

# Noncommutative Lebesgue decomposition and contiguity with applications in quantum statistics

AKIO FUJIWARA<sup>1</sup> and KOICHI YAMAGATA<sup>2</sup>

<sup>1</sup>Department of Mathematics, Osaka University, Toyonaka, Osaka 560-0043, Japan.  
E-mail: fujiwara@math.sci.osaka-u.ac.jp

<sup>2</sup>Graduate School of Informatics and Engineering, The University of Electro-Communications, Chofu, Tokyo 182-8585, Japan. E-mail: koichi.yamagata@uec.ac.jp

We herein develop a theory of contiguity in the quantum domain based upon a novel quantum analogue of the Lebesgue decomposition. The theory thus formulated is pertinent to the weak quantum local asymptotic normality introduced in the previous paper [Yamagata, Fujiwara, and Gill, *Ann. Statist.* **41** (2013) 2197–2217], yielding substantial enlargement of the scope of quantum statistics.

*Keywords:* contiguity; Lebesgue decomposition; likelihood ratio; local asymptotic normality; quantum statistics

## 1. Introduction

Quantum statistics is a rapidly growing field of research in quantum information science. When we consider the future direction of the field, we may learn much from the history of classical statistics. One of the deepest achievements in mathematical statistics is the theory of local asymptotic normality introduced by LeCam [22]. A sequence  $\{P_{\theta}^{(n)} \mid \theta \in \Theta \subset \mathbb{R}^d\}$  of  $d$ -dimensional parametric models, each comprising probability measures on a measurable space  $(\Omega^{(n)}, \mathcal{F}^{(n)})$ , is said to be *locally asymptotically normal* (LAN) at  $\theta_0 \in \Theta$  (in the “weak” sense) if there exist a sequence  $\Delta^{(n)} = (\Delta_1^{(n)}, \dots, \Delta_d^{(n)})$  of  $d$ -dimensional random vectors and a  $d \times d$  real symmetric positive definite matrix  $J$  such that  $\Delta^{(n)} \overset{0}{\rightsquigarrow} N(0, J)$  and

$$\log \frac{dP_{\theta_0+h/\sqrt{n}}^{(n)}}{dP_{\theta_0}^{(n)}} = h^i \Delta_i^{(n)} - \frac{1}{2} h^i h^j J_{ij} + o_{P_{\theta_0}^{(n)}}(1) \quad (h \in \mathbb{R}^d). \quad (1.1)$$

Here, the arrow  $\overset{h}{\rightsquigarrow}$  stands for the convergence in distribution under  $P_{\theta_0+h/\sqrt{n}}^{(n)}$ , the remainder term  $o_{P_{\theta_0}^{(n)}}(1)$  converges in probability to zero under  $P_{\theta_0}^{(n)}$ , and Einstein’s summation convention is used.

The notion of local asymptotic normality provides a useful tool to cope with various statistical models in a unified manner by reducing them to relevant Gaussian shift models in the asymptotic limit. Observe that the expansion (1.1) is similar in form to the log-likelihood ratio of the

Gaussian shift model:

$$\log \frac{dN(h, J^{-1})}{dN(0, J^{-1})}(X^1, \dots, X^d) = h^i (X^j J_{ij}) - \frac{1}{2} h^i h^j J_{ij}.$$

This similarity suggests a deep relationship between the models  $\{P_{\theta_0+h/\sqrt{n}}^{(n)} \mid h \in \mathbb{R}^d\}$  and  $\{N(h, J^{-1}) \mid h \in \mathbb{R}^d\}$ . In order to put the similarity to practical use, LeCam introduced the notion of contiguity [22]. A sequence  $Q^{(n)}$  of probability measures is called *contiguous* with respect to another sequence  $P^{(n)}$  of probability measures, denoted  $Q^{(n)} \triangleleft P^{(n)}$ , if  $P^{(n)}(A^{(n)}) \rightarrow 0$  implies  $Q^{(n)}(A^{(n)}) \rightarrow 0$  for any sequence  $A^{(n)}$  of measurable sets. An important conclusion pertinent to the notion of contiguity is the following theorem, which is usually referred to as LeCam’s third lemma: if  $Q^{(n)} \triangleleft P^{(n)}$  and

$$\left( X^{(n)}, \frac{dQ^{(n)}}{dP^{(n)}} \right) \overset{P^{(n)}}{\rightsquigarrow} (X, V),$$

then  $X^{(n)} \overset{Q^{(n)}}{\rightsquigarrow} L$ , where  $L$  is the law defined by  $L(B) := E[1_B(X)V]$ . Since the local asymptotic normality (1.1) entails mutual contiguity  $P_{\theta_0+h/\sqrt{n}}^{(n)} \triangleleft\triangleright P_{\theta_0}^{(n)}$ , LeCam’s third lemma proves that  $X^{(n)j} := (J^{-1})^{jk} \Delta_k^{(n)}$  exhibits  $X^{(n)} \overset{h}{\rightsquigarrow} N(h, J^{-1})$ . This gives a precise meaning of the statement that the model  $\{P_{\theta_0+h/\sqrt{n}}^{(n)} \mid h \in \mathbb{R}^d\}$  satisfying (1.1) is statistically similar to the Gaussian shift model  $\{N(h, J^{-1}) \mid h \in \mathbb{R}^d\}$ .

Note that such an interpretation is realized in the asymptotic framework. A measure theoretic counterpart of LeCam’s third lemma is the identity  $dQ = (dQ/dP)dP$ , which is valid when  $Q$  is absolutely continuous to  $P$ . In the nonasymptotic framework, the likelihood ratio  $dP_{\theta_0+h/\sqrt{n}}^{(n)}/dP_{\theta_0}^{(n)}$  carries full information about the measure  $P_{\theta_0+h/\sqrt{n}}^{(n)}$  only when  $P_{\theta_0+h/\sqrt{n}}^{(n)}$  is absolutely continuous to  $P_{\theta_0}^{(n)}$ . This fact demonstrates the differences between the contiguity and the absolute continuity, highlighting the notable flexibility and usefulness of the notion of contiguity when it is used in conjunction with the weak LAN.

Extending the notion of local asymptotic normality to the quantum domain was pioneered by Guřa and Kahn [11, 18]. They proved that, given a quantum parametric model  $\mathcal{S}(\mathbb{C}^D) = \{\rho_\theta > 0 \mid \theta \in \Theta \subset \mathbb{R}^{D^2-1}\}$  comprising the totality of faithful density operators on a  $D$ -dimensional Hilbert space and a point  $\theta_0$  on the parameter space  $\Theta$  such that  $\rho_{\theta_0}$  is nondegenerate (i.e., every eigenvalue of  $\rho_{\theta_0}$  is simple), there exist, for any compact subset  $K \subset \mathbb{R}^{D^2-1}$ , quantum channels  $S_n$  and  $T_n$  such that

$$\lim_{n \rightarrow \infty} \sup_{h \in K} \|\sigma_h - T_n(\rho_{\theta_0+h/\sqrt{n}}^{\otimes n})\|_1 = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \sup_{h \in K} \|S_n(\sigma_h) - \rho_{\theta_0+h/\sqrt{n}}^{\otimes n}\|_1 = 0,$$

where  $\{\sigma_h \mid h \in \mathbb{R}^{D^2-1}\}$  are density operators of a certain quantum Gaussian shift model represented on a suitably chosen separable Hilbert space. (See Appendix A for a brief account of quantum Gaussian states.)

Note that this formulation is not a direct analogue of the weak LAN defined by (1.1); 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 called a “strong” q-LAN (cf. [23], Chapter 10).

Guță and Kahn’s theorem in terms of the strong q-LAN was so powerful that it was applied to the study of asymptotic quantum parameter estimation problems in [35]. However, the strong q-LAN after Guță and Kahn is not fully satisfactory because it is applicable only to i.i.d. extensions of a quantum statistical model around a nondegenerate reference state  $\rho_{\theta_0}$ . It is natural to seek a more flexible formulation that is applicable to non-i.i.d. cases with possibly degenerate reference states. In [10], they tried a different approach to a “weak” q-LAN via the Connes co-cycle 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 (1.1) in the classical LAN.

A different approach to a weak q-LAN was put forward in [34], in which a sequence of quantum statistical models comprising mutually absolutely continuous density operators was treated. Here, density operators  $\rho$  and  $\sigma$  on a finite dimensional Hilbert space are said to be *mutually absolutely continuous*,  $\rho \sim \sigma$  in symbols, if there exists a Hermitian operator  $\mathcal{L}$  that satisfies

$$\sigma = e^{\frac{1}{2}\mathcal{L}} \rho e^{\frac{1}{2}\mathcal{L}}.$$

The operator  $\mathcal{L}$  satisfying this relation is called (a version of) the *quantum log-likelihood ratio*. When the reference states  $\rho$  and  $\sigma$  need to be specified,  $\mathcal{L}$  is denoted as  $\mathcal{L}(\sigma|\rho)$ , so that

$$\sigma = e^{\frac{1}{2}\mathcal{L}(\sigma|\rho)} \rho e^{\frac{1}{2}\mathcal{L}(\sigma|\rho)}.$$

For example, when both  $\rho$  and  $\sigma$  are strictly positive, the quantum log-likelihood ratio is uniquely given by

$$\mathcal{L}(\sigma|\rho) = 2 \log(\sigma \# \rho^{-1}).$$

Here,  $\#$  denotes the operator geometric mean [1,21]: for strictly positive operators  $A$  and  $B$ , the operator geometric mean  $A\#B$  is defined as the unique positive operator  $X$  that satisfies the equation  $B = XA^{-1}X$ , and is explicitly given by

$$A\#B = \sqrt{A} \sqrt{\sqrt{A^{-1}} B \sqrt{A^{-1}}} \sqrt{A}.$$

The theory of weak q-LAN developed in [34] was successfully applied to quantum statistical models satisfying only some mild regularity conditions, and clarified that the Holevo bound was asymptotically achievable. However, this formulation, too, is not fully satisfactory because it is applicable only to quantum statistical models that comprises mutually absolutely continuous density operators. This is in good contrast to the classical definition (1.1), in which mutual absolute continuity for the model was not assumed [32]. The key idea behind this classical formulation is the use of the Radon–Nikodym density, or more fundamentally, the use of the Lebesgue decomposition of  $P_{\theta_0+h/\sqrt{n}}^{(n)}$  with respect to  $P_{\theta_0}^{(n)}$ . Thus, in order to extend such a flexible formulation to the quantum domain, we must invoke an appropriate quantum counterpart of the Lebesgue decomposition. Several noncommutative analogues of the Lebesgue decomposition and/or the

Radon–Nikodym derivative have been devised, for example, [2,4–6,17,20,24–26,28–31]. However, each of them has its own scope, and to the best of our knowledge, no appropriate quantum counterpart that is applicable to the theory of weak q-LAN has been established.

The objective of the present paper is threefold: Firstly, we devise a novel quantum analogue of the Lebesgue decomposition that is pertinent to the framework of weak q-LAN introduced in the previous paper [34]. Secondly, we develop a theory of contiguity in the quantum domain based on the novel quantum Lebesgue decomposition. One of the remarkable achievements of the theory is the abstract version of LeCam’s third lemma (Theorem 6.1). Finally, we apply the theory of quantum contiguity to weak q-LAN, yielding substantial enlargement of the scope of q-LAN as compared with the previous paper [34].

The present paper is organized as follows. In Section 2, we extend the notions of absolute continuity and singularity to the quantum domain in order that they are fully consistent with the notion of mutual absolute continuity introduced in [34]. In Section 3, we formulate a quantum Lebesgue decomposition based on the quantum absolute continuity and singularity introduced in Section 2. In Section 4, we develop a theory of quantum contiguity by taking full advantage of the novel quantum Lebesgue decomposition established in Section 3. In Section 5, we introduce the notion of convergence in distribution in terms of the quasi-characteristic function, and prove a noncommutative version of the Lévy–Cramér continuity theorem under the “sandwiched” convergence in distribution, which plays a key role in the subsequent discussion. In Section 6, we prove a quantum counterpart of the LeCam third lemma. This achievement manifests the validity of the novel quantum Lebesgue decomposition and quantum contiguity as well as the notion of sandwiched convergence in distribution. In Section 7, we give some illustrative examples that demonstrate the flexibility and applicability of the present formulation in asymptotic quantum statistics, including a quantum contiguity version of the Kakutani dichotomy, and enlargement of the scope of q-LAN. Section 8 is devoted to brief concluding remarks. For the reader’s convenience, some additional material is presented in the Appendix, including the quantum Gaussian states, and a noncommutative Lévy–Cramér continuity theorem.

## 2. Absolute continuity and singularity

Given positive operators  $\rho$  and  $\sigma$  on a (finite dimensional) Hilbert space  $\mathcal{H}$  with  $\rho \neq 0$ , let  $\sigma \downarrow_{\text{supp } \rho}$  denote the *excision* of  $\sigma$  relative to  $\rho$  by the operator on the subspace  $\text{supp } \rho := (\ker \rho)^\perp$  of  $\mathcal{H}$  defined by

$$\sigma \downarrow_{\text{supp } \rho} := \iota_\rho^* \sigma \iota_\rho,$$

where  $\iota_\rho : \text{supp } \rho \hookrightarrow \mathcal{H}$  is the inclusion map. More specifically, let

$$\rho = \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_0 & \alpha \\ \alpha^* & \beta \end{pmatrix} \quad (2.1)$$

be a simultaneous block matrix representations of  $\rho$  and  $\sigma$ , where  $\rho_0 > 0$ . Then the excision  $\sigma \downarrow_{\text{supp } \rho}$  is nothing but the operator represented by the (1, 1)th block  $\sigma_0$  of  $\sigma$ . The notion of

excision was exploited in [34]. In particular, it was shown that  $\rho$  and  $\sigma$  are mutually absolutely continuous if and only if

$$\sigma \llcorner_{\text{supp } \rho} > 0 \quad \text{and} \quad \text{rank } \rho = \text{rank } \sigma,$$

or equivalently, if and only if

$$\sigma \llcorner_{\text{supp } \rho} > 0 \quad \text{and} \quad \rho \llcorner_{\text{supp } \sigma} > 0. \tag{2.2}$$

Now we introduce noncommutative analogues of the notions of absolute continuity and singularity that played essential roles in the classical measure theory. Given positive operators  $\rho$  and  $\sigma$ , we say  $\rho$  is *singular* with respect to  $\sigma$ , denoted  $\rho \perp \sigma$ , if

$$\sigma \llcorner_{\text{supp } \rho} = 0.$$

The following lemma implies that the relation  $\perp$  is symmetric; this fact allows us to say that  $\rho$  and  $\sigma$  are mutually singular, as in the classical case.

**Lemma 2.1.** *For nonzero positive operators  $\rho$  and  $\sigma$ , the following are equivalent:*

- (a)  $\rho \perp \sigma$ .
- (b)  $\text{supp } \rho \perp \text{supp } \sigma$ .
- (c)  $\text{Tr } \rho \sigma = 0$ .

**Proof.** Let us represent  $\rho$  and  $\sigma$  in the form (2.1). Then (a) is equivalent to  $\sigma_0 = 0$ . In this case, the positivity of  $\sigma$  entails that the off-diagonal blocks  $\alpha$  and  $\alpha^*$  of  $\sigma$  also vanish, and  $\sigma$  takes the form

$$\sigma = \begin{pmatrix} 0 & 0 \\ 0 & \beta \end{pmatrix}.$$

This implies (b). Next, (b)  $\Rightarrow$  (c) is obvious. Finally, assume (c). With the representation (2.1), this is equivalent to  $\text{Tr } \rho_0 \sigma_0 = 0$ . Since  $\rho_0 > 0$ , we have  $\sigma_0 = 0$ , proving (a).  $\square$

We next introduce the notion of absolute continuity. Given positive operators  $\rho$  and  $\sigma$ , we say  $\rho$  is *absolutely continuous* with respect to  $\sigma$ , denoted  $\rho \ll \sigma$ , if

$$\sigma \llcorner_{\text{supp } \rho} > 0.$$

Some remarks are in order. Firstly, the above definition of absolute continuity is consistent with the definition of mutual absolute continuity: in fact, as demonstrated in (2.2),  $\rho$  and  $\sigma$  are mutually absolutely continuous if and only if both  $\rho \ll \sigma$  and  $\sigma \ll \rho$  hold. Secondly,  $\rho \ll \sigma$  is a much weaker condition than  $\text{supp } \rho \subset \text{supp } \sigma$ : this makes a striking contrast to the classical measure theory. For example, pure states  $\rho = |\psi\rangle\langle\psi|$  and  $\sigma = |\xi\rangle\langle\xi|$  are mutually absolutely continuous if and only if  $\langle\xi|\psi\rangle \neq 0$  (see [34], Example 2.3).

The following lemma plays a key role in the next section, leading to a novel noncommutative Lebesgue decomposition.

**Lemma 2.2.** For nonzero positive operators  $\rho$  and  $\sigma$ , the following are equivalent:

- (a)  $\rho \ll \sigma$ .
- (b)  $\exists R > 0$  such that  $\sigma \geq R\rho R$ .
- (c)  $\exists R > 0$  such that  $\rho \leq R\sigma R$ .
- (d)  $\exists R \geq 0$  such that  $\rho = R\sigma R$ .
- (e)  $\exists R \geq 0$  such that  $\rho \geq R\sigma R$  and  $\text{Tr } \rho = \text{Tr } \sigma R^2$ .

**Proof.** We first prove (a)  $\Rightarrow$  (b). Let

$$\rho = \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_0 & \alpha \\ \alpha^* & \beta \end{pmatrix}$$

where  $\rho_0 > 0$ . Since  $\sigma_0 = \sigma \downarrow_{\text{supp } \rho} > 0$ , the matrix  $\sigma$  is further decomposed as

$$\sigma = E^* \begin{pmatrix} \sigma_0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} E, \quad E := \begin{pmatrix} I & \sigma_0^{-1} \alpha \\ 0 & I \end{pmatrix}.$$

Note that, since  $\sigma \geq 0$  and  $E$  is full-rank, we have

$$\beta - \alpha^* \sigma_0^{-1} \alpha \geq 0. \tag{2.3}$$

Now we set

$$R := E^* \begin{pmatrix} X & 0 \\ 0 & \gamma \end{pmatrix} E,$$

where  $X := \sigma_0 \# \rho_0^{-1}$ , and  $\gamma$  is an arbitrary strictly positive operator. Then

$$\begin{aligned} R\rho R &= E^* \begin{pmatrix} X & 0 \\ 0 & \gamma \end{pmatrix} E \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix} E^* \begin{pmatrix} X & 0 \\ 0 & \gamma \end{pmatrix} E \\ &= E^* \begin{pmatrix} X & 0 \\ 0 & \gamma \end{pmatrix} \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} X & 0 \\ 0 & \gamma \end{pmatrix} E \\ &= E^* \begin{pmatrix} X\rho_0 X & 0 \\ 0 & 0 \end{pmatrix} E \\ &= E^* \begin{pmatrix} \sigma_0 & 0 \\ 0 & 0 \end{pmatrix} E \\ &\leq E^* \begin{pmatrix} \sigma_0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} E = \sigma. \end{aligned}$$

Here, the inequality is due to (2.3). Since  $R > 0$ , we have (b).

We next prove (b)  $\Rightarrow$  (a). Due to assumption, there is a positive operator  $\tau \geq 0$  such that

$$\sigma = R\rho R + \tau.$$

Let

$$\rho = \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad R = \begin{pmatrix} R_0 & R_1 \\ R_1^* & R_2 \end{pmatrix}, \quad \tau = \begin{pmatrix} \tau_0 & \tau_1 \\ \tau_1^* & \tau_2 \end{pmatrix},$$

where  $\rho_0 > 0$ . Then

$$\sigma = \begin{pmatrix} R_0 \rho_0 R_0 + \tau_0 & R_0 \rho_0 R_1 + \tau_1 \\ R_1^* \rho_0 R_0 + \tau_1^* & R_1^* \rho_0 R_1 + \tau_2 \end{pmatrix}$$

and

$$\sigma \downarrow_{\text{supp } \rho} = R_0 \rho_0 R_0 + \tau_0.$$

Since  $R_0 > 0$  and  $\tau_0 \geq 0$ , we have  $\sigma \downarrow_{\text{supp } \rho} > 0$ .

For the proof of (a)  $\Rightarrow$  (d), let

$$\rho = \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_0 & \alpha \\ \alpha^* & \beta \end{pmatrix},$$

where  $\rho_0 > 0$ . Since  $\sigma_0 = \sigma \downarrow_{\text{supp } \rho} > 0$ ,

$$R := \begin{pmatrix} \rho_0 \# \sigma_0^{-1} & 0 \\ 0 & 0 \end{pmatrix}$$

is a well-defined positive operator satisfying

$$\rho = R \sigma R.$$

This proves (d).

For (d)  $\Rightarrow$  (a), let the positive operator  $R$  in  $\rho = R \sigma R$  be represented as

$$R = \begin{pmatrix} R_0 & 0 \\ 0 & 0 \end{pmatrix},$$

where  $R_0 > 0$ , and accordingly, let us represent  $\rho$  and  $\sigma$  as

$$\rho = \begin{pmatrix} \rho_0 & \rho_1 \\ \rho_1^* & \rho_2 \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_0 & \sigma_1 \\ \sigma_1^* & \sigma_2 \end{pmatrix}.$$

The relation  $\rho = R \sigma R$  is then reduced to

$$\begin{pmatrix} \rho_0 & \rho_1 \\ \rho_1^* & \rho_2 \end{pmatrix} = \begin{pmatrix} R_0 \sigma_0 R_0 & 0 \\ 0 & 0 \end{pmatrix}.$$

This implies that  $\text{supp } \rho = \text{supp } \rho_0$  and  $\rho_0 \sim \sigma_0$ . Consequently,

$$\sigma \downarrow_{\text{supp } \rho} = \sigma \downarrow_{\text{supp } \rho_0} = \sigma_0 \downarrow_{\text{supp } \rho_0} > 0.$$

In the last inequality, we used the fact that  $\rho_0 \sim \sigma_0$  implies  $\rho_0 \ll \sigma_0$ .

Now that (b)  $\Leftrightarrow$  (c) and (d)  $\Leftrightarrow$  (e) are obvious, the proof is complete. □

### 3. Lebesgue decomposition

In this section, we extend the Lebesgue decomposition to the quantum domain.

#### 3.1. Case 1: When $\sigma \gg \rho$

To elucidate our motivation, let us first treat the case when  $\sigma \gg \rho$ . In Lemma 2.2, we found the following characterization:

$$\sigma \gg \rho \iff \exists R > 0 \text{ such that } \sigma \geq R\rho R.$$

Note that such an operator  $R$  is not unique. For example, suppose that  $\sigma \geq R_1\rho R_1$  holds for some  $R_1 > 0$ . Then for any  $t \in (0, 1]$ , the operator  $R_t := tR_1$  is strictly positive and satisfies  $\sigma \geq R_t\rho R_t$ . It is then natural to seek, if any, the “maximal” operator of the form  $R\rho R$  that is packed into  $\sigma$ . Put differently, letting  $\tau := \sigma - R\rho R$ , we want to find the “minimal” positive operator  $\tau$  that satisfies

$$\sigma = R\rho R + \tau, \tag{3.1}$$

where  $R > 0$ . This question naturally leads us to a noncommutative analogue of the Lebesgue decomposition, in that a positive operator  $\tau$  satisfying (3.1) is regarded as minimal if  $\tau \perp \rho$ .

In the proof of Lemma 2.2, we found the following decomposition:

$$\begin{aligned} \sigma &= E^* \begin{pmatrix} \sigma_0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} E \\ &= E^* \begin{pmatrix} \sigma_0 & 0 \\ 0 & 0 \end{pmatrix} E + E^* \begin{pmatrix} 0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} E \\ &= R\rho R + \begin{pmatrix} 0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix}, \end{aligned}$$

where

$$\rho = \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_0 & \alpha \\ \alpha^* & \beta \end{pmatrix}, \quad E := \begin{pmatrix} I & \sigma_0^{-1} \alpha \\ 0 & I \end{pmatrix}, \quad R = E^* \begin{pmatrix} \sigma_0 \# \rho_0^{-1} & 0 \\ 0 & I \end{pmatrix} E$$

with  $\rho_0 > 0$  and  $\sigma_0 > 0$ . Since

$$\begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix} \perp \begin{pmatrix} 0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix},$$

we have the following decomposition:

$$\sigma = \sigma^{\text{ac}} + \sigma^\perp, \tag{3.2}$$



where

$$\sigma^{\text{ac}} := R\rho R = \begin{pmatrix} \sigma_0 & \alpha \\ \alpha^* & \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} \tag{3.3}$$

is the (mutually) *absolutely continuous part* of  $\sigma$  with respect to  $\rho$ , and

$$\sigma^\perp := \begin{pmatrix} 0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} \tag{3.4}$$

is the *singular part* of  $\sigma$  with respect to  $\rho$ .

We may call the decomposition (3.2) a *quantum Lebesgue decomposition* for the following reasons. Firstly, although (3.2) was defined by using a simultaneous block matrix representation of  $\rho$  and  $\sigma$ , which has an arbitrariness of unitary transformations of the form  $U_1 \oplus U_2$ , the matrices (3.3) and (3.4) are covariant under those unitary transformations, and hence the operators  $\sigma^{\text{ac}}$  and  $\sigma^\perp$  are well-defined regardless of the arbitrariness of the block matrix representation. Secondly, the decomposition (3.2) is unique, as the following lemma asserts.

**Lemma 3.1.** *Suppose  $\sigma \gg \rho$ . Then the decomposition*

$$\sigma = \sigma^{\text{ac}} + \sigma^\perp \quad (\sigma^{\text{ac}} \ll \rho, \sigma^\perp \perp \rho) \tag{3.5}$$

is uniquely given by (3.3) and (3.4).

**Proof.** We show that the decomposition

$$\sigma = R\rho R + \tau \quad (R \geq 0, \tau \geq 0, \tau \perp \rho) \tag{3.6}$$

is unique. Let

$$\rho = \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_0 & \alpha \\ \alpha^* & \beta \end{pmatrix}$$

with  $\rho_0 > 0$ . Due to assumption  $\rho \ll \sigma$ , we have  $\sigma_0 > 0$ . Let

$$E := \begin{pmatrix} I & \sigma_0^{-1} \alpha \\ 0 & I \end{pmatrix}.$$

Since  $E$  is invertible, the operator  $R$  appeared in (3.6) is represented in the form

$$R = E^* \begin{pmatrix} R_0 & R_1 \\ R_1^* & R_2 \end{pmatrix} E.$$

With this representation,

$$\begin{aligned} R\rho R &= E^* \begin{pmatrix} R_0 & R_1 \\ R_1^* & R_2 \end{pmatrix} E \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix} E^* \begin{pmatrix} R_0 & R_1 \\ R_1^* & R_2 \end{pmatrix} E \\ &= E^* \begin{pmatrix} R_0 \rho_0 R_0 & R_0 \rho_0 R_1 \\ R_1^* \rho_0 R_0 & R_1^* \rho_0 R_1 \end{pmatrix} E \end{aligned}$$

$$\begin{aligned} &\leq \sigma \\ &= E^* \begin{pmatrix} \sigma_0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} E. \end{aligned}$$

Here, the inequality is due to (3.6). Let us denote the singular part  $\tau$  as

$$\tau = \begin{pmatrix} 0 & 0 \\ 0 & \tau_0 \end{pmatrix} = E^* \begin{pmatrix} 0 & 0 \\ 0 & \tau_0 \end{pmatrix} E.$$

Then the decomposition (3.6) is equivalent to

$$\begin{pmatrix} \sigma_0 & 0 \\ 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} = \begin{pmatrix} R_0 \rho_0 R_0 & R_0 \rho_0 R_1 \\ R_1^* \rho_0 R_0 & R_1^* \rho_0 R_1 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \tau_0 \end{pmatrix}. \tag{3.7}$$

Comparison of the (1, 1)th blocks of both sides yields  $R_0 = \sigma_0 \# \rho_0^{-1}$ . Since this  $R_0$  is strictly positive, comparison of other blocks of (3.7) further yields

$$R_1 = 0 \quad \text{and} \quad \tau_0 = \beta - \alpha^* \sigma_0^{-1} \alpha.$$

Consequently, the singular part  $\tau$  is uniquely determined by (3.4). □

An immediate consequence of Lemma 3.1 is the following.

**Corollary 3.2.** *When  $\sigma \gg \rho$ , the absolutely continuous part  $\sigma^{\text{ac}}$  of the quantum Lebesgue decomposition (3.5) is in fact mutually absolutely continuous to  $\rho$ , that is,  $\sigma^{\text{ac}} \sim \rho$ .*

Note that the operator  $R_2$  appeared in the proof of Lemma 3.1 is arbitrary as long as it is positive. Because of this arbitrariness, we can take the operator  $R$  in (3.6) to be strictly positive. This gives an alternative view of Corollary 3.2.

### 3.2. Case 2: Generic case

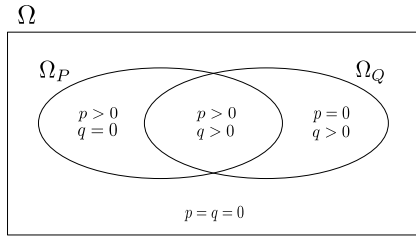
Let us extend the quantum Lebesgue decomposition (3.5) to a generic case when  $\rho$  is not necessarily absolutely continuous with respect to  $\sigma$ . When  $\rho$  and  $\sigma$  are mutually singular, we just let  $\sigma^{\text{ac}} = 0$  and  $\sigma^\perp = \sigma$ . We therefore assume in the rest of this section that  $\rho$  and  $\sigma$  are not mutually singular.

Given positive operators  $\rho$  and  $\sigma$  that satisfy  $\rho \not\perp \sigma$ , let  $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$  be the orthogonal direct sum decomposition defined by

$$\mathcal{H}_1 := \ker(\sigma \downarrow_{\text{supp } \rho}), \quad \mathcal{H}_2 := \text{supp}(\sigma \downarrow_{\text{supp } \rho}), \quad \mathcal{H}_3 := \ker \rho.$$

Then  $\rho$  and  $\sigma$  are represented in the form of block matrices as follows:

$$\rho = \begin{pmatrix} \rho_2 & \rho_1 & 0 \\ \rho_1^* & \rho_0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 & \alpha \\ 0 & \alpha^* & \beta \end{pmatrix}, \tag{3.8}$$



**Figure 1.** Schematic diagram of support sets of measures  $P$  and  $Q$  on a classical measure space  $(\Omega, \mathcal{F}, \mu)$  having densities  $p$  and  $q$ , respectively. Here,  $\Omega_P := \{\omega \in \Omega \mid p(\omega) > 0\}$  and  $\Omega_Q := \{\omega \in \Omega \mid q(\omega) > 0\}$ . The induced measures  $Q^{\text{ac}}(A) := Q(A \cap \{p > 0\})$  and  $Q^\perp(A) := Q(A \cap \{p = 0\})$  give the Lebesgue decomposition  $Q = Q^{\text{ac}} + Q^\perp$  with respect to  $P$ , in which  $Q^{\text{ac}} \ll P$  and  $Q^\perp \perp P$  (cf. [32], Chapter 6).

where

$$\begin{pmatrix} \rho_2 & \rho_1 \\ \rho_1^* & \rho_0 \end{pmatrix} > 0, \quad \sigma_0 > 0.$$

Note that when  $\sigma \gg \rho$  (Case 1), the subspace  $\mathcal{H}_1$  becomes zero; in this case, the first rows and columns in (3.8) should be ignored. Likewise, when  $\rho > 0$ , the subspace  $\mathcal{H}_3$  becomes zero; in this case, the third rows and columns in (3.8) should be ignored.

There is an obvious similarity between the block matrix structure in (3.8) and the diagram illustrated in Figure 1 that displays the support sets of two measures  $P$  and  $Q$  on a classical measure space  $(\Omega, \mathcal{F}, \mu)$  having densities  $p$  and  $q$ , respectively. However, it should be warned that

$$\mathcal{H}'_1 := \text{supp } \rho \cap \ker \sigma, \quad \mathcal{H}'_2 := \text{supp } \rho \cap \text{supp } \sigma$$

are different from  $\mathcal{H}_1$  and  $\mathcal{H}_2$ , respectively. This is most easily seen by considering the case when both  $\rho$  and  $\sigma$  are pure states: for pure states  $\rho = |\psi\rangle\langle\psi|$  and  $\sigma = |\xi\rangle\langle\xi|$ , we see that  $\mathcal{H}_2 \neq \{0\}$  if and only if  $\langle\xi|\psi\rangle \neq 0$  (cf. [34], Example 2.3), whereas  $\mathcal{H}'_2 \neq \{0\}$  if and only if  $\rho = \sigma$ .

Let us rewrite  $\sigma$  in the form

$$\sigma = E^* \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 & 0 \\ 0 & 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} E,$$

where

$$E := \begin{pmatrix} I & 0 & 0 \\ 0 & I & \sigma_0^{-1} \alpha \\ 0 & 0 & I \end{pmatrix}.$$

Since  $E$  is invertible and  $\sigma \geq 0$ , we see that

$$\beta - \alpha^* \sigma_0^{-1} \alpha \geq 0.$$

Now let

$$\sigma^{\text{ac}} := E^* \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 & 0 \\ 0 & 0 & 0 \end{pmatrix} E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 & \alpha \\ 0 & \alpha^* & \alpha^* \sigma_0^{-1} \alpha \end{pmatrix}$$

and let

$$\sigma^\perp := E^* \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix}.$$

Then it is shown that  $\sigma^{\text{ac}} \ll \rho$  and  $\sigma^\perp \perp \rho$ . In fact, the latter is obvious from Lemma 2.1. To prove the former, let

$$R := E^* \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 \# \rho_0^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix} E.$$

Then  $R$  is a positive operator satisfying

$$\begin{aligned} R\rho R &= E^* \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 \# \rho_0^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \rho_2 & \rho_1 & 0 \\ \rho_1^* & \rho_0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 \# \rho_0^{-1} & 0 \\ 0 & 0 & 0 \end{pmatrix} E \\ &= E^* \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 & 0 \\ 0 & 0 & 0 \end{pmatrix} E = \sigma^{\text{ac}}. \end{aligned}$$

It then follows from Lemma 2.2 that  $\sigma^{\text{ac}} \ll \rho$ .

In summary, given  $\rho$  and  $\sigma$  that satisfy  $\sigma \not\ll \rho$ , let

$$\rho = \begin{pmatrix} \rho_2 & \rho_1 & 0 \\ \rho_1^* & \rho_0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \sigma = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 & \alpha \\ 0 & \alpha^* & \beta \end{pmatrix} \tag{3.9}$$

be their simultaneous block matrix representations relative to the aforementioned direct sum decomposition  $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$ . Then

$$\sigma^{\text{ac}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0 & \alpha \\ 0 & \alpha^* & \alpha^* \sigma_0^{-1} \alpha \end{pmatrix}, \quad \sigma^\perp = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \beta - \alpha^* \sigma_0^{-1} \alpha \end{pmatrix} \tag{3.10}$$

give the following decomposition:

$$\sigma = \sigma^{\text{ac}} + \sigma^\perp \quad (\sigma^{\text{ac}} \ll \rho, \sigma^\perp \perp \rho) \tag{3.11}$$

with respect to  $\rho$ .

As in the previous subsection, we may call (3.11) a *quantum Lebesgue decomposition* for the following reasons. Firstly, although the simultaneous block representation (3.9) has arbitrariness of unitary transformations of the form  $U_1 \oplus U_2 \oplus U_3$ , the operators  $\sigma^{\text{ac}}$  and  $\sigma^\perp$  are well-defined because the matrices (3.10) are covariant under those unitary transformations. Secondly, the decomposition (3.11) is unique, as the following lemma asserts.

**Lemma 3.3.** *Given  $\rho$  and  $\sigma$  with  $\sigma \not\ll \rho$ , the decomposition*

$$\sigma = \sigma^{\text{ac}} + \sigma^\perp \quad (\sigma^{\text{ac}} \ll \rho, \sigma^\perp \perp \rho)$$

is uniquely given by (3.10).

**Proof.** We show that the decomposition

$$\sigma = R\rho R + \tau \quad (R \geq 0, \tau \geq 0, \tau \perp \rho) \tag{3.12}$$

is unique. Because of Lemma 3.1, it suffices to treat the case when  $\sigma \not\gg \rho$ , that is, when  $\mathcal{H}_1 \neq \{0\}$ .

Let  $\rho$  and  $\sigma$  be represented as (3.9). It then follows from (3.12) that, for any  $x \in \mathcal{H}_1$ ,

$$0 = \langle x | \sigma x \rangle \geq \langle x | R\rho R x \rangle = \langle R x | \rho R x \rangle.$$

This implies that  $Rx \in \ker \rho (= \mathcal{H}_3)$ : in particular,  $\langle x | R x \rangle = 0$ , so that the (1, 1)th block of  $R$  is zero. This fact, combined with the positivity of  $R$ , entails that  $R$  must have the form

$$R = \begin{pmatrix} 0 & 0 & 0 \\ 0 & R_0 & R_1 \\ 0 & R_1^* & R_2 \end{pmatrix}.$$

Consequently, the problem is reduced to finding the decomposition

$$\hat{\sigma} = \hat{R}\hat{\rho}\hat{R} + \hat{\tau} \quad (\hat{R} \geq 0, \hat{\tau} \geq 0, \hat{\tau} \perp \hat{\rho}), \tag{3.13}$$

where

$$\hat{\rho} = \begin{pmatrix} \rho_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \hat{\sigma} = \begin{pmatrix} \sigma_0 & \alpha \\ \alpha^* & \beta \end{pmatrix}, \quad \hat{R} = \begin{pmatrix} R_0 & R_1 \\ R_1^* & R_2 \end{pmatrix}.$$

Since  $\hat{\rho} \ll \hat{\sigma}$ , the uniqueness of the decomposition (3.13) immediately follows from Lemma 3.1. This completes the proof.  $\square$

Now that a quantum Lebesgue decomposition is established, we shall call the operator  $R$  satisfying (3.12) the *square-root likelihood ratio* of  $\sigma$  relative to  $\rho$ , and shall denote it as  $\mathcal{R}(\sigma | \rho)$ .

**Remark 3.4.** The square-root likelihood ratio  $R = \mathcal{R}(\sigma | \rho)$  is explicitly written as

$$R = \sqrt{\sigma}(\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}})^+ \sqrt{\sigma} + \gamma, \tag{3.14}$$

where  $A^+$  denotes the generalized inverse of an operator  $A$ , and  $\gamma$  is an arbitrary positive operator that is singular with respect to  $\rho$ . The proof is given in the Supplementary Material [8].

### 4. Contiguity

As we have seen in the [Introduction](#), the asymptotic version of absolute continuity called the contiguity played an important role in classical statistics [22,23,32]. In this section, we extend it to the quantum domain. There are several equivalent characterizations of the contiguity. Among others, the following characterization is particularly relevant to our purpose because it makes no use of the notion of measurable sets that are characteristic of classical measure theory. Let  $P^{(n)}$  and  $Q^{(n)}$  be sequences of probability measures on measurable spaces  $(\Omega^{(n)}, \mathcal{F}^{(n)})$ . Then  $Q^{(n)}$  is contiguous with respect to  $P^{(n)}$  if and only if the sequence  $dQ^{(n)}/dP^{(n)}$  of likelihood ratios is uniformly integrable under  $P^{(n)}$ , and  $\lim_{n \rightarrow \infty} E_{P^{(n)}}[dQ^{(n)}/dP^{(n)}] = 1$  (cf. [14], Lemma V.1.10).

Let  $\mathcal{H}^{(n)}$  be a sequence of finite dimensional Hilbert spaces, and let  $\rho^{(n)}$  and  $\sigma^{(n)}$  be quantum states on  $\mathcal{H}^{(n)}$ . Further, let  $R^{(n)}$  be (a version of) the square-root likelihood ratio  $\mathcal{R}(\sigma^{(n)} | \rho^{(n)})$ . Motivated by the above consideration, one may envisage that the sequence  $\sigma^{(n)}$  could be designated as ‘‘contiguous’’ with respect to  $\rho^{(n)}$  if:

- (i)  $\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} R^{(n)2} = 1$ , and
- (ii) the sequence  $R^{(n)2}$  is uniformly integrable under  $\rho^{(n)}$ ; that is, for any  $\varepsilon > 0$  there exist an  $M > 0$  such that

$$\sup_n \text{Tr} \rho^{(n)} R^{(n)2} (I - \mathbb{1}_M(R^{(n)})) < \varepsilon.$$

Here,  $\mathbb{1}_M$  is the truncation function:

$$\mathbb{1}_M(x) = \begin{cases} 1 & \text{if } |x| \leq M, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the operator  $\mathbb{1}_M(R^{(n)})$  is the orthogonal projection onto the subspace of  $\mathcal{H}^{(n)}$  spanned by the eigenvectors of  $R^{(n)}$  corresponding to the eigenvalues less than or equal to  $M$ .

However, such a naive definition fails, as the following example demonstrates.

**Example 4.1.** Let

$$\rho^{(n)} = \frac{1}{2n^3} \begin{pmatrix} 2n^3 - 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma^{(n)} = \frac{1}{2(n^2 + n + 1)} \begin{pmatrix} n^2 & n^2 + 1 \\ n^2 + 1 & n^2 + 2n + 2 \end{pmatrix}$$

be sequences of faithful states on a fixed Hilbert space  $\mathcal{H}^{(n)} = \mathbb{C}^2$ . For all  $n \in \mathbb{N}$ , they are mutually absolutely continuous. Moreover, the limiting states

$$\rho^{(\infty)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma^{(\infty)} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

are also mutually absolutely continuous since they are nonorthogonal pure states. Therefore, one would expect that  $\rho^{(n)}$  and  $\sigma^{(n)}$  should be contiguous. However, this does not follow from the

above naive definition. In fact, the square-root likelihood ratio  $R^{(n)} = \mathcal{R}(\sigma^{(n)} \mid \rho^{(n)})$  is uniquely given by

$$R^{(n)} = \frac{n}{\sqrt{2(n^2 + n + 1)}} \begin{pmatrix} 1 & 1 \\ 1 & 2n + 1 \end{pmatrix}.$$

Therefore, for any  $M > 1/\sqrt{2}$ ,

$$\lim_{n \rightarrow \infty} \mathbb{1}_M(R^{(n)}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$$

and

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} R^{(n)2} (I - \mathbb{1}_M(R^{(n)})) = \text{Tr} \sigma^{(\infty)} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}.$$

Namely,  $R^{(n)2}$  is not uniformly integrable under  $\rho^{(n)}$ .

The above strange phenomenon stems from the fact that the (2, 2)th entry of the square-root likelihood ratio  $R^{(n)}$  diverges as  $n \rightarrow \infty$ , although this entry is asymptotically inessential in that it corresponds to the singular part of the limiting reference state  $\rho^{(\infty)}$ . In other words, this divergence might be illusory in discussing the asymptotic behavior. This observation may lead us to a “modified” positive operator

$$\overline{R}^{(n)} = \frac{n}{\sqrt{2(n^2 + n + 1)}} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

which would contain essential information about asymptotic relationship between  $\rho^{(n)}$  and  $\sigma^{(n)}$ . In fact,

$$\overline{R}^{(n)} \rho^{(n)} \overline{R}^{(n)} = \frac{1}{2(n^2 + n + 1)} \begin{pmatrix} n^2 & n^2 \\ n^2 & n^2 \end{pmatrix}$$

approaches  $\sigma^{(\infty)}$  as  $n \rightarrow \infty$ , and the sequence  $\overline{R}^{(n)2}$  is uniformly integrable under  $\rho^{(n)}$ .

In order to formulate the idea presented in Example 4.1, we introduce a class of modifications that is asymptotically negligible. We say a sequence  $O^{(n)}$  of observables is *infinitesimal in  $L^2$*  (or simply  *$L^2$ -infinitesimal*) under  $\rho^{(n)}$ , denoted  $O^{(n)} = o_{L^2}(\rho^{(n)})$ , if

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} O^{(n)2} = 0.$$

It is easily verified that in Example 4.1, the operator  $O^{(n)} := \overline{R}^{(n)} - R^{(n)}$  is  $L^2$ -infinitesimal under  $\rho^{(n)}$ .

Now we introduce a quantum extension of the contiguity.

**Definition 4.2.** Let  $\mathcal{H}^{(n)}$  be a sequence of finite dimensional Hilbert spaces, and let  $\rho^{(n)}$  and  $\sigma^{(n)}$  be quantum states on  $\mathcal{H}^{(n)}$ . Further, let  $R^{(n)}$  be (a version of) the square-root likelihood

ratio  $\mathcal{R}(\sigma^{(n)} | \rho^{(n)})$ . The sequence  $\sigma^{(n)}$  is *contiguous* with respect to the sequence  $\rho^{(n)}$ , denoted  $\sigma^{(n)} \triangleleft \rho^{(n)}$ , if:

- (i)  $\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} R^{(n)2} = 1$ , and
- (ii) there is an  $L^2$ -infinitesimal sequence  $O^{(n)}$  of observables, each defined on  $\mathcal{H}^{(n)}$ , such that  $\bar{R}^{(n)} := R^{(n)} + O^{(n)}$  is positive and  $\bar{R}^{(n)2}$  is uniformly integrable under  $\rho^{(n)}$ .

We also use the notation  $\sigma^{(n)} \triangleleft_{O^{(n)}} \rho^{(n)}$  when  $O^{(n)}$  needs to be specified.

Several remarks are in order. Firstly, the above definition is independent of the choice of the square-root likelihood ratio  $R^{(n)}$ , since its arbitrariness (see Remark 3.4) does not affect condition (i), and is absorbed into the  $L^2$ -infinitesimal modification  $O^{(n)}$  in condition (ii). Secondly, condition (i) and the uniform integrability in (ii) can be merged into a single condition

$$\lim_{M \rightarrow \infty} \liminf_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \mathbb{1}_M(\bar{R}^{(n)}) \bar{R}^{(n)2} = 1$$

or

$$\lim_{M \rightarrow \infty} \liminf_{n \rightarrow \infty} \text{Tr} \sigma^{(n)ac} \mathbb{1}_M(\bar{R}^{(n)}) = 1.$$

Here,  $\sigma^{(n)ac} = R^{(n)} \rho^{(n)} R^{(n)}$  is the absolutely continuous part of  $\sigma^{(n)}$  with respect to  $\rho^{(n)}$ . Thirdly, the definition is unitarily covariant, in that

$$\sigma^{(n)} \triangleleft_{O^{(n)}} \rho^{(n)} \quad \text{if and only if} \quad U^{(n)} \sigma^{(n)} U^{(n)*} \triangleleft_{U^{(n)} O^{(n)} U^{(n)*}} U^{(n)} \rho^{(n)} U^{(n)*},$$

where  $U^{(n)}$  is an arbitrary unitary operator on  $\mathcal{H}^{(n)}$ . This fact could be useful in representing a state in a matrix form. Fourthly, the positivity of  $\bar{R}^{(n)}$  can be replaced with an asymptotic positivity; that is, the negative part of  $\bar{R}^{(n)}$  is  $L^2$ -infinitesimal under  $\rho^{(n)}$ . However, the positivity of  $\bar{R}^{(n)}$ , whether asymptotically or not, is indispensable as the following example illustrates.

**Example 4.3.** Let

$$\rho^{(n)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma^{(n)} = \frac{1}{1+n^2} \begin{pmatrix} 1 & n \\ n & n^2 \end{pmatrix}$$

be sequences of pure states on  $\mathcal{H}^{(n)} = \mathbb{C}^2$ . The square-root likelihood ratio  $\mathcal{R}(\sigma^{(n)} | \rho^{(n)})$  is given by

$$R^{(n)} = \frac{1}{\sqrt{1+n^2}} \begin{pmatrix} 1 & n \\ n & n^2 + \gamma \end{pmatrix},$$

where  $\gamma$  is an arbitrary nonnegative number. Now let

$$O^{(n)} = \frac{1}{\sqrt{1+n^2}} \begin{pmatrix} 0 & 0 \\ 0 & -n^2 \end{pmatrix}$$



and let  $\bar{R}^{(n)} = R^{(n)} + O^{(n)}$ . Then  $\bar{R}^{(n)}$  is uniformly bounded, and conditions (i) and (ii) in Definition 4.2, except the positivity of  $\bar{R}^{(n)}$ , are fulfilled. However, the limiting states

$$\rho^{(\infty)} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \sigma^{(\infty)} = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

are mutually singular.

The validity of Definition 4.2 is demonstrated by the following.

**Theorem 4.4.** *Let  $\rho^{(n)}$  and  $\sigma^{(n)}$  be sequences of quantum states on a fixed finite dimensional Hilbert space  $\mathcal{H}$ , and suppose that they have the limiting states  $\lim_{n \rightarrow \infty} \rho^{(n)} = \rho^{(\infty)}$  and  $\lim_{n \rightarrow \infty} \sigma^{(n)} = \sigma^{(\infty)}$ . Then  $\sigma^{(n)} \triangleleft \rho^{(n)}$  if and only if  $\sigma^{(\infty)} \ll \rho^{(\infty)}$ .*

When the reference states  $\rho^{(n)}$  are pure, there is a simple criterion for the contiguity.

**Theorem 4.5.** *Let  $\mathcal{H}^{(n)}$  be a sequence of finite dimensional Hilbert spaces, and let  $\rho^{(n)}$  and  $\sigma^{(n)}$  be quantum states on  $\mathcal{H}^{(n)}$ . Suppose that  $\rho^{(n)}$  is pure for all  $n \in \mathbb{N}$ . Then  $\sigma^{(n)} \triangleleft \rho^{(n)}$  if and only if  $\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} R^{(n)2} = 1$  and  $\liminf_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \sigma^{(n)} > 0$ , where  $R^{(n)}$  is (a version of) the square-root likelihood ratio  $\mathcal{R}(\sigma^{(n)} | \rho^{(n)})$ .*

The proofs of Theorems 4.4 and 4.5 are lengthy, and are deferred to the Supplementary Material [8].

## 5. Convergence in distribution

In this section, we introduce a quantum extension of the notion of convergence in distribution in terms of the ‘‘quasi-characteristic’’ function [16,34]. This mode of convergence turns out to be useful in asymptotic theory of quantum statistics. For a brief account of quantum Gaussian states, see Appendix A.

**Definition 5.1.** 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  $\phi$  be a normal state (represented by a linear functional) and  $X^{(\infty)} = (X_1^{(\infty)}, \dots, X_d^{(\infty)})$  be a list of observables on a possibly infinite dimensional Hilbert space  $\mathcal{H}^{(\infty)}$  such that  $\xi^i X_i^{(\infty)}$  is densely defined for every  $\xi = (\xi^i) \in \mathbb{R}^d$ . We say the sequence  $(X^{(n)}, \rho^{(n)})$  converges in distribution to  $(X^{(\infty)}, \phi)$ , in symbols

$$(X^{(n)}, \rho^{(n)}) \rightsquigarrow (X^{(\infty)}, \phi),$$

if

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \left( \prod_{i=1}^r e^{\sqrt{-1} \xi_i^i X_i^{(n)}} \right) = \phi \left( \prod_{i=1}^r e^{\sqrt{-1} \xi_i^i X_i^{(\infty)}} \right)$$

holds for any  $r \in \mathbb{N}$  and subset  $\{\xi_t\}_{t=1}^r$  of  $\mathbb{R}^d$ . When the limiting state  $\phi$  is a quantum Gaussian state, in that  $(X^{(\infty)}, \phi) \sim N(h, J)$ , we also use the abridged notation

$$X^{(n)} \overset{\rho^{(n)}}{\rightsquigarrow} N(h, J),$$

in accordance with the convention in classical statistics.

A slight generalization is the following mode of convergence, which plays an essential role in the present paper.

**Definition 5.2.** In addition to the setting for Definition 5.1, let  $Y^{(n)}$  and  $Y^{(\infty)}$  be observables on  $\mathcal{H}^{(n)}$  and  $\mathcal{H}^{(\infty)}$ , respectively, with  $Y^{(\infty)}$  being densely defined. If

$$\begin{aligned} & \lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} e^{\sqrt{-1}\eta_1 Y^{(n)}} \left\{ \prod_{t=1}^r e^{\sqrt{-1}\xi_t^i X_i^{(n)}} \right\} e^{\sqrt{-1}\eta_2 Y^{(n)}} \\ &= \phi \left( e^{\sqrt{-1}\eta_1 Y^{(\infty)}} \left\{ \prod_{t=1}^r e^{\sqrt{-1}\xi_t^i X_i^{(\infty)}} \right\} e^{\sqrt{-1}\eta_2 Y^{(\infty)}} \right) \end{aligned}$$

holds for any  $r \in \mathbb{N}$ , subset  $\{\xi_t\}_{t=1}^r$  of  $\mathbb{R}^d$ , and  $\eta_1, \eta_2 \in \mathbb{R}$ , then we denote

$$\langle \langle Y^{(n)}, X^{(n)}, Y^{(n)} \rangle, \rho^{(n)} \rangle \rightsquigarrow \langle \langle Y^{(\infty)}, X^{(\infty)}, Y^{(\infty)} \rangle, \phi \rangle$$

or

$$\langle Y^{(n)}, X^{(n)}, Y^{(n)} \rangle_{\rho^{(n)}} \rightsquigarrow \langle Y^{(\infty)}, X^{(\infty)}, Y^{(\infty)} \rangle_{\phi}.$$

We shall call this type of convergence a *sandwiched convergence in distribution* to emphasize that the observables  $Y^{(n)}$  and  $Y^{(\infty)}$  that appear at both ends of the quasi-characteristic function play special roles.

The sandwiched convergence in distribution will be used in conjunction with the following form of the quantum Lévy–Cramér continuity theorem.

**Lemma 5.3.** Let  $(X^{(n)}, Y^{(n)}, \rho^{(n)})$  and  $(X^{(\infty)}, Y^{(\infty)}, \phi)$  be as in Definition 5.2. If

$$\langle Y^{(n)}, X^{(n)}, Y^{(n)} \rangle_{\rho^{(n)}} \rightsquigarrow \langle Y^{(\infty)}, X^{(\infty)}, Y^{(\infty)} \rangle_{\phi},$$

then

$$\begin{aligned} & \lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} g_1(Y^{(n)}) \left\{ \prod_{t=1}^r f_t(\xi_t^i X_i^{(n)}) \right\} g_2(Y^{(n)}) \\ &= \phi \left( g_1(Y^{(\infty)}) \left\{ \prod_{t=1}^r f_t(\xi_t^i X_i^{(\infty)}) \right\} g_2(Y^{(\infty)}) \right) \end{aligned} \tag{5.1}$$

holds for any  $r \in \mathbb{N}$ , subset  $\{\xi_t\}_{t=1}^r$  of  $\mathbb{R}^d$ , bounded continuous functions  $f_1, \dots, f_r$ , and bounded Borel functions  $g_1, g_2$  on  $\mathbb{R}$  such that the set  $\mathcal{D}(g_i)$  of discontinuity points of  $g_i$  has  $\mu$ -measure zero for  $i = 1, 2$ , where  $\mu$  is the classical probability measure on  $\mathbb{R}$  having the characteristic function  $\varphi_\mu(\eta) := \phi(e^{\sqrt{-1}\eta} Y^{(\infty)})$ .

**Proof.** Let  $s := r + 2$ , and let  $J$  be an arbitrary natural number between 1 and  $s - 1$  (say  $J = 1$ ). Then the list of observables

$$Z^{(n)} = (Z_1^{(n)}, \dots, Z_s^{(n)}) := (Y^{(n)}, \xi_1^i X_i^{(n)}, \dots, \xi_r^i X_i^{(n)}, Y^{(n)})$$

fulfils conditions (B.3), (B.4) and (B.5) in the quantum Lévy–Cramér continuity Theorem B.1 cited in Appendix B. Furthermore, the functions  $g_1$  and  $g_2$  satisfy condition (B.6) in the theorem. Thus the claim is an immediate consequence of Theorem B.1.  $\square$

In classical statistics, if random variables  $X^{(n)}$  converge in distribution to a random variable  $X$ , and random variables  $O^{(n)}$  converge in  $L^2$  (and hence in probability) to 0, then  $X^{(n)} + O^{(n)}$  converge in distribution to  $X$  [32], Lemma 2.8. However, its obvious analogue in quantum statistics fails to be true, as the following example illustrates.

**Example 5.4.** Let

$$\rho^{(n)} := \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad X^{(n)} := \begin{pmatrix} 1 & n \\ n & 1 + n^2 \end{pmatrix}, \quad O^{(n)} := \begin{pmatrix} 0 & 0 \\ 0 & -n^2 \end{pmatrix}.$$

It is not difficult to verify that

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} e^{\sqrt{-1}\xi X^{(n)}} = 1$$

for all  $\xi \in \mathbb{R}$ , and  $O^{(n)} = o_{L^2}(\rho^{(n)})$ . However,

$$\text{Tr} \rho^{(n)} e^{\sqrt{-1}\xi(X^{(n)} + O^{(n)})} = e^{\sqrt{-1}\xi} \cos n\xi,$$

which has no limit as  $n \rightarrow \infty$ .

The above example shows that an  $L^2$ -infinitesimal sequence of observables is not always negligible in quasi-characteristic functions. We therefore introduce another class of infinitesimal objects pertinent to the convergence in distribution.

**Definition 5.5.** Let  $\mathcal{H}^{(n)}$  be a sequence of finite dimensional Hilbert spaces, and let  $Z^{(n)}$  and  $\rho^{(n)}$  be an observable and a state on  $\mathcal{H}^{(n)}$ . We say a sequence  $O^{(n)}$  of observables, each defined on  $\mathcal{H}^{(n)}$ , is *infinitesimal in distribution* (or simply *D-infinitesimal*) with respect to  $(Z^{(n)}, \rho^{(n)})$ , denoted  $O^{(n)} = o_D(Z^{(n)}, \rho^{(n)})$ , if

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \left\{ \prod_{t=1}^r e^{\sqrt{-1}(\xi_t Z^{(n)} + \eta_t O^{(n)})} \right\} = \lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \left\{ \prod_{t=1}^r e^{\sqrt{-1}\xi_t Z^{(n)}} \right\} \tag{5.2}$$

holds for any  $r \in \mathbb{N}$ , and subsets  $\{\xi_t\}_{t=1}^r$  and  $\{\eta_t\}_{t=1}^r$  of  $\mathbb{R}$ .

The following lemma asserts that a D-infinitesimal sequence is negligible in the sandwiched convergence.

**Lemma 5.6.** *If  $\langle Z^{(n)}, X^{(n)}, Z^{(n)} \rangle \overset{\rho^{(n)}}{\rightsquigarrow} \langle Z^{(\infty)}, X^{(\infty)}, Z^{(\infty)} \rangle$  and  $O^{(n)} = o_D(Z^{(n)}, \rho^{(n)})$  then*

$$\langle Z^{(n)} + O^{(n)}, X^{(n)}, Z^{(n)} + O^{(n)} \rangle \overset{\rho^{(n)}}{\rightsquigarrow} \langle Z^{(\infty)}, X^{(\infty)}, Z^{(\infty)} \rangle.$$

The proof of Lemma 5.6 is straightforward, and is deferred to the Supplementary Material [8].

## 6. LeCam’s third lemma

We are now ready to extend LeCam’s third lemma to the quantum domain. Our first result is the following abstract version of LeCam’s third lemma, a noncommutative analogue of [32], Theorem 6.6.

**Theorem 6.1.** *Given a sequence  $\mathcal{H}^{(n)}$  of finite dimensional Hilbert spaces, let  $\rho^{(n)}$  and  $\sigma^{(n)}$  be quantum states and let  $X^{(n)} = (X_1^{(n)}, \dots, X_d^{(n)})$  be a list of observables on  $\mathcal{H}^{(n)}$ . Further, let  $R^{(n)}$  be (a version of) the square-root likelihood ratio  $\mathcal{R}(\sigma^{(n)} | \rho^{(n)})$ . Suppose that:*

- (i) *there exists an  $L^2$ -infinitesimal sequence  $O^{(n)}$  of observables such that  $\sigma^{(n)} \triangleleft_{O^{(n)}} \rho^{(n)}$ , and*
- (ii) *there exist a normal state  $\phi$ , a list of observables  $X^{(\infty)} = (X_1^{(\infty)}, \dots, X_d^{(\infty)})$ , and a positive observable  $R^{(\infty)}$  on a possibly infinite dimensional Hilbert space  $\mathcal{H}^{(\infty)}$  such that*

$$\langle R^{(n)} + O^{(n)}, X^{(n)}, R^{(n)} + O^{(n)} \rangle_{\rho^{(n)}} \rightsquigarrow \langle R^{(\infty)}, X^{(\infty)}, R^{(\infty)} \rangle_{\phi}.$$

Then

$$\langle X^{(n)}, \sigma^{(n)} \rangle \rightsquigarrow \langle X^{(\infty)}, \psi \rangle,$$

where  $\psi$  is a normal state on  $\mathcal{H}^{(\infty)}$  defined by

$$\psi(A) := \phi(R^{(\infty)} A R^{(\infty)}) \tag{6.1}$$

for bounded operators  $A \in \mathcal{B}(\mathcal{H}^{(\infty)})$ .

In order to get a better understanding of Theorem 6.1, we give an informal interpretation. Let

$$(\sigma^{(n)})^{\text{ac}} = R^{(n)} \rho^{(n)} R^{(n)}$$

be the absolutely continuous part of  $\sigma^{(n)}$  with respect to  $\rho^{(n)}$ . Then, thanks to the contiguity (i) and the sandwiched convergence in distribution (ii), the absolutely continuous part  $(\sigma^{(n)})^{\text{ac}}$  converges (in a certain sense) to a density operator

$$\sigma^{(\infty)} := R^{(\infty)} \rho^{(\infty)} R^{(\infty)}$$

on  $\mathcal{H}^{(\infty)}$ , where  $\rho^{(\infty)}$  is the density operator of  $\phi$ , so that

$$\text{Tr } \sigma^{(\infty)} A = \text{Tr} (R^{(\infty)} \rho^{(\infty)} R^{(\infty)}) A = \text{Tr } \rho^{(\infty)} (R^{(\infty)} A R^{(\infty)}).$$

Letting  $\psi(A) := \text{Tr } \sigma^{(\infty)} A$ , we have (6.1). The proof of Theorem 6.1 is slightly complicated, and is deferred to the Supplementary Material [8].

A crucial application of Theorem 6.1 is the following theorem, which is a natural quantum counterpart of the standard LeCam third lemma [32], Example 6.7.

**Theorem 6.2 (Quantum LeCam third lemma).** *Given a sequence  $\mathcal{H}^{(n)}$  of finite dimensional Hilbert spaces, let  $\rho^{(n)}$  and  $\sigma^{(n)}$  be quantum states and let  $X^{(n)} = (X_1^{(n)}, \dots, X_d^{(n)})$  be a list of observables on  $\mathcal{H}^{(n)}$ . Further, let  $R^{(n)}$  be (a version of) the square-root likelihood ratio  $\mathcal{R}(\rho^{(n)} | \sigma^{(n)})$ . Suppose that there exist a sequence  $O^{(n)} = o_{L^2}(\rho^{(n)})$  satisfying  $R^{(n)} + O^{(n)} > 0$ , and a sequence  $\tilde{O}^{(n)} = o_D(\log(R^{(n)} + O^{(n)}), \rho^{(n)})$  satisfying*

$$\left( 2 \log(R^{(n)} + O^{(n)}) - \tilde{O}^{(n)} \right) \overset{\rho^{(n)}}{\rightsquigarrow} N \left( \begin{pmatrix} \mu \\ -\frac{1}{2}s^2 \end{pmatrix}, \begin{pmatrix} \Sigma & \kappa \\ \kappa^* & s^2 \end{pmatrix} \right). \tag{6.2}$$

Here,  $\mu \in \mathbb{R}^d$ ,  $s \in \mathbb{R}$ ,  $\kappa \in \mathbb{C}^d$  and  $\Sigma$  is a  $d \times d$  complex Hermitian positive semidefinite matrix. Then

$$\sigma^{(n)} \triangleleft \rho^{(n)} \tag{6.3}$$

and

$$X^{(n)} \overset{\sigma^{(n)}}{\rightsquigarrow} N(\mu + \text{Re}(\kappa), \Sigma). \tag{6.4}$$

**Proof.** Let  $(X_1, \dots, X_d, L)$  be the defining canonical observables of the algebra

$$\text{CCR} \left( \text{Im} \begin{pmatrix} \Sigma & \kappa \\ \kappa^* & s^2 \end{pmatrix} \right),$$

and let

$$\phi \sim N \left( \begin{pmatrix} \mu \\ -\frac{1}{2}s^2 \end{pmatrix}, \begin{pmatrix} \Sigma & \kappa \\ \kappa^* & s^2 \end{pmatrix} \right).$$

Further, let  $\bar{R}^{(n)} := R^{(n)} + O^{(n)}$ , and let  $L^{(n)} := 2 \log(\bar{R}^{(n)})$ . It then follows from (6.2) that

$$\langle L^{(n)} - \tilde{O}^{(n)}, X^{(n)}, L^{(n)} - \tilde{O}^{(n)} \rangle_{\rho^{(n)}} \rightsquigarrow \langle L, X, L \rangle_{\phi}.$$

With Lemma 5.6, this implies that

$$\langle L^{(n)}, X^{(n)}, L^{(n)} \rangle_{\rho^{(n)}} \rightsquigarrow \langle L, X, L \rangle_{\phi}. \tag{6.5}$$

Let us introduce a complex-valued bounded continuous function

$$f_\eta(x) := \exp\left[\sqrt{-1}\eta\left\{\exp\left(\frac{x}{2}\right)\right\}\right]$$

on  $\mathbb{R}$  having a real parameter  $\eta \in \mathbb{R}$ . It then follows from (6.5) and the sandwiched version of the quantum Lévy–Cramér continuity theorem (Lemma 5.3) that

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} f_{\eta_1}(L^{(n)}) \left\{ \prod_{i=1}^r e^{\sqrt{-1}\xi_i^i X_i^{(n)}} \right\} f_{\eta_2}(L^{(n)}) = \phi\left(f_{\eta_1}(L) \left\{ \prod_{i=1}^r e^{\sqrt{-1}\xi_i^i X_i} \right\} f_{\eta_2}(L)\right),$$

where  $\eta_1, \eta_2 \in \mathbb{R}$ . This equality is rewritten as

$$\langle e^{\frac{1}{2}L^{(n)}}, X^{(n)}, e^{\frac{1}{2}L^{(n)}} \rangle_{\rho^{(n)}} \rightsquigarrow \langle e^{\frac{1}{2}L}, X, e^{\frac{1}{2}L} \rangle_\phi,$$

or equivalently,

$$\langle \bar{R}^{(n)}, X^{(n)}, \bar{R}^{(n)} \rangle_{\rho^{(n)}} \rightsquigarrow \langle e^{\frac{1}{2}L}, X, e^{\frac{1}{2}L} \rangle_\phi.$$

Specifically,  $\bar{R}^{(n)} \stackrel{\rho^{(n)}}{\rightsquigarrow} e^{\frac{1}{2}L}$ , and Lemma 5.3 leads to

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \mathbb{1}_M(\bar{R}^{(n)}) \bar{R}^{(n)2} = \phi(\mathbb{1}_M(e^{\frac{1}{2}L})e^L) = E[\mathbb{1}_M(e^{\frac{1}{2}Z})e^Z],$$

where  $Z$  is a classical random variable that obeys the normal distribution  $N(-\frac{1}{2}s^2, s^2)$ , and the right-hand side converges to  $E[e^Z] = 1$  as  $M \rightarrow \infty$ . This implies that  $\sigma^{(n)} \triangleleft \rho^{(n)}$ , proving (6.3).

To prove (6.4), we need only evaluate the quasi-characteristic function of the state  $\psi$  defined by (6.1), that is,

$$\psi\left(\prod_{i=1}^r e^{\sqrt{-1}\xi_i^i X_i}\right) = \phi\left(e^{\frac{1}{2}L} \left\{ \prod_{i=1}^r e^{\sqrt{-1}\xi_i^i X_i} \right\} e^{\frac{1}{2}L}\right).$$

In calculating this function, it is convenient to introduce the following enlarged vectors and matrices:

$$\begin{aligned} \tilde{\mu} &:= \begin{pmatrix} \mu \\ 1 \\ -\frac{1}{2}s^2 \end{pmatrix}, & \tilde{\Sigma} &:= \begin{pmatrix} \Sigma & \kappa \\ \kappa^* & s^2 \end{pmatrix}, \\ \tilde{\xi}_0 = \tilde{\xi}_{r+1} &:= \begin{pmatrix} 0 \\ \sqrt{-1} \\ -\frac{1}{2} \end{pmatrix}, & \tilde{\xi}_t &:= \begin{pmatrix} \xi_t \\ 0 \end{pmatrix} \quad (1 \leq t \leq r). \end{aligned}$$

Then by using the quasi-characteristic function (A.1) of the quantum Gaussian state  $\phi$ , we have

$$\begin{aligned} &\psi\left(\prod_{i=1}^r e^{\sqrt{-1}\xi_i^i X_i}\right) \\ &= \phi\left(e^{\sqrt{-1}(-\frac{\sqrt{-1}}{2})L} \left\{ \prod_{i=1}^r e^{\sqrt{-1}\xi_i^i X_i} \right\} e^{\sqrt{-1}(-\frac{\sqrt{-1}}{2})L}\right) \\ &= \exp\left[\sum_{i=0}^{r+1} \left(\sqrt{-1}\tilde{\xi}_i^i \tilde{\mu}_i - \frac{1}{2} \sum_{i=0}^{r+1} \tilde{\xi}_i^i \tilde{\xi}_i^j \tilde{\Sigma}_{ji}\right) - \sum_{i=0}^{r+1} \sum_{u=i+1}^{r+1} \tilde{\xi}_i^i \tilde{\xi}_u^j \tilde{\Sigma}_{ji}\right] \\ &= \exp\left[\sum_{i=1}^r \left(\sqrt{-1}\xi_i^i (\mu_i + \text{Re}(\kappa_i)) - \frac{1}{2} \xi_i^i \xi_i^j \Sigma_{ji}\right) - \sum_{i=1}^r \sum_{s=i+1}^r \xi_i^i \xi_s^j \Sigma_{ji}\right]. \end{aligned}$$

This is identical to the quasi-characteristic function of the quantum Gaussian state  $N(\mu + \text{Re}(\kappa), \Sigma)$ . The assertion (6.4) now follows immediately from Theorem 6.1.  $\square$

## 7. Applications

In this section, we present three examples to demonstrate the validity, flexibility and applicability of our theory.

### 7.1. Contiguity without absolute continuity

For each  $n \in \mathbb{N}$ , let us consider quantum states

$$\rho^{(n)} = \begin{pmatrix} \rho_2^{(n)} & \rho_1^{(n)} & 0 \\ \rho_1^{(n)*} & \rho_0^{(n)} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \sigma^{(n)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0^{(n)} & \sigma_1^{(n)} \\ 0 & \sigma_1^{(n)*} & \sigma_2^{(n)} \end{pmatrix}$$

on  $\mathcal{H}^{(n)} \simeq \mathbb{C}^{2n+2}$ , where

$$\begin{aligned} \rho_0^{(n)} &= \frac{1}{4n^3} \begin{pmatrix} 2n^3 - 1 & 0 \\ 0 & 1 \end{pmatrix}, & \sigma_0^{(n)} &= \frac{1 - 1/(2n)}{2(n^2 + n + 1)} \begin{pmatrix} n^2 & n^2 + 1 \\ n^2 + 1 & n^2 + 2n + 2 \end{pmatrix}, \\ \rho_1^{(n)*} &= \frac{1}{(n+1)^3} \begin{pmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{pmatrix}, & \sigma_1^{(n)} &= \frac{1}{(n+1)^3} \begin{pmatrix} 1 & \cdots & 1 \\ 1 & \cdots & 1 \end{pmatrix}, \end{aligned}$$

and

$$\rho_2^{(n)} = \frac{1}{2n} I_n, \quad \sigma_2^{(n)} = \frac{1}{2n^2} I_n,$$

with  $I_n$  the  $n \times n$  identity matrix. Note that, for all  $n \in \mathbb{N}$ ,  $\sigma^{(n)}$  is not absolutely continuous to  $\rho^{(n)}$  because the singular part

$$\sigma^{(n)\perp} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \sigma_2^{(n)} - \sigma_1^{(n)*} \sigma_0^{(n)-1} \sigma_1^{(n)} \end{pmatrix}$$

is nonzero. However,  $\sigma^{(n)}$  is ‘‘asymptotically’’ absolutely continuous to  $\rho^{(n)}$  in that  $\lim_{n \rightarrow \infty} \sigma^{(n)\perp} = 0$ . Furthermore, the (2, 2)th blocks  $\rho_0^{(n)}$  and  $\sigma_0^{(n)}$  are identical, up to scaling, to the states studied in Example 4.1. Therefore, it is expected that  $\sigma^{(n)}$  would be contiguous to  $\rho^{(n)}$ . This expectation is justified by the following more general assertion.

**Theorem 7.1.** *For each  $n \in \mathbb{N}$ , let*

$$\rho^{(n)} = \begin{pmatrix} \rho_2^{(n)} & \rho_1^{(n)} & 0 \\ \rho_1^{(n)*} & \rho_0^{(n)} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \sigma^{(n)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_0^{(n)} & \sigma_1^{(n)} \\ 0 & \sigma_1^{(n)*} & \sigma_2^{(n)} \end{pmatrix}$$

*be quantum states on a Hilbert space  $\mathcal{H}^{(n)}$  represented by block matrices, where*

$$\begin{pmatrix} \rho_2^{(n)} & \rho_1^{(n)} \\ \rho_1^{(n)*} & \rho_0^{(n)} \end{pmatrix} > 0, \quad \begin{pmatrix} \sigma_0^{(n)} & \sigma_1^{(n)} \\ \sigma_1^{(n)*} & \sigma_2^{(n)} \end{pmatrix} > 0.$$

*Suppose that*

$$\liminf_{n \rightarrow \infty} \text{Tr} \rho_0^{(n)} > 0, \quad \lim_{n \rightarrow \infty} \text{Tr} \sigma_0^{(n)} = 1,$$

*and*

$$\frac{\rho_0^{(n)}}{\text{Tr} \rho_0^{(n)}} \triangleright \frac{\sigma_0^{(n)}}{\text{Tr} \sigma_0^{(n)}}.$$

*Then we have  $\rho^{(n)} \triangleright \sigma^{(n)}$ .*

The proof of Theorem 7.1 is deferred to the Supplementary Material [8].

### 7.2. Contiguity for tensor product states

Let us consider tensor product states

$$\rho^{(n)} := \bigotimes_{i=1}^n \rho_i, \quad \sigma^{(n)} := \bigotimes_{i=1}^n \sigma_i,$$



where  $\rho_i$  and  $\sigma_i$  are quantum states on a finite dimensional Hilbert space  $\mathcal{H}_i$ . Suppose that  $\sigma_i \ll \rho_i$  for all  $i$ . Then  $\sigma^{(n)} \ll \rho^{(n)}$  for all  $n \in \mathbb{N}$ . It is thus natural to enquire whether or not  $\sigma^{(n)}$  is contiguous with respect to  $\rho^{(n)}$ . The answer is given by the following.

**Theorem 7.2.** *Let  $\rho_i$  and  $\sigma_i$  be quantum states on a finite dimensional Hilbert space  $\mathcal{H}_i$  that satisfy  $\sigma_i \ll \rho_i$ , and let*

$$\rho^{(n)} := \bigotimes_{i=1}^n \rho_i, \quad \sigma^{(n)} := \bigotimes_{i=1}^n \sigma_i.$$

Then  $\sigma^{(n)} \triangleleft \rho^{(n)}$  if and only if

$$\prod_{i=1}^{\infty} \text{Tr} \rho_i R_i > 0, \tag{7.1}$$

or equivalently

$$\sum_{i=1}^{\infty} (1 - \text{Tr} \rho_i R_i) < \infty, \tag{7.2}$$

where  $R_i$  is (a version of) the square-root likelihood ratio  $\mathcal{R}(\sigma_i | \rho_i)$ .

The proof of Theorem 7.2 is deferred to the Supplementary Material [8].

**Remark 7.3.** Theorem 7.2 bears obvious similarities to Kakutani’s theorem for infinite product measures [19,33] and its noncommutative extension due to Bures [3]. In fact, by using Remark 3.4, conditions (7.1) and (7.2) are rewritten as

$$\prod_{i=1}^{\infty} \text{Tr} \sqrt{\sqrt{\sigma_i} \rho_i \sqrt{\sigma_i}} > 0 \quad \text{and} \quad \sum_{i=1}^{\infty} (1 - \text{Tr} \sqrt{\sqrt{\sigma_i} \rho_i \sqrt{\sigma_i}}) < \infty.$$

The summand in the latter condition is identical, up to a factor of 2, to the square of the Bures distance between  $\rho_i$  and  $\sigma_i$ . The main difference is that we are dealing with sequences of finite tensor product states rather than infinite tensor product states.

Let us give a simple example that demonstrates the criterion established in Theorem 7.2. Let

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_t = \frac{1}{4t^2 + 2} \begin{pmatrix} 2t^2 + 2t + 1 & 2t \\ 2t & 2t^2 - 2t + 1 \end{pmatrix},$$

where  $t$  is a parameter with  $t \geq 1$ , and let us consider three sequences of tensor product states:

$$\rho^{(n)} := \bigotimes_{i=1}^n \rho, \quad \sigma^{(n)} := \bigotimes_{i=1}^n \sigma_i, \quad \tilde{\sigma}^{(n)} := \bigotimes_{i=1}^n \sigma_{\sqrt{i}}.$$

Since  $\sigma_t \rightarrow \rho$  as  $t \rightarrow \infty$ , it is meaningful to enquire whether or not  $\sigma^{(n)}$  and  $\tilde{\sigma}^{(n)}$  are contiguous to  $\rho^{(n)}$ . As a matter of fact,  $\sigma^{(n)}$  is contiguous to  $\rho^{(n)}$ , whereas  $\tilde{\sigma}^{(n)}$  is not; this is proved as follows. The square-root likelihood ratio  $R_t = \mathcal{R}(\sigma_t | \rho)$  is

$$R_t = \frac{1}{\sqrt{4t^2 + 2}} \begin{pmatrix} 2t + 1 & 1 \\ 1 & 2t - 1 \end{pmatrix},$$

and thus

$$\text{Tr } \rho R_t = \sqrt{\frac{2t^2}{2t^2 + 1}}.$$

In view of the criterion (7.2), it suffices to verify that

$$\sum_{n=1}^{\infty} \left( 1 - \sqrt{\frac{2n^2}{2n^2 + 1}} \right) < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \left( 1 - \sqrt{\frac{2n}{2n + 1}} \right) = \infty,$$

and this is elementary. These results could be paraphrased by saying that the sequence  $\sigma_n$  converges to  $\rho$  quickly enough for  $\sigma^{(n)}$  to be contiguous with respect to  $\rho^{(n)}$ , whereas the sequence  $\sigma_{\sqrt{n}}$  does not.

### 7.3. Local asymptotic normality

In [34], we formulated a direct analogue of the weak LAN in the quantum domain. However, that formulation was not fully satisfactory because it was applicable only to quantum statistical models that comprise mutually absolutely continuous density operators. Here, we enlarge the scope of weak q-LAN to a much wider class of models by taking advantage of the quantum Lebesgue decomposition and quantum contiguity.

**Definition 7.4.** For each  $n \in \mathbb{N}$ , let  $\mathcal{S}^{(n)} = \{\rho_{\theta}^{(n)} \mid \theta \in \Theta \subset \mathbb{R}^d\}$  be a  $d$ -dimensional quantum statistical model on a finite dimensional Hilbert space  $\mathcal{H}^{(n)}$ , where  $\Theta$  is an open set. We say  $\mathcal{S}^{(n)}$  is *locally asymptotically normal* at  $\theta_0 \in \Theta$  if:

- (i) there exist a list  $\Delta^{(n)} = (\Delta_1^{(n)}, \dots, \Delta_d^{(n)})$  of observables on each  $\mathcal{H}^{(n)}$  that satisfies

$$\Delta^{(n)} \overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N(0, J), \tag{7.3}$$

where  $J$  is a  $d \times d$  Hermitian positive semidefinite matrix with  $\text{Re } J > 0$ , and

- (ii) the square-root likelihood ratio  $R_h^{(n)} = \mathcal{R}(\rho_{\theta_0+h/\sqrt{n}}^{(n)} | \rho_{\theta_0}^{(n)})$  is expanded in  $h \in \mathbb{R}^d$  as

$$R_h^{(n)} = \exp \left\{ \frac{1}{2} \left( h^i \Delta_i^{(n)} - \frac{1}{2} (J_{ij} h^i h^j) I^{(n)} + o_D(h^i \Delta_i^{(n)}, \rho_{\theta_0}^{(n)}) \right) \right\} - o_{L^2}(\rho_{\theta_0}^{(n)}),$$

where  $I^{(n)}$  is the identity operator on  $\mathcal{H}^{(n)}$ .

Note that, in contrast to the previous paper [34], we here define the local asymptotic normality in terms of the square-root likelihood ratio rather than the log-likelihood ratio; in particular, we do not assume that  $\rho_{\theta}^{(n)}$  is mutually absolutely continuous with respect to  $\rho_{\theta_0}^{(n)}$ . Moreover, the present definition is pertinent to the setting for the quantum LeCam third lemma (Theorem 6.2). In fact, we have the following.

**Corollary 7.5 (Quantum LeCam third lemma under q-LAN).** *Let  $\mathcal{S}^{(n)}$  be as in Definition 7.4, and let  $X^{(n)} = (X_1^{(n)}, \dots, X_{d'}^{(n)})$  be a list of observables on  $\mathcal{H}^{(n)}$ . Suppose that  $\mathcal{S}^{(n)}$  is locally asymptotically normal at  $\theta_0 \in \Theta$  and*

$$\begin{pmatrix} X^{(n)} \\ \Delta^{(n)} \end{pmatrix} \overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma & \tau \\ \tau^* & J \end{pmatrix} \right). \tag{7.4}$$

Here,  $\Sigma$  and  $J$  are Hermitian positive semidefinite matrices of size  $d' \times d'$  and  $d \times d$ , respectively, with  $\text{Re } J > 0$ , and  $\tau$  is a complex matrix of size  $d' \times d$ . Then

$$\rho_{\theta_0+h/\sqrt{n}}^{(n)} \triangleleft \rho_{\theta_0}^{(n)} \quad \text{and} \quad X^{(n)} \overset{\rho_{\theta_0+h/\sqrt{n}}^{(n)}}{\rightsquigarrow} N((\text{Re } \tau)h, \Sigma) \tag{7.5}$$

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

**Proof.** From the definition of q-LAN, the square-root likelihood ratio is written as

$$R_h^{(n)} = \exp \left\{ \frac{1}{2} \left( h^i \Delta_i^{(n)} - \frac{1}{2} J_{ij} h^i h^j I^{(n)} + \tilde{O}^{(n)} \right) \right\} - O^{(n)},$$

where  $\tilde{O}^{(n)} = o_D(h^i \Delta_i^{(n)}, \rho_{\theta_0}^{(n)})$  and  $O^{(n)} = o_{L^2}(\rho_{\theta_0}^{(n)})$ . Let

$$L^{(n)} := 2 \log(R_h^{(n)} + O^{(n)}) - \tilde{O}^{(n)} = h^i \Delta_i^{(n)} - \frac{1}{2} J_{ij} h^i h^j I^{(n)}.$$

Then (7.4) implies that

$$\begin{pmatrix} X^{(n)} \\ L^{(n)} \end{pmatrix} \overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N \left( \begin{pmatrix} 0 \\ -\frac{1}{2} {}^t h J h \end{pmatrix}, \begin{pmatrix} \Sigma & \tau h \\ (\tau h)^* & {}^t h J h \end{pmatrix} \right).$$

Thus, (7.5) immediately follows from Theorem 6.2. □

A prototype of Corollary 7.5 first appeared in [34], Theorem 2.9, under the assumptions that each model  $\mathcal{S}^{(n)}$  comprised mutually absolutely continuous density operators and the pairs  $(\mathcal{S}^{(n)}, X^{(n)})$  were jointly q-LAN. In contrast, Corollary 7.5 makes no use of such restrictive assumptions, and is a straightforward consequence of a much general result (Theorem 6.2). This is a notable achievement, demonstrating the advantages and usefulness of the present formulation based on the quantum Lebesgue decomposition and contiguity.

A remark on the quantum convergence in distribution is now in order. In classical statistics, marginal distributions for several random variables do not uniquely specify their joint distributions. In quantum statistics, a similar but different type of arbitrariness occurs. For example, when  $d = 1$ , the expression  $N(0, J)$  in the right-hand side of (7.3) represents a ‘‘classical’’ Gaussian state. Nevertheless, this does not imply that the limit state is classical, as demonstrated by (7.4) and (7.5) that identify the joint limiting distribution for the sequence  $(X^{(n)}, \Delta^{(n)})$ . This issue will be discussed afterwards with a concrete example.

Now we restrict ourselves to the i.i.d case. In classical statistics, it is known that the i.i.d. extension of a model  $\{P_\theta \mid \theta \in \Theta \subset \mathbb{R}^d\}$  on a measure space  $(\Omega, \mathcal{F}, \mu)$  having densities  $p_\theta$  with respect to  $\mu$  is LAN at  $\theta_0$  if the model is *differentiable in quadratic mean* at  $\theta_0$  [32], p. 93, that is, if there are random variables  $\ell_1, \dots, \ell_d$  that satisfy

$$\int_{\Omega} \left[ \sqrt{p_{\theta_0+h}} - \sqrt{p_{\theta_0}} - \frac{1}{2} h^i \ell_i \sqrt{p_{\theta_0}} \right]^2 d\mu = o(\|h\|^2)$$

as  $h \rightarrow 0$ . This condition is rewritten as

$$\int_{\Omega} \left[ \sqrt{\frac{p_{\theta_0+h}^{\text{ac}}}{p_{\theta_0}}} - 1 - \frac{1}{2} h^i \ell_i \right]^2 p_{\theta_0} d\mu + \int_{\Omega} p_{\theta_0+h}^{\perp} d\mu = o(\|h\|^2), \tag{7.6}$$

where

$$p_{\theta_0+h}^{\text{ac}}(\omega) := \begin{cases} p_{\theta_0+h}(\omega), & \omega \in \Omega_0, \\ 0, & \omega \notin \Omega_0, \end{cases}$$

and

$$p_{\theta_0+h}^{\perp}(\omega) := \begin{cases} 0, & \omega \in \Omega_0, \\ p_{\theta_0+h}(\omega), & \omega \notin \Omega_0, \end{cases}$$

with  $\Omega_0 := \{\omega \in \Omega \mid p_{\theta_0}(\omega) > 0\}$ . The first term in the left-hand side of (7.6) pertains to the differentiability of the (square-root) likelihood ratio at  $h = 0$ , while the second term to the negligibility of the singular part.

The quantum counterpart of this characterization is given by the following.

**Theorem 7.6 (q-LAN for i.i.d. models).** *Let  $\{\rho_\theta \mid \theta \in \Theta \subset \mathbb{R}^d\}$  be a quantum statistical model on a finite dimensional Hilbert space  $\mathcal{H}$ , and suppose that, for some  $\theta_0 \in \Theta$ , a version  $R_h$  of the square-root likelihood ratio  $\mathcal{R}(\rho_{\theta_0+h} \mid \rho_{\theta_0})$  is differentiable at  $h = 0$ , and the absolutely continuous part of  $\rho_{\theta_0+h}$  with respect to  $\rho_{\theta_0}$  satisfies*

$$\text{Tr } \rho_{\theta_0} R_h^2 = 1 - o(\|h\|^2). \tag{7.7}$$

*Then  $\{\rho_\theta^{\otimes n} \mid \theta \in \Theta \subset \mathbb{R}^d\}$  is locally asymptotically normal at  $\theta_0$ , in that*

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

satisfies (i) and (ii) in Definition 7.4. Here  $L_i$  is (a version of) the  $i$ th symmetric logarithmic derivative (SLD, for short, [12]) at  $\theta_0$ , and  $J = (J_{ij})$  is given by

$$J_{ij} := \text{Tr } \rho_{\theta_0} L_j L_i.$$

Further, given observables  $\{B_i\}_{1 \leq i \leq d'}$  on  $\mathcal{H}$  satisfying  $\text{Tr } \rho_{\theta_0} B_i = 0$  for  $i = 1, \dots, d'$ , let  $X^{(n)} = \{X_i^{(n)}\}_{1 \leq i \leq d'}$  be observables on  $\mathcal{H}^{\otimes n}$  defined by

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

Then we have

$$\rho_{\theta_0+h/\sqrt{n}}^{\otimes n} \triangleleft \rho_{\theta_0}^{\otimes n} \quad \text{and} \quad X^{(n)} \overset{\rho_{\theta_0+h/\sqrt{n}}^{\otimes n}}{\rightsquigarrow} N((\text{Re } \tau)h, \Sigma) \tag{7.8}$$

for  $h \in \mathbb{R}^d$ , where  $\Sigma$  is the  $d' \times d'$  positive semidefinite matrix defined by  $\Sigma_{ij} = \text{Tr } \rho_{\theta_0} B_j B_i$  and  $\tau$  is the  $d' \times d$  matrix defined by  $\tau_{ij} = \text{Tr } \rho_{\theta_0} L_j B_i$ .

The proof of Theorem 7.6 is deferred to the Supplementary Material [8].

Let us demonstrate the power of Theorem 7.6. First, we recall the following two-dimensional spin-1/2 pure state model treated in [34], Example 3.3:

$$\begin{aligned} \tilde{\rho}_\theta &:= e^{\frac{1}{2}(\theta^1 \sigma_1 + \theta^2 \sigma_2 - \psi(\theta))} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} e^{\frac{1}{2}(\theta^1 \sigma_1 + \theta^2 \sigma_2 - \psi(\theta))} \\ &= \frac{1}{2} \left\{ I + \frac{\tanh \|\theta\|}{\|\theta\|} (\theta^1 \sigma_1 + \theta^2 \sigma_2) + \frac{1}{\cosh \|\theta\|} \sigma_3 \right\}, \end{aligned}$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -\sqrt{-1} \\ \sqrt{-1} & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the Pauli matrices,  $\theta = (\theta^1, \theta^2) \in \mathbb{R}^2$  are parameters to be estimated, and  $\psi(\theta) := \log \cosh \|\theta\|$ . A version of the square-root likelihood ratio  $\mathcal{R}(\tilde{\rho}_\theta \mid \tilde{\rho}_0)$  is given by  $\tilde{R}_\theta = e^{\frac{1}{2}(\theta^1 \sigma_1 + \theta^2 \sigma_2 - \psi(\theta))}$ , and is expanded in  $\theta$  as

$$\tilde{R}_\theta = I + \frac{1}{2} L_i \theta^i + o(\|\theta\|),$$

where  $L_i := \sigma_i$  is a version of the  $i$ th SLD of the model  $\tilde{\rho}_\theta$  at  $\theta = 0$ . Let  $X^{(n)} = (X_1^{(n)}, X_2^{(n)})$  be defined by

$$X_i^{(n)} := \Delta_i^{(n)} := \frac{1}{\sqrt{n}} \sum_{k=1}^n I^{\otimes(k-1)} \otimes L_i \otimes I^{\otimes(n-k)}. \tag{7.9}$$

Then it is shown that  $\{\tilde{\rho}_\theta^{\otimes n}\}$  is locally asymptotically normal at  $\theta = 0$ , and

$$X^{(n)} \overset{\tilde{\rho}_{h/\sqrt{n}}^{\otimes n}}{\rightsquigarrow} N(h, J), \tag{7.10}$$

where

$$J = [\text{Tr } \tilde{\rho}_0 L_j L_i]_{ij} = \begin{pmatrix} 1 & -\sqrt{-1} \\ \sqrt{-1} & 1 \end{pmatrix}.$$

Incidentally, let us investigate what happens when the scaling factor  $1/\sqrt{n}$  of the parameter  $\theta = h/\sqrt{n}$  is replaced with  $1/g(n)$ , where  $g(n) > 0$  and  $\lim_{n \rightarrow \infty} g(n) = \infty$ . By direct computation, we have

$$\begin{aligned} \liminf_{n \rightarrow \infty} \text{Tr } \tilde{\rho}_0^{\otimes n} \tilde{\rho}_{h/g(n)}^{\otimes n} &= \liminf_{n \rightarrow \infty} \{\text{Tr } \tilde{\rho}_0 \tilde{\rho}_{h/g(n)}\}^n \\ &= \liminf_{n \rightarrow \infty} \left\{ \frac{1}{2} \left( 1 + \frac{1}{\cosh(\|h\|/g(n))} \right) \right\}^n \\ &= \liminf_{n \rightarrow \infty} \left\{ 1 - \frac{\|h\|^2}{4g(n)^2} + o\left(\frac{1}{g(n)^2}\right) \right\}^n \\ &= \liminf_{n \rightarrow \infty} \left\{ 1 - \frac{\|h\|^2}{4g(n)^2} + o\left(\frac{1}{g(n)^2}\right) \right\}^{g(n)^2 \frac{n}{g(n)^2}} \\ &= \liminf_{n \rightarrow \infty} e^{-\frac{\|h\|^2}{4} \frac{n}{g(n)^2}}. \end{aligned}$$

It then follows from Theorem 4.5 that  $\tilde{\rho}_{h/g(n)}^{\otimes n} \triangleleft \tilde{\rho}_0^{\otimes n}$  if and only if  $n/g(n)^2$  is bounded.

Now we consider a perturbed model

$$\rho_\theta := e^{-f(\theta)} \tilde{\rho}_\theta + (1 - e^{-f(\theta)}) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad (\theta \in \mathbb{R}^2),$$

where  $f(\theta)$  is a smooth function that is positive for all  $\theta \neq 0$  and  $f(0) = 0$ . Geometrically, this model is tangential to the Bloch sphere at the north pole  $\rho_0 (= \tilde{\rho}_0)$ , and has a singularity at  $\theta = 0$  in that the rank of the model drops there. Such a model was beyond the scope of our previous paper [34].

Since  $\rho_\theta \geq e^{-f(\theta)} \tilde{\rho}_\theta$ , we see from Lemma 2.2 that  $\rho_\theta \gg \rho_0$  for all  $\theta$ . It is also easily seen that the quantum Lebesgue decomposition  $\rho_\theta = \rho_\theta^{\text{ac}} + \rho_\theta^\perp$  with respect to  $\rho_0$  is given by

$$\rho_\theta^{\text{ac}} := e^{-f(\theta)} \tilde{\rho}_\theta, \quad \rho_\theta^\perp := (1 - e^{-f(\theta)}) \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Similarly, the quantum Lebesgue decomposition  $\rho_\theta^{\otimes n} = (\rho_\theta^{\otimes n})^{\text{ac}} + (\rho_\theta^{\otimes n})^\perp$  with respect to  $\rho_0^{\otimes n}$  is given by

$$(\rho_\theta^{\otimes n})^{\text{ac}} = (\rho_\theta^{\text{ac}})^{\otimes n}, \quad (\rho_\theta^{\otimes n})^\perp = \rho_\theta^{\otimes n} - (\rho_\theta^{\otimes n})^{\text{ac}}.$$

For a positive sequence  $g(n)$  satisfying  $\lim_{n \rightarrow \infty} g(n) = \infty$ , we have

$$\text{Tr}(\rho_{h/g(n)}^{\otimes n})^{\text{ac}} = e^{-nf(h/g(n))}$$

and

$$\begin{aligned} \liminf_{n \rightarrow \infty} \text{Tr} \rho_0^{\otimes n} (\rho_{h/g(n)}^{\otimes n})^{\text{ac}} &= \liminf_{n \rightarrow \infty} e^{-nf(h/g(n))} \left\{ \frac{1}{2} \left( 1 + \frac{1}{\cosh(\|h\|/g(n))} \right) \right\}^n \\ &= \liminf_{n \rightarrow \infty} e^{-nf(h/g(n)) - \frac{\|h\|^2}{4} \frac{n}{g(n)^2}}. \end{aligned}$$

It then follows from Theorem 4.5 that  $\rho_{h/g(n)}^{\otimes n} \triangleleft \rho_0^{\otimes n}$  if and only if  $nf(h/g(n))$  converges to zero and  $n/g(n)^2$  is bounded.

For the standard scaling  $g(n) = \sqrt{n}$ , the above observation shows that  $\rho_{h/\sqrt{n}}^{\otimes n} \triangleleft \rho_0^{\otimes n}$  if and only if  $f(\theta) = o(\|\theta\|^2)$ . Then the operator  $R_\theta := e^{-\frac{1}{2}f(\theta)} \tilde{R}_\theta$ , a version of the square-root likelihood ratio  $\mathcal{R}(\rho_\theta | \rho_0)$ , is expanded in  $\theta$  as

$$R_\theta = I + \frac{1}{2} L_i \theta^i + o(\|\theta\|),$$

where  $L_i := \sigma_i$  is a version of the  $i$ th SLD of the model  $\rho_\theta$  at  $\theta = 0$ . On the other hand, the singular part  $\rho_\theta^\perp$  exhibits  $\text{Tr} \rho_\theta^\perp = o(\|\theta\|^2)$ ; this ensures the condition (7.7). It then follows from Theorem 7.6 that  $\{\rho_\theta^{\otimes n}\}_\theta$  is locally asymptotically normal at  $\theta = 0$ , and the sequence  $X^{(n)}$  of observables defined by (7.9) exhibits

$$X^{(n)} \overset{\rho_{h/\sqrt{n}}^{\otimes n}}{\rightsquigarrow} N(h, J). \tag{7.11}$$

In summary, as far as the observables  $X^{(n)} = (X_1^{(n)}, X_2^{(n)})$  defined by (7.9) are concerned, the i.i.d. extension  $\{\rho_{h/\sqrt{n}}^{\otimes n} | h \in \mathbb{R}^2\}$  of the perturbed model  $\rho_\theta$  around the singular point  $\theta = 0$  is asymptotically similar to the quantum Gaussian shift model  $\{N(h, J) | h \in \mathbb{R}^2\}$  as shown in (7.11), and is also asymptotically similar to the i.i.d. extension  $\{\tilde{\rho}_{h/\sqrt{n}}^{\otimes n} | h \in \mathbb{R}^2\}$  of the unperturbed pure state model  $\tilde{\rho}_\theta$  around  $\theta = 0$  as shown in (7.10).

Now we touch upon the issue of limit states in quantum convergence in distribution. Let us consider a one-dimensional qubit rotation model  $\rho_\theta := U(\theta)\rho_0U(\theta)^*$ , where  $U(\theta) := e^{-\sqrt{-1}\theta\sigma_3}$  and  $\rho_0 := \frac{1}{2}(I + x\sigma_1 + z\sigma_3)$  with  $x \neq 0$  and  $x^2 + z^2 < 1$ . This model is faithful, and the SLD  $L$  of the model at  $\theta = 0$  is uniquely given by  $L = 2x\sigma_2$ . It is easily verified from Theorem 7.6 that the i.i.d. sequence  $\rho_\theta^{\otimes n}$  is locally asymptotically normal at  $\theta = 0$ , in that the sequence

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

satisfies conditions (i) and (ii) in Definition 7.4. In particular,

$$\Delta^{(n)} \overset{\rho_0^{\otimes n}}{\rightsquigarrow} N(0, J),$$

where  $J := \text{Tr} \rho_0 L^2 = 4x^2$  is the SLD Fisher information. Observe that the limit law  $N(0, J)$  is a classical Gaussian distribution. However, this does not imply that the limit state of  $\rho_\theta^{\otimes n}$  is classical. To see this, consider the following observables:

$$X_i^{(n)} = \frac{1}{\sqrt{n}} \sum_{k=1}^n I^{\otimes(k-1)} \otimes B_i \otimes I^{\otimes(n-k)} \quad (i = 1, 2, 3),$$

where

$$B_1 = \sigma_2, \quad B_2 = \frac{1}{r}(z\sigma_1 - x\sigma_3), \quad B_3 = \frac{1}{r}(x\sigma_1 + z\sigma_3) - rI$$

with  $r := \sqrt{x^2 + z^2}$ . It then follows from Theorem 7.6 that

$$\begin{pmatrix} X^{(n)} \\ \Delta^{(n)} \end{pmatrix} \overset{\rho_0^{(n)}}{\rightsquigarrow} N \left( \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sqrt{-1}r & 0 & 2x \\ -\sqrt{-1}r & 1 & 0 & -2\sqrt{-1}xr \\ 0 & 0 & 1-r^2 & 0 \\ 2x & 2\sqrt{-1}xr & 0 & 4x^2 \end{pmatrix} \right) \tag{7.12}$$

and

$$X^{(n)} \overset{\rho_{h/\sqrt{n}}^{\otimes n}}{\rightsquigarrow} N \left( \begin{pmatrix} 2xh \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sqrt{-1}r & 0 \\ -\sqrt{-1}r & 1 & 0 \\ 0 & 0 & 1-r^2 \end{pmatrix} \right) \tag{7.13}$$

for all  $h \in \mathbb{R}^d$ . Since the covariance matrices appeared in (7.12) and (7.13) have nonzero imaginary parts, we see that the limit state is nonclassical. Incidentally, the expression (7.13) is reducible, and is decomposed as

$$N \left( \begin{pmatrix} 2xh \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \sqrt{-1}r \\ -\sqrt{-1}r & 1 \end{pmatrix} \right) \otimes N(0, 1-r^2).$$

Here, the first factor represents an irreducible quantum Gaussian shift model having a local shift parameter  $h$ , and the second a classical Gaussian state that is independent of  $h$ . Such a display is comparable to [11].

We conclude this subsection with a short remark that, for any quantum statistical model that fulfils assumptions of Theorem 7.6, the Holevo bound [13] is asymptotically achievable at  $\theta_0$ . In fact, let  $\{B_i\}_{1 \leq i \leq d'}$  be a basis of the minimal  $\mathcal{D}$ -invariant extension of the SLD tangent space at  $\theta_0$ , where  $\mathcal{D}$  is the commutation operator [13]. Then the Holevo bound for the original model  $\{\rho_\theta\}_\theta$  at  $\theta = \theta_0$  coincides with that for the quantum Gaussian shift model  $N((\text{Re } \tau)h, \Sigma)$  at  $h = 0$ , and hence at any  $h$ . Thus the asymptotic property

$$X^{(n)} \overset{\rho_{\theta_0+h/\sqrt{n}}^{\otimes n}}{\rightsquigarrow} N((\text{Re } \tau)h, \Sigma)$$

enables us to construct a sequence of observables that asymptotically achieves the Holevo bound. For a concrete construction of estimators, see the proof of [34], Theorem 3.1. In order to make a



comparison between the Holevo bound and the standard Cramér–Rao-type bound, let us consider the three-dimensional qubit model  $\rho_\theta = \frac{1}{2}(I + \theta^1\sigma_1 + \theta^2\sigma_2 + \theta^3\sigma_3)$ . As shown in [34], Example 3.2, the Holevo bound  $C_{\theta_0}^H(\rho_\theta, W_\theta)$  and the Hayashi–Gill–Massar bound  $C_{\theta_0}^{\text{HGM}}(\rho_\theta, W_\theta)$  for the model with the SLD Fisher information weight  $W_\theta := J_\theta^{(S)}$  are given by

$$C_{\theta_0}^H(\rho_\theta, J_\theta^{(S)}) = 3 + 2\|\theta_0\|, \quad C_{\theta_0}^{\text{HGM}}(\rho_\theta, J_\theta^{(S)}) = 9.$$

Clearly, the former is much smaller than the latter for all values of the true parameter  $\theta_0$ . This implies that asymptotically we can extract more information about the parameter using collective measurements than when we use separable measurements such as the Pauli measurements. It is believed that under some regularity conditions, there are no “superefficient” estimators that breaks the Holevo bound. Therefore, the universal achievability of the Holevo bound is of fundamental importance in asymptotic quantum statistics.

## 8. Concluding remarks

In the present paper, we first formulated a novel quantum Lebesgue decomposition (Lemma 3.3), and then developed a theory of quantum contiguity (Definition 4.2). We further studied the notion of convergence in distribution in the quantum domain, and proved a noncommutative extension of the Lévy–Cramér continuity theorem under the sandwiched convergence in distribution (Lemma 5.3). Combining these key results, we arrived at our main result, the quantum LeCam third lemma (Theorems 6.1 and 6.2). The power and usefulness of our theory were demonstrated by several examples, including a quantum contiguity version of the Kakutani dichotomy (Theorem 7.2) and enlargement of the scope of q-LAN (Corollary 7.5).

We believe that the paper presented some notable progresses in asymptotic quantum statistics. Nevertheless, there are many open problems left to study in the future. Among others, it is not clear whether every sequence of positive operator-valued measures on a weak q-LAN model can be realized on the limiting quantum Gaussian shift model. In classical statistics, this question has been solved affirmatively by the representation theorem [32], which asserts that, given a weakly convergent sequence  $T^{(n)}$  of statistics on a LAN model  $\{P_{\theta_0+h/\sqrt{n}}^{(n)} \mid h \in \mathbb{R}^d\}$ , there exist a limiting statistics  $T$  on the Gaussian shift model  $\{N(h, J^{-1}) \mid h \in \mathbb{R}^d\}$  such that  $T^{(n)} \xrightarrow{h} T$ . The representation theorem is useful in proving, for example, the nonexistence of an estimator that can asymptotically do better than what can be achieved in the limiting Gaussian shift model. Moreover, the so-called convolution theorem and local asymptotic minimax theorem, which are the standard tools in discussing asymptotic lower bounds for estimation in LAN models, immediately follows [32]. Extending the representation theorem, convolution theorem, and local asymptotic minimax theorem to the quantum domain is one of the most important open problems.

In quantum information theory, the notion of sufficiency of a subalgebra plays an important role. Let  $\rho$  and  $\omega$  be faithful density operators on a matrix algebra  $\mathcal{A}$ . Further, let  $\mathcal{A}_0$  be a subalgebra of  $\mathcal{A}$ , and let  $\rho_0$  and  $\omega_0$  be the reduced density operators of  $\rho$  and  $\omega$  onto  $\mathcal{A}_0$ . Due

to Petz’s sufficiency theorem [24,27,28], the subalgebra  $\mathcal{A}_0$  is sufficient for  $\{\rho, \omega\}$  if and only if  $[D\rho, D\omega]_t = [D\rho_0, D\omega_0]_t$  for all  $t \in \mathbb{R}$ , where  $[D\rho, D\omega]_t$  is the Connes cocycle defined by

$$[D\rho, D\omega]_t := \rho^{\sqrt{-1}t} \omega^{-\sqrt{-1}t}.$$

The sufficiency is also equivalent to the conservation of the quantum relative entropy, and thus the notion of sufficiency plays an essential role in quantum hypothesis testing problems. On the other hand, the square-root likelihood ratio was introduced as a direct analogue of the classical likelihood ratio in which the absolutely continuous part relative to the dominating measure was taken into account, and was applicable even to non-faithful density operators in a quite natural way, playing an essential role in quantum estimation problems as demonstrated through the present paper. Investigating the relationship between the Connes cocycle and the square-root likelihood ratio is an interesting future problem.

It also remains to be investigated whether our asymptotically optimal statistical procedures 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  $\theta_0$  is unknown. Some authors [9,35] 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  $\hat{\theta}_1$ , of the parameter. One next applies the theory of q-LAN to compute the asymptotically optimal measurement scheme which corresponds to the situation  $\theta_0 = \hat{\theta}_1$ , and then proceeds to implement this measurement on the remaining  $n_2$  ( $:= n - n_1$ ) quantum systems collectively, estimating  $h$  in the model  $\theta = \hat{\theta}_1 + h/\sqrt{n_2}$ . However such procedures are inherently limited to within the scope of weak consistency. Studying the strong consistency and asymptotic efficiency [7] in the framework of collective quantum estimation scheme is also an important open problem.

## Appendix A: Quantum Gaussian state

Given a  $d \times d$  real skew-symmetric matrix  $S = [S_{ij}]$ , let  $\text{CCR}(S)$  denote the algebra generated by the observables  $X = (X_1, \dots, X_d)$  that satisfy the following canonical commutation relations (CCR):

$$\frac{\sqrt{-1}}{2} [X_i, X_j] = S_{ij} \quad (1 \leq i, j \leq d),$$

or more precisely

$$e^{\sqrt{-1}X_i} e^{\sqrt{-1}X_j} = e^{-\sqrt{-1}S_{ij}} e^{\sqrt{-1}(X_i+X_j)} \quad (1 \leq i, j \leq d).$$

A state  $\phi$  on  $\text{CCR}(S)$  is called a *quantum Gaussian state*, denoted  $\phi \sim N(h, J)$ , if the characteristic function  $\mathcal{F}_\xi\{\phi\} := \phi(e^{\sqrt{-1}\xi^i X_i})$  takes the form

$$\mathcal{F}_\xi\{\phi\} = e^{\sqrt{-1}\xi^i h_i - \frac{1}{2}\xi^i \xi^j V_{ij}},$$

where  $\xi = (\xi^i)_{i=1}^d \in \mathbb{R}^d$ ,  $h = (h_i)_{i=1}^d \in \mathbb{R}^d$ , and  $V = [V_{ij}]$  is a real symmetric matrix such that the Hermitian matrix  $J := V + \sqrt{-1}S$  is positive semidefinite. When the canonical observables  $X$  need to be specified, we also use the notation  $(X, \phi) \sim N(h, J)$ .

When we discuss relationships between a quantum Gaussian state  $\phi$  on a CCR and a state on another algebra, we need to use the *quasi-characteristic function* [16]

$$\phi\left(\prod_{i=1}^r e^{\sqrt{-1}\xi_i^i X_i}\right) = \exp\left(\sum_{i=1}^r\left(\sqrt{-1}\xi_i^i h_i - \frac{1}{2}\xi_i^i \xi_i^j J_{ji}\right) - \sum_{i=1}^r \sum_{u=i+1}^r \xi_i^i \xi_u^j J_{ji}\right) \tag{A.1}$$

of a quantum Gaussian state, where  $(X, \phi) \sim N(h, J)$  and  $\{\xi_i\}_{i=1}^r \subset \mathbb{R}^d$ . Note that (A.1) is analytically continued to  $\{\xi_i\}_{i=1}^r \subset \mathbb{C}^d$ .

## Appendix B: Quantum Lévy–Cramér continuity theorem

In [15], they derived a noncommutative version of the Lévy–Cramér continuity theorem. Let us first cite their main result in a form consistent with the present paper.

For each  $n \in \mathbb{N}$ , let  $\rho^{(n)}$  be a state (density operator) and  $Z^{(n)} = (Z_1^{(n)}, \dots, Z_s^{(n)})$  be observables on a finite dimensional Hilbert space  $\mathcal{H}^{(n)}$ . Further, let  $\phi$  be a normal state (linear functional) and  $Z^{(\infty)} = (Z_1^{(\infty)}, \dots, Z_s^{(\infty)})$  be densely defined observables on a possibly infinite dimensional Hilbert space  $\mathcal{H}^{(\infty)}$ . Assume that for all  $m \in \mathbb{N}$ ,  $\alpha = (\alpha_1, \dots, \alpha_m) \in \mathbb{R}^m$ , and  $j_1, \dots, j_m \in \{1, \dots, s\}$ , one has

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \prod_{i=1}^m e^{\sqrt{-1}\alpha_i Z_{j_i}^{(n)}} = \phi\left(\prod_{i=1}^m e^{\sqrt{-1}\alpha_i Z_{j_i}^{(\infty)}}\right). \tag{B.1}$$

Then it holds that

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} \prod_{i=1}^s f_i(Z_i^{(n)}) = \phi\left(\prod_{i=1}^s f_i(Z_i^{(\infty)})\right) \tag{B.2}$$

for any bounded continuous functions  $f_1, \dots, f_s$  on  $\mathbb{R}$ . Furthermore, (B.2) remains true for bounded Borel functions  $f_1, \dots, f_s$  on  $\mathbb{R}$  that enjoy certain measure conditions for the sets of discontinuity points (which will be stated below).

Now observe that assumption (B.1) requires every finite repetition and permutation of the given observables  $\{Z_i^{(\cdot)}\}_{1 \leq i \leq s}$ . Nevertheless, what Jakšić et al. elucidated was something stronger in that their proof did not make full use of assumption (B.1) and is effective under certain weaker assumptions. In particular, the following variant, in which assumption (B.1) is replaced with (B.3)–(B.5), plays a key role in the present paper.

**Theorem B.1.** *For  $n \in \mathbb{N} \cup \{\infty\}$ ,  $i \in \{1, \dots, s\}$ , and  $\alpha = (\alpha_1, \dots, \alpha_s) \in \mathbb{R}^s$ , let  $U_i^{-(n)}(\alpha)$  and  $U_i^{+(n)}(\alpha)$  be unitary operators defined by*

$$U_i^{-(n)}(\alpha) := \prod_{t=1}^i e^{\sqrt{-1}\alpha_t Z_t^{(n)}} \quad \text{and} \quad U_i^{+(n)}(\alpha) := \prod_{t=i}^s e^{\sqrt{-1}\alpha_t Z_t^{(n)}},$$

and let  $U_0^{-(n)}(\alpha)$  and  $U_{s+1}^{+(n)}(\alpha)$  be identity operators. Assume that there is a  $J \in \{0, 1, \dots, s\}$  such that, for all  $\alpha, \beta \in \mathbb{R}^s$ , the following three conditions are satisfied:

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} U_s^{-(n)}(\alpha) = \phi(U_s^{-(\infty)}(\alpha)), \tag{B.3}$$

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} U_J^{-(n)}(\alpha) U_J^{-(n)}(\beta)^* = \phi(U_J^{-(\infty)}(\alpha) U_J^{-(\infty)}(\beta)^*), \tag{B.4}$$

$$\lim_{n \rightarrow \infty} \text{Tr} \rho^{(n)} U_{J+1}^{+(n)}(\alpha)^* U_{J+1}^{+(n)}(\beta) = \phi(U_{J+1}^{+(\infty)}(\alpha)^* U_{J+1}^{+(\infty)}(\beta)). \tag{B.5}$$

Then (B.2) holds for any bounded continuous functions  $f_1, \dots, f_s$  on  $\mathbb{R}$ .

Furthermore, let  $f_1, \dots, f_s$  be bounded Borel functions on  $\mathbb{R}$ , and let  $\mathcal{D}(f_i)$  be the set of discontinuity points of  $f_i$ . Assume, in addition to (B.3)–(B.5), that one has

$$\mu_i^\alpha(\mathcal{D}(f_i)) = 0 \tag{B.6}$$

for all  $i \in \{1, \dots, s\}$  and  $\alpha \in \mathbb{R}^s$ , where  $\mu_i^\alpha$  is the classical probability measure having the characteristic function

$$\varphi_{\mu_i^\alpha}(\gamma) := \begin{cases} \phi(U_{i-1}^{-(\infty)}(\alpha) (e^{\sqrt{-1}\gamma Z_i^{(\infty)}}) U_{i-1}^{-(\infty)}(\alpha)^*) & \text{if } i \leq J, \\ \phi(U_{i+1}^{+(\infty)}(\alpha)^* (e^{\sqrt{-1}\gamma Z_i^{(\infty)}}) U_{i+1}^{+(\infty)}(\alpha)) & \text{if } i \geq J + 1. \end{cases} \tag{B.7}$$

Then (B.2) remains true.

The proof of Theorem B.1 is exactly the same as [15]. Note that when  $J \in \{1, \dots, s - 1\}$ , the characteristic functions (B.7) for  $i = 1$  and  $s$  are reduced to

$$\varphi_{\mu_1^\alpha}(\gamma) = \phi(e^{\sqrt{-1}\gamma Z_1^{(\infty)}}) \quad \text{and} \quad \varphi_{\mu_s^\alpha}(\gamma) = \phi(e^{\sqrt{-1}\gamma Z_s^{(\infty)}}).$$

In particular, they are independent of  $\alpha$ . This fact is exploited in our sandwiched-type continuity theorem (Lemma 5.3).

## Acknowledgments

The present study was supported by JSPS KAKENHI Grants JP22340019 and JP17H02861.

## Supplementary Material

**Supplementary material to “Noncommutative Lebesgue decomposition and contiguity with applications in quantum statistics”** (DOI: [10.3150/19-BEJ1185SUPP](https://doi.org/10.3150/19-BEJ1185SUPP); .pdf). This supplementary material is devoted to proofs of Remark 3.4, Theorem 4.4, Theorem 4.5, Lemma 5.6, Theorem 6.1, Theorem 7.1, Theorem 7.2 and Theorem 7.6.

## References

- [1] Bhatia, R. (1997). *Matrix Analysis. Graduate Texts in Mathematics* **169**. New York: Springer. MR1477662 <https://doi.org/10.1007/978-1-4612-0653-8>
- [2] Blackadar, B. (2006). *Operator Algebras: Theory of  $C^*$ -Algebras and von Neumann Algebras. Encyclopaedia of Mathematical Sciences* **122**. Operator Algebras and Non-commutative Geometry III. Berlin: Springer. MR2188261 <https://doi.org/10.1007/3-540-28517-2>
- [3] Bures, D. (1969). An extension of Kakutani's theorem on infinite product measures to the tensor product of semifinite  $w^*$ -algebras. *Trans. Amer. Math. Soc.* **135** 199–212. MR0236719 <https://doi.org/10.2307/1995012>
- [4] Connes, A. (1973). Une classification des facteurs de type III. *Ann. Sci. Éc. Norm. Supér.* (4) **6** 133–252. MR0341115
- [5] Connes, A. (1994). *Noncommutative Geometry*. San Diego, CA: Academic Press. MR1303779
- [6] Dye, H.A. (1952). The Radon–Nikodým theorem for finite rings of operators. *Trans. Amer. Math. Soc.* **72** 243–280. MR0045954 <https://doi.org/10.2307/1990754>
- [7] Fujiwara, A. (2006). Strong consistency and asymptotic efficiency for adaptive quantum estimation problems. *J. Phys. A: Math. Gen.* **39** 12489–12504. Corrigendum: *J. Phys. A: Math. Theor.* **44** Art. ID 079501. MR2265836 <https://doi.org/10.1088/0305-4470/39/40/014>
- [8] Fujiwara, A. and Yamagata, K. (2020). Supplement to “Noncommutative Lebesgue decomposition and contiguity with applications in quantum statistics.” <https://doi.org/10.3150/19-BEJ1185SUPP>
- [9] Gill, R.D. and Massar, S. (2000). State estimation for large ensembles. *Phys. Rev. A* **61** Art. ID 042312.
- [10] Guță, 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>
- [11] Guță, M. and Kahn, J. (2006). Local asymptotic normality for qubit states. *Phys. Rev. A* (3) **73** Art. ID 052108. MR2229156 <https://doi.org/10.1103/PhysRevA.73.052108>
- [12] Helstrom, C.W. (1976). *Quantum Detection and Estimation Theory*. New York: Academic Press.
- [13] Holevo, A.S. (1982). *Probabilistic and Statistical Aspects of Quantum Theory. North-Holland Series in Statistics and Probability* **1**. Amsterdam: North-Holland. MR0681693
- [14] Jacod, J. and Shiryaev, A.N. (2003). *Limit Theorems for Stochastic Processes*, 2nd ed. *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]* **288**. Berlin: Springer. MR1943877 <https://doi.org/10.1007/978-3-662-05265-5>
- [15] Jakšić, V., Pautrat, Y. and Pillet, C.-A. (2010). A non-commutative Lévy–Cramér continuity theorem. *Markov Process. Related Fields* **16** 59–78. MR2664336
- [16] 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** Art. ID 015208. MR2605841 <https://doi.org/10.1063/1.3285287>
- [17] Kadison, R.V. and Ringrose, J.R. (1997). *Fundamentals of the Theory of Operator Algebras, Vol. II: Advanced Theory. Graduate Studies in Mathematics* **16**. Providence, RI: Amer. Math. Soc. MR1468230 <https://doi.org/10.1090/gsm/016/01>
- [18] Kahn, J. and Guță, 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>
- [19] Kakutani, S. (1948). On equivalence of infinite product measures. *Ann. of Math.* (2) **49** 214–224. MR0023331 <https://doi.org/10.2307/1969123>
- [20] Kosaki, H. (1985). Lebesgue decomposition of states on a von Neumann algebra. *Amer. J. Math.* **107** 697–735. MR0789660 <https://doi.org/10.2307/2374374>
- [21] Kubo, F. and Ando, T. (1979/80). Means of positive linear operators. *Math. Ann.* **246** 205–224. MR0563399 <https://doi.org/10.1007/BF01371042>

- [22] Le Cam, L. (1960). Locally asymptotically normal families of distributions. Certain approximations to families of distributions and their use in the theory of estimation and testing hypotheses. *Univ. California Publ. Statist.* **3** 37–98. MR0126903
- [23] 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>
- [24] Ohya, M. and Petz, D. (2004). *Quantum Entropy and Its Use. Texts and Monographs in Physics.* Berlin: Springer. MR1230389 <https://doi.org/10.1007/978-3-642-57997-4>
- [25] Parthasarathy, K.R. (1996). Comparison of completely positive maps on a  $C^*$ -algebra and a Lebesgue decomposition theorem. In *Athens Conference on Applied Probability and Time Series Analysis, Vol. I* (1995) (C.C. Heyde, Y.V. Prohorov, R. Pyke, S.T. Rachev, eds.). *Lect. Notes Stat.* **114** 34–54. New York: Springer. MR1466706 [https://doi.org/10.1007/978-1-4612-0749-8\\_3](https://doi.org/10.1007/978-1-4612-0749-8_3)
- [26] Pedersen, G.K. and Takesaki, M. (1973). The Radon–Nikodym theorem for von Neumann algebras. *Acta Math.* **130** 53–87. MR0412827 <https://doi.org/10.1007/BF02392262>
- [27] Petz, D. (1988). Sufficiency of channels over von Neumann algebras. *Quart. J. Math. Oxford Ser. (2)* **39** 97–108. MR0929798 <https://doi.org/10.1093/qmath/39.1.97>
- [28] Petz, D. (2008). *Quantum Information Theory and Quantum Statistics. Theoretical and Mathematical Physics.* Berlin: Springer. MR2363070
- [29] Sakai, S. (1965). A Radon–Nikodým theorem in  $W^*$ -algebras. *Bull. Amer. Math. Soc.* **71** 149–151. MR0174992 <https://doi.org/10.1090/S0002-9904-1965-11265-4>
- [30] Takesaki, M. (1979). *Theory of Operator Algebras. I.* New York: Springer. MR0548728
- [31] Umegaki, H., Ohya, M. and Hiai, F. (1985). *Introduction to Operator Algebras: From Hilbert Spaces to von Neumann Algebras. Modern Mathematics* **23**. Tokyo: Kyoritsu Shuppan Co., Ltd. (in Japanese). MR0832874
- [32] van der Vaart, A.W. (1998). *Asymptotic Statistics. Cambridge Series in Statistical and Probabilistic Mathematics* **3**. Cambridge: Cambridge Univ. Press. MR1652247 <https://doi.org/10.1017/CBO9780511802256>
- [33] Williams, D. (1991). *Probability with Martingales. Cambridge Mathematical Textbooks.* Cambridge: Cambridge Univ. Press. MR1155402 <https://doi.org/10.1017/CBO9780511813658>
- [34] 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>
- [35] Yang, Y., Chiribella, G. and Hayashi, M. (2019). Attaining the ultimate precision limit in quantum state estimation. *Comm. Math. Phys.* **368** 223–293. MR3946409 <https://doi.org/10.1007/s00220-019-03433-4>

Received August 2019 and revised December 2019