SUPPLEMENTARY MATERIAL TO "LOCAL ASYMPTOTIC EQUIVALENCE OF PURE STATES ENSEMBLES AND QUANTUM GAUSSIAN WHITE NOISE"

BY CRISTINA BUTUCEA, MĂDĂLIN GUȚĂ AND MICHAEL NUSSBAUM *

CREST, ENSAE, Université Paris-Saclay; University of Nottingham and Cornell University

Equation numbers without section refer to the main part of the article.

APPENDIX A: FURTHER DISCUSSION

A.1. Classical models. Here we review several asymptotic normality results for classical models which are analogous to the quantum models investigated in the paper.

A classical statistical model is defined as a family of probability distributions $\mathcal{Q} = \{\mathbb{P}_f : f \in \mathcal{W}\}$ on a measurable space $(\mathcal{X}, \mathcal{A})$, indexed by an unknown, possibly infinite dimensional parameter f to be estimated, which belongs to a parameter space \mathcal{W} . In the asymptotic framework considered here we assume that we are given a (large) number n of independent, identically distributed samples X_1, \ldots, X_n from \mathbb{P}_f , from which we would like to estimate f. If $d: \mathcal{W} \times \mathcal{W} \to \mathbb{R}_+$ is a chosen loss function, then the risk of an estimator $\hat{f}_n = \hat{f}_n(X_1, \ldots, X_n)$ is

$$R(\hat{f}_n, f) = \mathbb{E}_f \left[d(\hat{f}_n, f)^2 \right].$$

In nonparametric statistics, the parameter of the model f is often a function that belongs to a smoothness class. We consider two classes \mathcal{W} : the periodic Sobolev class $\mathcal{S}^{\alpha}(L)$ of functions on [0,1] with smoothness $\alpha > 1/2$, and the Hölder class $\Lambda^{\alpha}(L)$, with smoothness $\alpha > 0$. For any $f \in \mathbb{L}_2[0,1]$, let $\{f_j, j \in \mathbb{Z}\}$ be the set of Fourier coefficients with respect to the standard trigonometric basis. The classes are defined as

$$\mathcal{S}^{\alpha}(L) := \left\{ f: [0,1] \to \mathbb{R} : \sum_{j \in \mathbb{Z}} \int |f_j|^2 |j|^{2\alpha} du \le L \right\}.$$

and

$$\Lambda^{\alpha}(L) := \{ f : [0,1] \to \mathbb{R} : |f(x) - f(y)| \le L|x - y|^{\alpha}, x, y \in [0,1] \}.$$

In addition, when densities f are considered, we will assume that \mathcal{W} includes an additional restriction to a class

$$\mathcal{D}_{\varepsilon} = \left\{ f : [0,1] \to [\varepsilon,\infty) : \int_{[0,1]} f(x) dx = 1 \right\}$$

for some $\varepsilon > 0$.

Density model. The classical density model consists of n observations X_1, \ldots, X_n which are independent, identically distributed (i.i.d.) with common probability density f

$$\mathcal{P}_n = \left\{ \mathbb{P}_f^{\otimes n} : f \in \mathcal{W} \right\}.$$

^{*}Supported in part by NSF Grant DMS-1407600

Gaussian regression model with fixed equidistant design. In this model, we observe $Y_1, ..., Y_n$ such that

$$Y_i = f^{1/2}\left(\frac{i}{n}\right) + \xi_i, \quad i = 1, ..., n,$$

where the errors $\xi_1, ..., \xi_n$ are i.i.d., standard Gaussian variables. Denote the Gaussian regression model by

$$\mathcal{R}_n = \left\{ \bigotimes_{i=1}^n \mathcal{N}\left(f^{1/2}\left(\frac{1}{n}\right), 1\right) : f \in \mathcal{W} \right\}.$$

Gaussian white noise model. In this model the square-root density $f^{1/2}$ is observed with Gaussian white noise of variance n^{-1} , i.e.

(A.1)
$$dY_t = f^{1/2}(t)dt + \frac{1}{\sqrt{n}}dW_t, \quad t \in [0,1].$$

If we denote by \mathbb{Q}_f the probability distribution of $\{Y(t) : t \in [0,1]\}$, the corresponding model is

$$\mathcal{F}_n := \{\mathbb{Q}_f : f \in \mathcal{W}\}.$$

Gaussian sequence model. In this model we observe a sequence of Gaussian random variables with means equal to the coefficients of $f^{1/2}$ in some orthonormal basis of $\mathbb{L}_2[0,1]$ for $f \in \mathcal{F}$

(A.2)
$$y_j = \theta_j(f^{1/2}) + \frac{1}{\sqrt{n}}\xi_j, \qquad i = 1, 2, \dots$$

where $\{\xi_i\}_{i\geq 1}$ are Gaussian i.i.d. random variables. We denote this model

$$\mathcal{N}_n = \left\{ \bigotimes_{j \ge 1} \mathcal{N}\left(\theta_j\left(f^{1/2}\right), \frac{1}{n}\right) : f \in \mathcal{W} \right\}.$$

In [16] it was shown that the sequences of models \mathcal{P}_n and \mathcal{F}_n are asymptotically equivalent in the sense that their Le Cam distance converges to zero as $n \to \infty$ when $\mathcal{W} = \Lambda^{\alpha}(L) \cap \mathcal{D}_{\varepsilon}$ with $\alpha > 1/2$; in [3], a similar result was established for \mathcal{R}_n and \mathcal{F}_n (more precisely, with $f^{1/2}$ any real valued function $f^{1/2} \in \Lambda^{\alpha}(L)$). Later, [20] showed that models \mathcal{F}_n and \mathcal{N}_n are asymptotically equivalent over periodic Sobolev classes $f^{1/2} \in \mathcal{S}^{\alpha}(L)$ with smoothness $\alpha >$ 1/2. Among many other results [8] considered generalized linear models, [2] regression models with random design and [18] multivariate and random design, [7] compared the stationary Gaussian process with the Gaussian white noise model \mathcal{F}_n . In [17] sharp rates of convergence are obtained for the equivalence of \mathcal{P}_n and \mathcal{F}_n , including also Poisson process models.

In all classical results, the underlying nonparametric function was assumed to belong to a smoothness class in order to establish asymptotic equivalence of models. In the quantum setup of pure states and Gaussian states that we discuss in Section 4, no such smoothness assumption is needed.

A.2. Quadratic Functionals

. The elbow phenomenon. The change of regime which occurs in the optimal MSE rate η_n^2 in (26) has been described as the elbow phenomenon in the literature [4]. In the classical

Gaussian sequence model, it takes the following shape. Consider observations introduced in (15):

$$y_j = \vartheta_j + n^{-1/2} \xi_j, \ j = 1, 2, \dots,$$

where $\{\xi_j\}$ are i.i.d. standard normal, and the parameter $\vartheta = (\vartheta_j)_{j=1}^{\infty}$ satisfies a restriction $\sum_{j=1}^{\infty} j^{2\alpha} \vartheta_j^2 \leq L$ for some $\alpha > 0$. For estimation of the quadratic functional $\tilde{F}(\vartheta) = \sum_{j=1}^{\infty} j^{2\beta} \vartheta_j^2$ with $\beta < \alpha$, the minimax MSE rate of convergence is

$$\tilde{\eta}_n^2 = \begin{cases} n^{-1} \text{ if } \alpha \ge 2\beta + 1/4 \\ n^{-2\frac{4(\alpha-\beta)}{4\alpha+1}} \text{ if } \beta < \alpha < 2\beta + 1/4 \end{cases} = n^{-2\tilde{r}} \text{ for } \tilde{r} = \min\left(\frac{1}{2}, \frac{4(\alpha-\beta)}{4\alpha+1}\right)$$

(cf [15] and references cited therein). The same rate holds for estimation of the squared L_2 norm of the β -th derivative of a density in an α -Hölder class, cf. [1]. Comparing with our rate η_n^2 in (26) which can be written $\eta_n^2 = n^{-2r}$ for $r = \min\left(\frac{1}{2}, \frac{4(\alpha-\beta)}{4\alpha}\right)$, we see that both rates exhibit the elbow phenomenon, but at different critical values for (α, β) , and the rate for the quantum case is slightly faster in the region $\alpha < 2\beta + 1/4$.

A tail functional of a discrete distribution. Our method of proof for the optimal rate $\eta_n^2 = n^{-2r}$ shows that it is also the optimal rate in the following non-quantum problem: suppose $P = \{p_j\}_{j=0}^{\infty} j^{2\alpha} p_j \leq L$, and the aim is to estimate the linear functional $F_0(P) = \sum_{j=0}^{\infty} j^{2\beta} p_j$ on the basis of n i.i.d. observations X_1, \ldots, X_n having law P. Indeed, Theorem 5.4 shows that the estimator $\hat{F}_n = \sum_{j=0}^{N} j^{2\beta} \hat{p}_j$ with $\hat{p}_j = n^{-1} \sum_{i=1}^{n} I(X_i = j)$ attains the rate η_n^2 for mean square error, for an appropriate choice of N. On the other hand, the observations X_1, \ldots, X_n are obtained from one specific measurement in the quantum model (18), in such a way that $p_j = |\psi_j|^2$ for $j \geq 0$ and $F_0(P) = F(\psi)$. If the rate η_n^2 is unimprovable in the quantum model then it certainly is in the present derived (less informative) classical model. In the latter model, we note that since $F_0(P)$ is linear and the law P is restricted to a convex body, optimality of the rate η_n^2 can be confirmed by standard methods, e.g. based on the concept of modulus of continuity [5]. The current problem is thus an example where the elbow phenomenon is present for estimation of a linear functional; a specific feature here is that the probability measure P is discrete.

Fuzzy quantum hypotheses. Our method of proof of the lower bound for quadratic functionals, which works in the approximating quantum Gaussian model, utilizes the well-known idea of setting up two prior distributions and then invoking a testing bound between simple hypotheses. This has been described as the method of fuzzy hypotheses in the literature [21]. A summary of the present quantum variant could be as follows. First, the Gaussian quantum model is represented in a fashion analogous to the classical sequence model (15) where the ϑ_j correspond to the displacement parameter u_j in certain Gaussian pure states (the coherent states). These displacement parameters are then assumed to be random as independent, nonidentically distributed normal, for $j = 1, \ldots, N$ where N = o(n). Now Gaussian averaging over the displacements u_j leads to certain non-pure Gaussian states, i.e. the thermal states as the alternative, which happen to commute with the vacuum pure state (corresponding to $u_j = 0$) as the null hypothesis. Even though both are again Gaussian states, by commutation the problem is reduced to testing between two ordinary discrete probability distributions, i.e. the point mass at 0 and a certain geometric distribution with parameter r_j , depending on j = 1, ..., N. The combined error probability for this classical testing problem with N independent observations gives the lower risk bound.

A.3. Nonparametric Testing

. The separation rate $n^{-1/2}$. Recall that for the classical Gaussian sequence model (15), for the testing problem

(A.3)
$$\begin{array}{ccc} H_0: & \vartheta = 0\\ H_1(\varphi_n): & \sum_{j=1}^{\infty} j^{2\alpha} \vartheta_j^2 \leq L \text{ and } \|\vartheta\|_2 \geq \varphi_n \end{array}$$

(Sobolev ellipsoid with an L_2 -ball removed), the separation rate is $\varphi_n = n^{-2\alpha/(4\alpha+1)}$ [13]. We established that $\varphi_n = n^{-1/2}$ is the separation rate for the quantum nonparametric testing problem (28) involving a pure state ρ . While this "parametric" rate for a nonparametric problem is somewhat surprising, it should be noted that there also exist testing problems for classical i.i.d. data with nonparametric alternative where that separation rate applies; cf [13], sec. 2.6.2.

In our case, the rate $n^{-1/2}$ appears to be related to the fast rate $\varphi_n^2 = n^{-1}$ in the following nonparametric classical problem: given n i.i.d. observations X_1, \ldots, X_n having law $P = \{p_j\}_{j=0}^{\infty}$ on the nonnegative integers, the hypotheses are

(A.4)
$$\begin{aligned} H_0: & P = \delta_0 \text{ (the degenerate law at 0)} \\ H_1(\varphi_n): & \|P - \delta_0\|_1 \geq \varphi_n^2. \end{aligned}$$

For that, note first that

$$||P - \delta_0||_1 = 1 - p_0 + \sum_{j=1}^{\infty} p_j = 2(1 - p_0).$$

The likelihood ratio test for δ_0 against any $P \in H_1(\varphi_n)$ rejects if $\max_{1 \le j \le n} X_j > 0$, thus it does not depend on P. The pertaining sum of error probabilities is

$$P\left(\max_{1 \le j \le n} X_j = 0\right) = p_0^n = \left(1 - \frac{1}{2} \|P - \delta_0\|_1\right)^n \le \left(1 - \frac{1}{2}\varphi_n^2\right)^n$$

and with a supremum over $P \in H_1(\varphi_n)$, the upper bound is attained. This means that for $\varphi_n = cn^{-1/2}$, the minimax sum of error probabilities tends to $\exp(-c^2/2)$, so that $\varphi_n^2 = n^{-1}$ is the separation rate here as claimed.

In fact there is a direct connection to the quantum nonparametric testing problem (28): in the latter, for n = 1, consider a measurement defined as follows. Let $\{|\tilde{e}_j\rangle\}_{j=0}^{\infty}$ be an orthonormal basis in \mathcal{H} such that $\rho_0 = |\tilde{e}_0\rangle \langle \tilde{e}_0|$ and consider the POVM $\{|\tilde{e}_j\rangle \langle \tilde{e}_j|\}_{j=0}^{\infty}$; the corresponding measurement yields a probability measure P on the nonnegative integers. Here the state ρ_0 is mapped into δ_0 and an alternative state ρ is mapped into $P = \{p_j\}_{j=0}^{\infty}$ such that $p_0 = \text{Tr}(\rho_0 \rho)$. Condition (27) on the distance of the two states implies (cp (10))

$$\varphi_n \le \|\rho - \rho_0\|_1 = 2\sqrt{1 - \operatorname{Tr}(\rho_0 \rho)} = 2\sqrt{1 - p_0} = \sqrt{2\|P - \delta_0\|_1}$$

so that up to a constant, the testing problem (A.4) is obtained.

In the quantum problem (28), we noted that the optimal test between ρ_0 and a specific alternative ρ depends on ρ , but found that the test (binary POVM) $M_n = \{\rho_0^{\otimes n}, I - \rho_0^{\otimes n}\}$ is

minimax optimal in the sense of the rate and also in the sense of a sharp risk asymptotics. The sharp minimax optimality seems to be a specific result for the quantum case. We note that the optimal test M_n can be realized via a measurement $\{|\tilde{e}_j\rangle \langle \tilde{e}_j|\}_{j=0}^{\infty}$ as described above, applied separately to each component of $\rho^{\otimes n}$, resulting in independent identically distributed r.v.'s X_1, \ldots, X_n . The test M_n then amounts to rejecting H_0 if $\max_{1 \leq j \leq n} X_j > 0$. Note that this measurement is incompatible with the one (B.8) providing the optimal rate for state estimation.

Other separation rates. In our proof of the lower bound for quadratic functionals, we formulate the nonparametric testing problem for pure states (B.17) where the alternative includes the restriction $\sum_{j\geq 0} |\psi_j|^2 j^{2\beta} \geq \eta_n$, and establish that the rate $\eta_n = n^{-1+\beta/\alpha}$ is unimprovable there. Introduce a seminorm

$$\|\psi\|_{2,\beta} = \left(\sum_{j\geq 1} |\psi_j|^2 \, j^{2\beta}\right)^{1/2}$$

(excluding the term for j = 0) and write the restriction as

(A.5)
$$\|\psi\|_{2,\beta} \ge \varphi_n = \eta_n^{1/2};$$

then the case $\beta = 0$ gives (cp (10))

$$\varphi_n^2 \le \sum_{j\ge 1} |\psi_j|^2 = 1 - |\psi_0|^2 = 1 - |\langle \psi|e_0\rangle|^2 = \frac{1}{4} ||e_0\rangle \langle e_0| - |\psi\rangle \langle \psi||_1^2,$$

in other words, for $\rho_0 = |e_0\rangle \langle e_0|$ and $\rho = |\psi\rangle \langle \psi|$, the restriction (A.5) is equivalent to $\|\rho - \rho_0\|_1 \ge 2\varphi_n$. In that sense, the testing problems (28) and (B.17) in are equivalent up to a constant, if $\beta = 0$ and $\rho_0 = |e_0\rangle \langle e_0|$. For $\beta > 0$, the testing problem (B.17) in is a quantum pure state analog of the generalization of the classical problem (A.3) where $\|\vartheta\|_2 \ge \varphi_n$ is replaced by $\|\vartheta\|_{2,\beta} \ge \varphi_n$ (α -ellipsoid with a β -ellipsoid removed); the separation rate in the latter is $\varphi_n = n^{-2(\alpha-\beta)/(4\alpha+1)}$, cf. [13], sec. 6.2.1. In (B.17) the separation rate is $\varphi_n = n^{-1/2+\beta/2\alpha}$, i.e. of the more typical nonparametric form as well.

APPENDIX B: PROOFS

PROOF OF THEOREM 4.1. The direct map channel T_n is defined as an isometric embedding

$$T_n: \mathcal{T}_1(\mathcal{H}^{\otimes_s n}) \to \mathcal{T}_1(\mathcal{F}(\mathcal{H}_0))$$
$$\rho \mapsto V_n \rho V_n^*.$$

where $V_n : \mathcal{H}^{\otimes_s n} \to \mathcal{F}(\mathcal{H}_0)$ is an isometry defined below. Since we deal with pure states, it suffices to prove that

(B.1)
$$\limsup_{n \to \infty} \sup_{|\psi_0\rangle \in \mathcal{H}} \sup_{\|u\| \le \gamma_n} \left\| V_n \psi_u^{\otimes n} - G(\sqrt{n}u) \right\| = 0.$$

We now define the isometric embedding V_n by showing its explicit action on the vectors of an ONB. For any permutation $\sigma \in S_n$, let

$$U_{\sigma}: |u_1\rangle \otimes \cdots \otimes |u_n\rangle \mapsto |u_{\sigma^{-1}(1)}\rangle \otimes \cdots \otimes |u_{\sigma^{-1}(n)}\rangle$$

be the unitary action on $\mathcal{H}^{\otimes n}$ by tensor permutations. Then $P_s := \frac{1}{n!} \sum_{\sigma \in S_n} U_{\sigma}$ is the orthogonal projector onto the subspace of symmetric tensors $\mathcal{H}^{\otimes_s n}$. We construct an orthonormal basis in $\mathcal{H}^{\otimes_s n}$ as follows.

Let $B_0 := \{|e_1\rangle, |e_2\rangle, \dots\}$ be an orthonormal basis in \mathcal{H}_0 . Let $\tilde{\mathbf{n}} = (n_0, \mathbf{n}) = (n_0, n_1, \dots)$ be an infinite sequence of integers such that $\sum_{i\geq 0} n_i = n$, and note that only a finite number of n_i s are different from zero. Then the symmetric vectors

$$|\tilde{\mathbf{n}}\rangle = |n_0, n_1, n_2, \dots\rangle := \sqrt{\frac{n!}{n_0! \cdot n_1! \cdot \dots}} P_s \left[|\psi_0\rangle^{\otimes n_0} \otimes \bigotimes_{i \ge 1} |e_i\rangle^{\otimes n_i} \right]$$

form an ONB of $\mathcal{H}^{\otimes_s n}$.

As discussed in section 2.2.2 the Fock space $\mathcal{F}(\mathcal{H}_0)$ can be identified with the infinite tensor product of one-mode Fock spaces $\bigotimes_{i\geq 1} \mathcal{F}(\mathbb{C}|e_i\rangle)$ which has an orthonormal number basis (or Fock basis) consisting of products of number basis vectors of individual modes

$$|\mathbf{n}\rangle := \bigotimes_{i\geq 1} |n_i\rangle$$

where $n_i \neq 0$ only for a finite number of indices. We define $V_n : \mathcal{H}^{\otimes_s n} \to \mathcal{F}(\mathcal{H}_0)$ as follows

$$V_n: |\tilde{\mathbf{n}}\rangle \mapsto |\mathbf{n}\rangle.$$

Its image consists of states with at most n "excitations", with $|\psi_0\rangle^{\otimes n}$ being mapped to the vacuum state $|\mathbf{0}\rangle$. We would like to show that the embedded state $V_n |\psi_u\rangle^{\otimes n}$ are well approximated by the coherent states $|G(\sqrt{nu})\rangle$ uniformly over the local neighbourhood $||u|| \leq \gamma_n$. For this we will make use of the covariance and functorial properties of the second quantisation construction in order to reduce the non-parametric LAE statement to the corresponding one for 2-dimensional systems.

Let $|u\rangle \in \mathcal{H}_0$ be a fixed unit vector. Let $j: \mathbb{C}^2 \mapsto \mathcal{H}$ be the isometric embedding

$$j:|0
angle\mapsto|\psi_0
angle,\qquad j:|1
angle\mapsto|u
angle$$

and let $j_0 : \mathbb{C}|1\rangle \to \mathcal{H}_0$ be the restriction of j to the one dimensional subspace $\mathbb{C}|1\rangle$. Since second quantisation is functorial under contractive maps, there is a corresponding isometric embedding $J_0 = \Gamma(j_0)$ satisfying

(B.2)
$$J_0: \mathcal{F}(\mathbb{C}|1\rangle) \to \mathcal{F}(\mathcal{H}_0)$$
$$|G(\alpha)\rangle \mapsto |G(j_0(\alpha))\rangle = |G(\alpha u)\rangle.$$

Let $\tilde{V}_n : (\mathbb{C}^2)^{\otimes_s n} \to \mathcal{F}(\mathbb{C}|1\rangle)$ be the isometry constructed in the same way as V_n , where $|0\rangle$ plays the role of $|\psi_0\rangle$ and $\mathbb{C}|1\rangle$ is the analogue of \mathcal{H}_0 . As before, let $|\tilde{\psi}_{\alpha}\rangle = \sqrt{1 - |\alpha|^2}|0\rangle + \alpha|1\rangle$, with $|\alpha| \leq 1$. Then by the properties of the embedding map V_n we have

(B.3)
$$J_0 \tilde{V}_n |\tilde{\psi}_\alpha\rangle^{\otimes n} = V_n |\psi_{\alpha u}\rangle^{\otimes n}$$

From equations (B.2) and (B.3) we find

$$\sup_{|\alpha| \le \gamma_n} \left\| V_n \psi_{\alpha u}^{\otimes n} - G(\sqrt{n\alpha u}) \right\| = \sup_{|\alpha| \le \gamma_n} \left\| \tilde{V}_n \tilde{\psi}_{\alpha}^{\otimes n} - G(\sqrt{n\alpha}) \right\|$$

Since the right-hand side of the above equality is independent of $|u\rangle$ the same equality holds with supremum on the left side taken over all $|u\rangle \in \mathcal{H}_0$ with ||u|| = 1, which is the same as the supremum in equation (B.1). Therefore the LAE for the non-parametric models has been reduced to that of a two-dimensional (qubit) model. This approximation has been established in the more general case of mixed states in [10, 9], but the current case of pure states allows an improvement in rate. The product state $|\tilde{\psi}_{\alpha}\rangle^{\otimes n}$ is mapped into the following pure state on the Fock space $\mathcal{F}(\mathbb{C}|1\rangle)$

$$\tilde{V}_n |\tilde{\psi}_\alpha\rangle^{\otimes n} = \sum_{k=0}^n c_{k,n}(\alpha) |k\rangle, \quad c_{k,n}(\alpha) = \alpha^k (1 - |\alpha|^2)^{(n-k)/2} \sqrt{\binom{n}{k}}.$$

On the other hand, in view of (2) the coherent state can be written as

$$G(\sqrt{n\alpha}) = \sum_{k} c_k(\sqrt{n\alpha}) |k\rangle, \quad c_k(\sqrt{n\alpha}) := \exp(-n|\alpha|^2/2) \frac{(\sqrt{n\alpha})^k}{\sqrt{k!}}.$$

Set $\alpha = \phi_{\alpha} |\alpha|$ where ϕ_{α} is a phase; then it follows that $c_{k,n}(\alpha) = \phi_{\alpha}^{k} c_{k,n}(|\alpha|)$ and $c_{k}(\sqrt{n\alpha}) = \phi_{\alpha}^{k} c_{k}(\sqrt{n}|\alpha|)$. With this we have

(B.4)
$$\left\|\tilde{V}_{n}\tilde{\psi}_{\alpha}^{\otimes n} - G\left(\sqrt{n\alpha}\right)\right\|^{2} = \sum_{k=0}^{\infty} \left|c_{k,n}\left(\alpha\right) - c_{k}\left(\sqrt{n\alpha}\right)\right|^{2}$$
$$= \sum_{k=0}^{\infty} \left|c_{k,n}\left(|\alpha|\right) - c_{k}\left(\sqrt{n}\left|\alpha\right|\right)\right|^{2}.$$

Let X be a binomial r.v. with parameters $n, |\alpha|^2$ and Y be a Poisson r.v. with parameter $n |\alpha|^2$. Note that $c_{k,n}(|\alpha|) = P(X = k)^{1/2}$ and $c_k(\sqrt{n} |\alpha|) = P(Y = k)^{1/2}$, and that therefore (B.4) is the squared Hellinger distance between these two laws. According to Theorem 1.3.1 (ii) in [19] we have

$$\sum_{k=0}^{\infty} \left| c_{k,n} \left(|\alpha| \right) - c_k \left(\sqrt{n} |\alpha| \right) \right|^2 \le 3 \left| \alpha \right|^4.$$

Since $|\alpha| \leq \gamma_n = o(1)$, we have shown the first part of LAE in which the i.i.d. and Gaussian models are expressed in terms of the local parameter $|u\rangle$

(B.5)
$$\limsup_{n \to \infty} \sup_{|\psi_0\rangle \in \mathcal{H}} \sup_{\|u\| \le \gamma_n} \left\| V_n \psi_u^{\otimes n} - G(\sqrt{n}u) \right\| = 0.$$

Conversely, we define the reverse channel $S_n : \mathcal{T}_1(\mathcal{F}(\mathcal{H}_0)) \to \mathcal{T}_1(\mathcal{H}^{\otimes_s n})$ as follows. Let P_n denote the orthogonal projection in $\mathcal{F}(\mathcal{H}_0)$ onto the image space of V_n , i.e. the subspace with total excitation number at most n

$$\mathcal{F}_{\leq n}(\mathcal{H}_0) := \operatorname{Lin}\{|n_1, n_2, \dots\rangle : \sum_{i \geq 1} n_i \leq n\}.$$

Let $R_n : \mathcal{F}(\mathcal{H}_0) \to \mathcal{H}^{\otimes_s n}$ be a right inverse of V_n , i.e. $R_n V_n = \mathbf{1}$. Then the reverse channel is defined as

$$S_n(\rho) = R_n P_n \rho P_n R_n^* + \operatorname{Tr}(\rho(1-P_n)) |\psi_0\rangle \langle \psi_0 |^{\otimes n}$$

Operationally, the action of S_n consists of two steps. We first perform a projection measurement with projections P_n and $(\mathbf{1} - P_n)$; if the first outcome occurs the conditional state of the system is $P_n\rho P_n/\text{Tr}(P_n\rho)$, while if the second outcome occurs the state is $(\mathbf{1} - P_n)\rho(\mathbf{1} - P_n)/\text{Tr}((\mathbf{1} - P_n)\rho)$. In the second stage, if the first outcome was obtained we map the projected state through the map R_n into a state in $\mathcal{H}^{\otimes_s n}$, while if the second outcome was obtained, we prepare the fixed state $|\psi_0\rangle\langle\psi_0|^{\otimes n}$.

When applied to the pure Gaussian states $|G(\sqrt{n}u)\rangle$, the output of S_n is the mixed state

$$S_n(|G(\sqrt{n}u)\rangle\langle G(\sqrt{n}u)|) = p_u^n |\phi_u^n\rangle\langle \phi_u^n| + (1-p_u^n)|\psi_0\rangle\langle \psi_0|^{\otimes n}$$

where

$$|\phi_u^n\rangle := \frac{R_n P_n |G(\sqrt{n}u)\rangle}{\sqrt{p_u^n}}, \quad p_u^n = \|P_n G(\sqrt{n}u)\|^2.$$

The key observation is that the Gaussian states are almost completely supported by the subspace $\mathcal{F}_{\leq n}(\mathcal{H}_0)$, uniformly with respect to the ball $||u|| \leq \gamma_n$. Indeed, since $V_n \psi_u^{\otimes n}$ is in $\mathcal{F}_{\leq n}(\mathcal{H}_0)$, from (B.5) and the properties of projections it follows

$$\limsup_{n \to \infty} \sup_{\|\psi_0\rangle} \sup_{\|u\| \le \gamma_n} \left\| P_n G\left(\sqrt{n}u\right) - G\left(\sqrt{n}u\right) \right\| = 0,$$

so that

(B.6)
$$\limsup_{n \to \infty} \sup_{|\psi_0\rangle} \sup_{\|u\| \le \gamma_n} (1 - p_u^n) = 0.$$

Now again from (B.5) and the fact that R_n is the inverse of V_n it follows

$$\limsup_{n \to \infty} \sup_{|\psi_0\rangle} \sup_{\|u\| \le \gamma_n} \left\| \psi_u^{\otimes n} - R_n P_n G\left(\sqrt{n}u\right) \right\| = 0,$$

which in conjunction with (B.6) implies

$$\limsup_{n \to \infty} \sup_{|\psi_0\rangle} \sup_{\|u\| \le \gamma_n} \left\| S_n(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)|) - |\psi_u\rangle \langle \psi_u|^{\otimes n} \right\|_1 = 0.$$

This completes the proof of (23).

PROOF OF THEOREM 5.1. According to inequalities (12) and (13) the two distances are equivalent on pure states, so it suffices to prove the upper bound for the trace-norm distance.

Firstly, a projective operation is applied to each of the *n* copies separately, whose aim is to truncate the state to a finite dimensional subspace of dimension $d_n = [n^{1/(2\alpha+1)}]+1$. Let P_n be the projection onto the subspace \mathcal{H}_n spanned by the first d_n basis vectors $\{|e_0\rangle, \ldots, |e_{d_n-1}\rangle\}$. For a given state $|\psi\rangle$ the operation consists of randomly projecting the state with P_n or $(1 - P_n)$, which produces i.i.d. outcomes $O_i \in \{0, 1\}$ with $\mathbb{P}(O_i = 1) = p_n = ||P_n\psi||^2$. The posterior state conditioned on the measurement outcome is

$$|\psi\rangle\langle\psi|\mapsto \begin{cases} |\psi^{(n)}\rangle\langle\psi^{(n)}| := \frac{P_n|\psi\rangle\langle\psi|P_n}{p_n} & \text{with probability} & p_n\\ \\ \frac{(\mathbf{1}-P_n)|\psi\rangle\langle\psi|(\mathbf{1}-P_n)}{1-p_n} & \text{with probability} & 1-p_n \end{cases}$$

Since $|\psi\rangle\langle\psi|\in S^{\alpha}(L)$, the probability $1-p_n$ is bounded as

(B.7)
$$1 - p_n = \sum_{i=d_n}^{\infty} |\psi_i|^2 = \sum_{i=d_n}^{\infty} i^{-2\alpha} i^{2\alpha} |\psi_i|^2 \le d_n^{-2\alpha} \sum_{i=1}^{\infty} i^{2\alpha} |\psi_i|^2 = n^{-2\alpha/(2\alpha+1)} L$$

Let $\tilde{n} = \sum_{i=1}^{n} O_i$ be the number of systems for which the outcome was equal to 1, so that \tilde{n} has binomial distribution $\operatorname{Bin}(n, p_n)$. Then $\mathbb{E}(\tilde{n}/n) = p_n$ and $\operatorname{Var}(\tilde{n}/n) = p_n(1-p_n)/n = O(1/n)$. Therefore $\tilde{n}/n \to 1$ in probability.

In the second step we discard the systems for which the outcome was 0, and we collect those with outcome 1, so that the joint state is $|\psi^{(n)}\rangle\langle\psi^{(n)}|^{\otimes \tilde{n}}$ which is supported by the symmetric subspace $\mathcal{H}_n^{\otimes_s \tilde{n}}$. In order to estimate the truncated state $|\psi^{(n)}\rangle$ (and by implication $|\psi\rangle$), we perform a *covariant* measurement M_n [11] whose space of outcomes is the space of pure states $\hat{\rho}_n = |\hat{\psi}_n\rangle\langle\hat{\psi}_n|$ over \mathcal{H}_n , and the infinitesimal POVM element is

(B.8)
$$M_n(d\hat{\rho}) = \begin{pmatrix} \tilde{n} + d_n - 1 \\ d_n - 1 \end{pmatrix} \hat{\rho}^{\otimes n} d\hat{\rho}.$$

The covariance property means that the unitary group has a covariant action on states and their corresponding probability distributions

$$\mathbb{P}_{U\rho U^*}^{M_n}(d\hat{\rho}) = \operatorname{Tr}(U\rho U^* \cdot d\hat{\rho}) = \mathbb{P}_{\rho}^{M_n}(d(U^*\hat{\rho}U)).$$

Recall that the trace-norm distance squared for pure states is given by $d_1^2(\rho, \rho') := \|\rho - \rho'\|_1^2 = 4(1 - |\langle \psi | \psi' \rangle|^2)$. In [11] it has been shown that, conditionally on \tilde{n} , the risk of the estimator $\hat{\rho}$ with respect to the trace-norm square distance is¹

$$\mathbb{E}^{\tilde{n}}\left[d_1^2(\hat{\rho}_n,\rho^{(n)})\right] = \frac{4(d_n-1)}{d_n+\tilde{n}}$$

Using the triangle inequality we have $d_1^2(\hat{\rho}_n, \rho) \leq 2(d_1^2(\hat{\rho}_n, \rho^{(n)}) + d_1^2(\rho, \rho^{(n)}))$. Since $|\psi^{(n)}\rangle = P_n |\psi\rangle / \sqrt{p_n}$, the bias term is $d_1^2(\rho, \rho^{(n)}) = 4(1-p_n)$, which by (B.7) is bounded by $4n^{-2\alpha/(2\alpha+1)}L$. Therefore

$$\mathbb{E}\left[d_b^2(\hat{\rho}_n,\rho)\right] \le 8\mathbb{E}\left[\frac{(d_n-1)}{d_n+\tilde{n}}\right] + 8n^{-2\alpha/(2\alpha+1)}L.$$

For an arbitrary small $\varepsilon > 0$, we have

$$\mathbb{E}\left[\frac{(d_n-1)}{d_n+\tilde{n}}\right] \le P\left[\frac{\tilde{n}}{n} < 1-\varepsilon\right] + \mathbb{E}\left[\frac{(d_n-1)}{d_n+n\cdot\tilde{n}/n} \cdot I(\frac{\tilde{n}}{n} \ge 1-\varepsilon)\right] \le O\left(\frac{1}{n}\right) + C\frac{d_n}{n}.$$

Putting together the last two upper bounds concludes the proof.

PROOF OF THEOREM 5.2. Let us denote by $R_n^E = \inf_{\widehat{\psi}_n} \sup_{\psi \in S^{\alpha}(L)} \mathbb{E}_{\psi} \left[\|\widehat{\psi}_n - \psi\|_2^2 \right]$ the minimax risk.

The first step is to reduce the set of states $S^{\alpha}(L)$ to a finite hypercube denoted $S_{1:N}^{\alpha}(L)$ consisting of certain "truncated" vectors $|\psi\rangle = \sum_{1 \le i \le N} \psi_i |e_i\rangle$ which have $N \asymp n^{1/(2\alpha+1)}$ non-zero coefficients with respect to the standard basis. This will provide a lower bound to the minimax risk. The coefficients are chosen as

$$\psi_j = \pm \frac{\sigma_j}{\sqrt{n}}, \quad \sigma_j^2 = \lambda (1 - (j/N)^{2\alpha}), \quad j = 1, \dots, N, \text{ for some fixed } \lambda > 0$$

 $^{^1\}mathrm{Reference}$ [11] uses a fidelity distance erroneously called "Bures distance" , which for pure states coincides with the trace-norm distance up to a constant

and we check that they satisfy the ellipsoid constraint

$$\sum_{j\geq 1} |\psi_j|^2 j^{2\alpha} = \frac{\lambda}{n} \sum_{j=1}^N (j^{2\alpha} - j^{4\alpha} N^{-2\alpha}) \le \frac{N^{2\alpha+1}}{n} \frac{2\alpha\lambda}{(2\alpha+1)(4\alpha+1)} (1+o(1)) \le L$$

for an appropriate choice of $\lambda > 0$.

Using the factorisation property (8) we can identify the corresponding Gaussian states with the N-mode state defined by $|\phi\rangle = \bigotimes_{j=1}^{N} |G(\sqrt{n}\psi_j)\rangle$, where the remaining modes are in the vacuum state and can be ignored.

Thus

$$\begin{aligned} R_n^E &\geq \inf_{\widehat{\psi}} \sup_{\psi \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\psi} \left[\| \widehat{\psi} - \psi \|_2^2 \right] \\ &= \inf_{\widehat{\psi}} \sup_{\psi \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\psi} \left[\sum_{j=1}^N | \widehat{\psi}_j - \psi_j |^2 \right] \end{aligned}$$

The supremum over the finite hypercube $S_{1:N}^{\alpha}(L)$ is bounded from below by the average over all its elements. This turns the previous maximal risk into a Bayesian risk, that we can further bound from below as follows:

$$R_n^E \geq \inf_{\widehat{\psi}} \frac{1}{2^N} \sum_{\psi \in S_{1:N}^{\alpha}(L)} \sum_{j=1}^N \mathbb{E}_{\psi} \left[|\widehat{\psi}_j - \psi_j|^2 \right]$$
$$= \inf_{\widehat{\psi}} \sum_{j=1}^N \frac{1}{2^N} \sum_{\psi \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\psi} \left[|\widehat{\psi}_j - \psi_j|^2 \right]$$
$$\geq \sum_{j=1}^N \inf_{\widehat{\psi}_j} \frac{1}{2^N} \sum_{\psi \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\psi} \left[|\widehat{\psi}_j - \psi_j|^2 \right].$$

In the second line $\hat{\psi}$ is the result of an arbitrary measurement and estimation procedure of the state $|G(\sqrt{n}\psi)\rangle$. In the third line each infimum is over procedures for estimating the component ψ_j only; since such procedure may not be compatible with a single measurement, the third line is upper bounded by the second.

The second major step in the proof of the lower bounds is to reduce the risk over all measurements, to testing two simple hypotheses. Let us bound from below the term (B.9) for arbitrary fixed j between 1 and N:

$$\frac{1}{2^{N}} \sum_{\psi \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\psi} \left[|\hat{\psi}_{j} - \psi_{j}|^{2} \right]$$
$$= \frac{1}{2} \left\{ \frac{1}{2^{N-1}} \sum_{\psi \in S_{(j+)}^{\alpha}(L)} \mathbb{E}_{\psi} \left[|\hat{\psi}_{j} - \sigma_{j}/\sqrt{n}|^{2} \right] + \frac{1}{2^{N-1}} \sum_{\psi \in S_{(j-)}^{\alpha}(L)} \mathbb{E}_{\psi} \left[|\hat{\psi}_{j} - (-\sigma_{j}/\sqrt{n})|^{2} \right] \right\}$$
$$(B.10) \qquad = \frac{1}{2} \left\{ \mathbb{E}_{\rho_{j}^{+}} \left[|\hat{\psi}_{j} - \sigma_{j}/\sqrt{n}|^{2} \right] + \mathbb{E}_{\rho_{j}^{-}} \left[|\hat{\psi}_{j} - (-\sigma_{j}/\sqrt{n})|^{2} \right] \right\},$$

where the sum over $\psi \in S^{\alpha}_{(j\pm)}(L)$ means that the j^{th} coordinate is fixed to $\pm \sigma_j/\sqrt{n}$ and all k^{th} coordinates, for $k \neq j$, take values in $\{\sigma_k/\sqrt{n}, -\sigma_k/\sqrt{n}\}$. In the third line, we denote by ρ_j^{\pm} the average state over states in $S^{\alpha}_{(j\pm)}(L)$.

Let us define the testing problem of the two hypotheses $H_0: \rho = \rho_j^+$ against $H_1: \rho = \rho_j^-$. For a given estimator $\hat{\psi}_j$ we construct the test

$$\Delta = I\left(\left|\widehat{\psi}_j - \frac{\sigma_j}{\sqrt{n}}\right| > \left|\widehat{\psi}_j - (-\frac{\sigma_j}{\sqrt{n}})\right|\right),\,$$

and decide H_1 or H_0 , if Δ equals 1 or 0, respectively. By the Markov inequality, we get that

$$\mathbb{E}_{\rho_j^{\pm}}\left[\left|\widehat{\psi}_j - (\pm \frac{\sigma_j}{\sqrt{n}})\right|^2\right] \geq \frac{\sigma_j^2}{n} \mathbb{P}_{\rho_j^{\pm}}\left(\left|\widehat{\psi}_j - (\pm \frac{\sigma_j}{\sqrt{n}})\right| \geq \frac{\sigma_j}{\sqrt{n}}\right).$$

On the one hand,

(B.11)
$$\mathbb{P}_{\rho_j^+}\left(|\widehat{\psi}_j - \sigma_j/\sqrt{n}| \ge \frac{\sigma_j}{\sqrt{n}}\right) \ge \mathbb{P}_{\rho_j^+}(\Delta = 1).$$

Indeed, under $\mathbb{P}_{\rho_j^+}$, the event $\Delta = 1$ implies that $|\widehat{\psi}_j - \frac{\sigma_j}{\sqrt{n}}| > |\widehat{\psi}_j + \frac{\sigma_j}{\sqrt{n}}|$, which further implies by the triangular inequality that

$$\left|\widehat{\psi}_{j} - \frac{\sigma_{j}}{\sqrt{n}}\right| \geq \frac{2\sigma_{j}}{\sqrt{n}} - \left|\widehat{\psi}_{j} + \frac{\sigma_{j}}{\sqrt{n}}\right| \geq \frac{2\sigma_{j}}{\sqrt{n}} - \left|\widehat{\psi}_{j} - \frac{\sigma_{j}}{\sqrt{n}}\right|$$

giving $|\widehat{\psi}_j - \psi_j| \ge \frac{\sigma_j}{\sqrt{n}}$. By a similar reasoning for the $\mathbb{P}_{\rho_j^-}$ distribution we get

(B.12)
$$\mathbb{P}_{\rho_j^-}\left(|\widehat{\psi}_j + \sigma_j/\sqrt{n}| \ge \frac{\sigma_j}{\sqrt{n}}\right) \ge \mathbb{P}_{\rho_j^-}(\Delta = 0)$$

By using (B.11) and (B.12) in (B.10)

$$\frac{1}{2} \left\{ \mathbb{E}_{\rho_j^+} \left[\left| \widehat{\psi}_j - \sigma_j / \sqrt{n} \right|^2 \right] + \mathbb{E}_{\rho_j^-} \left[\left| \widehat{\psi}_j - \left(-\sigma_j / \sqrt{n} \right) \right|^2 \right] \right\} \ge \frac{\sigma_j^2}{2n} \left(\mathbb{P}_{\rho_j^+} (\Delta = 1) + \mathbb{P}_{\rho_j^-} (\Delta = 0) \right).$$

To summarise, we have lower bounded the MSE by the probability of error for testing between the states ρ_j^{\pm} . At closer inspection, these states are of the form $|G(\sigma_j)\rangle\langle G(\sigma_j)|\otimes\rho$ and $|G(-\sigma_j)\rangle\langle G(-\sigma_j)|\otimes\rho$ where ρ is a fixed state obtained by averaging the coherent states of all the modes except j. Recall that the optimal testing error in (9) gives a further bound from below

$$\mathbb{P}_{\rho_j^+}(\Delta = 1) + \mathbb{P}_{\rho_j^-}(\Delta = 0) \ge 1 - \frac{1}{2} \|\rho_j^+ - \rho_j^-\|_1$$

Moreover, the state ρ can be dropped without changing the optimal testing error

$$\|\rho_j^+ - \rho_j^-\|_1 = \||G(\sigma_j)\rangle \langle G(\sigma_j)| - |G(-\sigma_j)\rangle \langle G(-\sigma_j)|\|_1 = 2(1 - \exp(-2\sigma_j^2)).$$

We conclude that

$$\inf_{\widehat{\psi}_j} \frac{1}{2} \left\{ \mathbb{E}_{\rho_j^+} \left[\left| \widehat{\psi}_j - \sigma_j / \sqrt{n} \right|^2 \right] + \mathbb{E}_{\rho_j^-} \left[\left| \widehat{\psi}_j - (-\sigma_j / \sqrt{n}) \right|^2 \right] \right\} \ge \frac{\sigma_j^2}{2n} \cdot \exp(-2\sigma_j^2)$$

and we further use this in (B.10) to get

$$R_n^E \ge \sum_{j=1}^N \frac{\sigma_j^2}{2n} \cdot \exp(-2\sigma_j^2) = \frac{N}{n} \cdot \frac{\lambda}{2N} \sum_{j=1}^N \left(1 - (\frac{j}{N})^{2\alpha}\right) \exp\left(-2 \cdot \lambda (1 - (\frac{j}{N})^{2\alpha})\right) \ge c\frac{N}{n}.$$

Indeed, the average over j is the Riemann sum associated to the integral of a positive function and can be bounded from below by some constant c > 0 depending on α . Moreover, $N/n \approx n^{-2\alpha/(2\alpha+1)}$ and thus we finish the proof of the theorem.

PROOF OF THEOREM 5.3. Let $\tilde{R}_n^E = \inf_{|\hat{\psi}_n\rangle} \sup_{|\psi\rangle \in S^{\alpha}(L)} \mathbb{E}_{\rho} \left[d(\hat{\rho}_n, \rho)^2 \right]$ be the minimax risk for \mathcal{Q}_n .

We bound from below the risk by restricting to (pure) states in a neighbourhood $\Sigma_n(e_0)$ of the basis vector $|e_0\rangle$ defined as follows. As in (20) we write the state and the estimator in terms of their corresponding local vectors

$$|\psi\rangle = \sqrt{1 - ||u||^2} |e_0\rangle + |u\rangle, \qquad |\hat{\psi}\rangle = \sqrt{1 - ||\hat{u}||^2} |e_0\rangle + |\hat{u}\rangle, \qquad |u\rangle, |\hat{u}\rangle \perp |e_0\rangle.$$

Then the neighbourhood is given by $\Sigma_n(e_0) := \{ |\psi_u\rangle : ||u|| \le \gamma_n \}$; we choose $\gamma_n = (\log n)^{-1}$. Such states are described by the local model $\mathcal{Q}_n(e_0, \gamma_n)$, cf. equation (21). The risk is bounded from below by

$$\tilde{R}_n^E \ge \inf_{|\hat{\psi}_n\rangle} \sup_{|\psi\rangle \in S^{\alpha}(L) \cap \Sigma_n(e_0)} \mathbb{E}_{\rho} \left[d(\hat{\rho}_n, \rho)^2 \right].$$

By using the triangle inequality we can assume that $\hat{\psi} \in \Sigma_n(e_0)$, while incurring at most a factor 2 in the risk. By using the quadratic approximation (24) we find that

(B.13)
$$d^{2}(\hat{\rho}_{n},\rho) = k \|u - \hat{u}\|^{2} + o(n^{-1})$$

where k = 1 or k = 4 depending on which distance we use. Since n^{-1} decreases faster than $n^{-2\alpha/(2\alpha+1)}$, the second term does not contribute to the asymptotic rate and can be neglected, so that the problem has been reduced to that of estimating the local parameter u with respect to the Hilbert space distance. To study the latter, we further restrict the set of states to a hypercube similar to the one in the proof of Theorem 5.2, consisting of states $|\psi_u\rangle$ with "truncated" local vectors $|u\rangle = \sum_{1 \le i \le N} u_i |e_i\rangle$ belonging to $S_{1:N}^{\alpha}(L)$. As before, there are $N \simeq n^{1/(2\alpha+1)}$ non-zero coefficients of the form

$$u_j = \pm \frac{\sigma_j}{\sqrt{n}}, \quad \sigma_j^2 = \lambda (1 - (j/N)^{2\alpha}), \quad j = 1, \dots, N.$$

It has been already shown that such vectors belong to the ellipsoid $S^{\alpha}(L)$. Additionally, we show that they also belong to the local ball $\Sigma_n(e_0)$. Indeed

$$||u||^{2} = \sum_{j=1}^{N} |u_{j}|^{2} = \frac{1}{n} \sum_{j=1}^{N} \sigma_{j}^{2} = \frac{1}{n} \sum_{j=1}^{N} \lambda \left(1 - (j/N)^{2\alpha} \right)$$
$$= \frac{N}{n} \left(\frac{1}{N} \sum_{j=1}^{N} \lambda \left(1 - (j/N)^{2\alpha} \right) \right) \le C_{1} \frac{N}{n},$$

where we used that as $N \to \infty$ the expression between the parentheses tens to a finite integral. As N scales as $n^{1/(2\alpha+1)}$, the upper bound becomes

$$||e_0 - \psi_u||^2 \le C_2 n^{-2\alpha/(2\alpha+1)} = o(\gamma_n^2)$$

and the state $|\psi_u\rangle$ belongs to the local ball $\Sigma_n(e_0)$. Taking into account (B.13) the risk is therefore lower bounded as

$$\tilde{R}_n^E \ge \inf_{\hat{u}} \sup_{u \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\rho_u} \left[\|u - \hat{u}\|^2 \right] + o(n^{-1}).$$

where $\rho_u = |\psi_u\rangle\langle\psi_u|$, and the infimum is now taken over the local component $|\hat{u}\rangle$ of an estimator $|\hat{\psi}\rangle = \sqrt{1 - ||\hat{u}||^2} |e_0\rangle + |\hat{u}\rangle$. The first term is further lower bounded by passing to the Bayes risk for the uniform distribution over $S_{1:N}^{\alpha}(L)$, similarly to the proof of Theorem 5.2

$$\tilde{R}_{n}^{E} \geq \sum_{j=1}^{N} \inf_{\widehat{u}_{j}} \frac{1}{2^{N}} \sum_{u \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\psi_{u}} \left[|\widehat{u}_{j} - u_{j}|^{2} \right] + o(n^{-1}).$$

By following the same steps we get

$$\frac{1}{2^{N}} \sum_{u \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\rho_{u}} \left[|\widehat{u}_{j} - u_{j}|^{2} \right] = \frac{1}{2} \left\{ \mathbb{E}_{\tau_{j}^{+}} \left[|\widehat{\psi}_{j} - \sigma_{j}/\sqrt{n}|^{2} \right] + \mathbb{E}_{\tau_{j}^{-}} \left[|\widehat{\psi}_{j} - (-\sigma_{j}/\sqrt{n})|^{2} \right] \right\},$$
(B.14)
$$\geq \frac{\sigma_{j}^{2}}{2n} \left(\mathbb{P}_{\tau_{j}^{+}}(\Delta = 1) + \mathbb{P}_{\tau^{-}}(\Delta = 0) \right) \geq \frac{\sigma_{j}^{2}}{2n} \cdot \left(1 - \frac{1}{2} \|\tau_{j}^{+} - \tau_{j}^{-}\|_{1}\right)$$

where we denote by τ_j^{\pm} the average state over states $|\psi_u\rangle\langle\psi_u|^{\otimes n}$ with $u \in S^{\alpha}_{(j\pm)}(L)$, and Δ is a test for the hypotheses $H_0: \tau = \tau_j^+$ and $H_1: \tau = \tau_j^-$. In the last inequality we used the Helstrom bound [12] which expresses the optimal average error probability for two states discrimination in terms of the norm-one distance between states.

We now make use of the local asymptotic equivalence result in Theorem 4.1. From (23) we know that there exist quantum channels S_n such that

$$\delta_n := \max_{u \in S_{1:N}^{\alpha}(L)} \left\| |\psi_u\rangle \langle \psi_u|^{\otimes n} - S_n \left(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)| \right) \right\|_1 \le \Delta(\mathcal{Q}_n, \mathcal{G}_n) = o(1).$$

By Lemma 3.1 we get

$$\|\tau_j^+ - \tau_j^-\|_1 \le \|\rho_j^+ - \rho_j^-\|_1 + 2\delta_n$$

where ρ_j^{\pm} are the corresponding mixtures in the Gaussian model as defined in the proof of Theorem 5.2. From (B.14) we then get

$$\frac{1}{2^N} \sum_{u \in S_{1:N}^{\alpha}(L)} \mathbb{E}_{\rho_u} \left[|\widehat{u}_j - u_j|^2 \right] \ge \frac{\sigma_j^2}{2n} \cdot \left(1 - \frac{1}{2} \|\rho_j^+ - \rho_j^-\|_1 - \delta_n\right) \ge \frac{\sigma_j^2}{2n} \cdot \left(\exp(-2\sigma_j^2) - \delta_n\right)$$

The rest of the proof follows as in the proof of Theorem 5.2, with the additional remark that

$$\min_{j} \exp(-2\sigma_{j}^{2}) = \lambda(1 - N^{-\alpha}) \asymp \lambda$$

and infinitely larger than δ_n , for *n* large enough.

PROOF OF THEOREM 5.4. The usual bias-variance decomposition yields

$$\mathbb{E}_{\psi}\left(\widehat{F}_{n}-F(\psi)\right)^{2}=\left(\mathbb{E}_{\psi}\widehat{F}_{n}-F(\psi)\right)^{2}+Var_{\psi}\left(\widehat{F}_{n}\right).$$

The bias can be upper bounded as

$$\left| F(\psi) - \mathbb{E}_{\psi} \widehat{F}_n \right| = \left| F(\psi) - \sum_{j=1}^N p_j \cdot j^{2\beta} \right| = \sum_{j \ge N+1} p_j \cdot j^{2\beta} \le N^{-2(\alpha-\beta)} \sum_{j \ge N+1} p_j \cdot j^{2\alpha} \le L N^{-2(\alpha-\beta)}.$$

For the variance, let us note that the vector

$$\widehat{V} = n \cdot (\widehat{p}_1, \dots, \widehat{p}_N, \widehat{p}_{N+1}^*), \text{ with } \widehat{p}_{N+1}^* = n^{-1} \sum_{k=1}^n I(X_k \ge N+1),$$

has a multinomial distribution with parameters n and probability vector $V := (p_1, \ldots, p_N, p_{N+1}^* = \sum_{j \ge N+1} p_j)^\top$. The covariance matrix of a multinomial vector writes $n \cdot (\text{Diag}(V) - V \cdot V^\top)$, where Diag(V) denotes the diagonal matrix with entries from V. In particular, if $\hat{p} := (\hat{p}_1, \ldots, \hat{p}_N)^\top$, $p := (p_1, \ldots, p_N)^\top$ and $B := (1, 2^{2\beta}, \ldots, N^{2\beta})^\top$ then

$$Cov_{\psi}(\widehat{F}_n) = Cov_{\psi}(B^{\top} \cdot \widehat{p}) = B^{\top} \cdot Cov_{\psi}(\widehat{p}) \cdot B = \frac{1}{n} \cdot B^{\top} \cdot (\operatorname{Diag}(p) - p \cdot p^{\top}) \cdot B.$$

This gives

$$Cov_{\psi}(\widehat{F}_n) \leq \frac{1}{n} \cdot B^{\top} \cdot \operatorname{Diag}(p) \cdot B = \frac{1}{n} \sum_{j=1}^{N} p_j \cdot j^{4\beta}.$$

The bound of this last term and the resulting bound of the risk is treated separately for the two cases.

a) Case $\alpha \geq 2\beta$. In that case,

$$\sum_{j=1}^{N} p_j \cdot j^{4\beta} \le \sum_{j=1}^{N} p_j \cdot j^{2\alpha} \le L \text{ implying that } Var(\widehat{F}_n) \le \frac{L}{n}.$$

The upper bound of the risk is, in this case,

$$\mathbb{E}_{\psi}\left(\widehat{F}_n - F(\psi)\right)^2 \le L^2 N^{-4(\alpha-\beta)} + \frac{L}{n}.$$

If we choose $N \simeq n^{1/(4(\alpha-\beta))}$ or larger, then the parametric rate is attained for the risk:

$$\mathbb{E}_{\psi}\left(\widehat{F}_n - F(\psi)\right)^2 = O(1) \cdot n^{-1}.$$

b) Case $\beta < \alpha < 2\beta$. Here we have,

$$Cov_{\psi}(\widehat{F}_n) \leq \frac{1}{n} \sum_{j=1}^N p_j \cdot j^{4\beta} \leq \frac{1}{n} \sum_{j=1}^N p_j \cdot j^{4\beta-2\alpha} j^{2\alpha} p_j \leq \frac{N^{4\beta-2\alpha}}{n} L.$$

The upper bound of the risk becomes

$$\mathbb{E}_{\psi}\left(\widehat{F}_n - F(\psi)\right)^2 \le L^2 N^{-4(\alpha-\beta)} + \frac{N^{4\beta-2\alpha}}{n}L.$$

The optimal choice of the parameter N that balances the two previous terms is $N \simeq n^{1/(2\alpha)}$, giving the attainable rate for the quadratic risk

$$\mathbb{E}_{\psi}\left(\widehat{F}_n - F(\psi)\right)^2 = O(1) \cdot n^{-2(1-\beta/\alpha)}.$$

Cases a) and b) together prove that the rate η_n^2 is attainable.

PROOF OF THEOREM 5.5. Denote by $R_n^F = \inf_{\widehat{F}_n} \sup_{\psi \in S^{\alpha}(L)} \eta_n^{-2} \cdot \mathbb{E}_{\psi} \left(\widehat{F}_n - F(\psi) \right)^2$ the minimax risk.

The case **a**) where $\alpha \geq 2\beta$ reduces to the Cramér-Rao bound that proves that the parametric rate 1/n is always a lower bound for the mean square error for estimating $F(\psi)$.

We prove that in the case **b**) where $\beta < \alpha < 2\beta$, this bound from below increases to $n^{-2(1-\beta/\alpha)}$ (up to constants). By the Markov inequality,

(B.15)
$$\eta_n^{-2} \cdot \mathbb{E}_{\psi} \left(\widehat{F}_n - F(\psi) \right)^2 \ge \frac{1}{4} \cdot \mathbb{P}_{\psi} \left(|\widehat{F}_n - F(\psi)| \ge \frac{\eta_n}{2} \right).$$

Let us restrict the set of pure states $S^{\alpha}(L)$ to its intersection with the local model $Q_n(e_0, \gamma_n)$ (see equation (21)) where $|\psi_u\rangle = \sqrt{1 - ||u||^2} \cdot |e_0\rangle + |u\rangle$ is such that $||u|| \leq \gamma_n$, with $\gamma_n = (\log n)^{-1}$. In other words, u belongs to the set

$$s^{\alpha}(L,\gamma_n) = \left\{ u \in \ell_2(\mathbb{N}^*) : \sum_{j \ge 1} |u_j|^2 j^{2\alpha} \le L \text{ and } ||u|| \le \gamma_n \right\}.$$

Using the fact that $F(e_0) = 0$, we have

$$\sup_{\psi \in S^{\alpha}(L)} \frac{1}{4} \cdot \mathbb{P}_{\psi} \left(|\widehat{F}_{n} - F(\psi)| \geq \frac{\eta_{n}}{2} \right) \\
\geq \frac{1}{4} \max \left\{ \mathbb{P}_{e_{0}} \left(|\widehat{F}_{n}| \geq \frac{\eta_{n}}{2} \right), \sup_{u \in s^{\alpha}(L,\gamma_{n}), F(\psi_{u}) \geq \eta_{n}} \mathbb{P}_{\psi_{u}} \left(|\widehat{F}_{n} - F(\psi_{u})| \geq \frac{\eta_{n}}{2} \right) \right\} \\
\geq \frac{1}{8} \left\{ \mathbb{P}_{e_{0}} \left(|\widehat{F}_{n}| \geq \frac{\eta_{n}}{2} \right) + \sup_{u \in s^{\alpha}(L,\gamma_{n}), F(\psi_{u}) \geq \eta_{n}} \mathbb{P}_{\psi_{u}} \left(|\widehat{F}_{n} - F(\psi_{u})| \geq \frac{\eta_{n}}{2} \right) \right\} \\$$
(B.16)
$$\geq \frac{1}{8} \left\{ \mathbb{P}_{e_{0}} \left(|\widehat{F}_{n}| \geq \frac{\eta_{n}}{2} \right) + \sup_{u \in s^{\alpha}(L,\gamma_{n}), F(\psi_{u}) \geq \eta_{n}} \mathbb{P}_{\psi_{u}} \left(|\widehat{F}_{n}| < \frac{\eta_{n}}{2} \right) \right\}$$

where in the last inequality we used that $|\hat{F}_n| < \eta_n/2$ and $F(\psi_u) \ge \eta_n$ imply $|\hat{F}_n - F(\psi_u)| \ge \eta_n/2$. Note also that $F(\psi_u) = F(u)$ for $|u\rangle \in \mathcal{H}_0$; we now consider the testing problem with hypotheses

(B.17)
$$\begin{cases} H_0: & |u\rangle = |0\rangle \\ H_1(\alpha, L, \gamma_n, \eta_n): & |u\rangle, \text{ with } u \in s^{\alpha}(L, \gamma_n) \text{ and } F(u) \ge \eta_n. \end{cases}$$

Let $\Delta = \Delta(\eta_n) = I(|\hat{F}_n| \ge \eta_n/2)$ be the test that accepts the null hypothesis when $\Delta = 0$ and rejects the null hypothesis when $\Delta = 1$. Then the right-hand side of (B.16) is lower bounded by the sum of the error probability of type I and of the maximal error probability

15

of type II of Δ . We can describe Δ as a binary POVM $M = (M_0, M_1)$, depending on η_n : $M(\eta_n) = (M_0(\eta_n), M_1(\eta_n))$. Thus,

(B.18)
$$\mathbb{P}_{e_0}\left(|\widehat{F}_n| \ge \frac{\eta_n}{2}\right) = \operatorname{Tr}(|e_0\rangle \langle e_0|^{\otimes n} \cdot M_1)$$

and

(B.19)
$$\mathbb{P}_{\psi_u}\left(|\widehat{F}_n| < \frac{\eta_n}{2}\right) = \operatorname{Tr}(|\psi_u\rangle\langle\psi_u|^{\otimes n} \cdot M_0).$$

By putting together (B.15)-(B.19), we get that the minimax risk has the lower bound

$$R_n^F \ge \frac{1}{8} \inf_M \left(\langle e_0^{\otimes n} | M_1 | e_0^{\otimes n} \rangle + \sup_{u \in s^{\alpha}(L, \gamma_n), F(u) \ge \eta_n} \langle \psi_u^{\otimes n} | M_0 | \psi_u^{\otimes n} \rangle \right).$$

Now, using the local asymptotic equivalence Theorem 4.1 with respect to the state $|\psi_0\rangle := |e_0\rangle$ we map the i.i.d. ensemble $|\psi_u\rangle^{\otimes n}$ to the Gaussian state $|G(u)\rangle \in \mathcal{F}(\mathcal{H}_0)$. The lower bound becomes

(B.20)
$$R_n^F \ge \frac{1}{8} \inf_M \left(\langle 0|M_1|0\rangle + \sup_{u \in s^\alpha(L,\gamma_n), F(u) \ge \eta_n} \langle G(\sqrt{n}u)|M_0|G(\sqrt{n}u)\rangle \right) + o(1)$$

where the infimum is taken over tests $M = (M_0, M_1)$ and the o(1) terms stems from the vanishing Le Cam distance $\Delta(\mathcal{Q}_n(e_0, \gamma_n), \mathcal{G}_n(e_0, \gamma_n))$. The lower bound has been transformed into a testing problem for the Gaussian model.

In order to bound from below the maximal error probability of type II, we define a prior distribution on the set of alternatives and average over the whole set with respect to this a priori distribution. Similarly to the classical proofs of lower bounds, our construction will lead to a test of simple hypotheses: the former null and the constructed averaged state. Assume that $\{u_j\}_{j\geq 1}$ are all independently distributed, such that u_j has a complex (bivariate) Gaussian distribution $N_2(0, \frac{1}{2}\sigma_j^2 \cdot I_2)$ for all j from 1 to N, and that $u_j = 0$ for all j > N, where I_2 is the 2 × 2 identity matrix. The σ_j^2 are defined as

(B.21)
$$\sigma_j^2 = \lambda \left(1 - \left(\frac{j}{N}\right)^{2\alpha} \right)_+,$$

where $\lambda, N > 0$ are selected such that

(B.22)
$$\sum_{j\geq 1} j^{2\alpha} \sigma_j^2 = L(1-\varepsilon) \text{ and } \sum_{j\geq 1} j^{2\beta} \sigma_j^2 = n^{-1+\beta/\alpha} (1+\varepsilon),$$

for an arbitrary $\varepsilon > 0$. Let us denote by Π the joint prior distribution of $\{u_j\}_{j \ge 1}$.

Such a choice of the prior distribution was first introduced in [6] for establishing sharp minimax risk bounds for nonparametric testing in the Gaussian white noise model. This construction represents an analog of the prior distribution used in Pinsker's theory for sharp estimation of functions. In our case, using a Gaussian prior as an alternative hypothesis leads to the well-known Gaussian thermal state.

The essence of this construction is that the random vectors $u = \{u_j\}_{j\geq 1}$ concentrate asymptotically, with probability tending to 1, on the spherical segment

$$\{u \in \ell_2(\mathbb{N}) : C \, n^{-1} \le \|u\|^2 \le C \, n^{-1}(1+2\varepsilon')\},\$$

for $\varepsilon' > 0$ depending on ε and some constant C > 0 depending on α and β described later on, and on the alternative set of hypothesis, $H_1(\alpha, L, \gamma_n, \eta_n)$. Note that the spherical segment is included in the set $||u|| \leq \gamma_n$, as $\gamma_n = (\log n)^{-1} \gg n^{-1/2}$. The asymptotic concentration is proved by the following lemma.

LEMMA B.1. A unique solution (λ, N) of (B.21), (B.22), exists for n large enough and admits an asymptotic expansion with respect to n

$$\lambda \sim n^{-1-1/2\alpha} C_{\lambda} \frac{(1+\varepsilon)^{(\alpha+1/2)/(\alpha-\beta)}}{(1-\varepsilon)^{(\beta+1/2)/(\alpha-\beta)}}, \ C_{\lambda} = \frac{((2\beta+1)(2\beta+2\alpha+1))^{(\alpha+1/2)/(\alpha-\beta)}}{2\alpha(L(2\alpha+1)(4\alpha+1))^{(\beta+1/2)/(\alpha-\beta)}}$$
$$N \sim n^{1/2\alpha} C_N \left(\frac{1-\varepsilon}{1+\varepsilon}\right)^{1/(2(\alpha-\beta))}, \ C_N = \left(\frac{L(2\alpha+1)(4\alpha+1)}{(2\beta+1)(2\beta+2\alpha+1)}\right)^{1/(2(\alpha-\beta))}.$$

The independent complex Gaussian random variables $u_j \sim N_2(0, \frac{1}{2}\sigma_j^2 I_2)$, with σ_j 's and (λ, N) given in (B.21), (B.22), are such that, for an arbitrary $\varepsilon > 0$,

(B.23)
$$\mathbb{P}\left(C \, n^{-1} \le \sum_{j=1}^{N} |u_j|^2 \le C \, n^{-1} (1+2\varepsilon')\right) \to 1,$$

(B.24)
$$\mathbb{P}\left(\sum_{j=1}^{N} j^{2\alpha} |u_j|^2 \le L\right) \to 1,$$

(B.25)
$$\mathbb{P}\left(\sum_{j=1}^{N} j^{2\beta} |u_j|^2 \ge n^{-1+\beta/\alpha}\right) \to 1,$$

where $C = C_{\lambda} \cdot C_N \cdot 2\alpha/(2\alpha + 1)$ is a positive constant depending on α and β , and $\varepsilon' > 0$ depends only on ε .

PROOF OF LEMMA B.1. The solution of the problem (B.21), (B.22) can be found in [6] (see also [14], Lemma A.1) for $\beta = 0$; a similar reasoning applies here. Let us prove that the random variables $\{u_j\}_{j=1,\dots,N}$ satisfy (B.23) to (B.25). We have

(B.26)
$$\sum_{j=1}^{N} \sigma_j^2 = \lambda \sum_{j=1}^{N} \left(1 - \left(\frac{j}{N}\right)^{2\alpha} \right) \sim \lambda N \frac{2\alpha}{2\alpha + 1} \sim C_\lambda C_N \frac{2\alpha}{2\alpha + 1} n^{-1} (1 + \varepsilon)^{\alpha/(\alpha - \beta)} (1 - \varepsilon)^{-\beta/(\alpha - \beta)} = C n^{-1} (1 + \varepsilon'),$$

where we denote $\varepsilon' = (1 + \varepsilon)^{\alpha/(\alpha-\beta)}(1 - \varepsilon)^{-\beta/(\alpha-\beta)} - 1$ which is positive for all $\varepsilon \in (0, 1)$. Note that $E |u_j|^2 = \sigma_j^2$ and $Var\left(|u_j|^2\right) = \sigma_j^4$. We have

$$\mathbb{P}\left(C n^{-1} \le \sum_{j=1}^{N} |u_j|^2 \le C n^{-1} (1+2\varepsilon')\right) = 1 - \mathbb{P}\left(\sum_{j=1}^{N} |u_j|^2 < C n^{-1}\right) - \mathbb{P}\left(|u_j|^2 > C n^{-1} (1+2\varepsilon')\right)$$

Now, by the Markov inequality,

$$\begin{split} \mathbb{P}\left(\sum_{j=1}^{N} |u_{j}|^{2} < C \, n^{-1}\right) &= \mathbb{P}\left(\sum_{j=1}^{N} (|u_{j}|^{2} - \sigma_{j}^{2}) < C \, n^{-1} - C \, n^{-1}(1 + \varepsilon' + o(1))\right) \\ &\leq \mathbb{P}\left(\sum_{j=1}^{N} (\sigma_{j}^{2} - |u_{j}|^{2}) > C \, n^{-1}(\varepsilon' + o(1))\right) \\ &\leq \frac{\sum_{j=1}^{N} Var(|u_{j}|^{2})}{C^{2} \, n^{-2} \varepsilon'^{2}/2} \leq \frac{2\sum_{j=1}^{N} \sigma_{j}^{4}}{C^{2} \, n^{-2} \varepsilon'^{2}} \\ &\asymp \frac{\lambda^{2} N}{C^{2} \, n^{-2} \varepsilon'^{2}} \asymp n^{-1/2\alpha} = o(1). \end{split}$$

Moreover,

$$\mathbb{P}\left(\sum_{j=1}^{N} |u_j|^2 > C \, n^{-1} (1+2\varepsilon')\right) = \mathbb{P}\left(\sum_{j=1}^{N} (|u_j|^2 - \sigma_j^2) > C \, n^{-1} (\varepsilon' + o(1))\right) = o(1),$$

which finishes the proof of (B.23).

Also, in view of (B.22), we have

$$\begin{split} \mathbb{P}\left(\sum_{j=1}^{N} j^{2\alpha} |u_j|^2 > L\right) &= \mathbb{P}\left(\sum_{j=1}^{N} j^{2\alpha} (|u_j|^2 - \sigma_j^2) > L \varepsilon\right) \\ &\leq \frac{\sum_{j=1}^{N} j^{4\alpha} Var\left(|u_j|^2\right)}{L^2 \varepsilon^2} = \frac{\sum_{j=1}^{N} j^{4\alpha} \sigma_j^4}{L^2 \varepsilon^2} \\ &\approx \frac{\lambda^2 N^{4\alpha+1}}{L^2 \varepsilon^2} \asymp n^{-1/2\alpha} = o(1), \end{split}$$

proving (B.24). Also,

$$\mathbb{P}\left(\sum_{j=1}^{N} j^{2\beta} |u_{j}|^{2} < n^{-1+\beta/\alpha}\right) \leq \mathbb{P}\left(\sum_{j=1}^{N} j^{2\beta} (|u_{j}|^{2} - \sigma_{j}^{2}) < -n^{-1+\beta/\alpha}\varepsilon\right) \\
\leq \frac{\sum_{j=1}^{N} j^{4\beta} Var(|u_{j}|^{2})}{n^{-2+2\beta/\alpha}\varepsilon^{2}} = \frac{\sum_{j=1}^{N} j^{4\beta}\sigma_{j}^{4}}{n^{-2+2\beta/\alpha}\varepsilon^{2}} \\
\approx \frac{\lambda^{2} N^{4\beta+1}}{n^{-2+2\beta/\alpha}\varepsilon^{2}} \asymp n^{-1/2\alpha} = o(1),$$
B.25)

proving (B.25).

Let us go back to (B.20) and bound from below the maximal error probability of type II by the averaged risk, with respect to our prior measure Π :

$$\sup_{u \in s^{\alpha}(L), F(u) \ge \eta_{n}} \langle G(\sqrt{n} \, u) | M_{0} | G(\sqrt{n} \, u) \rangle \ge \int_{H_{1}(\alpha, L, \gamma_{n}, \eta_{n})} \operatorname{Tr}(|G(\sqrt{n} \, u)) \langle G(\sqrt{n} \, u)| \cdot M_{0}) \Pi(du)$$

$$= \operatorname{Tr}\left(\int |G(\sqrt{n} \, u)) \langle G(\sqrt{n} \, u)| \Pi(du) \cdot M_{0}\right) - \int_{H_{1}(\alpha, L, \gamma_{n}, \eta_{n})^{C}} \operatorname{Tr}(|G(\sqrt{n} \, u)) \langle G(\sqrt{n} \, u)| \cdot M_{0}) \Pi(du)$$

$$\ge \operatorname{Tr}\left(\int |G(\sqrt{n} \, u)) \langle G(\sqrt{n} \, u)| \Pi(du) \cdot M_{0}\right) - \Pi(H_{1}(\alpha, L, \gamma_{n}, \eta_{n})^{C}).$$

In the last inequality we used that $\operatorname{Tr}(|G(\sqrt{n}\,u)\rangle\langle G(\sqrt{n}\,u)|\cdot M_0) \leq 1$. By Lemma B.1, $\Pi(H_1(\alpha, L, \gamma_n, \eta_n)^C) = o(1)$ and thus we deduce from (B.20) that

$$R_n^F \ge \frac{1}{8} \inf_M \left(\operatorname{Tr} \left(|G(0)\rangle \langle G(0)| \cdot M_1 \right) + \operatorname{Tr} \left(\int |G(\sqrt{n}\,u)\rangle \langle G(\sqrt{n}\,u)|\Pi(du) \cdot M_0 \right) \right) + o(1).$$

We recognize in the previous line the sum of error probabilities of type I and II for testing two simple quantum hypotheses, i.e. the underlying state is either $|G(0)\rangle$ or the mixed state

$$\Phi := \int |G(\sqrt{n}\,u)\rangle \langle G(\sqrt{n}\,u)|\Pi(du).$$

As a last step of the proof, we characterize more precisely the previous mixed Gaussian state as a thermal state and use classical results from quantum testing of two simple hypotheses to give the bound from below of the testing risk. Recall from Section 2.2.2, equation (8) that coherent states $|G(\sqrt{n} u)\rangle$ factorize as tensor product of one-mode coherent states with displacements u_j , i.e. $\bigotimes_{j\geq 1}|G(\sqrt{n}u_j)\rangle$. A coherent state with displacement z = x + iy with $x, y \in \mathbb{R}$ is fully characterized by its Wigner function given by equation (3). Since the prior is Gaussian, our mixed state Φ is Gaussian and can be written

$$\begin{split} \int |G(\sqrt{n}\,u)\rangle\langle G(\sqrt{n}\,u)|\Pi(du) &= \left(\bigotimes_{j=1}^N \int |G(\sqrt{n}\,u_j)\rangle\langle G(\sqrt{n}\,u_j)|\Pi_j(du_j)\right) \otimes \left(\bigotimes_{j\ge N+1}|0\rangle\langle 0|\right) \\ &:= \bigotimes_{j=1}^N \Phi_j \otimes \left(\bigotimes_{j\ge N+1}|0\rangle\langle 0|\right) \end{split}$$

where Π_j represents the bivariate centred Gaussian distribution with covariance matrix $\sigma_j^2/2 \cdot I_2$ over the complex plane $u_j = x_j + iy_j$. Using equation (5), and setting $\sigma^2 = n\sigma_j^2/2$ there, we find that the individual modes with index $j \leq N$ are centred Gaussian thermal states $\Phi_j = \Phi(r_j)$ (cf. definition (4)) with $r_j = n\sigma_j^2/(n\sigma_j^2 + 1)$.

In order to bound from below the right-hand side term in (B.20) we use the theory of quantum testing of two simple hypotheses

$$H_0: \otimes_{j\geq 1} \Phi(0)$$
 against $H_1: \otimes_{j=1}^N \Phi(r_j) \otimes_{j\geq N+1} \Phi(0).$

Using (9), it is easy to see that this testing problem is equivalent to

$$H_0: (\Phi(0))^{\otimes N}$$
 against $H_1: \bigotimes_{j=1}^N \Phi(r_j)$

As the vacuum and the thermal state are both diagonalized by the Fock basis, they commute, which reduces the problem to a classical test between the *N*-fold products of discrete distributions $H_0: \{\mathcal{G}(0)\}^{\otimes N}$ and $H_1: \{\otimes_{j=1}^N \mathcal{G}(r_j)\}$. In view of the form (4) of the thermal state, $\mathcal{G}(r_j)$ is the geometric distribution $\{(1-r_j)r_j^k\}_{k=0}^{\infty}$ and $\mathcal{G}(0)$ is the degenerate distribution concentrated at 0. The optimal testing error is given by the maximum likelihood test which decides H_0 if and only if all observations are 0. The type I error is 0 and the type II error is

$$\prod_{j=1}^{N} (1 - r_j) = \prod_{j=1}^{N} \frac{1}{n\sigma_j^2 + 1} \ge \exp\left(-n\sum_{j=1}^{N} \sigma_j^2\right) \ge \exp(-c),$$

for some c > 0, where in the last inequality we used (B.26). Using this in (B.20), we get as a lower bound

$$R_n^F \ge \exp(-c) + o(1) \ge c_0,$$

where $c_0 > 0$ is some constant depending on c. This finishes the proof.

PROOF OF THEOREM 5.6. Let $\varphi_n = c_n n^{-1/2}$ for a positive sequence c_n . Let $M_n = (\rho_0^{\otimes n}, I - \rho_0^{\otimes n})$ be the well-known projection test for the problem (28). Then

$$R_n^T(M_n) = \operatorname{Tr}(\rho^{\otimes n} \cdot \rho_0^{\otimes n}) + \operatorname{Tr}(\rho_0^{\otimes n} \cdot (I - \rho_0^{\otimes n}))$$
$$= (\operatorname{Tr}(\rho \cdot \rho_0))^n = |\langle \psi | \psi_0 \rangle|^{2n}.$$

Let us recall that for any pure states $\rho = |\psi\rangle\langle\psi|$ and $\rho_0 = |\psi_0\rangle\langle\psi_0|$, we have

(B.27)
$$\|\rho - \rho_0\|_1 = 2\sqrt{1 - |\langle \psi | \psi_0 \rangle|^2} ,$$

thus $|\langle \psi | \psi_0 \rangle|^2 = 1 - \frac{1}{4} \| \rho - \rho_0 \|_1^2$ and hence

$$R_n^T(M_n) = \left(1 - \frac{1}{4} \|\rho - \rho_0\|_1^2\right)^n.$$

For any ρ satisfying the alternative hypothesis $H_1(\varphi_n)$, we have $\|\rho - \rho_0\|_1 \ge \varphi_n$ and consequently

$$\mathbb{P}_{e}^{M_{n}}\left(\varphi_{n}\right) \leq \left(1 - \frac{1}{4}\varphi_{n}^{2}\right)^{n} = \left(1 - \frac{c_{n}^{2}}{4}n^{-1}\right)^{n}$$
$$\leq \left(\exp\left(-\frac{c_{n}^{2}}{4}n^{-1}\right)\right)^{n} = \exp\left(-\frac{c_{n}^{2}}{4}\right)$$

If now $\varphi_n/\varphi_n^* \to \infty$ then $c_n \to \infty$ and $\mathbb{P}_e^{M_n}(\varphi_n) \to 0$, so that the second relation in (29) is fulfilled.

Consider now the case $\varphi_n/\varphi_n^* \to 0$ so that $c_n \to 0$. For any vector $v \in \mathcal{H}$ define

(B.28)
$$||v||_{\alpha}^{2} = \sum_{j=0}^{\infty} |\langle e_{j}|v\rangle|^{2} j^{2\alpha};$$

then $\|v\|_{\alpha}$ is a seminorm on the space of v fulfilling $\|v\|_{\alpha}^2 < \infty$. The assumption that $\rho_0 = |\psi_0\rangle\langle\psi_0| \in S^{\alpha}(L')$ means that $\|\psi_0\|_{\alpha}^2 \leq L' < L$. For some N > 0, consider the linear space

$$\mathcal{H}_{0,N} = \{ u \in \mathcal{H} : \langle u | \psi_0 \rangle = 0, \langle u | e_j \rangle = 0, j > N \};$$

it is nonempty if $N \ge 1$. Let $u \in \mathcal{H}_{0,N}$, ||u|| = 1 be an unit vector; and for $\varepsilon > 0$ consider

(B.29)
$$\psi_{u,\varepsilon} = \psi_0 \sqrt{1 - \varepsilon^2} + \varepsilon u.$$

Then $\|\psi_{u,\varepsilon}\| = 1$, $\rho_{u,\varepsilon} = |\psi_{u,\varepsilon}\rangle\langle\psi_{u,\varepsilon}|$ is a pure state, and

$$|\langle \psi_{u,\varepsilon} | \psi_0 \rangle|^2 = 1 - \varepsilon^2$$

According to (B.27) we then have

$$\left\|\rho_{u,\varepsilon} - \rho_0\right\|_1 = 2\sqrt{1 - |\langle \psi_{u,\varepsilon} | \psi_0 \rangle|^2} = 2\varepsilon$$

so for a choice $\varepsilon = c_n n^{-1/2}/2$ it follows $\|\rho_{u,\varepsilon} - \rho_0\|_1 = \varphi_n$ and $\rho_{u,\varepsilon} \in B(\varphi_n)$. On the other hand, by (B.29) and the triangle inequality

$$\left\|\psi_{u,\varepsilon}\right\|_{\alpha} \leq \sqrt{1-\varepsilon^2} \left\|\psi_0\right\|_{\alpha} + \varepsilon \left\|u\right\|_{\alpha}.$$

Now $||u||_{\alpha} < \infty$ for $u \in \mathcal{H}_{0,N}$, and by assumption $||\psi_0||_{\alpha} < L^{1/2}$, so for sufficiently large n

$$\|\psi_{u,\varepsilon}\|_{\alpha} \leq L^{1/2}$$

and thus $\rho_{u,\varepsilon} \in S^{\alpha}(L)$. Thus $\rho_{u,\varepsilon} \in S^{\alpha}(L) \cap B(\varphi_n)$ for sufficiently large *n*. By (9) the optimal error probability for testing between states $\rho_{u,\varepsilon}$ and ρ_0 fulfills

$$\inf_{M \text{ binary POVM}} R_n^T(\rho_0^{\otimes n}, \rho_{u,\varepsilon}^{\otimes n}, M) = 1 - \frac{1}{2} \left\| \rho_0^{\otimes n} - \rho_{u,\varepsilon}^{\otimes n} \right\|_1$$

$$= 1 - \sqrt{1 - |\langle \psi_0^{\otimes n} | \psi_{u,\varepsilon}^{\otimes n} \rangle|^2} = 1 - \sqrt{1 - |\langle \psi_0 | \psi_{u,\varepsilon} \rangle|^{2n}}$$
$$= 1 - \sqrt{1 - (1 - \varepsilon^2)^n} = 1 - \sqrt{1 - (1 - c_n^2 n^{-1}/4)^n}.$$

Obviously if $c_n^2 \to 0$ then $(1 - c_n^2 n^{-1}/4)^n \to 1$ so that

$$\inf_{M \text{ binary POVM}} R_n^T(\rho_0^{\otimes n}, \rho_{u,\varepsilon}^{\otimes n}, M) \ge 1 + o(1).$$

But since $\rho_{u,\varepsilon} \in S^{\alpha}(L) \cap B(\varphi_n)$ we have

$$\mathbb{P}_{e}^{*}\left(\varphi_{n}\right) \geq \inf_{M \text{ binary POVM}} R_{n}^{T}(\rho_{0}^{\otimes n}, \rho_{u,\varepsilon}^{\otimes n}, M) \geq 1 + o\left(1\right),$$

so that the first relation in (29) is shown.

PROOF OF THEOREM 5.7. It suffices to prove that if $\varphi_n = c_n n^{-1/2}$ with $c_n \to c > 0$ then $\mathbb{P}_e^*(\varphi_n) \to \exp(-c^2/4)$. In view of the upper bound (30), if suffices to prove

(B.30)
$$\mathbb{P}_e^*(\varphi_n) \ge \exp\left(-c^2/4\right)\left(1+o\left(1\right)\right).$$

Recall (cf. (B.27)) that for any pure states $\rho = |\psi\rangle\langle\psi|$ and $\rho_0 = |\psi_0\rangle\langle\psi_0|$, the condition $\|\rho - \rho_0\|_1 \ge \varphi_n$ in $H_1(\varphi_n)$ is equivalent to a condition for the fidelity $F^2(\rho, \rho_0) = |\langle\psi|\psi_0\rangle|^2 \le 1 - \varphi_n^2/4$.

Let $\mathcal{H}_0 \subset \mathcal{H}$ be the orthogonal complement of $\mathbb{C}|\psi_0\rangle$ in \mathcal{H} . Consider the vector

$$\psi_u = \sqrt{1 - \|u\|^2} \cdot \psi_0 + u, \quad u \in \mathcal{H}_0$$

and the corresponding pure state $|\psi_u\rangle\langle\psi_u|$ defined in terms of the local vector u. We restrict the alternative hypothesis to a smaller set of states such that $||u|| \leq \gamma_n$, with $\gamma_n = (\log n)^{-1}$.

Since the fidelity is given by $F^2(\rho_0, |\psi_u\rangle\langle\psi_u|) = |\langle\psi_u|\psi_0\rangle|^2 = 1 - ||u||^2$, the restricted hypothesis is characterised by

$$1 - \gamma_n^2 \le F^2(\rho_0, |\psi_u\rangle\langle\psi_u|) \le 1 - \varphi_n^2/4, \text{ or } \varphi_n^2/4 \le ||u||^2 \le \gamma_n^2.$$

and additionally by $\|\psi_u\|_{\alpha}^2 \leq L$ where $\|\cdot\|_{\alpha}$ is given by (B.28).

Consider again the linear space $\mathcal{H}_{0,N}$ defined in the proof of Theorem 5.7 for a choice $N = N_n \sim \log \log n$. Since $\mathcal{H}_{0,N} \subset \mathcal{H}_0$, we can further restrict the local vector u to $u \in \mathcal{H}_{0,N}$. Note that for $u \in \mathcal{H}_{0,N}$ and $||u|| \leq \gamma_n$ we have

$$||u||_{\alpha}^{2} = \sum_{j=0}^{N} |\langle e_{j} | u \rangle|^{2} j^{2\alpha} \le N^{2\alpha} ||u||^{2} \le N^{2\alpha} \gamma_{n}^{2}$$

 $\sim (\log \log n)^{2\alpha} (\log n)^{-2} = o(1).$

It follows that

$$\|\psi_u\|_{\alpha} \le \sqrt{1 - \|u\|^2} \|\psi_0\|_{\alpha} + \|u\|_{\alpha} \le L^{1/2}$$

for sufficiently large n, thus $\psi_u \in S^{\alpha}(L)$. We can now write the test problem with restricted alternative as

$$\begin{aligned} H_0: & \rho = \rho_0 \\ H'_1(\varphi_n): & \rho = |\psi_u\rangle \langle \psi_u|: u \in \mathcal{H}_{0,N}, \, \varphi_n/2 \le ||u|| \le \gamma_n. \end{aligned}$$

By the strong approximation proven in Theorem 4.1 we get that the models

$$\{|\psi_u\rangle\langle\psi_u|^{\otimes n}, \|u\|\leq\gamma_n\}$$
 and $\{|G(\sqrt{n}u)\rangle\langle G(\sqrt{n}u)|, \|u\|\leq\gamma_n\}$

are asymptotically equivalent, where $G(\sqrt{n}u)$ is the coherent vector in the Fock space $\Gamma_s(\mathcal{H}_0)$ pertaining to $\sqrt{n}u$. Note that this proof is very similar to the previous proofs of lower bounds, with a major difference: the reduced set of states under the alternative hypothesis is defined with repect to ρ_0 given by the null hypothesis H_0 instead of an arbitrary state previously.

In the asymptotically equivalent Gaussian white noise model, the modified hypotheses concern Gaussian states which can be written in terms of their coherent vectors as

$$\begin{aligned} H_0: & |G(0)\rangle \\ H_1(\varphi_n): & |G(\sqrt{n}u)\rangle: u \in \mathcal{H}_{0,N}, \, \varphi_n/2 \le \|u\| \le \gamma_n. \end{aligned}$$

In order to prove the theorem it is sufficient to prove that

(B.31)
$$\inf_{M_n} \sup_{\varphi_n/2 \le ||u|| \le \gamma_n, \ u \in \mathcal{H}_{0,N}} R_n^T(|G(0)\rangle \langle G(0)|, |G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)|, M_n)$$

(B.32)
$$\ge \exp\left(-c^2/4\right) + o(1)$$

as $n \to \infty$.

Note that dim $\mathcal{H}_{0,N} = N$; let $\{g_j, j = 1, ..., N\}$ be an orthogonal basis of $\mathcal{H}_{0,N}$ and let $|u\rangle = \sum_{j=1}^{N} u_j |g_j\rangle$. The quantum Gaussian white noise model $\{|G(\sqrt{n}u)\rangle, u \in \mathcal{H}_{0,N}, ||u|| \leq \gamma_n\}$ is then equivalent to the quantum Gaussian sequence model $\{\otimes_{j=1}^{N} |G(\sqrt{n}u_j)\rangle, ||u|| \leq \gamma_n\}$. From now on $|G(z)\rangle$ denotes the coherent vector in the Fock space $\mathcal{F}(\mathbb{C})$ pertaining to $z := x + iy \in \mathbb{C}$. Recall that such a state is fully characterized by its Wigner function $W_{G(z)}$, which in the case of coherent states is the density function of a bivariate Gaussian distribution.

We shall bound from below the maximal type 2 error probability in the risk $R_n^T(M_n)$ in (B.31)

(B.33)
$$\sup_{\varphi_n/2 \le \|u\| \le \gamma_n, \ u \in \mathcal{H}_{0,N}} \operatorname{Tr} \left(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)| \cdot M_{n,0} \right)$$

by an average over u, where the average is taken with respect to a prior distribution defined as follows. Assume that u_j , j = 1, ..., N are independently distributed following a complex centered Gaussian law with variance $\frac{\sigma^2}{2}I_2$, where $\sigma^2 = \frac{c^2}{4n}\frac{1+\varepsilon}{N}$, for some fixed and arbitrary small $\varepsilon > 0$, and I_2 is the 2 by 2 identity matrix.

LEMMA B.2. Let Π be the distribution of independent complex random variables u_j , for j = 1, ..., N, each one distributed as

$$N\left(0,\frac{\sigma^2}{2}I_2\right), \quad \sigma^2 = \frac{c^2}{4n}\frac{1+\varepsilon}{N},$$

for fixed $\varepsilon > 0$ and $N \sim \log \log n$. Then as $n \to \infty$

$$\mathbb{P}\left(\frac{c_n^2}{4n} \le \|u\|^2 \le \frac{c_n^2}{4n}(1+\varepsilon)^2\right) \to 1, \quad as \ n \to \infty,$$

and in particular if $\gamma_n = (\log n)^{-1}$ then $\mathbb{P}(\varphi_n/2 \le ||u|| \le \gamma_n) \to 1$, as $n \to \infty$.

PROOF. We have

$$\mathbb{P}\left(\|u\|^{2} < \frac{c_{n}^{2}}{4n}\right) = \mathbb{P}\left(\sum_{j=1}^{N} (|u_{j}|^{2} - \sigma^{2}) < \frac{c_{n}^{2}}{4n} - N\frac{c^{2}}{4n}\frac{1+\varepsilon}{N}\right)$$

$$\leq \frac{\operatorname{Var}(\sum_{j=1}^{N} |u_{j}|^{2})}{(c_{n}^{2} - c^{2}(1+\varepsilon))^{2}/16n^{2}} = \frac{N\sigma^{4}}{(c^{2}\varepsilon + o(1))^{2}/16n^{2}}$$

$$= \frac{Nc^{4}(1+\varepsilon)^{2}/16n^{2}N^{2}}{(c^{2}\varepsilon + o(1))^{2}/16n^{2}} = \left(\frac{1+\varepsilon}{\varepsilon + o(1)}\right)^{2}\frac{1}{N} = o(1),$$

since $N \sim \log \log n \to \infty$. Similarly, as $(1 + \varepsilon)^2 > 1 + \varepsilon$, one shows that

$$\mathbb{P}\left(\|u\|^2 > \frac{c_n^2}{4n}(1+\varepsilon)^2\right) \to 0,$$

as $n \to \infty$ and thus we get

$$\mathbb{P}\left(\frac{c_n^2}{4n} \le \|u\|^2 \le \frac{c_n^2}{4n}(1+\varepsilon)^2\right) \to 1.$$

As $\gamma_n^2 = (\log n)^{-2}$ decays slower than c_n^2/n , and $\varphi_n/2 = c_n n^{-1/2}/2$, we deduce that

$$\mathbb{P}\left(\varphi_n/2 \le \|u\| \le \gamma_n\right) \to 1$$

as $n \to \infty$ which ends the proof of the lemma.

Let us denote by Π the prior distribution introduced in Lemma B.2. Let us go back to (B.33) and bound the expression from below as follows:

$$\sup_{\varphi_n/2 \le ||u|| \le \gamma_n, u \in \mathcal{H}_{0,N}} \operatorname{Tr} \left(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)| \cdot M_{n,0} \right)$$
$$\ge \int_{\varphi_n/2 \le ||u|| \le \gamma_n} \operatorname{Tr} \left(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)|M_{n,0} \right) \Pi(du)$$

$$\geq \int \operatorname{Tr}(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)|M_{n,0})\Pi(du) - \int_{\{\varphi_n/2 \le \|u\| \le \gamma_n\}^c} \operatorname{Tr}(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)|M_{n,0})\Pi(du) \\ \geq \int \operatorname{Tr}(|G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)|M_{n,0})\Pi(du) - \Pi\left(\{\varphi_n/2 \le \|u\| \le \gamma_n\}^c\right).$$

By Lemma B.2, we get for (B.31)

$$\sup_{\varphi_n/2 \le \|u\| \le \gamma_n, \ u \in \mathcal{H}_{0,N}} R_n^T(G(0), G(\sqrt{n}u), M_n)$$

(B.34)
$$\geq \operatorname{Tr}(|G(0)\rangle\langle G(0)|M_{n,1}) + \operatorname{Tr}\left(\int |G(\sqrt{n}u)\rangle\langle G(\sqrt{n}u)|\Pi(du)\cdot M_{n,0}\right) + o(1).$$

The integral on the right side is a mixed state which can be written as

$$\Phi := \int |G(\sqrt{n}u)\rangle \langle G(\sqrt{n}u)|\Pi(du) = \bigotimes_{j=1}^N \int |G(\sqrt{n}u_j)\rangle \langle G(\sqrt{n}u_j)| \cdot \Pi_j(du_j).$$

Similarly to the proof of Theorem 5.5 we use equation (5) to show that each of the Gaussian integrals above produces a thermal (Gaussian) state

$$\Phi(r) = (1-r)\sum_{k=0}^{\infty} r^k |k\rangle \langle k|, \qquad r = \frac{n\sigma^2}{n\sigma^2 + 1}.$$

Since $|G(0)\rangle\langle G(0)| = \Phi(0)$, the main terms in (B.34) are the sum of error probabilities for testing two simple hypothesis $H_0 : \Phi(0)^{\otimes N}$ against $H_1 : \Phi(r)^{\otimes N}$. Moreover, we have two commuting product states under the two simple hypotheses, which reduces the problem to a classical test between the N-fold products of discrete distributions $H_0 : {\mathcal{G}(0)}^{\otimes N}$ and $H_1 : {\mathcal{G}(r)}^{\otimes N}$. Here $\mathcal{G}(r)$ is the geometric distribution ${(1-r)r^k}_{k=0}^{\infty}$; in particular s $\mathcal{G}(0)$ is the degenerate distribution concentrated at 0. The optimal testing error is given by the maximum likelihood test which decides H_0 if and only if all observations are 0. The type 1 error is 0 and the type 2 error is

$$(1-r)^N = (n\sigma^2 + 1)^{-N} \ge \exp(-N \cdot n\sigma^2)$$
$$= \exp\left(-Nn\frac{c^2}{4n}\frac{1+\varepsilon}{N}\right) = \exp\left(-\frac{c^2(1+\varepsilon)}{4}\right).$$

Since $\varepsilon > 0$ was arbitrary, this establishes the lower bound (B.32) and thus (B.30).

REFERENCES

- BICKEL, P. J. and RITOV, Y. (1988). Estimating integrated squared density derivatives: sharp best order of convergence estimates. Sankhyā Ser. A 50 381–393. MR1065550
- [2] BROWN, L. D., CAI, T. T., LOW, M. G. and ZHANG, C.-H. (2002). Asymptotic equivalence theory for nonparametric regression with random design. Ann. Statist. 30 688–707. MR1922538
- BROWN, L. D. and LOW, M. G. (1996). Asymptotic equivalence of nonparametric regression and white noise. Ann. Statist. 24 2384–2398. MR1425958
- [4] CAI, T. T. and LOW, M. G. (2005). Nonquadratic estimators of a quadratic functional. Ann. Statist. 33 2930–2956.
- [5] DONOHO, D. L. and LIU, R. C. (1991). Geometrizing rates of convergence II, III. Ann. Statist. 19 633–667, 668–701.
- [6] ERMAKOV, M. S. (1990). Minimax detection of a signal in Gaussian white noise. Teor. Veroyatnost. i Primenen. 35 704–715. MR1090496
- [7] GOLUBEV, G. K., NUSSBAUM, M. and ZHOU, H. H. (2010). Asymptotic equivalence of spectral density estimation and Gaussian white noise. Ann. Statist. 38 181–214. MR2589320
- [8] GRAMA, I. and NUSSBAUM, M. (1998). Asymptotic equivalence for nonparametric generalized linear models. Probab. Theory Related Fields 111 167–214. MR1633574
- [9] GUŢĂ, M., JANSSENS, B. and KAHN, J. (2008). Optimal estimation of qubit states with continuous time measurements. *Commun. Math. Phys.* 277 127-160.
- [10] GUŢĂ, M. and KAHN, J. (2006). Local asymptotic normality for qubit states. Phys. Rev. A 73 052108.
- [11] HAYASHI, M. (1998). Asymptotic estimation theory for a finite-dimensional pure state model. Journal of Physics A: Mathematical and General 31.
- [12] HELSTROM, C. W. (1976). Quantum Detection and Estimation Theory. Academic Press, New York.
- [13] INGSTER, Y. I. and SUSLINA, I. A. (2003). Nonparametric Goodness-of-Fit Testing under Gaussian Models. Lecture Notes in Statistics 169. Springer-Verlag, New York. MR1991446
- [14] JI, P. and NUSSBAUM, M. (2017). Sharp minimax adaptation over Sobolev ellipsoids in nonparametric testing. *Electron. J. Stat.* **11** 4515-4562.
- [15] KLEMELÄ, J. (2006). Sharp adaptive estimation of quadratic functionals. Probab. Theory Related Fields 134 539–564.
- [16] NUSSBAUM, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. Ann. Statist. 24 2399–2430. MR1425959
- [17] RAY, K. and SCHMIDT-HIEBER, J. (2017). The Le Cam distance between density estimation, Poisson processes and Gaussian white noise. arXiv:1608.01824 [math.ST].
- [18] REISS, M. (2008). Asymptotic equivalence for nonparametric regression with multivariate and random design. Ann. Statist. 36 1957–1982. MR2435461
- [19] REISS, R. D. (1993). A Course on Point Processes. Springer-Verlag, New York. MR1199815
- [20] ROHDE, A. (2004). On the asymptotic equivalence and rate of convergence of nonparametric regression and Gaussian white noise. *Statist. Decisions* 22 235–243. MR2125610
- [21] TSYBAKOV, A. B. (2009). Introduction to Nonparametric Estimation. Springer, New York.

CREST, ENSAE, UNIVERSITÉ PARIS-SACLAY 5, AVENUE HENRY LE CHATELIER, 91120 PALAISEAU, FRANCE E-MAIL: cristina.butucea@ensae.fr UNIVERSITY OF NOTTINGHAM UNIVERSITY OF NOTTINGHAM

> DEPARTMENT OF MATHEMATICS MALOTT HALL, CORNELL UNIVERSITY ITHACA, NY 14853 USA E-MAIL: nussbaum@math.cornell.edu