

13 March 2024

Recent Developments in Quantum Statistics

Akio Fujiwara (Osaka University)

Joint Work with K. Yamagata and R. D. Gill

Plan of talk

Motivation

- Part I: Review of quantum estimation theory
- Part II: Review of classical local asymptotic normality
- Part III: Local asymptotic normality in the quantum domain

Part I: Review of quantum
estimation theory

Classical estimation

- Statistical model:

$$\{P_\theta\}_{\theta \in \Theta \subset \mathbb{R}^d} \text{ on } (\Omega, \mathcal{F}, \mu) \xrightarrow{P_\theta \ll \mu} \left\{ p_\theta := \frac{dP_\theta}{d\mu} \right\}_{\theta \in \Theta}$$

- Estimator: $\hat{\theta} : \Omega \longrightarrow \Theta$

- Unbiased estimator:

$$E_{P_\theta}[\hat{\theta}] := \int_{\Omega} \hat{\theta}(\omega) p_\theta(\omega) \mu(d\omega) = \theta \quad (\forall \theta \in \Theta)$$

• Fisher information matrix $J_\theta = [(J_\theta)_{ij}]$

$$(J_\theta)_{ij} := E_{P_\theta} [(\partial_i \log p_\theta)(\partial_j \log p_\theta)]$$

$$(\partial_i := \partial / \partial \theta^i)$$

Logarithmic Derivative

• Cramér-Rao inequality:

$$V_{P_\theta}[\hat{\theta}] \geq (J_\theta)^{-1}$$

for any unbiased estimator $\hat{\theta}$

Quantum estimation

1) Quantum statistical model

$$\mathcal{S} = \{\rho_\theta (> 0) : \theta = (\theta^1, \dots, \theta^d) \in \Theta \subset \mathbb{R}^d\}$$

2) Symmetric Logarithmic Derivative (SLD):

$$\partial_i \rho_\theta = \frac{1}{2}(\rho_\theta L_i + L_i \rho_\theta)$$

3) SLD Fisher information matrix $J_\theta^S = [(J_\theta^S)_{ij}]$

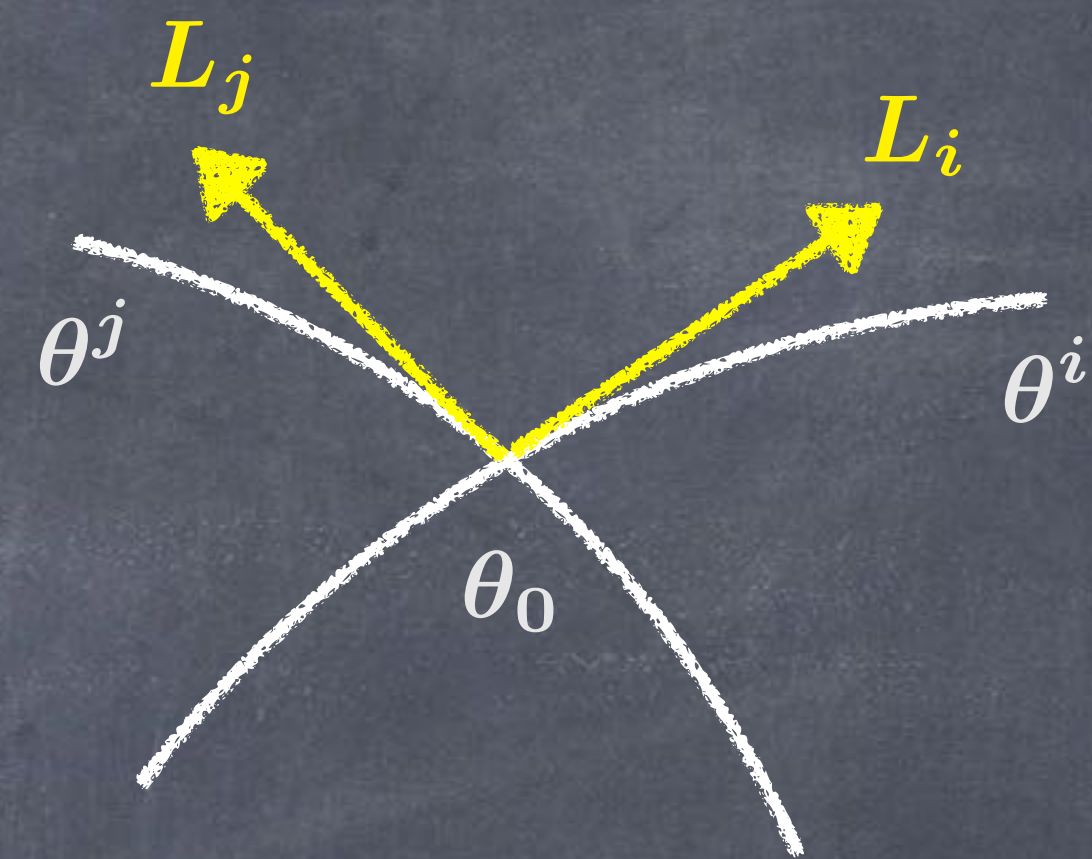
$$(J_\theta^S)_{ij} = \text{Re}(\text{Tr} \rho_\theta L_j L_i)$$

4) SLD Cramér-Rao inequality:

$$V_{\rho_{\theta}}[M] \geq (J_{\theta}^S)^{-1}$$

with equality iff SLDs commute

and M the simultaneous spectral measure



5) It is thus customary to switch the target to

$$\min_M \text{tr} W V_{\rho_{\theta}}[M]$$

given a weight matrix $W > 0$

6) Holevo bound

$$\text{tr } W V_{\rho_{\theta_0}} [M] \geq C_{\theta_0}^H(\rho_{\theta}, W)$$

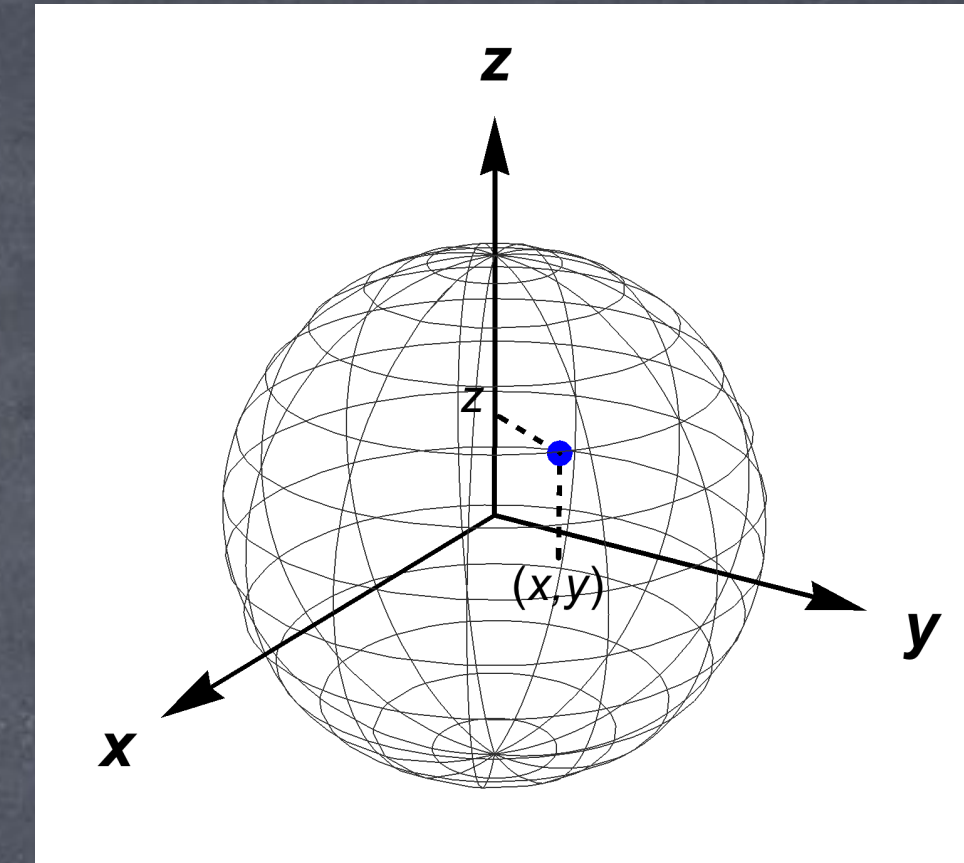
where

$$C_{\theta_0}^H(\rho_{\theta}, W) := \min_{V, B} \{ \text{tr } W V : V \geq [\text{Tr } \rho_{\theta_0} B_j B_i],$$

$$\text{Re Tr } \rho_{\theta_0} L_j B_i = \delta_{ij} \}$$

with V a real symmetric matrix and
 B_1, \dots, B_d selfadjoint operators

Eg. qubit estimation



Qubit state model:

$$\mathcal{S} = \left\{ \rho_{(x,y,z)} = \frac{1}{2}(I + x\sigma_x + y\sigma_y + z\sigma_z) : x^2 + y^2 + z^2 < 1 \right\}$$

Hayashi-Gill-Massar bound:

$$\min_M \text{tr} W V_\theta(M) = C_\theta^{\text{HGM}}(\rho_\theta, W) := \left(\text{tr} \sqrt{(J_\theta^S)^{-1/2} W (J_\theta^S)^{-1/2}} \right)^2$$

When $W := J_\theta^S$

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

separable

$$\begin{array}{ccccccc} \rho_\theta & \otimes & \rho_\theta & \otimes & \cdots & \otimes & \rho_\theta \\ \uparrow & & \uparrow & & & & \uparrow \\ M_1 & \otimes & M_2 & \otimes & \cdots & \otimes & M_n \end{array}$$

Motivation

collective

$$\underbrace{\rho_\theta \otimes \rho_\theta \otimes \cdots \otimes \rho_\theta}_{M^{(n)}}$$

- Suppose we have n copies of a quantum system each in the same state
- We are allowed to use "collective" measurement
- What is the best we can do as $n \rightarrow \infty$?
- We study this problem by extending the theory of LAN to the quantum domain

Main result (i.i.d. case)

- [Direct part]

The Holevo bound is achievable.

- [Converse part]

The Holevo bound is unbreakable.

Main result (generic case)

- [Direct part]

The asymptotic representation bound is achievable.

- [Converse part]

The asymptotic representation bound is unbreakable.

Part II: Review of classical
local asymptotic normality

Local Asymptotic Normality

A sequence of models $\{P_{\theta}^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$

is called LAN at $\theta_0 \in \Theta$ if

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

where

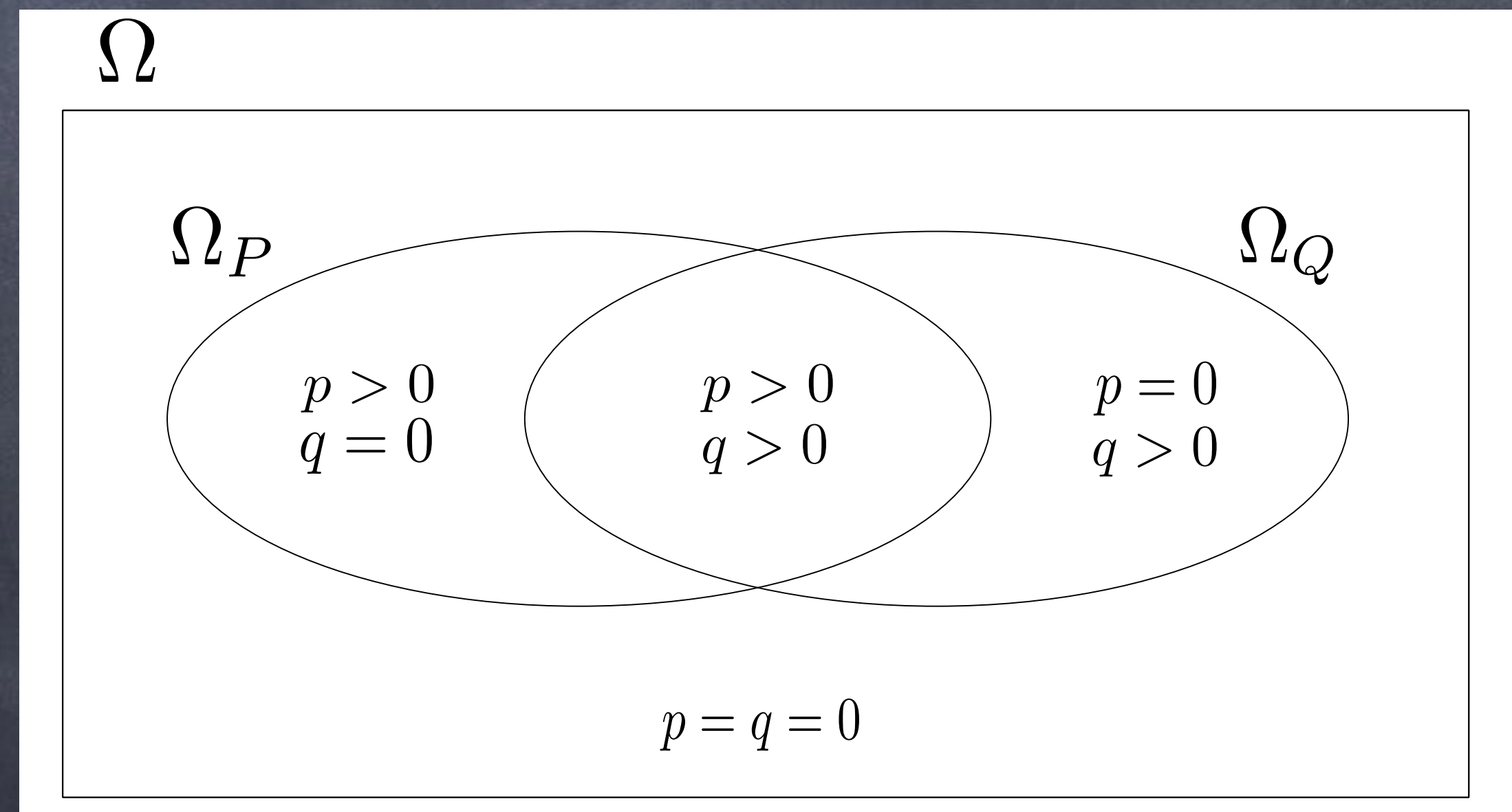
$$\Delta^{(n)} \overset{P_{\theta_0}^{(n)}}{\rightsquigarrow} N(0, J)$$

NB. Given P and Q , let

$$Q = Q^{ac} + Q^\perp \quad (Q^{ac} \ll P, \quad Q^\perp \perp P)$$

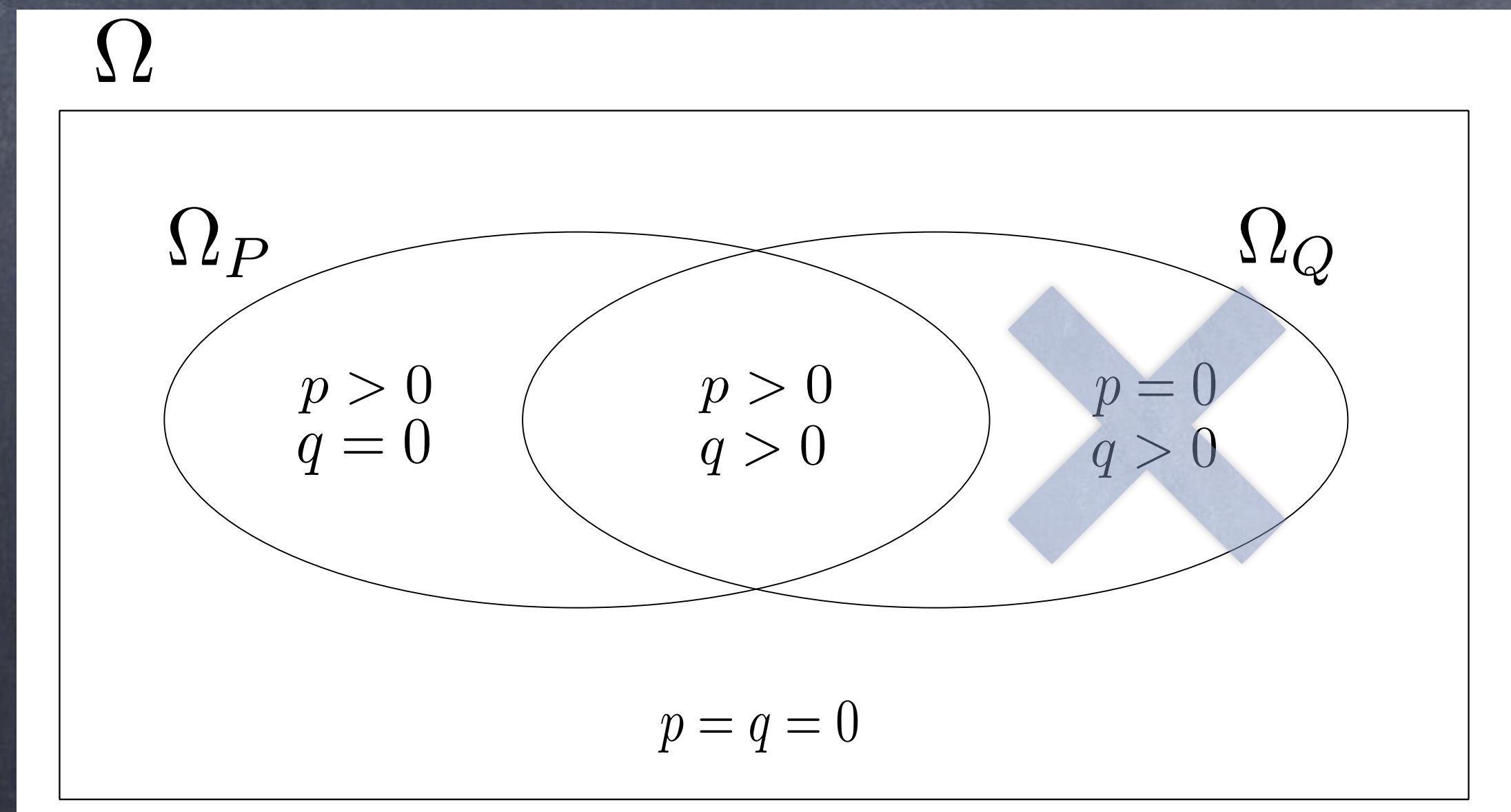
be the Lebesgue decomposition. Then

$$\begin{aligned} \frac{dQ}{dP}(\omega) &:= \frac{dQ^{ac}}{dP}(\omega) \\ &= \frac{q^{ac}(\omega)}{p(\omega)} \end{aligned}$$



Radon-Nikodym Theorem

$$Q \ll P \implies dQ = \frac{dQ}{dP} dP$$



Local Asymptotic Normality

A sequence of models $\{P_{\theta}^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$

is called LAN at $\theta_0 \in \Theta$ if

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

where

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

Prototype of LAN: $p_{\theta}^{(n)} = p_{\theta}^{\otimes n}$

$$\begin{aligned} & \log \frac{p_{\theta_0 + h/\sqrt{n}}^{\otimes n}(X_1, \dots, X_n)}{p_{\theta_0}^{\otimes n}} \\ &= h^i \underbrace{\left\{ \frac{1}{\sqrt{n}} \sum_{k=1}^n \partial_i \log p_{\theta_0}(X_k) \right\}}_{\Delta_i^{(n)}(X_1, \dots, X_n)} - \frac{1}{2} h^i h^j \underbrace{\left\{ -\frac{1}{n} \sum_{k=1}^n \partial_i \partial_j \log p_{\theta_0}(X_k) \right\}}_{J_{ij} + o_{p_{\theta_0}}(1)} + o\left(\frac{1}{n}\right) \end{aligned}$$

where

$$\Delta^{(n)} \stackrel{p_{\theta_0}^{\otimes n}}{\rightsquigarrow} N(0, J)$$

with J being the Fisher information matrix

Similarity to Gaussian shift model

LAN

$$\log \frac{dP_{\theta_0+h/\sqrt{n}}^{(n)}}{dP_{\theta_0}^{(n)}} = h^i \Delta_i^{(n)} - \frac{1}{2} h^i h^j J_{ij} + o_{p_{\theta_0}}(1) \quad (1)$$

Gaussian shift model

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

Contiguity

A sequence $Q^{(n)}$ of probability measures is called contiguous to another sequence $P^{(n)}$ of probability measures, denoted $Q^{(n)} \triangleleft P^{(n)}$, if

$$P^{(n)}(A^{(n)}) \rightarrow 0 \implies Q^{(n)}(A^{(n)}) \rightarrow 0$$

contiguous = asymptotically absolutely continuous

Le Cam's Third Lemma

Theorem (Radon-Nikodym)

$$Q \ll P \implies dQ = \frac{dQ}{dP} dP$$

weak convergence analogue

Theorem (Le Cam)

If $Q^{(n)} \triangleleft P^{(n)}$ and $\left(X^{(n)}, \frac{dQ^{(n)}}{dP^{(n)}} \right) \overset{P^{(n)}}{\rightsquigarrow} (X, V)$, then

$$X^{(n)} \overset{Q^{(n)}}{\rightsquigarrow} \mathcal{L}(\bullet) := E[\mathbb{1}_{\bullet}(X)V]$$

Le Cam's Third Lemma

(Gaussian version)

Theorem

If

$$\begin{pmatrix} X^{(n)} \\ \log \frac{dQ^{(n)}}{dP^{(n)}} \end{pmatrix} \stackrel{P^{(n)}}{\rightsquigarrow} N \left(\begin{pmatrix} 0 \\ -\frac{1}{2}\sigma^2 \end{pmatrix}, \begin{pmatrix} \Sigma & \tau \\ \tau^\top & \sigma^2 \end{pmatrix} \right)$$

then $(Q^{(n)} \triangleleft P^{(n)})$ and

$$X^{(n)} \stackrel{Q^{(n)}}{\rightsquigarrow} N(\tau, \Sigma)$$

Third Lemma under LAN

Suppose

$$\log \frac{dP_{\theta_0+h/\sqrt{n}}^{(n)}}{dP_{\theta_0}^{(n)}} = h^i \Delta_i^{(n)} - \frac{1}{2} h^i h^j J_{ij} + o_{p_{\theta_0}}(1)$$

and

$$\begin{pmatrix} X^{(n)} \\ \Delta^{(n)} \end{pmatrix} \overset{P_{\theta_0}^{(n)}}{\rightsquigarrow} N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma & \tau \\ \tau^\top & J \end{pmatrix} \right)$$

Then

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

The moral:
LAN model is
locally
asymptotically
similar to
Gaussian shift
model

Asymptotic Representation Theorem

Assume that

1) $\{P_\theta^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ is LAN at $\theta_0 \in \Theta$

2) Seq. of statistics $T^{(n)}$ on $P_{\theta_0+h/\sqrt{n}}^{(n)}$ is weakly convergent for each h (i.e. $T^{(n)} \xrightarrow{h} \exists \mathcal{L}_h$)

Then there exists a statistic T on $N(Jh, J)$ s.t.

$$T^{(n)} \xrightarrow{h} T$$

for all h

$$T = T(\Delta, Z) \quad (\Delta \sim N(Jh, J))$$

Part III: LAN in the quantum domain

[Annals of Statistics, 41, 2197-2217 (2013)]

[Bernoulli, 26, 2105-2142 (2020)]

[Annals of Statistics, 51, 1159-1182 (2023)]

Difficulties in extending LAN to the quantum domain

$$\text{LAN: } \log \frac{dP_{\theta_0+h/\sqrt{n}}^{(n)}}{dP_{\theta_0}^{(n)}} = h^i \Delta_i^{(n)} - \frac{1}{2} h^i h^j J_{ij} + o_{P_{\theta_0}}(1)$$

$(\Delta^{(n)} \xrightarrow{\theta_0} N(0, J))$

What is the quantum counterpart of

- 1) Radon-Nikodym derivative?
- 2) contiguity and third lemma?
- 3) weak convergence?

History of quantum LAN

- Guta and Kahn's strong q-LAN (2006, 2009)

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

and

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

Drawbacks

- iid
- nondegeneracy

- Guta and Jencova's weak q-LAN (2007)

Radon-Nikodym derivative in the quantum domain?

f -divergence

- Classical

$$D_f(P||Q) = \int_{\Omega} f\left(\frac{dQ}{dP}\right) dP$$

- Quantum

$$D_f(\rho||\sigma) = \text{Tr} \rho \left\{ f\left(\frac{\mathbf{L}_{\sigma}}{\mathbf{R}_{\rho}}\right) I \right\}$$

Sufficiency

• $\mathcal{G} (\subset \mathcal{F})$ is **sufficient for $\{P, Q\}$**

$$\forall A \in \mathcal{F}, \exists h \in m\mathcal{G}, \forall \mu \in \{P, Q\} \\ E_\mu[\mathbb{1}_A | \mathcal{G}] = h \quad (\mu\text{-a.e.})$$

$$\iff D(P \| Q) = D(P |_{\mathcal{G}} \| Q |_{\mathcal{G}})$$

$$\iff \frac{dP}{dQ} \in m\mathcal{G}$$

• $\mathcal{A} (\subset B(\mathcal{H}))$ is **sufficient for $\{\rho, \sigma\}$**

$$\exists \Gamma : \mathcal{A} \rightarrow B(\mathcal{H}) : \text{CPTP} \\ \forall \tau \in \{\rho, \sigma\}, \Gamma(\tau |_{\mathcal{A}}) = \tau$$

$$\iff D(\rho \| \sigma) = D(\rho |_{\mathcal{A}} \| \sigma |_{\mathcal{A}})$$

$$\iff \underbrace{\rho^{\sqrt{-1}t} \sigma^{-\sqrt{-1}t}}_{[D\rho, D\sigma]_t} \in \mathcal{A} \quad (\forall t \in \mathbb{R})$$

Simple hypothesis testing

[Hiai-Petz (1991), Ogawa-Nagaoka (2000)]

- Classical likelihood ratio test (p vs. q)

$$\frac{1}{n} \log \frac{q^{\otimes n}}{p^{\otimes n}} > \lambda$$

- Quantum likelihood ratio test (ρ vs. σ)

$$\sigma^{\otimes n} - e^{n\lambda} \rho^{\otimes n} > 0$$

What about quantum parameter estimation?

- One may conceive a q-likelihood ratio via

$$\sigma - \lambda\rho = \sqrt{\rho} \left(\sqrt{\rho^{-1}\sigma\rho^{-1}} - \lambda I \right) \sqrt{\rho}$$

- Unfortunately, this quantity is not useful in quantum estimation.
- A completely new idea is needed.

A breakthrough: quantum information geometry

$$\rho_{\theta} = e^{\frac{1}{2}[\theta F - \psi(\theta)]} \rho_0 e^{\frac{1}{2}[\theta F - \psi(\theta)]}$$

SLD geodesic = quantum exp. family

Nice properties:

- 1) quantum Cramér-Rao lower bound is uniformly achievable
- 2) can be extended to non-faithful states

Absolute continuity & singularity

- 1) σ is absolutely continuous to ρ , denoted $\sigma \ll \rho$, if there is a positive operator $R (\geq 0)$ that satisfies $\sigma = R\rho R$
- 2) σ is singular to ρ , denoted $\sigma \perp \rho$, if $\text{supp } \sigma \perp \text{supp } \rho$

An explicit expression

- For $\rho, \sigma > 0$

operator geometric mean

$$\sigma = R \rho R \iff R = \sigma \# \rho^{-1} = \sqrt{\sigma} \left(\sqrt{\sqrt{\sigma} \rho \sqrt{\sigma}} \right)^{-1} \sqrt{\sigma}$$

- For a generic case, the inverse is replaced by the generalized inverse.

Remark

- Classical view for absolute continuity:

$$\text{supp } \sigma \subset \text{supp } \rho$$

$$\text{eg) } R = \sqrt{\rho^{-1}} \sigma \sqrt{\rho^{-1}} \iff \sigma = \sqrt{\rho} R \sqrt{\rho}$$

- For pure states $\rho = |\xi\rangle\langle\xi|$ and $\sigma = |\eta\rangle\langle\eta|$

1) $\text{supp } \sigma \subset \text{supp } \rho \iff \sigma = \rho$

2) $\sigma \ll \rho \iff \langle\xi|\eta\rangle \neq 0$

$$\sigma = R\rho R$$

Quantum Lebesgue decomposition

Theorem. Given quantum states $\rho, \sigma \in \mathcal{S}(\mathcal{H})$,
the decomposition

$$\sigma = \underbrace{R\rho R}_{\sigma^{ac}} + \underbrace{\tau}_{\sigma^\perp} \quad (R \geq 0, \tau \geq 0, \tau \perp \rho)$$

uniquely exists.

In fact, let $\mathcal{H} = \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$ with

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

and let

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

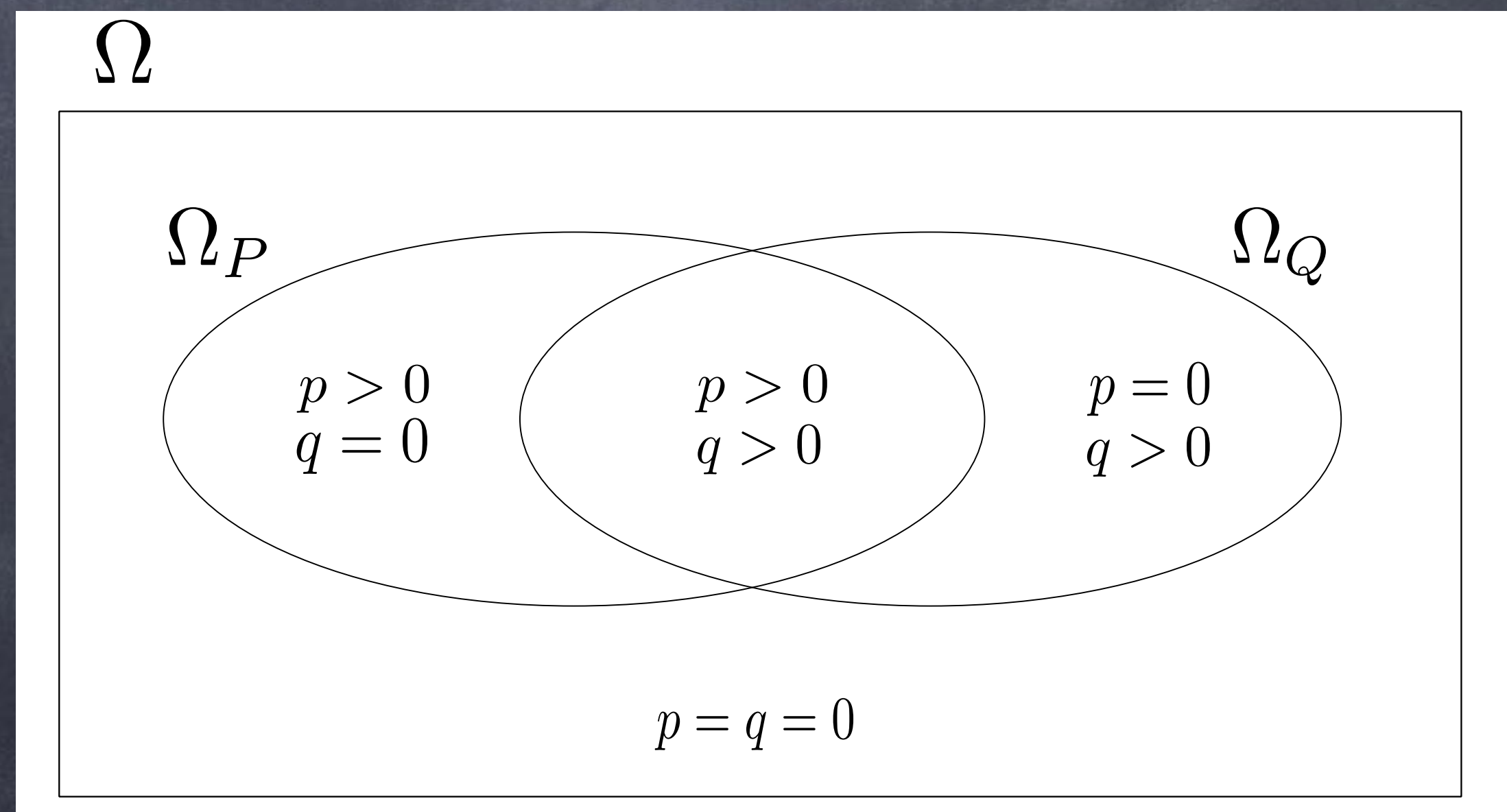
$(\sigma|_{\text{supp } \rho} := \iota_\rho^* \sigma \iota_\rho)$

$\sigma|_{\text{supp } \rho}$

Then

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

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



Remark

In general

$$\mathcal{H}_1 \neq \text{supp } \rho \cap \ker \sigma \quad \text{and} \quad \mathcal{H}_2 \neq \text{supp } \rho \cap \text{supp } \sigma$$

In fact, for $\rho = |\xi\rangle\langle\xi|$ and $\sigma = |\eta\rangle\langle\eta|$

$$1) \quad \text{supp } \rho \cap \text{supp } \sigma \neq \{0\} \iff \rho = \sigma$$

$$2) \quad \mathcal{H}_2 \neq \{0\} \iff \langle\xi|\eta\rangle \neq 0$$

Quantum Likelihood ratio

We call the positive operator R that satisfies

$$\sigma = R\rho R + \tau \quad (R \geq 0, \tau \geq 0, \tau \perp \rho)$$

the square-root likelihood ratio, denoting $\mathcal{R}(\sigma|\rho)$

This is our solution to the problem of finding the noncommutative Radon-Nikodym derivative pertinent to quantum estimation theory. In fact, it allows us to extend the notions of contiguity, weak convergence, etc., to the quantum domain, establishing a theory of quantum local asymptotic normality.

Quantum LAN

$\mathcal{S}^{(n)} = \{\rho_{\theta}^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ is called q-LAN at $\theta_0 \in \Theta$ if

$R_h^{(n)} := \mathcal{R}\left(\rho_{\theta_0+h/\sqrt{n}}^{(n)} \middle| \rho_{\theta_0}^{(n)}\right)$ is expanded in h as

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

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

Third Lemma under q-LAN

Suppose $\mathcal{S}^{(n)} = \{\rho_{\theta}^{(n)} : \theta \in \Theta \subset \mathbb{R}^d\}$ is q-LAN at $\theta_0 \in \Theta$ and

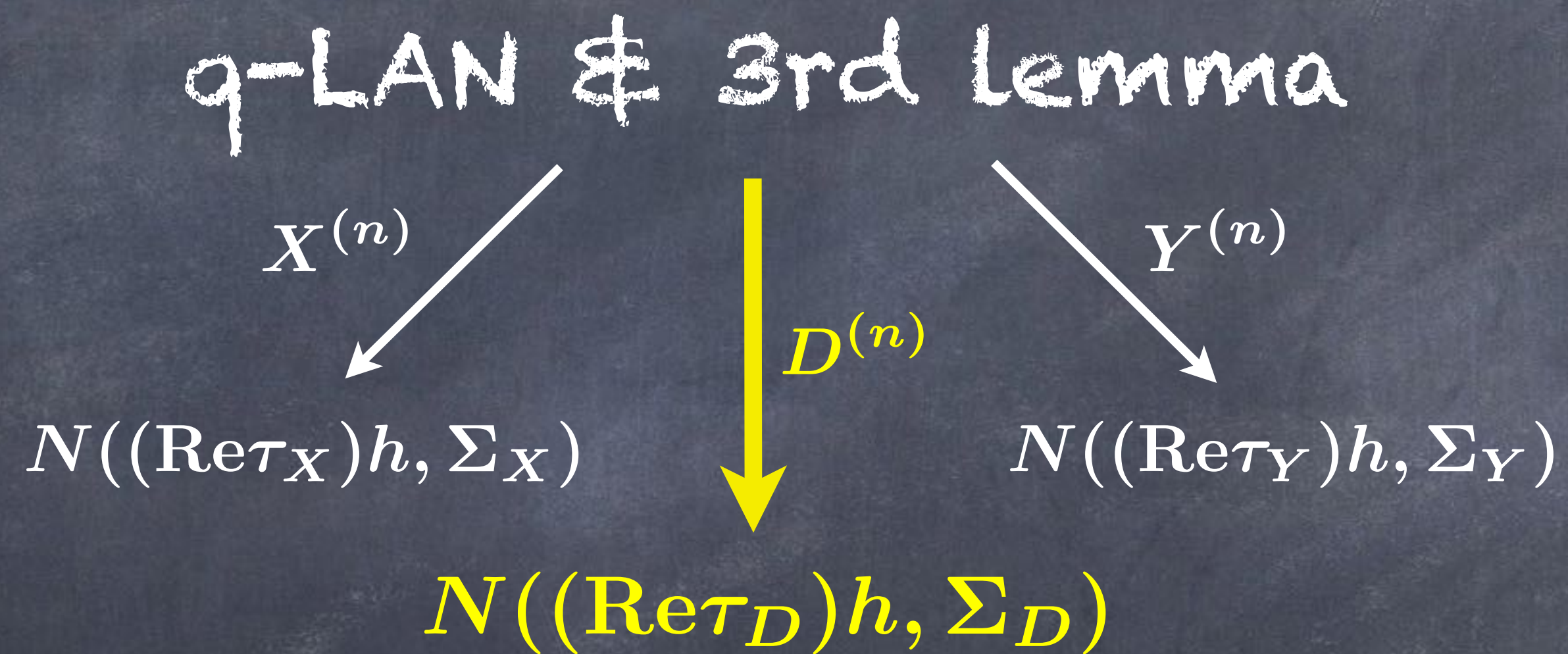
$$\begin{pmatrix} X^{(n)} \\ \Delta^{(n)} \end{pmatrix} \underset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma & \tau \\ \tau^* & J \end{pmatrix} \right)$$

Then $(\rho_{\theta_0+h/\sqrt{n}}^{(n)} \triangleleft \rho_{\theta_0}^{(n)} \text{ and })$

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

The moral:
q-LAN model is
locally
asymptotically
similar to q-
Gaussian shift
model

Embedding $\rho_\theta^{\otimes n}$ into Gaussian shift models



Choose $D^{(n)}$ suitably so that the Holevo bound of the resulting q-Gaussian is identical to that of ρ_θ

Asymptotic q-Repr. Theorem

Assume that [AoS, §1 (2023) 1159] $\begin{pmatrix} X^{(n)} \\ \Delta^{(n)} \end{pmatrix} \overset{\rho_{\theta_0}^{(n)}}{\rightsquigarrow} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma & \tau \\ \tau^* & J \end{pmatrix}\right)$

1) $\{\rho_{\theta}^{(n)}\}$ is q-LAN and \mathcal{D} -extendible at θ_0

2) Seq. of POVMs $M^{(n)} = \{M^{(n)}(B)\}_{B \in \mathcal{B}(\mathbb{R}^s)}$ enjoys

$$\text{Tr} \rho_{\theta_0 + h/\sqrt{n}}^{(n)} M^{(n)} \overset{h}{\rightsquigarrow} \exists \mathcal{L}_h$$

Then $\exists M^{(\infty)} = \{M^{(\infty)}(B)\}_{B \in \mathcal{B}(\mathbb{R}^s)}$ on CCR($\text{Im } \Sigma$) s.t.

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

where $\phi_h \sim N((\text{Re } \tau) h, \Sigma)$

Applications

- asymptotic quantum representation theorem converts a statistical problem for the local parameter h into another one for the limiting q -Gaussian shift model
- asymptotic representation bound beyond iid
- asymptotic regularity and asymptotic minimax theorems that exclude quantum superefficiency

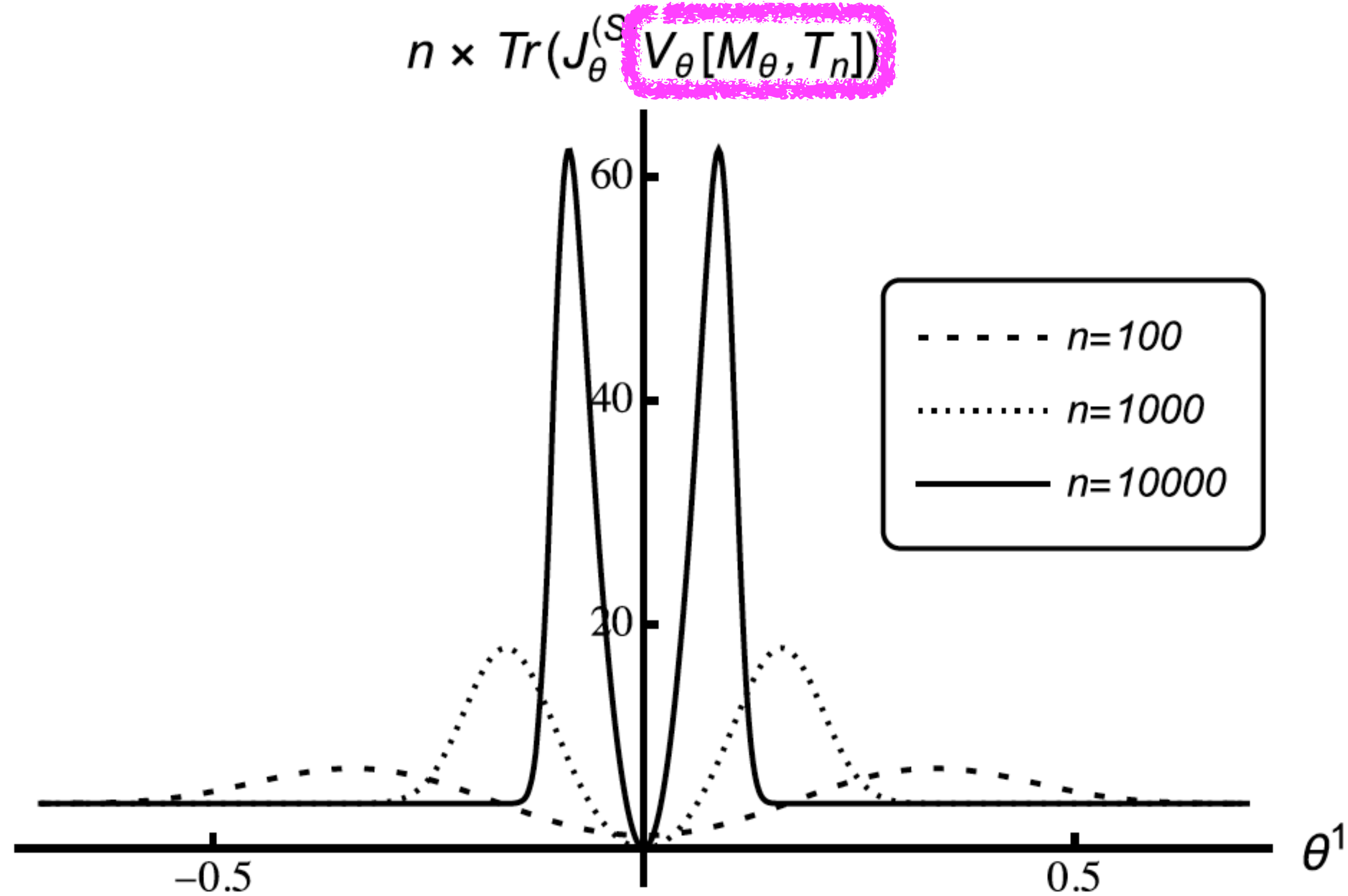


FIG. 1. Weighted trace of covariance matrix of the quantum Hodes estimator T_n with the weight $J_\theta^{(S)}$ for the spin coherent state model \mathcal{S} , based on the means of samples of size 100 (dashed), 1000 (dotted), and 10,000 (solid) observations. For reference, the corresponding Holevo bound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

Summary

- Quantum information geometry naturally led us to the quantum Lebesgue decomposition
- With additional notions (such as quantum weak convergence and quantum contiguity), we derived quantum LeCam third lemma and asymptotic quantum representation theorem
- They establish solid foundations of the theory of quantum local asymptotic normality, providing a powerful tool to cope with asymptotics in the quantum domain

"Thank you for your attention"



- Akio Fujiwara