^{d_q}$, $\xi_a, \eta_a \in \mathbb{R}^{d_a}$. Then

$$V(\varphi) \in \mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes I_a.$$

PROOF. See Section D of the Supplementary Material [4]. □

[4]The subscripts c, q, and a stand for the classical, quantum, and ancillary systems, respectively.

[5]A von Neumann subalgebra M of $B(\mathcal{H})$ that satisfies $M' = M$ is called a *maximal abelian subalgebra* (MASA). The name comes from the fact that if N is an abelian von Neumann algebra such that $M \subset N \subset B(\mathcal{H})$, then $M = N$. In fact, since $M \subset N$, we have $(M \subset N \subset)\, N' \subset M' = M$ so that $M = N$.

3.3. *Sandwiched quantum Lévy–Cramér continuity theorem.* In this subsection we generalize the quantum Lévy–Cramér continuity theorem [13] and quantum Le Cam third lemma [3] in forms suitable for our discussion. Throughout this subsection we use the following notation. For each $n \in \mathbb{N}$, let $\rho^{(n)}$ be a quantum state and $X^{(n)} = (X_1^{(n)}, \dots, X_d^{(n)})$ be a list of observables on a finite dimensional Hilbert space $\mathcal{H}^{(n)}$. Further, let $X^{(\infty)} = (X_1^{(\infty)}, \dots, X_d^{(\infty)})$ be the canonical observables for a quantum Gaussian state $\rho^{(\infty)} \sim N(h, J)$ with $J_{ij} = \operatorname{Tr} \rho^{(\infty)} X_j^{(\infty)} X_i^{(\infty)}$.

The following lemma is a variant of the noncommutative Lévy–Cramér continuity theorem [13, 14].

LEMMA 3.5 (Sandwiched Lévy–Cramér continuity theorem). *Assume that*

$$\left(X^{(n)}, \rho^{(n)}\right) \rightsquigarrow N(h, J), \tag{9}$$

and that a uniformly bounded sequence $\{A^{(n)}\}_{n \in \mathbb{N} \cup \{\infty\}}$ of observables satisfies

$$\lim_{n \to \infty} \operatorname{Tr} \rho^{(n)} e^{\sqrt{-1}\xi^i X_i^{(n)}} A^{(n)} e^{\sqrt{-1}\eta^i X_i^{(n)}} = \operatorname{Tr} \rho^{(\infty)} e^{\sqrt{-1}\xi^i X_i^{(\infty)}} A^{(\infty)} e^{\sqrt{-1}\eta^i X_i^{(\infty)}} \tag{10}$$

for all $\xi, \eta \in \mathbb{Q}^d$. Then for any $\{\xi_s\}_{s=1}^{r_1}, \{\eta_t\}_{t=1}^{r_2} \subset \mathbb{R}^d$ and any real-valued bounded Borel functions $\{f_s\}_{s=1}^{r_1}, \{g_t\}_{t=1}^{r_2}$ whose discontinuity points form Lebesgue null sets, the following identity holds:

$$\begin{aligned} &\lim_{n \to \infty} \operatorname{Tr} \rho^{(n)} \left\{ \prod_{s=1}^{r_1} f_s\left(\xi_s^i X_i^{(n)}\right) \right\} A^{(n)} \left\{ \prod_{t=1}^{r_2} g_t\left(\eta_t^i X_i^{(n)}\right) \right\}^* \\ &\quad = \operatorname{Tr} \rho^{(\infty)} \left\{ \prod_{s=1}^{r_1} f_s\left(\xi_s^i X_i^{(\infty)}\right) \right\} A^{(\infty)} \left\{ \prod_{t=1}^{r_2} g_t\left(\eta_t^i X_i^{(\infty)}\right) \right\}^*. \end{aligned} \tag{11}$$

PROOF. See Section D of the Supplementary Material [4]. □

When $A^{(n)} = I^{(n)}$ for all $n \in \mathbb{N} \cup \{\infty\}$, Lemma 3.5 is subsumed by [3], Lemma 5.3. In this sense Lemma 3.5 is a slight generalization of [3], Lemma 5.3. However, the assumption of boundedness for functions f_s and g_t in Lemma 3.5 sometimes causes inconvenience in applications. We, therefore, further aim for generalizing Lemma 3.5 to unbounded functions. The key to the generalization is the notion of uniform integrability [3].

For quantum states $\{\rho^{(n)}\}_{n \in \mathbb{N}}$ and observables $\{B^{(n)}\}_{n \in \mathbb{N}}$ on Hilbert spaces $\{\mathcal{H}^{(n)}\}_{n \in \mathbb{N}}$, we say that $B^{(n)}$ is *uniformly integrable* under $\rho^{(n)}$ if for all $\varepsilon > 0$, there exists $L > 0$ that satisfies

$$\operatorname{Tr} \rho^{(n)} \left| B^{(n)} - h_L\left(B^{(n)}\right) \right| < \varepsilon$$

for all n, where the function h_L is defined by

$$h_L(x) = \begin{cases} x & (|x| \le L), \\ 0 & (|x| > L). \end{cases}$$

Using the notion of uniform integrability, Lemma 3.5 is generalized as follows.

LEMMA 3.6 (Sandwiched Lévy–Cramér continuity theorem: An extended version). *Under the same setting as in Lemma 3.5, except that the functions f_1 and g_1 can be unbounded, assume further that both $\{f_1(\xi_1^i X_i^{(n)} + o_1^{(n)})^2\}_{n \in \mathbb{N} \cup \{\infty\}}$ and $\{g_1(\eta_1^i X_i^{(n)} + o_2^{(n)})^2\}_{n \in \mathbb{N} \cup \{\infty\}}$*

are uniformly integrable under $\{\rho^{(n)}\}_{n\in\mathbb{N}\cup\{\infty\}}$, *where* $o_1^{(n)} = o_D(\xi_1^i X_i^{(n)}, \rho^{(n)})$, *and* $o_2^{(n)} = o_D(\eta_1^i X_i^{(n)}, \rho^{(n)})$ *for* $n \in \mathbb{N}$ *and* $o_1^{(\infty)} = o_2^{(\infty)} = 0$. *Then the following identity holds*:

$$\lim_{n\to\infty} \operatorname{Tr} \rho^{(n)} f_1(\xi_1^i X_i^{(n)} + o_1^{(n)}) \left\{\prod_{s=2}^{r_1} f_s(\xi_s^i X_i^{(n)})\right\} A^{(n)} \left\{\prod_{t=2}^{r_2} g_t(\eta_t^i X_i^{(n)})\right\}^* g_1(\eta_1^i X_i^{(n)} + o_2^{(n)})$$
$$= \operatorname{Tr} \rho^{(\infty)} \left\{\prod_{s=1}^{r_1} f_s(\xi_s^i X_i^{(\infty)})\right\} A^{(\infty)} \left\{\prod_{t=1}^{r_2} g_t(\eta_t^i X_i^{(\infty)})\right\}^*.$$

PROOF. See Section D of the Supplementary Material [4]. □

By using Lemma 3.6, we can further generalize quantum Le Cam third lemma under q-LAN [3], Corollary 7.5, as follows.

COROLLARY 3.7 (Sandwiched Le Cam third lemma under D-extendibility). *Assume that a sequence* $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ *of quantum statistical models is q-LAN and D-extendible at* $\theta_0 \in \Theta$, *as in Definition* 2.2. *Then*

$$(X^{(n)}, \rho^{(n)}_{\theta_0+h/\sqrt{n}}) \overset{h}{\rightsquigarrow} N\big((\operatorname{Re}\tau)h, \Sigma\big), \tag{12}$$

where $\tau = \Sigma F$.

Assume further that a uniformly bounded sequence $A^{(n)}$ *of observables for* $n \in \mathbb{N} \cup \{\infty\}$ *satisfies*

$$\lim_{n\to\infty} \operatorname{Tr} \rho_{\theta_0}^{(n)} e^{\sqrt{-1}\xi^i X_i^{(n)}} A^{(n)} e^{\sqrt{-1}\eta^i X_i^{(n)}} = \operatorname{Tr} \rho_0^{(\infty)} e^{\sqrt{-1}\xi^i X_i^{(\infty)}} A^{(\infty)} e^{\sqrt{-1}\eta^i X_i^{(\infty)}}$$

for all $\xi, \eta \in \mathbb{Q}^r$, *where* $\rho_0^{(\infty)} \sim N(0, \Sigma)$ *and* $X^{(\infty)} = (X_1^{(\infty)}, \dots, X_r^{(\infty)})$ *are the canonical observables. Then it holds that*

$$\lim_{n\to\infty} \operatorname{Tr} \rho_{\theta_0}^{(n)} R_{h_1}^{(n)} A^{(n)} R_{h_2}^{(n)} = \operatorname{Tr} \rho_0^{(\infty)} R_{h_1}^{(\infty)} A^{(\infty)} R_{h_2}^{(\infty)} \tag{13}$$

for any $h_1, h_2 \in \mathbb{R}^d$, *where* $R_h^{(n)}$ *are square-root likelihood ratios defined by*

$$R_h^{(n)} = \mathcal{R}(\rho^{(n)}_{\theta_0+h/\sqrt{n}} \mid \rho^{(n)}_{\theta_0}) \quad \text{and} \quad R_h^{(\infty)} = \exp\left[\frac{1}{2}\left((Fh)^i X_i^{(\infty)} - \frac{1}{2}(h^\top F^\top \Sigma F h) I^{(\infty)}\right)\right].$$

Specifically,

$$\lim_{n\to\infty} \operatorname{Tr} \rho^{(n)}_{\theta_0+h/\sqrt{n}} A^{(n)} = \lim_{n\to\infty} \operatorname{Tr} \rho_{\theta_0}^{(n)} R_h^{(n)} A^{(n)} R_h^{(n)} = \operatorname{Tr} \rho_h^{(\infty)} A^{(\infty)} \tag{14}$$

for any $h \in \mathbb{R}^d$, *where* $\rho_h^{(\infty)} \sim N((\operatorname{Re}\tau)h, \Sigma)$.

PROOF. See Section D of the Supplementary Material [4]. □

Finally, the following asymptotic version of the Weyl CCR will turn out to be useful.

LEMMA 3.8 (Asymptotic Weyl CCR). *Let* $W^{(n)}(\xi) := e^{\sqrt{-1}\xi^i X_i^{(n)}}$ *for* $\xi \in \mathbb{R}^d$, *and assume that* $(X^{(n)}, \rho^{(n)}) \rightsquigarrow N(0, J)$. *Then*

$$\lim_{n\to\infty} \big\| W^{(n)}(\xi) W^{(n)}(\eta) \sqrt{\rho^{(n)}} - e^{\sqrt{-1}\xi^\top S\eta} W^{(n)}(\xi+\eta) \sqrt{\rho^{(n)}} \big\|_{\mathrm{HS}} = 0$$

for all $\xi, \eta \in \mathbb{R}^d$, *where* $S := \operatorname{Im} J$, *and* $\|\cdot\|_{\mathrm{HS}}$ *denotes the Hilbert–Schmidt norm.*

PROOF. See Section D of the Supplementary Material [4]. □

4. Proof of Theorem 2.4. Since the proof is somewhat lengthy, we first outline the proof. By choosing a suitable regular $r \times r$ matrix K, one finds another D-extension $X'^{(n)}_i := K_i^j X_j^{(n)}$ of $\Delta^{(n)}$ that exhibits

$$X'^{(n)} \overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N(0, \Sigma_c \oplus \Sigma_q),$$

where Σ_c is a real $r_c \times r_c$ matrix and Σ_q is a complex $r_q \times r_q$ matrix with $r_c + r_q = r$ so that the imaginary part $S_q := \operatorname{Im} \Sigma_q$ of Σ_q is invertible.[6] In what follows, we always adopt such a D-extension and simply denote $X'^{(n)}$ as $X^{(n)}$, omitting the prime. Further, we label the elements of $X^{(n)}$ as

$$X^{(n)}_{c,1}, \dots, X^{(n)}_{c,r_c}, X^{(n)}_{q,1}, \dots, X^{(n)}_{q,r_q}$$

in accordance with the decomposition $\Sigma = \Sigma_c \oplus \Sigma_q$.

We need to show that there exists a POVM M on $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$ that exhibits $\phi_h(M(B)) = \mathcal{L}_h(B)$ for any $B \in \mathcal{B}(R^s)$, where $\phi_h \sim N((\operatorname{Re}\tau)h, \Sigma)$. To this end, we first formally enlarge $N((\operatorname{Re}\tau)h, \Sigma)$ to $N((\operatorname{Re}\hat{\tau})h, \hat{\Sigma})$, where

$$\hat{\tau} := \begin{pmatrix} \tau \\ 0 \end{pmatrix}, \qquad \hat{\Sigma} := \begin{pmatrix} \Sigma_c & 0 & 0 \\ 0 & \Sigma_q & \Sigma_q \# \Sigma_q^\top \\ 0 & \Sigma_q \# \Sigma_q^\top & \Sigma_q^\top \end{pmatrix}.$$

We next construct a POVM on $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes \mathrm{CCR}(-S_q)$ and prove that it is a POVM on $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes I_a$. Finally, we prove that the POVM thus constructed enjoys the desired property.

PROOF. As stated above, we divide the proof into three steps. In Step 1 we define a Hilbert space on which $N((\operatorname{Re}\hat{\tau})h, \hat{\Sigma})$ is represented, and designate a fiducial cyclic vector for $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes \mathrm{CCR}(-S_q)$. In Step 2 we construct (a precursor of) a POVM on $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes \mathrm{CCR}(-S_q)$ and prove that it defines a POVM M on $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$. In Step 3 we prove that M enjoys the desired property.

Step 1. As a similar way to the prescription that precedes Lemma 3.4, we introduce a Hilbert space $\mathcal{H} = \mathcal{H}_c \otimes \mathcal{H}_q \otimes \mathcal{H}_a$ on which the von Neumann algebra $\mathrm{CCR}(\hat{S}) = \mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q) \otimes \mathrm{CCR}(S_a)$ is represented, where

$$\hat{S} = \operatorname{Im} \hat{\Sigma}, \qquad O_c = \operatorname{Im} \Sigma_c \,(= 0), \qquad S_q = \operatorname{Im} \Sigma_q, \quad \text{and} \quad S_a = \operatorname{Im} \Sigma_q^\top = -S_q.$$

The canonical observables are

$$\{\hat{X}_{c,i} := X_{c,i} \otimes I_q \otimes I_a\}_i \cup \{\hat{X}_{q,j} := I_c \otimes X_{q,j} \otimes I_a\}_j \cup \{\hat{X}_{a,k} := I_c \otimes I_q \otimes X_{a,k}\}_k,$$

where $\{X_{c,i}\}_i$, $\{X_{q,j}\}_j$, and $\{X_{a,k}\}_k$ are the canonical observables of $\mathrm{CCR}(O_c)$, $\mathrm{CCR}(S_q)$, and $\mathrm{CCR}(S_a)$, respectively.

In order to invoke the sandwiched coherent state representation for $\mathrm{CCR}(\hat{S})$, we need a cyclic vector ψ on $\mathcal{H}$. We first designate a cyclic vector $\psi_c \in \mathcal{H}_c$ for $\mathrm{CCR}(O_c) = L^\infty(\mathbb{R}^{r_c})$ in which each $\xi \in L^\infty(\mathbb{R}^{r_c})$ is identified with the bounded operator $T_\xi \in B(\mathcal{H}_c)$ defined by

$$(T_\xi \varphi)(x) := \xi(x)\varphi(x), \quad (\varphi \in \mathcal{H}_c, x \in \mathbb{R}^{r_c}).$$

[6]In fact, for any $\Sigma \geq 0$ with $\operatorname{Re}\Sigma > 0$, we have $(\operatorname{Re}\Sigma)^{-1}\Sigma(\sqrt{\operatorname{Re}\Sigma})^{-1} = I + \sqrt{-1}S$, where $S := (\sqrt{\operatorname{Re}\Sigma})^{-1}(\operatorname{Im}\Sigma)(\sqrt{\operatorname{Re}\Sigma})^{-1}$ is a real skew-symmetric matrix. Further, by choosing a suitable real orthogonal matrix P, the matrix S is transformed into the form $P^\top S P = 0 \oplus S_q$ with $\det S_q \neq 0$.

Let $\psi_c := \sqrt{g(x)}$, where $g(x)$ is the density function of the (classical) Gaussian distribution $N(0, \Sigma_c)$. Then any function $f \in L^2(\mathbb{R}^{r_c})$ can be approximated by a series of functions $f_n := T_{\xi_n}\psi_c$, where

$$\xi_n := \frac{f(x)}{\sqrt{g(x)}}\mathbb{1}_{B_n}(x),$$

with $\mathbb{1}_{B_n}$ being the indicator function of the ball B_n of radius $n \in \mathbb{N}$ centered at the origin of $\mathbb{R}^{r_c}$. As a consequence, ψ_c is a cyclic vector of $\mathcal{H}_c$.

We next specify a cyclic vector $\psi_{qa} \in \mathcal{H}_q \otimes \mathcal{H}_a$ for $\mathrm{CCR}(S_q) \otimes \mathrm{CCR}(S_a)$. Recall that

$$N\left(0, \begin{pmatrix} \Sigma_q & \Sigma_q \# \Sigma_q^\top \\ \Sigma_q \# \Sigma_q^\top & \Sigma_q^\top \end{pmatrix}\right)$$

is a pure state on $\mathcal{H}_q \otimes \mathcal{H}_a$ (Corollary 3.2). Let $\psi_{qa} \in \mathcal{H}_q \otimes \mathcal{H}_a$ be a unit vector that corresponds to the above state. Then it is well known in the theory of coherent states that ψ_{qa} is a cyclic vector for $\mathrm{CCR}(S_q) \otimes \mathrm{CCR}(S_a)$.

Now we arrive at a cyclic vector $\psi := \psi_c \otimes \psi_{qa} \in \mathcal{H}_c \otimes \mathcal{H}_q \otimes \mathcal{H}_a$ for $\mathrm{CCR}(\hat{S})$. This cyclic vector has the following nice property. Let

$$\hat{\Delta}_i := F_i^j \hat{X}_j$$

for $1 \le i \le d$, where F is the $r \times d$ real matrix introduced in Definition 2.2, and let

$$\hat{R}_h := \exp\left[\frac{1}{2}\left(h^i \hat{\Delta}_i - \frac{1}{2}(h^\top F^\top \Sigma F h)\hat{I}\right)\right]$$

for $h \in \mathbb{R}^d$, where $\hat{I}$ is the identity on $\mathrm{CCR}(\hat{S})$. Then, for any $A \in \mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$, the following identity holds:

$$\langle \psi, \hat{R}_h (A \otimes I_a) \hat{R}_h \psi \rangle = \phi_h(A), \tag{15}$$

where $\phi_h \sim N((\mathrm{Re}\,\tau)h, \Sigma)$ with $\tau := \Sigma F$. This relation will be used as a variant of the quantum Le Cam third lemma that goes back and forth between $\mathrm{CCR}(\hat{S})$ and $\mathrm{CCR}(S)$.

To prove (15), let

$$\overline{X}_i := \begin{cases} X_{c,i} \otimes I_q & \text{if } 1 \le i \le r_c, \\ I_c \otimes X_{q,i-r_c} & \text{if } r_c + 1 \le i \le r_c + r_q \end{cases} \tag{16}$$

be canonical observables of $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$. Then by a direct computation using the quasi-characteristic function of the vector state[7]

$$|\psi\rangle\langle\psi| \sim N(0, \hat{\Sigma}),$$

we can verify that

$$\langle \psi, \hat{R}_h (e^{\sqrt{-1}\xi^i \overline{X_i}} \otimes I_a) \hat{R}_h \psi \rangle = \langle \psi, \hat{R}_h\, e^{\sqrt{-1}\sum_{i=1}^r \xi^i \hat{X}_i}\, \hat{R}_h \psi \rangle = e^{\sqrt{-1}\xi^\top (\mathrm{Re}\,\tau) h - \frac{1}{2}\xi^\top \Sigma \xi}.$$

Since the last side is the characteristic function of $\phi_h \sim N((\mathrm{Re}\,\tau)h, \Sigma)$, we have

$$\langle \psi, \hat{R}_h (e^{\sqrt{-1}\xi^i \overline{X_i}} \otimes I_a) \hat{R}_h \psi \rangle = \phi_h(e^{\sqrt{-1}\xi^i \overline{X_i}}).$$

Finally, since $\{e^{\sqrt{-1}\xi^i \overline{X_i}}\}_{\xi \in \mathbb{R}^r}$ is SOT-dense in $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$, the identity (15) is proved.

[7] If $\det \hat{S} = 0$, the Hilbert space $\mathcal{H}_c$ is reducible under the action of $\mathrm{CCR}(O_c)$, and thus the vector state $|\psi\rangle\langle\psi|$ is a mixed state.

Step 2. Given a pair of vectors $(\lambda, \mu) \in \mathbb{R}^{r_c} \times \mathbb{R}^{r_q}$, let

$$W(\lambda, \mu) := e^{\sqrt{-1}(\lambda^i X_{c,i} + \mu^j X_{q,j})}$$

be the corresponding Weyl operator on $\mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$. By analogy to this operator, we introduce a unitary operator

$$W^{(n)}(\lambda, \mu) := e^{\sqrt{-1}(\lambda^i X_{c,i}^{(n)} + \mu^j X_{q,j}^{(n)})}$$

on each $\mathcal{H}^{(n)}$. We further define, for each $\xi = (\xi_c, \xi_q, \xi_a) \in \mathbb{R}^{r_c} \times \mathbb{R}^{r_q} \times \mathbb{R}^{r_q}$, operators

$$A^{(n)}(\xi) := W^{(n)}(\xi_c, \xi_q)\sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(0, \xi_a)$$

and

$$B^{(n)}(\xi) := W^{(n)}(0, \xi_q)\sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(\xi_c, \xi_a).$$

Note that these operators are asymptotically identified in that

$$\lim_{n\to\infty} \| A^{(n)}(\xi) - B^{(n)}(\xi)\|_{\mathrm{HS}} = 0. \tag{17}$$

This is proved by observing

$$\| A^{(n)}(\xi) - B^{(n)}(\xi)\|_{\mathrm{HS}}^2 = 2 - 2\,\mathrm{Re}\,\mathrm{Tr}\, A^{(n)}(\xi)^* B^{(n)}(\xi)$$

and

$$\begin{aligned}
&\lim_{n\to\infty} \mathrm{Tr}\, A^{(n)}(\xi)^* B^{(n)}(\xi)\\
&\quad = \lim_{n\to\infty} \mathrm{Tr}\, W^{(n)}(0, \xi_a)^* \sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(\xi_c, \xi_q)^* W^{(n)}(0, \xi_q)\sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(\xi_c, \xi_a)\\
&\quad = \lim_{n\to\infty} \mathrm{Tr}\{W^{(n)}(\xi_c, \xi_a) W^{(n)}(0, -\xi_a)\sqrt{\rho_{\theta_0}^{(n)}}\}\{W^{(n)}(-\xi_c, -\xi_q) W^{(n)}(0, \xi_q)\sqrt{\rho_{\theta_0}^{(n)}}\}\\
&\quad = \lim_{n\to\infty} \mathrm{Tr}\left\{\exp\left[\sqrt{-1}\begin{pmatrix}\xi_c\\ \xi_a\end{pmatrix}^{\top}(O_c \oplus S_q)\begin{pmatrix}0\\ -\xi_a\end{pmatrix}\right] W^{(n)}(\xi_c, 0)\sqrt{\rho_{\theta_0}^{(n)}}\right\}\\
&\qquad\times\left\{\exp\left[\sqrt{-1}\begin{pmatrix}-\xi_c\\ -\xi_q\end{pmatrix}^{\top}(O_c \oplus S_q)\begin{pmatrix}0\\ \xi_q\end{pmatrix}\right] W^{(n)}(-\xi_c, 0)\sqrt{\rho_{\theta_0}^{(n)}}\right\}\\
&\quad = \lim_{n\to\infty} \mathrm{Tr}\, W^{(n)}(\xi_c, 0)\sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(-\xi_c, 0)\sqrt{\rho_{\theta_0}^{(n)}}\\
&\quad = \exp\left\{-\frac{1}{2}\begin{pmatrix}\xi_c\\ 0\\ -\xi_c\\ 0\end{pmatrix}^{\top}\begin{pmatrix}\Sigma_c & 0 & \Sigma_c & 0\\ 0 & \Sigma_q & 0 & \Sigma_q\#\Sigma_q^{\top}\\ \Sigma_c & 0 & \Sigma_c & 0\\ 0 & \Sigma_q\#\Sigma_q^{\top} & 0 & \Sigma_q^{\top}\end{pmatrix}\begin{pmatrix}\xi_c\\ 0\\ -\xi_c\\ 0\end{pmatrix}\right\}\\
&\quad = 1.
\end{aligned}$$

Here the asymptotic Weyl CCR (Lemma 3.8) was used in the third equality, and condition (5) for D-extendibility was used in the second last equality.

Now, given a POVM $M^{(n)}$ on $\mathcal{H}^{(n)}$ whose outcomes take values in $\mathbb{R}^s$, let

$$M_t^{(n)} := M^{(n)}\big((-\infty, t]\big)$$

be the associated resolution of identity, where $t = (t_1, t_2, \dots, t_s) \in \mathbb{R}^s$ and $(-\infty, t]$ is the shorthand of the set

$$(-\infty, t_1] \times (-\infty, t_2] \times \cdots \times (-\infty, t_s].$$

By using the resolution of identity $M_t^{(n)}$, we define the following function:

$$\varphi_t^{(n)}(\xi;\eta) := \operatorname{Tr} A^{(n)}(\xi)^* M_t^{(n)} A^{(n)}(\eta)$$

for $\xi = (\xi_c, \xi_q, \xi_a), \eta = (\eta_c, \eta_q, \eta_a) \in \mathbb{R}^{r_c} \times \mathbb{R}^{r_q} \times \mathbb{R}^{r_q}$. Since $\varphi_t^{(n)}(\xi;\eta)$ is uniformly bounded in that $|\varphi_t^{(n)}(\xi;\eta)| \le 1$ for all $t \in \mathbb{R}^s$, $\xi, \eta \in \mathbb{R}^{r_c+2r_q}$, and $n \in \mathbb{N}$, the diagonal sequence trick [20] tells us that there is a subsequence $\{n_m\}_{m\in\mathbb{N}} \subset \{n\}_{n\in\mathbb{N}}$ such that $\varphi_\alpha^{(n_m)}(\xi;\eta)$ are convergent for all countably many arguments $\alpha \in \mathbb{Q}^s$ and $\xi, \eta \in \mathbb{Q}^{r_c+2r_q}$, defining a limiting function

$$\varphi_\alpha(\xi;\eta) := \lim_{m\to\infty} \varphi_\alpha^{(n_m)}(\xi;\eta). \tag{18}$$

We shall prove that this limiting function φ_α is the sandwiched coherent state representation of some operator $\tilde{M}_\alpha$ on $\mathcal{H}_c \otimes \mathcal{H}_q$.

First, we formally introduce the function φ_∞ by

$$\begin{aligned}\varphi_\infty(\xi;\eta) &:= \langle e^{\sqrt{-1}\xi^i \hat{X}_i}\psi, e^{\sqrt{-1}\eta^j \hat{X}_j}\psi\rangle \\ &= e^{-\sqrt{-1}\xi^\top \hat{S}\eta}\langle \psi, e^{\sqrt{-1}(\eta-\xi)^i \hat{X}_i}\psi\rangle \\ &= e^{-\sqrt{-1}\xi^\top \hat{S}\eta}\, e^{-\frac{1}{2}(\eta-\xi)^\top \hat{\Sigma}(\eta-\xi)}.\end{aligned}$$

Then it is shown that

$$\varphi_\infty(\xi;\eta) = \lim_{n\to\infty} \varphi_\infty^{(n)}(\xi;\eta),$$

where $\varphi_\infty^{(n)}(\xi;\eta) := \lim_{t\to\infty} \varphi_t^{(n)}(\xi;\eta) = \operatorname{Tr} A^{(n)}(\xi)^* A^{(n)}(\eta)$. In fact,

$$\begin{aligned}
&\lim_{n\to\infty} \varphi_\infty^{(n)}(\xi;\eta) \\
&\quad = \lim_{n\to\infty} \operatorname{Tr} A^{(n)}(\xi)^* A^{(n)}(\eta) \\
&\quad = \lim_{n\to\infty} \operatorname{Tr} W^{(n)}(0,\xi_a)^* \sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(\xi_c,\xi_q)^* W^{(n)}(\eta_c,\eta_q) \sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(0,\eta_a) \\
&\quad = \lim_{n\to\infty} \operatorname{Tr}\{W^{(n)}(0,\eta_a) W^{(n)}(0,-\xi_a)\sqrt{\rho_{\theta_0}^{(n)}}\}\{W^{(n)}(-\xi_c,-\xi_q) W^{(n)}(\eta_c,\eta_q)\sqrt{\rho_{\theta_0}^{(n)}}\} \\
&\quad = \lim_{n\to\infty} \operatorname{Tr}\left\{\exp\left[\sqrt{-1}\begin{pmatrix}0\\ \eta_a\end{pmatrix}^\top (O_c \oplus S_q)\begin{pmatrix}0\\ -\xi_a\end{pmatrix}\right] W^{(n)}(0,\eta_a-\xi_a)\sqrt{\rho_{\theta_0}^{(n)}}\right\} \\
&\qquad \times \left\{\exp\left[\sqrt{-1}\begin{pmatrix}-\xi_c\\ -\xi_q\end{pmatrix}^\top (O_c \oplus S_q)\begin{pmatrix}\eta_c\\ \eta_q\end{pmatrix}\right] W^{(n)}(\eta_c-\xi_c,\eta_q-\xi_q)\sqrt{\rho_{\theta_0}^{(n)}}\right\} \\
&\quad = e^{\sqrt{-1}(-\eta_a^\top S_q \xi_a - \xi_q^\top S_q \eta_q)} \lim_{n\to\infty} \operatorname{Tr} W^{(n)}(0,\eta_a-\xi_a)\sqrt{\rho_{\theta_0}^{(n)}}\, W^{(n)}(\eta_c-\xi_c,\eta_q-\xi_q)\sqrt{\rho_{\theta_0}^{(n)}} \\
&\quad = e^{-\sqrt{-1}\xi^\top \hat{S}\eta} \exp\left\{-\frac{1}{2}\begin{pmatrix}0\\ \eta_a-\xi_a\\ \eta_c-\xi_c\\ \eta_q-\xi_q\end{pmatrix}^\top \begin{pmatrix}\Sigma_c & 0 & \Sigma_c & 0\\ 0 & \Sigma_q & 0 & \Sigma_q \# \Sigma_q^\top\\ \Sigma_c & 0 & \Sigma_c & 0\\ 0 & \Sigma_q \# \Sigma_q^\top & 0 & \Sigma_q^\top\end{pmatrix}\begin{pmatrix}0\\ \eta_a-\xi_a\\ \eta_c-\xi_c\\ \eta_q-\xi_q\end{pmatrix}\right\}
\end{aligned}$$

$$= e^{-\sqrt{-1}\xi^\top \hat{S}\eta} e^{-\frac{1}{2}(\eta-\xi)^\top \hat{\Sigma}(\eta-\xi)}$$
$$= \varphi_\infty(\xi;\eta).$$

As a consequence, by taking the limit $m \to \infty$ in $0 \prec \varphi_\alpha^{(n_m)} \prec \varphi_\infty^{(n_m)}$, which follows from $0 \le M_\alpha^{(n_m)} \le I^{(n_m)}$, we have

$$0 \prec \varphi_\alpha \prec \varphi_\infty, \quad (\forall \alpha \in \mathbb{Q}^d). \tag{19}$$

We can also prove the following identity:

$$\varphi_\alpha(\xi_c, \xi_q, \xi_a; \eta_c, \eta_q, \eta_a) = e^{-\sqrt{-1}\xi_a^\top S_a \eta_a} \varphi_\alpha(\xi_c - \eta_c, \xi_q, \xi_a - \eta_a; 0, \eta_q, 0). \tag{20}$$

In fact, by using the asymptotic identifiability of $A^{(n)}(\xi)$ and $B^{(n)}(\xi)$, established in (17),

$$\begin{aligned}
&\varphi_\alpha(\xi;\eta)\\
&= \lim_{m\to\infty} \operatorname{Tr} A^{(n_m)}(\xi)^* M_\alpha^{(n_m)} A^{(n_m)}(\eta)\\
&= \lim_{m\to\infty} \operatorname{Tr} B^{(n_m)}(\xi)^* M_\alpha^{(n_m)} B^{(n_m)}(\eta)\\
&= \lim_{m\to\infty} \operatorname{Tr} W^{(n_m)}(\xi_c, \xi_a)^* \sqrt{\rho_{\theta_0}^{(n_m)}}\, W^{(n_m)}(0, \xi_q)^* M_\alpha^{(n_m)} W^{(n_m)}(0, \eta_q) \sqrt{\rho_{\theta_0}^{(n_m)}}\, W^{(n_m)}(\eta_c, \eta_a)\\
&= \lim_{m\to\infty} \operatorname{Tr}\{W^{(n_m)}(\eta_c, \eta_a) W^{(n_m)}(-\xi_c, -\xi_a) \sqrt{\rho_{\theta_0}^{(n_m)}}\}\\
&\quad \times \{W^{(n_m)}(0, -\xi_q) M_\alpha^{(n_m)} W^{(n_m)}(0, \eta_q) \sqrt{\rho_{\theta_0}^{(n_m)}}\}\\
&= \lim_{m\to\infty} \operatorname{Tr}\left\{\exp\left[\sqrt{-1}\begin{pmatrix}\eta_c\\ \eta_a\end{pmatrix}^\top (O_c \oplus S_q)\begin{pmatrix}-\xi_c\\ -\xi_a\end{pmatrix}\right] W^{(n_m)}(\eta_c - \xi_c, \eta_a - \xi_a) \sqrt{\rho_{\theta_0}^{(n_m)}}\right\}\\
&\quad \times \{W^{(n_m)}(0, -\xi_q) M_\alpha^{(n_m)} W^{(n_m)}(0, \eta_q) \sqrt{\rho_{\theta_0}^{(n_m)}}\}\\
&= e^{-\sqrt{-1}\eta_a^\top S_q \xi_a} \lim_{m\to\infty} \operatorname{Tr} W^{(n_m)}(\xi_c - \eta_c, \xi_a - \eta_a)^*\\
&\quad \times \sqrt{\rho_{\theta_0}^{(n_m)}}\, W^{(n_m)}(0, \xi_q)^* M_\alpha^{(n_m)} W^{(n_m)}(0, \eta_q) \sqrt{\rho_{\theta_0}^{(n_m)}}\\
&= e^{-\sqrt{-1}\eta_a^\top S_q \xi_a} \lim_{m\to\infty} \operatorname{Tr} B^{(n_m)}(\xi_c - \eta_c, \xi_q, \xi_a - \eta_a)^* M_\alpha^{(n_m)} B^{(n_m)}(0, \eta_q, 0)\\
&= e^{-\sqrt{-1}\eta_a^\top S_q \xi_a} \lim_{m\to\infty} \operatorname{Tr} A^{(n_m)}(\xi_c - \eta_c, \xi_q, \xi_a - \eta_a)^* M_\alpha^{(n_m)} A^{(n_m)}(0, \eta_q, 0)\\
&= e^{-\sqrt{-1}\xi_a^\top S_a \eta_a} \varphi_\alpha(\xi_c - \eta_c, \xi_q, \xi_a - \eta_a; 0, \eta_q, 0).
\end{aligned}$$

Now that (19) and (20) are verified, Lemmas 3.3 and 3.4 prove that there is a unique operator $\tilde{M}_\alpha \in \mathrm{CCR}(O_c) \otimes \mathrm{CCR}(S_q)$ satisfying $0 \le \tilde{M}_\alpha \le I_c \otimes I_q$ and

$$\varphi_\alpha(\xi;\eta) = \langle e^{\sqrt{-1}\xi^i \hat{X}_i}\psi, (\tilde{M}_\alpha \otimes I_a) e^{\sqrt{-1}\eta^j \hat{X}_j}\psi\rangle \tag{21}$$

for all $\alpha \in \mathbb{Q}^s$ and $\xi, \eta \in \mathbb{R}^{r_c + 2r_q}$.

We are now ready to construct a POVM $M = \{M(B) : B \in \mathcal{B}(\mathbb{R}^s)\}$ from $\{\tilde{M}_\alpha\}_{\alpha \in \mathbb{Q}^s}$. Since $\tilde{M}_\alpha$ is monotone in $\alpha \in \mathbb{Q}^s$, we can define, for each $t \in \mathbb{R}^s$,

$$\bar{M}_t := \inf_{\alpha > t, \alpha \in \mathbb{Q}^s} \tilde{M}_\alpha,$$

where the infimum is taken in the weak operator topology (WOT). Since $t \mapsto \bar{M}_t$ is right-continuous, it uniquely determines a POVM $\bar{M} = \{\bar{M}(B) : B \in \mathcal{B}(\bar{\mathbb{R}}^s)\}$ over the extended reals $\bar{\mathbb{R}}^s$. Finally, we transfer the "measure at infinity" $\bar{M}(\bar{\mathbb{R}}^s \setminus \mathbb{R}^s)$ to the origin so as to obtain

$$M(B) := \bar{M}(B) + \delta_0(B)\bar{M}(\bar{\mathbb{R}}^s \setminus \mathbb{R}^s), \quad (B \in \mathcal{B}(\mathbb{R}^s)),$$

where δ_0 is the Dirac measure concentrated at the origin.

Step 3. We prove that the POVM M, constructed in Step 2, is the desired one we have sought. Setting $\xi_a = \eta_a = 0$ in (18) and (21), we have

$$\begin{aligned}\varphi_\alpha(\xi_c, \xi_q, 0; \eta_c, \eta_q, 0) &= \lim_{m\to\infty} \operatorname{Tr} \sqrt{\rho_{\theta_0}^{(n_m)}} e^{-\sqrt{-1}\xi^i X_i^{(n_m)}} M_\alpha^{(n_m)} e^{\sqrt{-1}\eta^i X_i^{(n_m)}} \sqrt{\rho_{\theta_0}^{(n_m)}} \\ &= \langle \psi, e^{-\sqrt{-1}\xi^i (\overline{X}_i \otimes I_a)} (\tilde{M}_\alpha \otimes I_a) e^{\sqrt{-1}\eta^i (\overline{X}_i \otimes I_a)} \psi \rangle,\end{aligned}$$

or equivalently,

$$\lim_{m\to\infty} \operatorname{Tr} \rho_{\theta_0}^{(n_m)} e^{-\sqrt{-1}\xi^i X_i^{(n_m)}} M_\alpha^{(n_m)} e^{\sqrt{-1}\eta^i X_i^{(n_m)}} = \langle \psi, (e^{-\sqrt{-1}\xi^i \overline{X}_i} \tilde{M}_\alpha e^{\sqrt{-1}\eta^i \overline{X}_i} \otimes I_a)\psi \rangle.$$

Due to (15), this is further equal to

$$\phi_0(e^{-\sqrt{-1}\xi^i \overline{X}_i} \tilde{M}_\alpha e^{\sqrt{-1}\eta^i \overline{X}_i}).$$

Therefore, the sandwiched Le Cam third lemma (Corollary 3.7) yields

$$\lim_{m\to\infty} \mathcal{L}_h^{(n_m)}(-\infty, \alpha] = \lim_{m\to\infty} \operatorname{Tr} \rho_{\theta_0 + h/\sqrt{n_m}}^{(n_m)} M_\alpha^{(n_m)} = \phi_h(\tilde{M}_\alpha) \tag{22}$$

for all $h \in \mathbb{R}^d$ and $\alpha \in \mathbb{Q}^s$.

Fix $h \in \mathbb{R}^d$ arbitrarily. Due to assumption, $\mathcal{L}_h^{(n)}$ weakly converges to $\mathcal{L}_h$. Therefore, for any continuity point $t \in \mathbb{R}^s$ of $t \mapsto \mathcal{L}_h(-\infty, t]$,

$$\begin{aligned}\mathcal{L}_h(-\infty, t] &= \lim_{m\to\infty} \mathcal{L}_h^{(n_m)}(-\infty, t] \\ &\le \inf_{\alpha > t, \alpha \in \mathbb{Q}^s} \lim_{m\to\infty} \mathcal{L}_h^{(n_m)}(-\infty, \alpha] \\ &\le \inf_{\alpha > t, \alpha \in \mathbb{Q}^s} \mathcal{L}_h(-\infty, \alpha] \\ &= \mathcal{L}_h(-\infty, t].\end{aligned}$$

In the second inequality, we used the portmanteau lemma. It then follows from (22) that

$$\mathcal{L}_h(-\infty, t] = \inf_{\alpha > t, \alpha \in \mathbb{Q}^s} \lim_{m\to\infty} \mathcal{L}_h^{(n_m)}(-\infty, \alpha] = \inf_{\alpha > t,\, \alpha \in \mathbb{Q}^s} \phi_h(\tilde{M}_\alpha) = \phi_h(\bar{M}_t),$$

and thus $\mathcal{L}_h(B) = \phi_h(\bar{M}(B))$ for all $B \in \mathcal{B}(\mathbb{R}^s)$: in particular,

$$\phi_h(\bar{M}(\mathbb{R}^s)) = \mathcal{L}_h(\mathbb{R}^s) = 1.$$

Since $\mathcal{L}_h(\bar{\mathbb{R}}^s \setminus \mathbb{R}^s) = 0$, we have $\phi_h(\bar{M}(B)) = \phi_h(M(B))$ for all $B \in \mathcal{B}(\mathbb{R}^s)$.

In summary,

$$\mathcal{L}_h(B) = \phi_h(M(B)) \quad (\forall B \in \mathcal{B}(\mathbb{R}^s)).$$

This completes the proof of Theorem 2.4. □

5. Applications. In this section we apply the asymptotic representation Theorem 2.4 to the analysis of asymptotic efficiency for sequences of quantum estimators.

5.1. *Quantum Hodges estimator.* In order to motivate ourselves to study asymptotic efficiency in the quantum domain, let us touch upon the issue of quantum superefficiency first. In classical statistics there was a well-known superefficient estimator called the Hodges estimator that asymptotically breaks the Cramér–Rao bound [21]. An analogous estimator can be constructed in the quantum domain that asymptotically breaks the Holevo bound.

Let us consider the pure state model

$$\mathcal{S} = \left\{ \rho_\theta = \frac{1}{2}\left(I + \theta^1\sigma_1 + \theta^2\sigma_2 + \sqrt{1-(\theta^1)^2-(\theta^2)^2}\,\sigma_3\right) : \theta \in \mathbb{R}^2,\ (\theta^1)^2 + (\theta^2)^2 < 1 \right\},$$

on $\mathcal{H} = \mathbb{C}^2$ having two-dimensional parameter $\theta = (\theta^1, \theta^2)$, where σ_1, σ_2, σ_3 are the Pauli matrices. It is well known [11] that the weighted trace of the covariant matrix $V_\theta[M, \hat{\theta}]$ for a locally unbiased estimator $(M, \hat{\theta})$ with a weight matrix $G > 0$ is bounded from below by the Holevo bound $c_G^{(H)}$ as

$$\operatorname{Tr} G V_\theta[M, \hat{\theta}] \geq c_G^{(H)}.$$

If we set G to be the SLD Fisher information matrix $J_\theta^{(S)}$, the Holevo bound $c^{(H)}_{J_\theta^{(S)}}$ is reduced to 4, which is independent of θ, and is achieved when and only when $V_\theta[M, \hat{\theta}] = (J_\theta^{(S)}/2)^{-1}$; specifically, it is achievable by a randomized measurement scheme without invoking any collective measurement [22].

Now we construct a sequence of estimators that asymptotically breaks the Holevo bound. It is known that for the i.i.d. model $\mathcal{S}^{(n)} = \{\rho_\theta^{\otimes n}\}_\theta$, there is an adaptive estimation scheme $(\hat{M}^{(n)}, \hat{\theta}^{(n)})$ in which $\sqrt{n}(\hat{\theta}^{(n)} - \theta)$ weakly converges to the (classical) normal distribution $N(0, (J_\theta^{(S)}/2)^{-1})$ for every θ [2]. Introduce a second estimator T_n by

$$T_n := \begin{cases} \hat{\theta}^{(n)} & \text{if } \|\hat{\theta}^{(n)}\| \geq 1/\sqrt[4]{n}, \\ 0 & \text{if } \|\hat{\theta}^{(n)}\| < 1/\sqrt[4]{n}. \end{cases} \tag{23}$$

Then $\sqrt{n}(T_n - \theta)$ converges to $N(0, (J_\theta^{(S)}/2)^{-1})$ in distribution if $\theta \neq 0$, whereas it converges to 0 in probability if $\theta = 0$. At first sight, T_n is an improvement on $\hat{\theta}^{(n)}$. However, as demonstrated below, this reasoning is a bad use of asymptotics [21].

In order to evaluate the asymptotic behavior of $\sqrt{n}(T_n - \theta)$ in more detail, we assume the following situation: through the first stage of estimation, the adaptive measurement $\hat{M}^{(k)}$ has converged to a measurement M_θ that is optimal at the true value of θ [2]. Now we proceed to the second stage: fix the measurement to be the one that has been obtained through the first stage, that is, M_θ, and take $\hat{\theta}^{(n)}$ to be the sample average of outcomes over n-i.i.d. experiments, each being distributed as $N(\theta, V_\theta)$, where $V_\theta = (J_\theta^{(S)}/2)^{-1}$ so that $\hat{\theta}^{(n)} \sim N(\theta, V_\theta/n)$. Under this situation the weighted trace $\operatorname{Tr} J_\theta^{(S)} V_\theta[M_\theta, T_n]$ of covariance matrix of the quantum Hodges estimator T_n can be evaluated as follows. Because of the rotational symmetry of the model $\mathcal{S}$ around the origin of the parameter space, we can assume, without loss of generality, that the true parameter θ lies on the plane $\theta^2 = 0$. In this case

$$\operatorname{Tr} J_\theta^{(S)} V_\theta[M_\theta, T_n] = \int_0^{2\pi} d\phi \int_{1/\sqrt[4]{n}}^{\infty} w_\theta(r, \phi) q_\theta(r, \phi) r\, dr + \int_0^{2\pi} d\phi \int_0^{1/\sqrt[4]{n}} w_\theta(0, \phi) q_\theta(r, \phi) r\, dr,$$

where

$$w_\theta(r, \phi) := \frac{(r\cos\phi - \theta^1)^2}{1 - (\theta^1)^2} + r^2 \sin^2\phi$$

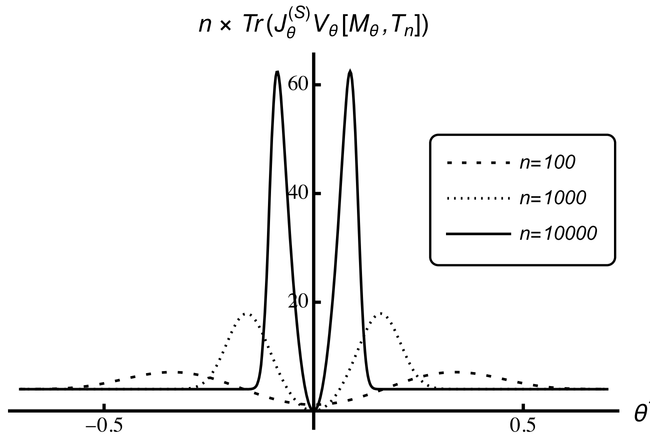

FIG. 1. *Weighted trace of covariance matrix of the quantum Hodges estimator T_n with the weight $J_\theta^{(S)}$ for the spin coherent state model $\mathcal{S}$, based on the means of samples of size* 100 (*dashed*), 1000 (*dotted*), *and* 10,000 (*solid*) *observations. For reference, the corresponding Holevo bound is* $c^{(H)}_{J_\theta^{(S)}} = 4$.

is the weighted sum of squared errors and

$$q_\theta(r,\phi) r\, dr\, d\phi := \frac{n}{4\pi\sqrt{1-(\theta^1)^2}} \exp\left[-\frac{n}{4} w_\theta(r,\phi)\right] r\, dr\, d\phi$$

is the probability density of $\hat{\theta}^{(n)} - \theta \sim N(0, V_\theta/n)$ in the polar coordinate system.

Figure 1 shows the graph of $n \times \operatorname{Tr} J_\theta^{(S)} V_\theta[M_\theta, T_n]$ for three different values of n. These functions are close to the Holevo bound $c^{(H)}_{J_\theta^{(S)}} = 4$ on most of the domain but possess peaks close to zero. As $n \to \infty$, the location and widths of the peaks converge to zero but their heights to infinity. Because the values of θ at which T_n behaves badly differ from n to n, the pathological behavior of $\sqrt{n}(T_n - \theta)$ is not visible in the pointwise limit distributions under fixed θ, as in the classical case [21].

5.2. *Quantum regular estimator.* In classical statistics it is customary to restrict ourselves to a certain class of estimators in order to avoid pathological behavior like the Hodges estimator. In this section we shall extend such a strategy to the quantum domain.

We begin with a standard estimation problem for a quantum Gaussian shift model. Our problem is to estimate the parameter $h \in \mathbb{R}^d$ of the quantum Gaussian shift model $\phi_h \sim N((\operatorname{Re}\tau)h, \Sigma)$, where Σ is an $r \times r$ complex nonnegative matrix ($r \ge d$) with $\operatorname{Re}\Sigma > 0$, and τ is an $r \times d$ complex matrix with $\operatorname{rank}(\operatorname{Re}\tau) = d$.

An estimator for the model ϕ_h is represented by a POVM M over $\mathbb{R}^d$. For each $h \in \mathbb{R}^d$, let $M - h$ denote the shifted POVM in which the outcome x of M is transformed into $y = x - h$. It is formally defined by

$$\int_B f(y)\, \phi_h\big((M-h)(dy)\big) := \int_{B+h} f(x-h)\, \phi_h\big(M(dx)\big) \quad (\forall B \in \mathcal{B}(\mathbb{R}^d)).$$

An estimator M for ϕ_h is called *equivalent in law* if the probability distribution of the outcomes of the shifted POVM $M - h$ applied to ϕ_h is independent of $h \in \mathbb{R}^d$ in that

$$\phi_h\big((M-h)(B)\big) = \phi_0\big(M(B)\big)$$

holds for all $B \in \mathcal{B}(\mathbb{R}^d)$. The following result is standard.

LEMMA 5.1. *Assume that an estimator M for the shift parameter h of a quantum Gaussian shift model $\phi_h \sim N((\mathrm{Re}\,\tau)h, \Sigma)$ is equivalent in law. Then, for any $d \times d$ weight matrix $G > 0$,*

$$\int_{\mathbb{R}^d} G_{ij}(x-h)^i(x-h)^j \phi_h\big(M(dx)\big) \geq c_G^{(H)},$$

where $c_G^{(H)}$ is the Holevo bound.

PROOF. See Section E of the Supplementary Material [4]. □

Now we introduce the notion of regular estimators[8] for q-LAN models. Suppose that we are given a sequence $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ of quantum statistical models that is q-LAN at $\theta_0 \in \Theta$. A sequence $M^{(n)}$ of POVMs is called *regular at* $\theta_0 \in \Theta$ if the classical distribution $\mathcal{L}^{(n)h}$ of outcomes of the shifted POVM

$$M^{(n)h} := \sqrt{n}\{M^{(n)} - (\theta_0 + h/\sqrt{n})\}$$

under $\rho^{(n)}_{\theta_0+h/\sqrt{n}}$ converges to a classical distribution $\mathcal{L}$ that is independent of h,

$$\mathcal{L}^{(n)h} \rightsquigarrow \mathcal{L} \quad (\forall h \in \mathbb{R}^d) \tag{24}$$

Note that $M^{(n)h}$ is a measurement in which the outcome $\hat{\theta} \in \mathbb{R}^d$ of $M^{(n)}$ is transformed into $\sqrt{n}\{\hat{\theta} - (\theta_0 + h/\sqrt{n})\}$. Since

$$\sqrt{n}\{\hat{\theta} - (\theta_0 + h/\sqrt{n})\} \leq t \quad \Longleftrightarrow \quad \hat{\theta} \leq \theta_0 + (h+t)/\sqrt{n},$$

we see that

$$M^{(n)h}(-\infty, t] = M^{(n)}\left(-\infty, \theta_0 + \frac{h+t}{\sqrt{n}}\right].$$

When a sequence $\mathcal{S}^{(n)}$ of quantum statistical models is q-LAN and D-extendible at $\theta_0 \in \Theta$, the next theorem is an immediate consequence of the asymptotic representation Theorem 2.4 and Lemma 5.1.

THEOREM 5.2 (Bound for quantum regular estimator). *Let $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ be a sequence of quantum statistical models that is q-LAN and D-extendible at $\theta_0 \in \Theta$. For any estimator $M^{(n)}$ that is regular at θ_0 and a $d \times d$ weight matrix $G > 0$, we have*

$$\int_{\mathbb{R}^d} G_{ij} x^i x^j \mathcal{L}(dx) \geq c_G^{(\mathrm{rep})}, \tag{25}$$

and hence

$$\liminf_{n\to\infty} \int_{\mathbb{R}^d} G_{ij}(x-h)^i(x-h)^j \operatorname{Tr} \rho^{(n)}_{\theta_0+h/\sqrt{n}} M^{(n)}(dx) \geq c_G^{(\mathrm{rep})}, \tag{26}$$

where $\mathcal{L}$ is the limit distribution of $M^{(n)h}$ under $\rho^{(n)}_{\theta_0+h/\sqrt{n}}$, and $c_G^{(\mathrm{rep})}$ is the asymptotic representation bound defined by (7).

PROOF. See Section E of the Supplementary Material [4]. □

It is natural to inquire whether there exists a regular estimator $M^{(n)}$ that achieves the lower bound $c_G^{(\mathrm{rep})}$ in Theorem 5.2. The answer is given by the following.

[8] In classical statistics regularity is also called *asymptotically equivalent in law*.

THEOREM 5.3 (Achievability of asymptotic representation bound). *Assume that a quantum statistical model* $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ *is q-LAN and D-extendible at* $\theta_0 \in \Theta$. *Given a* $d \times d$ *weight matrix* $G > 0$, *there exist a regular estimator* $M_\star^{(n)}$ *and a* $d \times d$ *real strictly positive matrix* $V_\star$ *such that*

$$(M_\star^{(n)}, \rho^{(n)}_{\theta_0+h/\sqrt{n}}) \overset{h}{\rightsquigarrow} N(h, V_\star)$$

and

$$\operatorname{Tr} G V_\star = c_G^{(\mathrm{rep})}$$

for all $h \in \mathbb{R}^d$.

PROOF. See Section E of the Supplementary Material [4]. □

Theorem 5.3 implies that the asymptotic representation bound $c_G^{(\mathrm{rep})}$ is achievable in that

$$\sup_{L>0} \lim_{n\to\infty} \int_{\mathbb{R}^d} L \wedge \{G_{ij}(x-h)^i(x-h)^j\} \operatorname{Tr} \rho^{(n)}_{\theta_0+h/\sqrt{n}} M_\star^{(n)}(dx) = c_G^{(\mathrm{rep})}.$$

Moreover, in combination with Theorem 5.2, Theorem 5.3 tells us that the bound $c_G^{(\mathrm{rep})}$ gives the ultimate limit of estimation precision. This fact has the following important consequence: since an achievable "scalar" lower bound for an estimation problem is necessarily unique, the bound $c_G^{(\mathrm{rep})}$ is uniquely determined. More precisely, we have the following.

COROLLARY 5.4 (Well-definedness of asymptotic representation bound). *For each* $d \times d$ *weight matrix* $G > 0$, *the asymptotic representation bound* $c_G^{(\mathrm{rep})}$ *is independent of the choice of a D-extension.*

It should be emphasized here that Theorem 5.3 is valid for *all* $h \in \mathbb{R}^d$. This is a remarkable refinement of the former result [23], Theorem 3.1, in which the Holevo bound $c_G^{(H)}$ for an i.i.d. model was achieved only on a *countable dense subset* of $\mathbb{R}^d$.

5.3. *Quantum minimax theorem.* We can also study efficiency in terms of minimax criteria. Let us begin with a minimax theorem for a quantum Gaussian shift model.

THEOREM 5.5 (Minimax theorem for quantum Gaussian shift model). *Suppose that we are given a quantum Gaussian shift model* $\phi_h \sim N((\operatorname{Re}\tau)h, \Sigma)$. *Then, for any estimator* M *and a weight matrix* $G > 0$,

$$\sup_{h\in\mathbb{R}^d} \int_{\mathbb{R}^d} G_{ij}(x-h)^i(x-h)^j \phi_h\big(M(dx)\big) \ge c_G^{(H)}.$$

PROOF. See Section E of the Supplementary Material [4]. □

By using the asymptotic representation Theorem 2.4 as well as Theorem 5.5, we can prove the following.

THEOREM 5.6 (Local asymptotic minimax theorem). *Let* $\mathcal{S}^{(n)} = \{\rho_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ *be a sequence of quantum statistical models that is q-LAN and D-extendible at* $\theta_0 \in \Theta$. *Then,* *for any sequence* $M^{(n)}$ *of estimators and* $d \times d$ *weight matrix* $G > 0$,

$$\begin{aligned}
&\lim_{\delta\to 0}\liminf_{n\to\infty}\sup_{\|h\|\le\delta\sqrt{n}}\int_{\mathbb{R}^d} G_{ij}(x-h)^i(x-h)^j \operatorname{Tr}\rho^{(n)}_{\theta_0+h/\sqrt{n}}M^{(n)}(dx)\\
&\quad\ge \sup_{H}\liminf_{n\to\infty}\sup_{h\in H}\int_{\mathbb{R}^d} G_{ij}(x-h)^i(x-h)^j \operatorname{Tr}\rho^{(n)}_{\theta_0+h/\sqrt{n}}M^{(n)}(dx)\\
&\quad\ge \sup_{L>0}\sup_{H}\liminf_{n\to\infty}\sup_{h\in H}\int_{\mathbb{R}^d} L\wedge\{G_{ij}(x-h)^i(x-h)^j\}\operatorname{Tr}\rho^{(n)}_{\theta_0+h/\sqrt{n}}M^{(n)}(dx)\\
&\quad\ge c_G^{(\mathrm{rep})}.
\end{aligned}\tag{27}$$

Here $a\wedge b := \min\{a,b\}$, *and* H *runs over all finite subsets of* $\mathbb{R}^d$. *Moreover, the last inequality is tight.*

PROOF. See Section E of the Supplementary Material [4]. □

Note that the quantities appeared in the first and second lines of (27) correspond to the minimax theorems due to Háyak [9] and in van der Vaart's book [21], respectively.

5.4. *Quantum James–Stein estimator.* As the final topic of this section, we touch upon a superefficient estimator that uniformly breaks the asymptotic representation bound $c_G^{(\mathrm{rep})}$.

Let us consider the i.i.d. quantum statistical model $\mathcal{S}^{(n)} := \{\rho_\theta^{\otimes n}\}$ with the base model

$$\rho_\theta = \frac{1}{2}(I+\theta^1\sigma_1+\theta^2\sigma_2+\theta^3\sigma_3),\quad (\theta=(\theta^1,\theta^2,\theta^3), \|\theta\|^2<1)$$

on $\mathcal{H}=\mathbb{C}^2$. We see from Section C of the Supplementary Material [4] that $\mathcal{S}^{(n)}$ is q-LAN and D-extendible at every point θ. In fact, since the linear span of SLDs at each θ is $\mathcal{D}_{\rho_\theta}$-invariant, the set of SLDs itself gives a D-extension.

Here we focus our attention on the local asymptotic estimation at around the origin $\theta=0$. The SLDs at $\theta=0$ are σ_i $(i=1,2,3)$, and the corresponding SLD Fisher information matrix J^S is the identity matrix. Let

$$\Delta_i^{(n)} := \frac{1}{\sqrt{n}}\sum_{k=1}^n I^{\otimes(k-1)}\otimes\sigma_i\otimes I^{\otimes(n-k)}.$$

Then the asymptotic representation Theorem 2.4 allows us to convert the problem of estimating the local parameter h of $\rho_{h/\sqrt{n}}^{\otimes n}$ into that of estimating the shift parameter h of the limiting (classical) Gaussian shift model

$$\{N(h,I): h\in\mathbb{R}^3\}.\tag{28}$$

Specifically, for any regular POVM $M^{(n)}$ that satisfies

$$(M^{(n)},\rho_{h/\sqrt{n}}^{\otimes n}) \rightsquigarrow \exists\mathcal{L}_h,$$

we see from Theorem 5.2 and (7) that

$$\int_{\mathbb{R}^d}\|x-h\|^2\mathcal{L}_h(dx)\ge c_I^{(\mathrm{rep})} = \operatorname{Tr} I = 3,$$

where we have taken the weight G to be the SLD Fisher information matrix $J^S = I$.

Now we demonstrate that, if one discards the requirement of regularity, one can construct an estimator that breaks the above inequality for all h. An estimator on the classical Gaussian shift model (28) that changes observed data $x \in \mathbb{R}^3$ into

$$y = \left(1 - \frac{1}{\|x\|}\right)x \tag{29}$$

is called the James–Stein estimator [15]. Letting $\mathcal{L}_h^{(JS)}$ be the corresponding probability distribution of y, it is well known that

$$\int_{\mathbb{R}^d} \|y-h\|^2 \mathcal{L}_h^{(JS)}(dy) < 3$$

for all $h \in \mathbb{R}^3$. Now we see from Theorem 5.3 that there is a regular POVM $N^{(n)}$ that exhibits

$$\left(N^{(n)}, \rho_{h/\sqrt{n}}^{\otimes n}\right) \rightsquigarrow N(h, I).$$

For each n, let $N^{(JS,n)}$ be a POVM that changes the outcome $x \in \mathbb{R}^3$ of $N^{(n)}$ into $y \in \mathbb{R}^3$ as (29). Then

$$\left(N^{(JS,n)}, \rho_{h/\sqrt{n}}^{\otimes n}\right) \rightsquigarrow \mathcal{L}_h^{(JS)},$$

and thus $N^{(JS,n)}$ asymptotically breaks the asymptotic representation bound $c_I^{(\mathrm{rep})}$ for all $h \in \mathbb{R}^3$.

6. Conclusions. In this paper we derived a noncommutative analogue of asymptotic representation theorem for a D-extendible q-LAN model (Theorem 2.4). This theorem converts an estimation problem for a local model $\{\rho_{\theta_0+h/\sqrt{n}}^{(n)} : h \in \mathbb{R}^d\}$ into another for the limiting quantum Gaussian shift model $\{N((\mathrm{Re}\,\tau)h, \Sigma) : h \in \mathbb{R}^d\}$. As a corollary, we arrived at a new bound $c_G^{(\mathrm{rep})}$ defined by the Holevo bound for the limiting model. This bound turned out to have universal importance in asymptotic quantum statistics. For example, it gave the ultimate limit of estimation precision for regular estimators (Theorems 5.2 and 5.3) and minimax estimators (Theorem 5.6). Note that, since the bound $c_G^{(\mathrm{rep})}$ for an i.i.d. model is reduced to the standard Holevo bound for the base model, the achievability theorem (Theorem 5.3) gives a substantial refinement of the former result [23], Theorem 3.1, in which the Holevo bound was achieved only on a countable dense subset of the parameter space.

The key ingredient of Theorem 2.4 was the notion of D-extendibility. Its importance is first realized in the present paper; however, its trace can be found elsewhere. Guţă and Kahn [8, 17] and Lahiry and Nussbaum [19] treated i.i.d. extensions of a quantum statistical model that has $\mathcal{D}_{\rho_\theta}$-invariant SLD-tangent space at every θ from the outset, and thus they did not need to care about the D-extendibility (Remark 2.3). In their framework the difficulty demonstrated in Example 2.1 is automatically avoided by regarding the model as a submodel of its ambient full pure state model. Yamagata *et al.* [23] introduced the notion of joint q-LAN for $(X^{(n)}, \Delta^{(n)})$. In view of the present paper, this was a forerunner of the D-extension $X^{(n)}$ of SLDs $\Delta^{(n)}$, whereby the achievability of the Holevo bound was proved. The notion of D-extendibility made it possible to generalize the Holevo bound to non-i.i.d. models, providing a proper perspective on the achievability of the asymptotic representation bound $c_G^{(\mathrm{rep})}$.

We believe that the paper has established a solid foundation of the theory of (weak) quantum local asymptotic normality. Nevertheless, its application has just begun, and many open problems are left for future study. For example, it is not clear whether the D-extendibility condition can be replaced with a weaker one. One would convince oneself that there are quantum statistical models that are not i.i.d. but are, nevertheless, q-LAN and D-extendible.

Imagine a convergent sequence $\sigma_\theta^{(n)} \to \sigma_\theta^{(\infty)}$ of quantum statistical models on a fixed finite dimensional Hilbert space $\mathcal{H}$. Then the tensor product models $\rho_\theta^{(n)} := \bigotimes_{k=1}^n \sigma_\theta^{(k)}$ would be q-LAN and D-extendible because they are "almost" i.i.d. in the asymptotic limit. In fact, it is not difficult to realize this idea with some additional conditions (cf. Section C of the Supplementary Material [4]). In this way the D-extendibility condition enables us to study quantum asymptotics beyond the i.i.d. assumption. In view of applications, however, it would be nice if there were a more tractable weaker condition that establishes an asymptotic representation theorem.

It also remains to be investigated whether an asymptotically optimal statistical procedure for the local model, indexed by the parameter $\theta_0 + h/\sqrt{n}$, can be translated into useful statistical procedures for the real world case in which θ_0 is unknown. Some authors [6] advocated two-step estimation procedures in which one first measures a small portion of the quantum system, in number n_1 say, using some standard measurement scheme and constructs an initial estimate, say $\tilde{\theta}_1$, of the parameter. One next applies the theory of q-LAN to compute the asymptotically optimal measurement scheme, which corresponds to the situation $\theta_0 = \tilde{\theta}_1$, and then proceeds to implement this measurement on the remaining n_2 $(:= n - n_1)$ quantum systems collectively, estimating h in the model $\theta = \tilde{\theta}_1 + h/\sqrt{n_2}$. However, such procedures are inherently limited to within the scope of weak consistency. Studying the strong consistency and asymptotic efficiency [2] in the framework of collective quantum estimation scheme is an important open problem.

Funding. The present study was supported by JSPS KAKENHI Grant Numbers JP17H02861, JP22K03466, and MEXT Quantum Leap Flagship Program (MEXT Q-LEAP) Grant Number JPMXS0120351339.

SUPPLEMENTARY MATERIAL

Supplementary Material to "Efficiency of estimators for locally asymptotically normal quantum statistical models" (DOI: 10.1214/23-AOS2285SUPP; .pdf). Section A gives an alternative view for the asymptotic representation theorem for classical LAN models as well as its comprehensible proof. Section B gives a brief account of degenerate canonical commutation relation (CCR) and hybrid classical/quantum Gaussian states. Section C gives a detailed account of the notion of D-extendibility, including a proof of D-extendibility of i.i.d. models and an example of non-i.i.d. model that is q-LAN and D-extendible. Section D is devoted to detailed proofs of lemmas presented in Section 3. Section E is devoted to detailed proofs of theorems presented in Section 5.

REFERENCES

[1] ANDO, T. (1978). *Topics on Operator Inequalities*. Hokkaido Univ., Research Institute of Applied Electricity, Division of Applied Mathematics, Sapporo. MR0482378

[2] FUJIWARA, A. (2006). Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. *J. Phys. A* **39** 12489–12504. MR2265836 https://doi.org/10.1088/0305-4470/39/40/014

[3] FUJIWARA, A. and YAMAGATA, K. (2020). Noncommutative Lebesgue decomposition and contiguity with applications in quantum statistics. *Bernoulli* **26** 2105–2142. MR4091103 https://doi.org/10.3150/19-BEJ1185

[4] FUJIWARA, A. and YAMAGATA, K. (2023). Supplement to "Efficiency of estimators for locally asymptotically normal quantum statistical models." https://doi.org/10.1214/23-AOS2285SUPP

[5] 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. IMS, Beachwood, OH. MR3186752 https://doi.org/10.1214/12-IMSCOLL909

[6] GILL, R. D. and MASSAR, S. (2000). State estimation for large ensembles. *Phys. Rev. A* **61** 042312.

[7] 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

[8] GUŢĂ, M. and KAHN, J. (2006). Local asymptotic normality for qubit states. *Phys. Rev. A* (3) **73** 052108, 15. MR2229156 https://doi.org/10.1103/PhysRevA.73.052108

[9] HÁJEK, J. (1972). Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* (*Univ. California*, *Berkeley*, *Calif.*, 1970/1971), *Vol. I*: *Theory of Statistics* 175–194. Univ. California Press, Berkeley, CA. MR0400513

[10] HIAI, F. (2021). *Lectures on Selected Topics in Von Neumann Algebras*. *EMS Series of Lectures in Mathematics*. EMS Press, Berlin. MR4331436 https://doi.org/10.4171/ELM/32

[11] HOLEVO, A. (2011). *Probabilistic and Statistical Aspects of Quantum Theory*, 2nd ed. *Quaderni/Monographs* **1**. Edizioni della Normale, Pisa. MR2797301 https://doi.org/10.1007/978-88-7642-378-9

[12] HUSIMI, K. (1940). Some formal properties of the density matrix. *Proc. Phys.-Math. Soc. Jpn.* **22** 264–314.

[13] JAKŠIĆ, V., PAUTRAT, Y. and PILLET, C.-A. (2010). A non-commutative Lévy–Cramér continuity theorem. *Markov Process. Related Fields* **16** 59–78. MR2664336

[14] JAKŠIĆ, V., PAUTRAT, Y. and PILLET, C.-A. (2010). A quantum central limit theorem for sums of independent identically distributed random variables. *J. Math. Phys.* **51** 015208, 8. MR2605841 https://doi.org/10.1063/1.3285287

[15] JUDGE, G. G. and BOCK, M. E. (1978). *The Statistical Implications of Pre-Test and Stein-Rule Estimators in Econometrics*. *Studies in Mathematical and Managerial Economics*, *Vol.* 25. North-Holland, Amsterdam–Oxford. MR0483199

[16] KADISON, R. V. and RINGROSE, J. R. (1997). *Fundamentals of the Theory of Operator Algebras. Vol. II*: *Advanced Theory*. *Graduate Studies in Mathematics* **16**. Amer. Math. Soc., Providence, RI. MR1468230 https://doi.org/10.1090/gsm/016

[17] 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

[18] KUBO, F. and ANDO, T. (1979/80). Means of positive linear operators. *Math. Ann.* **246** 205–224. MR0563399 https://doi.org/10.1007/BF01371042

[19] LAHIRY, S. and NUSSBAUM, M. (2021). Local asymptotic normality and optimal estimation of low-rank quantum systems. Available at arXiv:2111.03279.

[20] REED, M. and SIMON, B. (1979). *Methods of Modern Mathematical Physics. I*: *Functional Analysis*. Academic Press, New York.

[21] VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge Univ. Press, Cambridge. MR1652247 https://doi.org/10.1017/CBO9780511802256

[22] YAMAGATA, K. (2011). Efficiency of quantum state tomography for qubits. *Int. J. Quantum Inf.* **9** 1167–1183. MR2931457 https://doi.org/10.1142/S0219749911007551

[23] 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