

The van Trees inequality in the spirit of Hájek and Le Cam

Elisabeth Gassiat and Gilles Stoltz

Abstract. In honor of the 100th birth anniversary of Lucien Le Cam (November 18, 1924 – April 24, 2000), we work out a version of the van Trees inequality in a Hájek–Le Cam spirit, i.e., under minimal assumptions that, in particular, involve no direct pointwise regularity assumptions on densities but rather almost-everywhere differentiability in quadratic mean of the model. Surprisingly, it suffices that the latter differentiability holds along canonical directions—not along all directions. Also, we identify a (slightly stronger) version of the van Trees inequality as a very instance of a Cramér–Rao bound, i.e., the van Trees inequality is not just a Bayesian analog of the Cramér–Rao bound. We provide, as an illustration, an elementary proof of the local asymptotic minimax theorem for quadratic loss functions, again assuming differentiability in quadratic mean only along canonical directions.

Key words and phrases: van Trees inequality, Cramér–Rao bound, Differentiability in quadratic mean, Local asymptotic minimax theorem.

1. INTRODUCTION

Every statistician knows about the Cramér–Rao inequality but fewer knew about the van Trees inequality (van Trees, 1968, page 72) before Gill and Levit (1995) drew attention to some of its statistical uses. In their landmark article, they present the van Trees inequality as offering a Bayesian Cramér–Rao bound, to be applied in cases involving convergence of experiments to bypass the beautiful but sophisticated Hájek–Le Cam theory of convergence of experiments. Gill and Levit (1995) derived the van Trees inequality under precise analytic conditions, involving, in particular, smoothness assumptions on the densities; so did also later contributions, including the ones by Lenstra (2005), Jupp (2010), and Letac (2022). However, as summarized by Pollard (2001; 2005), who in turn refers to Bickel et al. (1993, page 12) and Lehmann and Romano (2005, Chapter 12), Le Cam and Hájek advocated resorting rather to conditions that are intrinsic; of particular interest, is the concept of differentiability in quadratic mean of a statistical model.

We provide a version of the van Trees inequality in the spirit promoted by Le Cam and Hájek, and aim for the weakest possible assumptions. In the one-dimensional

case (Section 2.1), on top of the assumptions merely ensuring the existence of the quantities involved in the inequality (which includes the almost-everywhere differentiability of the model), we only require that the prior vanishes at finite boundary points of the parameter space Θ (which is an arbitrary, not necessarily bounded, open subset of \mathbb{R}), together with some technical condition on the model that is weaker than its differentiability everywhere. We discuss these extremely mild assumptions (Section 2.2) by comparing them to the classic regularity assumptions proposed by Gill and Levit (1995). Our proof (Section 2.3) also exploits the same separation of x and θ variables as in Gill and Levit (1995), but we perform integrations in the reverse order, first over x then over θ , thus effectively avoiding pointwise regularity assumptions on densities. It turns out (Section 2.4) that the van Trees inequality is not only a Bayesian analog of the Cramér–Rao bound, as pointed out by van Trees (1968, page 72) and Gill and Levit (1995), but that it is exactly, at least in a slightly stronger form, an instance of a Cramér–Rao bound for a suitably chosen location model.

The rest of this contribution focuses on a multivariate version of the van Trees inequality. We provide (Section 3) weak conditions that only involve differentiability in quadratic mean of the model along canonical directions, not all directions. We illustrate (Section 4) the application of this multivariate version to establish a local asymptotic minimax theorem for quadratic loss functions.

Elisabeth Gassiat is Professor and Gilles Stoltz is CNRS Senior research fellow; Université Paris-Saclay, CNRS, Laboratoire de mathématiques d'Orsay, 91405, Orsay, France (e-mail: elisabeth.gassiat@universite-paris-saclay.fr; gilles.stoltz@universite-paris-saclay.fr).

2. ONE-DIMENSIONAL VERSION

We consider a statistical model $\mathcal{P} = (\mathbb{P}_\theta)_{\theta \in \Theta}$, defined on a measurable space $(\mathcal{X}, \mathcal{F})$ and indexed by an open subset Θ of \mathbb{R} (not necessarily an interval). We assume that \mathcal{P} is dominated by a σ -finite measure μ , with densities $f_\theta = d\mathbb{P}_\theta/d\mu$ such that $(\theta, x) \mapsto f_\theta(x)$ is measurable. Let $\xi_\theta = \sqrt{f_\theta} \in \mathbb{L}^2(\mu)$ be the square roots of these densities.

In the sequel, $\|\cdot\|_\mu$ refers to the Euclidean norm in $\mathbb{L}^2(\mu)$, i.e., for a function $g : \mathcal{X} \rightarrow \mathbb{R}$ in $\mathbb{L}^2(\mu)$,

$$\|g\|_\mu = \sqrt{\int_{\mathcal{X}} g^2 d\mu}.$$

DEFINITION 1 (Differentiability in \mathbb{L}_2). The μ -dominated statistical model \mathcal{P} is differentiable in $\mathbb{L}_2(\mu)$ at $\theta_0 \in \Theta$ if there exists a function $\dot{\xi}_{\theta_0} \in \mathbb{L}_2(\mu)$, called the $\mathbb{L}_2(\mu)$ -derivative of the model at θ_0 , such that

$$\|\xi_\theta - \xi_{\theta_0} - (\theta - \theta_0)\dot{\xi}_{\theta_0}\|_\mu = o(\|\theta - \theta_0\|) \text{ as } \theta \rightarrow \theta_0.$$

The Fisher information $\mathcal{I}_{\mathcal{P}}(\theta_0)$ of the model at θ_0 is then defined as

$$\mathcal{I}_{\mathcal{P}}(\theta_0) = 4 \int_{\mathcal{X}} (\dot{\xi}_{\theta_0})^2 d\mu.$$

DEFINITION 2 (Well-behaved prior). We call a probability measure \mathbb{Q} that concentrates on the open set $\Theta \subseteq \mathbb{R}$ a well-behaved prior if \mathbb{Q} has a density q with respect to the Lebesgue measure on Θ that is absolutely continuous on Θ , with almost-sure derivative q' satisfying

$$\mathcal{I}_{\mathbb{Q}} \stackrel{\text{def}}{=} \int_{\Theta} (q'(\theta))^2 \frac{\mathbf{1}_{\{q(\theta) > 0\}}}{q(\theta)} d\theta < \infty.$$

We denote by $\text{Supp}(q) = \{q > 0\}$ the open support of q .

A standard result (see, e.g., [Lehmann and Romano, 2005](#), Corollary 12.2.1) states that a location model based on a well-behaved prior \mathbb{Q} is differentiable in $\mathbb{L}^2(\lambda)$, where λ denotes the Lebesgue measure, with derivative at 0 equal to $q' \mathbf{1}_{\{q > 0\}} / (2\sqrt{q})$, and hence, with Fisher information $\mathcal{I}_{\mathbb{Q}}$.

2.1 Statement

The van Trees inequality lower bounds the Bayesian squared error of any, possibly biased, statistic $S : \mathcal{X} \rightarrow \mathbb{R}$ for the estimation of a functional $\psi(\theta)$, where we assume that ψ is an absolutely continuous function, with almost-everywhere derivative denoted by ψ' . More precisely, denoting by \mathbb{E}_θ the expectation under \mathbb{P}_θ , the one-dimensional version of the van Trees inequality reads (vT1)

$$\int_{\Theta} \mathbb{E}_\theta \left[(S - \psi(\theta))^2 \right] d\mathbb{Q}(\theta) \geq \frac{\left(\int_{\Theta} \psi'(\theta) d\mathbb{Q}(\theta) \right)^2}{\mathcal{I}_{\mathbb{Q}} + \int_{\Theta} \mathcal{I}_{\mathcal{P}}(\theta) d\mathbb{Q}(\theta)}.$$

Our version of the van Trees inequality requires two series of assumptions. The first series, stated in [Assumption 3](#) merely ensures that all quantities involved are defined and that the inequality has a meaning. The second series of assumptions are ‘‘real’’ assumptions and may be found in [Theorem 4](#).

ASSUMPTION 3 (ensuring definitions and meaning). The set Θ is any open subset of \mathbb{R} . The probability measure \mathbb{Q} is a well-behaved prior on Θ . The statistical model $\mathcal{P} = (\mathbb{P}_\theta)_{\theta \in \Theta}$ is dominated by a σ -finite measure μ , with densities $f_\theta = d\mathbb{P}_\theta/d\mu$ such that $(\theta, x) \mapsto f_\theta(x)$ is measurable. The model \mathcal{P} is differentiable in $\mathbb{L}^2(\mu)$ almost everywhere on $\Theta \cap \text{Supp}(q)$. The function $\psi : \Theta \rightarrow \mathbb{R}$ is absolutely continuous. Both ψ^2 and ψ' are \mathbb{Q} -integrable and

$$\int_{\Theta} \mathbb{E}_\theta [S^2] d\mathbb{Q}(\theta) < +\infty, \quad \int_{\Theta} \mathcal{I}_{\mathcal{P}}(\theta) d\mathbb{Q}(\theta) < +\infty.$$

THEOREM 4. *The one-dimensional van Trees inequality (vT1) holds with $\mathcal{I}_{\mathbb{Q}} > 0$ under [Assumption 3](#) and the following additional assumptions:*

- for all $A \in \mathcal{F}$, the functions $\theta \in \Theta \cap \text{Supp}(q) \mapsto \mathbb{P}_\theta(A)$ are absolutely continuous;
- $q(\theta) \rightarrow 0$ as θ approaches any finite boundary point of Θ .

The first assumption holds in particular if the model \mathcal{P} is differentiable in $\mathbb{L}^2(\mu)$ at all points of $\Theta \cap \text{Supp}(q)$, not just almost everywhere.

2.2 Comparison to classic regularity assumptions

We compare [Theorem 4](#) to the version under classic regularity assumptions by [Gill and Levit \(1995\)](#) based on [van Trees \(1968\)](#). With no loss of generality (on the contrary) and no change in their proof, we only replace their closed interval Θ by any open set Θ , possibly intersected with $\text{Supp}(q)$. The key additional assumption required is stated next.

ASSUMPTION 5 (main regularity assumption). In the μ -dominated model \mathcal{P} , the densities $f_\theta = d\mathbb{P}_\theta/d\mu$ are such that for μ -almost all x , the function $\theta \in \Theta \cap \text{Supp}(q) \mapsto f_\theta(x)$ is absolutely continuous, with almost-everywhere derivative denoted by $f'_\theta(x)$.

In that setting with classic regularity assumptions, the Fisher information is defined, where f'_θ exists, i.e., almost-everywhere, by

$$\tilde{\mathcal{I}}_{\mathcal{P}}(\theta) = \int_{\mathcal{X}} \left(\frac{f'_\theta}{f_\theta} \right)^2 d\mathbb{P}_\theta = \int_{\mathcal{X}} \frac{(f'_\theta)^2}{f_\theta} \mathbf{1}_{\{f_\theta > 0\}} d\mu.$$

A finite denominator in the right-hand side of the van Trees inequality entails (see the argument in the last lines

of Section 2.3.5) that $\tilde{\mathcal{I}}_{\mathcal{P}}$ is locally integrable around each $\theta \in \Theta \cap \text{Supp}(q)$, and thus, that almost all points of $\Theta \cap \text{Supp}(q)$ are Lebesgue points for $\tilde{\mathcal{I}}_{\mathcal{P}}$. Based on this and on Assumption 5, we apply a slight extension of Bickel et al. (1993, Proposition 1) or Lehmann and Romano (2005, Theorem 12.2.1), whose proofs show that continuity of $\tilde{\mathcal{I}}_{\mathcal{P}}$ is actually not required and that a Lebesgue-point assumption is sufficient; we obtain that the model \mathcal{P} is differentiable in $\mathbb{L}^2(\mu)$ almost everywhere on $\Theta \cap \text{Supp}(q)$, with $\mathbb{L}^2(\mu)$ -derivatives given by $\dot{\xi}_{\theta} = f'_{\theta} \mathbf{1}_{\{f_{\theta} > 0\}} / \sqrt{f_{\theta}}$. We also have $\tilde{\mathcal{I}}_{\mathcal{P}} = \mathcal{I}_{\mathcal{P}}$ almost everywhere on $\Theta \cap \text{Supp}(q)$.

Now, Gill and Levit (1995) prove the van Trees inequality under the boundary conditions on q and $q\psi$ stated in Theorem 4, under Assumption 5 and all of Assumption 3 except the almost everywhere $\mathbb{L}^2(\mu)$ -differentiability of \mathcal{P} . The other condition in Theorem 4, namely, that for all $A \in \mathcal{F}$, the function $\theta \in \Theta \cap \text{Supp}(q) \mapsto \mathbb{P}_{\theta}(A)$ is absolutely continuous, is a direct consequence of Assumption 5, by the Fubini–Tonelli theorem and the characterization of absolute continuity in terms of equality to the integral of the derivative. We therefore proved the following fact.

FACT. The regularity assumptions considered by Gill and Levit (1995) to prove the one-dimensional van Trees inequality (vT1) are more stringent than the Hájek–Le Cam-type assumptions considered in Theorem 4.

2.3 Proof of Theorem 4

The key lemma for our approach and its proof are extracted from the lecture notes by Pollard (2001; 2005), who adapted a result by Ibragimov and Has'minskii (1981, Lemma 7.2, page 67). The lemma stated in Pollard (2001; 2005) is actually stronger as it only requires local boundedness of T in $\mathbb{L}^2(\mathbb{P}_{\theta})$ around θ_0 .

LEMMA 6 (Pollard, 2001; 2005). *Let the μ -dominated model $\mathcal{P} = (\mathbb{P}_{\theta})_{\theta \in \Theta}$ be differentiable in $\mathbb{L}_2(\mu)$ at θ_0 . Consider a uniformly bounded statistic $T : \mathcal{X} \rightarrow \mathbb{R}$, i.e., there exists $M > 0$ with $|T| \leq M$ μ -a.s. Then, $\gamma_T : \theta \in \Theta \mapsto \mathbb{E}_{\theta}[T]$ is differentiable at θ_0 , with derivative*

$$\gamma'_T(\theta_0) = 2 \int_{\mathcal{X}} \dot{\xi}_{\theta_0} \xi_{\theta_0} T \, d\mu.$$

PROOF. Let $r_{\theta} = \xi_{\theta} - \xi_{\theta_0} - (\theta - \theta_0) \dot{\xi}_{\theta_0}$, so that

$$\begin{aligned} & \overbrace{(\xi_{\theta_0} + (\theta - \theta_0) \dot{\xi}_{\theta_0} + r_{\theta})^2}^{= \xi_{\theta}^2 = f_{\theta}} - \xi_{\theta_0}^2 - 2(\theta - \theta_0) \dot{\xi}_{\theta_0} \xi_{\theta_0} \\ & = (\theta - \theta_0)^2 \dot{\xi}_{\theta_0}^2 + r_{\theta}^2 + 2r_{\theta} \xi_{\theta_0} + 2(\theta - \theta_0) \dot{\xi}_{\theta_0} r_{\theta} \end{aligned}$$

The $\mathbb{L}^1(\mu)$ -norms of the first two terms in the right-hand side is of order $(\theta - \theta_0)^2$. The $\mathbb{L}^1(\mu)$ -norms of the last

two terms above are (by the Cauchy–Schwarz inequality) of order $\|r_{\theta}\|_{\mu}$, thus are $o(|\theta - \theta_0|)$. Multiplying both sides of the display above by the bounded T and integrating over μ , we obtain

$$\gamma_T(\theta) - \gamma_T(\theta_0) - 2(\theta - \theta_0) \int_{\mathcal{X}} \dot{\xi}_{\theta_0} \xi_{\theta_0} T \, d\mu = o(|\theta - \theta_0|). \quad \square$$

2.3.1 Overview of the proof. We introduce

$$\Delta : (x, \theta) \mapsto q'(\theta) \frac{\mathbf{1}_{\{q(\theta) > 0\}}}{2\sqrt{q(\theta)}} \xi_{\theta}(x) + \sqrt{q(\theta)} \dot{\xi}_{\theta}(x),$$

which is well-defined for almost all $\theta \in \Theta \cap \text{Supp}(q)$, and vanishes for $\theta \notin \text{Supp}(q)$. Let \mathfrak{m} denote the Lebesgue measure. We will show that

$$\begin{aligned} (2.1) \quad & 2 \int_{\Theta \times \mathcal{X}} \Delta(x, \theta) \sqrt{q(\theta)} \xi_{\theta}(x) (S(x) - \psi(\theta)) \, d\theta \, d\mu(x) \\ & = \int_{\Theta} \psi'(\theta) \, d\mathbb{Q}(\theta). \end{aligned}$$

We prove the equality (2.1) above in a direct way, and the van Trees inequality then follows by an application of the Cauchy–Schwarz inequality. Section 2.4 explains that (2.1) can actually be interpreted, under stronger assumptions, as a consequence of Lemma 6 with $T(x, \theta) = S(x) - \psi(\theta)$ and a well-chosen location model. Actually, a close look at the proof by Gill and Levit (1995, page 61) shows that they also exactly prove (2.1), though under additional regularity assumptions, like the $\theta \mapsto f_{\theta}(x)$ being absolutely continuous, and by first integrating in the left-hand side over θ then over x . We take the reverse order and first integrate over x , thanks to applications of Lemma 6, and then over θ .

2.3.2 Preparations. It suffices to prove (vT1) for statistics S given by finite linear combinations of indicator functions, the case of general statistics following by taking limits given the bounded second moment stated in Assumption 3. Similarly, the sequence of absolutely continuous functions $\psi_n = \max\{-n, \min\{\psi, n\}\}$ satisfies $\psi_n \rightarrow \psi$ and $\psi'_n \rightarrow \psi'$ almost-surely; by dominated convergence, it also suffices to prove (vT1) for bounded ψ with bounded derivatives.

The first assumption of Theorem 4 ensures that the function γ_S is absolutely continuous on $\Theta \cap \text{Supp}$. In addition, Lemma 6, based on the fact that \mathcal{P} is differentiable at almost all $\theta \in \Theta \cap \text{Supp}(q)$ and that S is in particular uniformly bounded, provides a closed-form expression for the almost-everywhere derivative γ'_S .

Finally, all integrands below belong to $\mathbb{L}^1(\mathfrak{m} \otimes \mu)$, as follows from applications of the Cauchy–Schwarz inequality. Hence, integrals of sums equal sums of integrals and Fubini's theorem may be applied to exchange orders of integration. We use the short-hand notation $\mu[f]$ for the expectation of a function $f : \mathcal{X} \rightarrow \mathbb{R}$ under μ .

2.3.3 *Proof of (2.1)*. Let $\Theta_q = \Theta \cap \text{Supp}(q)$. The integrals in (2.1) may be equivalently taken over Θ or Θ_q . The left-hand side of (2.1) consists of four terms, namely,

$$\begin{aligned} & \int_{\Theta_q} q'(\theta) \mu[f_\theta S] d\theta = \int_{\Theta_q} q'(\theta) \gamma_S(\theta) d\theta \\ & - \int_{\Theta_q} \psi(\theta) q'(\theta) \mu[f_\theta] d\theta = - \int_{\Theta_q} \psi(\theta) q'(\theta) d\theta, \\ & 2 \int_{\Theta_q} q(\theta) \mu[\dot{\xi}_\theta \xi_\theta S] d\theta = \int_{\Theta_q} q(\theta) \gamma'_S(\theta) d\theta, \\ & -2 \int_{\Theta_q} \psi(\theta) q(\theta) \mu[\dot{\xi}_\theta \xi_\theta] d\theta = 0. \end{aligned}$$

The fourth equality follows from Lemma 6 with $T \equiv 1$, which entails that $\mu[\dot{\xi}_\theta \xi_\theta] = 0$ for almost all $\theta \in \Theta_q$. Now, the functions γ_S , q and ψ are absolutely continuous on Θ_q , so that an integration by parts (Titchmarsh, 1939, page 375, §12.11) ensures that on any compact subinterval $[c, d] \subset \Theta_q$,

$$\begin{aligned} & \int_{[c,d]} (q'(\theta) \gamma_S(\theta) + q(\theta) \gamma'_S(\theta)) d\theta = [q(\theta) \gamma_S(\theta)]_c^d, \\ & \int_{[c,d]} \psi(\theta) q'(\theta) d\theta = [\psi(\theta) q(\theta)]_c^d - \int_{[c,d]} \psi'(\theta) q(\theta) d\theta. \end{aligned}$$

We write $\Theta_q = \Theta \cap \text{Supp}(q)$ as a countable union of disjoint intervals (a_τ, b_τ) , indexed by $\tau \in \mathcal{T}$. Each finite boundary point of Θ_q is either a finite boundary point of Θ , or lies in the interior of Θ and is a finite boundary point of $\text{Supp}(q)$; in the latter case, by continuity, ψ is bounded and q vanishes thereat. Therefore, by boundedness of γ_S and ψ and by the Θ -boundary assumptions on q , the quantities $\gamma_S(\theta) q(\theta)$ and $\psi(\theta) q(\theta)$ vanish as θ approaches any finite boundary point a_τ or b_τ of Θ_q . When $\pm\infty$ is a boundary point of Θ_q , given that q is integrable over Θ , the liminf of $q(\theta)$ is null as θ tends to $\pm\infty$. Therefore, by boundedness of γ_S and ψ again, for each τ , by letting $c \rightarrow a_\tau$ and $d \rightarrow b_\tau$ in a suitable manner and by dominated convergence, we have

$$\begin{aligned} & \int_{(a_\tau, b_\tau)} (q'(\theta) \gamma_S(\theta) + q(\theta) \gamma'_S(\theta)) d\theta = 0, \\ & \int_{(a_\tau, b_\tau)} \psi(\theta) q'(\theta) d\theta = - \int_{(a_\tau, b_\tau)} \psi'(\theta) q(\theta) d\theta. \end{aligned}$$

By dominated convergence, summing these inequalities over $\tau \in \mathcal{T}$ yields (2.1).

2.3.4 *Conclusion by a Cauchy–Schwarz inequality*. The van Trees inequality (vT1) follows by applying the Cauchy–Schwarz inequality to (2.1) together with the fact that

$$4 \int_{\Theta \times \mathcal{X}} \Delta(x, \theta)^2 d\theta d\mu(x) = \mathcal{I}_Q + \int_{\Theta} \mathcal{I}_P(\theta) dQ(\theta).$$

The equality above follows from the definitions of Fisher information (for the integrals of square terms) and the fact that the following integral (corresponding to the cross term) is null, since $\mu[\dot{\xi}_\theta \xi_\theta] = 0$ for almost all θ , as already noted above:

$$\int_{\Theta \times \mathcal{X}} q'(\theta) \mathbf{1}_{\{q(\theta) > 0\}} \dot{\xi}_\theta(x) \xi_\theta(x) d\theta d\mu(x) = 0.$$

That $\mathcal{I}_Q > 0$ follows from the impossibility of q to be a uniform distribution, because of the vanishing-at-the-border constraints. This concludes the proof of the first part of Theorem 4 and we now move to its last statement.

2.3.5 *Special case*. We finally show that when the model \mathcal{P} is $\mathbb{L}^2(\mu)$ -differentiable at all points of Θ_q , not just almost everywhere, the first assumption of Theorem 4 holds, namely, that for all events $A \in \mathcal{F}$, the functions $\gamma_A : \theta \mapsto \mathbb{P}_\theta(A)$ are absolutely continuous on Θ_q . Indeed, by Titchmarsh (1939, page 368, §11.83), it suffices to note that γ_A is differentiable everywhere on Θ_q (by Lemma 6 together with the assumption that the model \mathcal{P} is $\mathbb{L}^2(\mu)$ -differentiable everywhere), with a derivative γ'_A that is finite everywhere and locally integrable on Θ_q : by the Cauchy–Schwarz inequality,

$$|\gamma'_A(\theta)| = \left| 2 \int_{\mathcal{X}} \dot{\xi}_\theta \xi_\theta \mathbf{1}_A d\mu \right| \leq \sqrt{\mathcal{I}_P(\theta)} < +\infty.$$

The claimed local integrability follows from the bound above, the local integrability of $\mathcal{I}_P q$ (by Assumption 3), and the fact that by absolute continuity, $q \geq \delta$ for some $\delta > 0$ on any open interval $(a, b) \subseteq \Theta_q$.

2.4 The van Trees inequality as a Cramér–Rao bound

The Cramér–Rao bound (for possibly biased statistics T) is obtained as a corollary of Lemma 6. By applying the Cauchy–Schwarz inequality to the equality

$$\gamma'_T(\theta_0) = 2 \int_{\mathcal{X}} \dot{\xi}_{\theta_0} \xi_{\theta_0} T d\mu,$$

we get indeed, when $\mathcal{I}_P(\theta_0) > 0$,

$$\mathbb{E}_{\theta_0}[T^2] \geq \frac{(\gamma'_T(\theta_0))^2}{\mathcal{I}_P(\theta_0)}.$$

Actually, replacing in the argument above T by $T - c$, with $c = \mathbb{E}_{\theta_0}[T]$, yields the desired Cramér–Rao bound:

$$\text{Var}_{\theta_0}(T) = \mathbb{E}_{\theta_0}[(T - c)^2] \geq \frac{(\gamma'_T(\theta_0))^2}{\mathcal{I}_P(\theta_0)}.$$

Now, the van Trees inequality was obtained in Section 2.3 by an application of the Cauchy–Schwarz inequality to the equality (2.1), which was claimed to be a consequence of Lemma 6; this indicates that the van Trees inequality is exactly an instance of a Cramér–Rao bound (for the location model \mathcal{M} described below), at

least in the (slightly stronger) form of Corollary 7 below. The latter is an automatic improvement of Theorem 4, as its proof merely consists of applying Theorem 4 with $S - c$ and ψ (or, alternatively, with S and $\psi + c$) for a well-chosen c .

COROLLARY 7. *Under the assumptions of Theorem 4, we actually have the stronger lower bound*

$$\int_{\Theta} \mathbb{E}_{\theta} \left[(S - \psi(\theta))^2 \right] d\mathbb{Q}(\theta) \geq \left(\int_{\Theta} \mathbb{E}_{\theta} [S - \psi(\theta)] d\mathbb{Q}(\theta) \right)^2 + \frac{\left(\int_{\Theta} \psi'(\theta) d\mathbb{Q}(\theta) \right)^2}{\mathcal{I}_{\mathbb{Q}} + \int_{\Theta} \mathcal{I}_{\mathcal{P}}(\theta) d\mathbb{Q}(\theta)}.$$

Put differently, the van Trees inequality of Corollary 7 is not (only) to be understood as a Bayesian Cramér–Rao bound, as advocated by Gill and Levit (1995), it is exactly a Cramér–Rao bound. Similarly, van Trees (1968, page 72) underlines that he mimics the derivation of the Cramér–Rao bound to obtain his inequality, but does not see the latter as a very instance of the former.

We conclude this section by detailing our claim that the equality (2.1) may be seen, under suitable conditions (not required for our direct proof of Theorem 4), as a consequence of Lemma 6. We assume, in particular, that the support of q is δ -away from the border of Θ , i.e., that for all $\theta \in \Theta$ with $q(\theta) > 0$ and all $x \in [-\delta, \delta]$, one has $\theta + x \in \Theta$. This assumption ensures that the location model $\mathcal{M} = (\mathbb{M}_{\alpha})_{\alpha \in (-\delta, \delta)}$ is well defined, where \mathbb{M}_{α} is the distribution over $\Theta \times \mathcal{X}$ with density $(x, \theta) \mapsto q(\theta + \alpha) f_{\theta + \alpha}(x)$ with respect to $\mu \otimes \mathfrak{m}$. Under suitable conditions (not detailed), we may apply the same theorem as in Section 2.2 (Bickel et al., 1993, Proposition 1 or Lehmann and Romano, 2005, Theorem 12.2.1) establishing the $\mathbb{L}^2(\mu \otimes \mathfrak{m})$ -differentiability of \mathbb{M}_{α} at $\alpha_0 = 0$ and identifying its $\mathbb{L}^2(\mu \otimes \mathfrak{m})$ -derivative at $\alpha_0 = 0$, which we denote by Δ , with the pointwise derivative of $(x, \theta) \mapsto \sqrt{q(\theta + \alpha) f_{\theta + \alpha}(x)}$ at $\alpha_0 = 0$:

$$\Delta : (x, \theta) \mapsto q'(\theta) \frac{\mathbf{1}_{\{q(\theta) > 0\}}}{2\sqrt{q(\theta)}} \xi_{\theta}(x) + \sqrt{q(\theta)} \dot{\xi}_{\theta}(x).$$

For a bounded statistic S and an absolutely continuous and bounded target function ψ , whose derivative ψ' is also bounded, we consider the statistic $J(x, \theta) = S(x) - \psi(\theta)$. Its expectation under some \mathbb{M}_{α} equals

$$\begin{aligned} \gamma_J(\alpha) &= \mathbb{E}_{\mathbb{M}_{\alpha}} [J] \\ &= \int_{\mathcal{X} \times \Theta} (S(x) - \psi(\theta)) q(\theta + \alpha) f_{\theta + \alpha}(x) d\mu(x) d\theta \\ &= \int_{\Theta} \mathbb{E}_{\theta} [S] q(\theta) d\theta - \int_{\Theta} \psi(\theta - \alpha) q(\theta) d\theta. \end{aligned}$$

Differentiating the above equality at $\alpha_0 = 0$, we obtain, as claimed, the equality (2.1), whose left-hand side may be identified to $\gamma'_J(\alpha)$ thanks to Lemma 6, and whose right-hand side is obtained by differentiating under the integral sign.

3. MULTIVARIATE VERSION

There exist several ways to extend the van Trees inequality for multivariate estimation; see Gill and Levit (1995), who in turn refer to van Trees (1968) and Bobrovsky, Mayer-Wolf and Zakai (1987). We focus here on the elegant matrix-wise version by Letac (2022).

Let the statistical model $\mathcal{P} = (\mathbb{P}_{\theta})_{\theta \in \Theta}$ be indexed by an open set $\Theta \subseteq \mathbb{R}^p$, where $p \geq 2$. The estimation target will be some $\psi(\theta)$, where $\psi : \Theta \rightarrow \mathbb{R}^s$, and we consider some statistic $S : \mathcal{X} \rightarrow \mathbb{R}^s$ to that end. We still assume that \mathcal{P} is dominated by a σ -finite measure μ , with densities $f_{\theta} = d\mathbb{P}_{\theta}/d\mu$ such that $(\theta, x) \mapsto f_{\theta}(x)$ is measurable. In the sequel, $\|\cdot\|$ refers to the Euclidean norm in some \mathbb{R}^d space (with $d = p$ or $d = s$), and $\|\cdot\|_{\mu}$ denotes the Euclidean norm in $\mathbb{L}^2(\mu)$, i.e., for a function $g : \mathcal{X} \rightarrow \mathbb{R}^d$ in $\mathbb{L}^2(\mu)$,

$$\|g\|_{\mu} = \sqrt{\int_{\mathcal{X}} \|g\|^2 d\mu}.$$

3.1 Comparison to classic regularity assumptions

Both Gill and Levit (1995) and Letac (2022) assume some smoothness on the functions $\theta \mapsto f_{\theta}(x)$, for μ -almost all x , and also possibly on the border of Θ . These assumptions are useful to extend the integrations by parts performed in Section 2.3.3 to the multivariate case, via Stokes' theorem. More precisely, Letac (2022) assumes (this is what he calls a “regular Fisher model”) that the functions $\theta \mapsto f_{\theta}(x)$ are even C^1 -smooth but does not put any constraint on the boundary of Θ . Gill and Levit (1995) assume, in particular, that Θ is compact with a piecewise- C^1 -smooth boundary; as for the functions $\theta \mapsto f_{\theta}(x)$, they assume that they are “nice” for μ -almost all x in the sense of Definition 8 (which is actually a property that Sobolev functions enjoy, see Evans and Gariepy, 1992, Section 4.9). For $u = (u_1, \dots, u_d) \in \mathbb{R}^d$ and $i \in \{1, \dots, d\}$, we let u_{-i} denote the $(d - 1)$ -dimensional vector of all components of u but the i -th one, so that, by an abuse of notation, $u = (u_i, u_{-i})$. We introduce the projection of a subset $D \subseteq \mathbb{R}^d$ ignoring the i -th coordinates:

$$D_{-i} = \{u_{-i} \in \mathbb{R}^{d-1} : \exists u_i \in \mathbb{R} \text{ s.t. } (u_i, u_{-i}) \in D\}.$$

DEFINITION 8 (nice functions). Let $D \subseteq \mathbb{R}^d$ be an open domain, where $d \geq 2$. A function $\varphi : D \rightarrow \mathbb{R}$ is nice if for all $i \in \{1, \dots, d\}$, for almost all $u_{-i} \in D_{-i}$, the functions $u_i \mapsto \varphi(u_i, u_{-i})$ are absolutely continuous in the classic one-dimensional sense on the open domain $D(u_{-i}) = \{u_i \in \mathbb{R} : (u_i, u_{-i}) \in D\}$.

In particular, a function $\varphi : D \rightarrow \mathbb{R}$ that is nice admits at almost all $u \in D$ partial derivatives along canonical directions, which we denote by $\partial_1\varphi, \dots, \partial_d\varphi$. By an abuse of notation, we denote by $\nabla\varphi = (\partial_1\varphi, \dots, \partial_d\varphi)$ the vector of partial derivatives.

A vector-valued function is nice if each of its component is nice.

As in Section 2.3.3, we avoid issuing regularity assumptions on the functions $\theta \mapsto f_\theta(x)$ and replace them by $\mathbb{L}^2(\mu)$ -differentiability assumptions. Our version of the van Trees inequality only requires such an $\mathbb{L}^2(\mu)$ -differentiability to hold along canonical directions, not all directions. For the sake of a simpler exposition, and as in the second part of Theorem 4, we restrict our attention to a model that is $\mathbb{L}^2(\mu)$ -differentiable along canonical directions at all points. We denote by $u \otimes v = uv^\top$ the outer product of two vectors u and v (possibly of different lengths).

DEFINITION 9 (Differentiability in \mathbb{L}_2 along canonical directions). The μ -dominated statistical model \mathcal{P} indexed by an open subset $\Theta \subseteq \mathbb{R}^p$ is differentiable in $\mathbb{L}_2(\mu)$ at $\theta_0 \in \Theta$ along canonical directions if there exist scalar functions $\xi_{\theta_0,1}, \dots, \xi_{\theta_0,p} \in \mathbb{L}_2(\mu)$, called the $\mathbb{L}_2(\mu)$ -partial derivatives of the model at θ_0 , such that, for all $i \in \{1, \dots, p\}$, as $\theta_i \rightarrow \theta_{0,i}$,

$$\|\xi_{(\theta_i, \theta_{0,-i})} - \xi_{\theta_0} - (\theta_i - \theta_{0,i})\dot{\xi}_{\theta_0,i}\|_\mu = o(|\theta_i - \theta_{0,i}|).$$

Let $\dot{\xi}_{\theta_0} = (\dot{\xi}_{\theta_0,1}, \dots, \dot{\xi}_{\theta_0,p})$. The Fisher information $\mathcal{I}_{\mathcal{P}}(\theta_0)$ of the model at θ_0 is then defined as the $p \times p$ matrix

$$\mathcal{I}_{\mathcal{P}}(\theta_0) = 4 \int_{\mathcal{X}} \dot{\xi}_{\theta_0} \otimes \dot{\xi}_{\theta_0} d\mu.$$

While we avoid at all costs direct regularity assumptions on the functions $\theta \mapsto f_\theta(x)$, as we have no control on the model \mathcal{P} , we may be more lenient when it comes to the prior \mathbb{Q} , which the statistician chooses. Gill and Levit (1995) impose, among others, the following assumption on \mathbb{Q} , which generalizes Definition 2.

DEFINITION 10 (Well-behaved prior, multivariate version). We call a probability measure \mathbb{Q} that concentrates on the open set $\Theta \subseteq \mathbb{R}^p$ a well-behaved prior if \mathbb{Q} has a density q with respect to the Lebesgue measure on Θ that is nice on Θ , and whose vector of partial derivatives ∇q is such that $\|\nabla q\|_2^2 \mathbf{1}_{\{q>0\}}/q$ is Lebesgue-integrable. We define

$$\mathcal{I}_{\mathbb{Q}} \stackrel{\text{def}}{=} \int_{\Theta} \nabla q(\theta) \otimes \nabla q(\theta) \frac{\mathbf{1}_{\{q(\theta)>0\}}}{q(\theta)} d\theta.$$

3.2 Statement

The multivariate version of the van Trees inequality proposed by Letac (2022), as well as a consequence thereof (in terms of Schur complement) is stated in (vTm). Therein, where $M \succcurlyeq 0$ and $M \succ 0$ denote the fact that a symmetric matrix M is positive semi-definite and positive definite, respectively. Also, $\nabla\psi(\theta)$ denotes the $p \times s$ matrix whose component (i, j) equals $\nabla\psi(\theta)_{i,j} = \partial_i\psi_j(\theta)$.

The multivariate counterpart of Assumption 3 is stated next. It does not target generality and aims to ease exposition: as a consequence, it requires differentiability of the model at all points of $\Theta \cap \text{Supp}(q)$, not just almost everywhere, and also imposes that the density q is continuous (which does not follow from Definition 8).

ASSUMPTION 11 (for the multivariate case). The set Θ is any open subset of \mathbb{R}^p . The probability measure \mathbb{Q} is a well-behaved prior on Θ , with a continuous density q . The statistical model $\mathcal{P} = (\mathbb{P}_\theta)_{\theta \in \Theta}$ is dominated by a σ -finite measure μ , with densities $f_\theta = d\mathbb{P}_\theta/d\mu$ such that $(\theta, x) \mapsto f_\theta(x)$ is measurable. The model \mathcal{P} is differentiable in $\mathbb{L}^2(\mu)$ along canonical dimensions at all points of $\Theta \cap \text{Supp}(q)$. The function $\psi : \Theta \rightarrow \mathbb{R}^s$ is nice. Both $\|\psi\|^2$ and $\|\nabla\psi\|$ are \mathbb{Q} -integrable and both

$$\int_{\Theta} \mathbb{E}_\theta[\|S\|^2] d\mathbb{Q}(\theta), \int_{\Theta} \text{Tr}(\mathcal{I}_{\mathcal{P}}(\theta)) d\mathbb{Q}(\theta) < +\infty,$$

where Tr denotes the trace.

THEOREM 12. *The multivariate van Trees inequality (vTm) holds with $\mathcal{I}_{\mathbb{Q}} \succ 0$ under Assumption 11 and the fact that $q(\theta) \rightarrow 0$ as θ approaches any boundary point of Θ with finite norm along some canonical direction.*

3.3 Proof of Theorem 12

Up to resorting to dominated-convergence arguments (as in Section 2.3.2), we may restrict our attention to statistics S and to target functions ψ that are uniformly bounded.

3.3.1 Elements to perform integration by parts. The key to extend the univariate proof to a multivariate setting is the following lemma of integration by parts, which follows from a version of Stokes' theorem tailored to our needs. Its proof and some comments may be found in appendix.

LEMMA 13. *Let $D \subseteq \mathbb{R}^d$ be an open domain, where $d \geq 2$, and let $f, g : D \rightarrow \mathbb{R}$ be two functions that are nice on D , with g also being continuous, such that, for some $i \in \{1, \dots, d\}$,*

$$\int_D |f g| dm < +\infty \text{ and } \int_D |\partial_i f g + f \partial_i g| dm < +\infty,$$

$$(vTm) \quad \left[\begin{array}{cc} \int_{\Theta} \mathbb{E}_{\theta} [(S - \psi(\theta)) \otimes (S - \psi(\theta))] d\mathbb{Q}(\theta) & \left(\int_{\Theta} \nabla \psi(\theta) d\mathbb{Q}(\theta) \right)^{\top} \\ \int_{\Theta} \nabla \psi(\theta) d\mathbb{Q}(\theta) & \mathcal{I}_{\mathbb{Q}} + \int_{\Theta} \mathcal{I}_{\mathcal{P}}(\theta) d\mathbb{Q}(\theta) \end{array} \right] \succcurlyeq 0,$$

thus, whenever $\mathcal{I}_{\mathbb{Q}} \succ 0$,

$$\int_{\Theta} \mathbb{E}_{\theta} [(S - \psi(\theta)) \otimes (S - \psi(\theta))] d\mathbb{Q}(\theta) - \left(\int_{\Theta} \nabla \psi(\theta) d\mathbb{Q}(\theta) \right)^{\top} \left(\mathcal{I}_{\mathbb{Q}} + \int_{\Theta} \mathcal{I}_{\mathcal{P}}(\theta) d\mathbb{Q}(\theta) \right)^{-1} \left(\int_{\Theta} \nabla \psi(\theta) d\mathbb{Q}(\theta) \right) \succcurlyeq 0.$$

and $f(u)g(u) \rightarrow 0$ as u approaches any boundary point of D with finite norm along the i -th canonical direction. Then

$$\int_{D \cap \{q \neq 0\}} (\partial_i f g + f \partial_i g) dm = 0.$$

Denote by $\psi = (\psi_1, \dots, \psi_s)$ and $S = (S_1, \dots, S_s)$ the components of ψ and S . Given the assumptions of Theorem 12 and the boundedness of ψ , we may directly apply Lemma 13 to $D = \Theta$ and the pairs $f = \psi_j$ and $g = q$, where $j \in \{1, \dots, s\}$.

We wish to also do so with $D = \Theta \cap \text{Supp}(q)$ and the $f = \gamma_{S_j}$, where $\gamma_{S_j} : \theta \in \Theta \mapsto \mathbb{E}_{\theta}[S_j]$. The boundary of $\Theta \cap \text{Supp}(q)$ is included in the union of the boundaries of Θ and $\text{Supp}(q)$, and q vanishes when it approaches any of them. Together with the uniform boundedness of S_j , the boundary assumption of Lemma 13 is satisfied on $D = \Theta \cap \text{Supp}(q)$. It only remains to show that $f = \gamma_{S_j}$ is nice. To do so, we mimic and adapt arguments used in Section 2.3.5. Given that S_j is uniformly bounded, and given the $\mathbb{L}^2(\mu)$ -differentiability assumptions on the model, we may apply Lemma 6 along any canonical direction and get that the γ_{S_j} are differentiable in the i -th coordinate at all $\theta \in D$, with partial derivatives given by

$$(3.1) \quad \partial_i \gamma_{S_j}(\theta) = 2 \int \xi_{\theta,i} \xi_{\theta,i} S_j d\mu.$$

Denoting by B a uniform bound on the S_j , the Cauchy-Schwarz inequality guarantees that

$$|\partial_i \gamma_{S_j}(\theta)| \leq B \sqrt{\text{Tr}(\mathcal{I}_{\mathcal{P}}(\theta))} \leq B \left(1 + \text{Tr}(\mathcal{I}_{\mathcal{P}}(\theta)) \right).$$

Given the final integrability condition in Assumption 11 and the fact that q is nice, by Fubini's theorem, at almost all θ_{-i} , the function

$$\theta_i \in D \mapsto \text{Tr}(\mathcal{I}_{\mathcal{P}}(\theta_i, \theta_{-i})) q(\theta_i, \theta_{-i})$$

is integrable and $\theta_i \in D \mapsto q(\theta_i, \theta_{-i})$ is (absolutely) continuous, thus locally larger than some $\delta > 0$; recall indeed that $D = \Theta \cap \text{Supp}(q)$ here. Thus, $\theta_i \in D \mapsto \text{Tr}(\mathcal{I}_{\mathcal{P}}(\theta_i, \theta_{-i}))$ is locally integrable. Therefore, at these θ_{-i} , the function $\theta_i \in D(\theta_{-i}) \mapsto \gamma_{S_j}(\theta_i, \theta_{-i})$ is differentiable everywhere, with a derivative that is finite everywhere and locally integrable, thus (see again Titchmarsh, 1939, page 368, §11.83), it is absolutely continuous. This exactly corresponds to the fact that γ_{S_j} is nice on D .

3.3.2 Brief rest of proof of Theorem 12. We follow the same methodology as in Section 2.3, and introduce

$$\Delta(x, \theta) = \nabla q(\theta) \frac{\mathbf{1}_{\{q(\theta) > 0\}}}{2\sqrt{q(\theta)}} \xi_{\theta}(x) + \sqrt{q(\theta)} \dot{\xi}_{\theta}(x).$$

All integrands in the sequel belong to $\mathbb{L}^1(\mathfrak{m} \otimes \mu)$, as follows from applications of the Cauchy-Schwarz inequality. Hence, integrals of sums equal sums of integrals and Fubini's theorem may be applied to exchange orders of integration. We use again the short-hand notation $\mu[f]$ for the expectation of a function $f : \mathcal{X} \rightarrow \mathbb{R}^d$ under μ .

We show below that the multivariate van Trees inequality (vTm) corresponds to

$$\int_{\mathcal{X} \times \Theta} (V(x, \theta) \otimes V(x, \theta)) d\mu(x) d\theta \succcurlyeq 0,$$

$$\text{where } V(x, \theta) = \begin{bmatrix} (S(x) - \psi(\theta)) \xi_{\theta}(x) \sqrt{q(\theta)} \\ 2\Delta(x, \theta) \end{bmatrix}.$$

We start with the cross-products. As explained above, Lemma 6 may be applied along all canonical directions $i \in \{1, \dots, p\}$ to yield (3.1) as well as $\mu[\xi_{\theta,i} \xi_{\theta,i}] = 0$ for all $\theta \in \Theta \cap \text{Supp}(q)$. We therefore obtain the following extension of the four equalities of the beginning of Section 2.3.3: with the short-hand notation $\Theta_q = \Theta \cap \text{Supp}(q)$,

$$\begin{aligned} & 2 \int_{\mathcal{X} \times \Theta} \left(\Delta(x, \theta) \otimes (S(x) - \psi(\theta)) \right) \xi_{\theta}(x) \sqrt{q(\theta)} d\mu(x) d\theta \\ &= \int_{\Theta_q} \nabla q(\theta) \otimes \gamma_S(\theta) d\theta + \int_{\Theta_q} \nabla \gamma_S(\theta) q(\theta) d\theta \\ &\quad - \int_{\Theta_q} \nabla q(\theta) \otimes \psi(\theta) d\theta + (0, \dots, 0)^{\top}, \end{aligned}$$

where $\gamma_S = (\gamma_{S_j})_{1 \leq j \leq s}$ and $\nabla \gamma_S(\theta)$ is the $p \times s$ matrix whose component (i, j) equals $\nabla \gamma_S(\theta)_{i,j} = \partial_i \gamma_{S_j}(\theta)$. The results of Section 3.3.1 hold for all pairs (i, j) and thus guarantee that

$$\begin{aligned} \int_{\Theta_q} \nabla q(\theta) \otimes \gamma_S(\theta) d\theta &= - \int_{\Theta_q} \nabla \gamma_S(\theta) q(\theta) d\theta, \\ - \int_{\Theta_q} \nabla q(\theta) \otimes \psi(\theta) d\theta &= \int_{\Theta_q} \nabla \psi(\theta) q(\theta) d\theta. \end{aligned}$$

On the other hand, using again that $\mu[\dot{\xi}_{\theta,i} \xi_{\theta,i}] = 0$ for all $\theta \in \Theta_q$, we have that

$$\int_{\Theta} \nabla q(\theta) \otimes \mu[\dot{\xi}_{\theta} \dot{\xi}_{\theta}] \mathbf{1}_{\{q(\theta) > 0\}} d\theta = (0, \dots, 0)^T,$$

so that the bottom-right term in the multivariate van Trees inequality (**vTm**) corresponds to

$$\begin{aligned} & 4 \int_{\mathcal{X} \times \Theta} (\Delta(x, \theta) \otimes \Delta(x, \theta)) d\mu(x) d\theta \\ &= \int_{\Theta} \frac{\nabla q(\theta) \otimes \nabla q(\theta)}{q(\theta)} \mathbf{1}_{\{q(\theta) > 0\}} \mu[f_{\theta}] d\theta \\ &+ \int_{\Theta} 4 \mu[\dot{\xi}_{\theta} \otimes \dot{\xi}_{\theta}] q(\theta) d\theta \\ &= \mathcal{I}_{\mathbb{Q}} + \int_{\Theta} \mathcal{I}_{\mathcal{P}}(\theta) q(\theta) d\theta, \end{aligned}$$

where $\mathcal{I}_{\mathbb{Q}} \succ 0$ as q cannot be a uniform density due to the boundary conditions.

4. DIRECT PROOF OF LAM LOWER BOUNDS

Gill and Levit (1995, Section 3) provide a derivation of a version of the the Hájek–Le Cam convolution theorem (Hájek, 1970) based on the van Trees inequality. In the exact same vein, including the same techniques, we propose a version of the Hájek–Le Cam local asymptotic minimax [LAM] theorem (Hájek, 1972): see Theorem 16 below. We state it in a Hájek–Le Cam spirit, avoiding any classic regularity assumption (contrary to Gill and Levit, 1995, Section 3).

Its derivation is elementary and bypasses the typical arguments of the Hájek–Le Cam theory of convergence of experiments. However, our version requires, on many aspects, stronger assumptions than the original references, except for the differentiability of the model, which we only require along canonical directions (and not in all directions). See the comments after the statement of Theorem 16 for more detail.

Setting. We still consider an open subset $\Theta \subseteq \mathbb{R}^p$. For $n \geq 1$, we denote by $\mathbb{P}_{\theta}^{\otimes n}$ the law of a n -sample of observations based on some \mathbb{P}_{θ} , and $\mathcal{P}^{\otimes n} = (\mathbb{P}_{\theta}^{\otimes n})_{\theta \in \Theta}$ the associated statistical product model. When the base statistical model \mathcal{P} is differentiable in $\mathbb{L}_2(\mu)$ at some $\theta_0 \in \Theta$ along canonical directions, then so is $\mathcal{P}^{\otimes n}$, with a vector of $\mathbb{L}_2(\mu)$ -partial derivatives given by

$$(x_1, \dots, x_n) \mapsto \left(\sum_{k=1}^n \dot{\xi}_{\theta_0,i}(x_k) \prod_{k' \neq k} \xi_{\theta_0}(x_{k'}) \right)_{1 \leq i \leq p}.$$

In particular, the Fisher information of the product model $\mathcal{P}^{\otimes n}$ at θ_0 equals $\mathcal{I}_{\mathcal{P}^{\otimes n}}(\theta_0) = n \mathcal{I}_{\mathcal{P}}(\theta_0)$.

Consider some sequence of statistics $S_n : \mathcal{X}^n \rightarrow \mathbb{R}^s$ and fix for now some vector $U \in \mathbb{R}^s$. We assume the following.

ASSUMPTION 14. For a neighborhood N of $\theta_0 \in \Theta$, on the one hand, \mathcal{P} is differentiable in $\mathbb{L}_2(\mu)$ along canonical directions at all $\theta \in N$, and on the other hand, the \mathbb{R}^s -valued target function ψ is nice and bounded on N , with $\nabla \psi$ also bounded on N .

Derivation. For any distribution \mathbb{H} on \mathbb{R}^p , we denote by $\mathbb{Q}_{\theta_0, r}$ the distribution of $\theta_0 + rH$, where H is a random variable with distribution \mathbb{H} . There exist sufficiently regular priors \mathbb{H} on \mathbb{R}^p , with support in the unit ball \mathcal{B} , so that, for all $c > 0$, all assumptions of Theorem 12 are satisfied with $\mathbb{Q} = \mathbb{Q}_{\theta_0, c/\sqrt{n}}$, at least for n large enough (depending on \mathbb{H} and c), except maybe the finiteness of the two integrals stated in Assumption 11 (without which the inequality holds also but is pointless). Also, the Fisher information of $\mathbb{Q}_{\theta_0, c/\sqrt{n}}$ equals (n/c^2) times the Fisher information $\mathcal{I}_{\mathbb{Q}_{\theta_0, 1}}$ of $\mathbb{Q}_{\theta_0, 1}$.

Therefore, for such priors and for n large enough,

$$\begin{aligned} & \int_{\mathcal{B}} \mathbb{E}_{\theta_0 + ch/\sqrt{n}}^{\otimes n} \left[\left(U^T \left(S_n - \psi(\theta_0 + ch/\sqrt{n}) \right) \right)^2 \right] d\mathbb{H}(h) \\ & \geq \frac{1}{n} U^T G(\theta_0, c, n)^T I(\theta_0, c, n)^{-1} G(\theta_0, c, n) U, \end{aligned}$$

where we introduced the $p \times s$ and $p \times p$ matrices

$$\begin{aligned} G(\theta_0, c, n) &= \int_{\mathcal{B}} \nabla \psi(\theta_0 + ch/\sqrt{n}) d\mathbb{H}(h), \\ I(\theta_0, c, n) &= \frac{1}{c^2} \mathcal{I}_{\mathbb{Q}_{\theta_0, 1}} + \int_{\mathcal{B}} \mathcal{I}_{\mathcal{P}}(\theta_0 + ch/\sqrt{n}) d\mathbb{H}(h). \end{aligned}$$

Now, any positive quadratic form $\ell : \mathbb{R}^s \rightarrow [0, +\infty)$ can be decomposed as follows: there exists an orthogonal basis U_1, \dots, U_s of \mathbb{R}^s and nonnegative real numbers $\lambda_1, \dots, \lambda_s \geq 0$ such that for all $v \in \mathbb{R}^s$,

$$\ell(v) = \sum_{k=1}^s \lambda_k (U_k^T v)^2 = \sum_{k=1}^s \lambda_k U_k^T v v^T U_k.$$

This decomposition entails that for all $s \times s$ symmetric positive semi-definite matrices Γ , denoting by $\mathcal{N}([0], \Gamma)$ the Gaussian distribution over \mathbb{R}^s centered at $[0] = (0, \dots, 0)^T$ and with covariance matrix Γ ,

$$\int_{\mathbb{R}^s} \ell(v) d\mathcal{N}([0], \Gamma)(v) = \sum_{k=1}^s \lambda_k U_k^T \Gamma U_k.$$

Linear combinations of the applications above of the van Trees inequality thus yield

$$\begin{aligned} & \int_{\mathcal{B}} \mathbb{E}_{\theta_0 + ch/\sqrt{n}}^{\otimes n} \left[\ell \left(\sqrt{n} \left(S_n - \psi(\theta_0 + ch/\sqrt{n}) \right) \right) \right] d\mathbb{H}(h) \\ & \geq \int_{\mathbb{R}^s} \ell(v) d\mathcal{N}([0], \Gamma_{\theta_0, c, n})(v), \end{aligned}$$

where $\Gamma_{\theta_0, c, n} = G(\theta_0, c, n)^\top I(\theta_0, c, n)^{-1} G(\theta_0, c, n)$. By lower bounding a supremum by an integral, we obtain the desired LAM lower bound (4.1) below as soon as $\Gamma_{\theta_0, c, n}$ converges in the following sense. We recall that we do not aim for minimal assumptions in this section, but for elementary arguments.

ASSUMPTION 15. We have the component-wise convergence

$$\lim_{c \rightarrow +\infty} \lim_{n \rightarrow \infty} \Gamma_{\theta_0, c, n} = \nabla \psi(\theta_0)^\top \mathcal{I}_{\mathcal{P}}(\theta_0)^{-1} \nabla \psi(\theta_0).$$

It holds, in particular, as soon as $\nabla \psi$ and $\mathcal{I}_{\mathcal{P}}$ are continuous at θ_0 , with $\mathcal{I}_{\mathcal{P}}(\theta_0)$ being nonsingular.

THEOREM 16. *Under Assumptions 14 and 15, for all positive quadratic forms $\ell : \mathbb{R}^s \rightarrow [0, +\infty)$, for all sequences of statistics $S_n : \mathcal{X}^n \rightarrow \mathbb{R}^s$,*

$$(4.1) \quad \liminf_{c \rightarrow +\infty} \liminf_{n \rightarrow +\infty} \sup_{\theta: \sqrt{n} \|\theta - \theta_0\| \leq c} \mathbb{E}_{\theta}^{\otimes n} \left[\ell \left(\sqrt{n} (S_n - \psi(\theta)) \right) \right] \geq \int_{\mathbb{R}^s} \ell(v) d\mathcal{N}([0], \nabla \psi(\theta_0)^\top \mathcal{I}_{\mathcal{P}}(\theta_0)^{-1} \nabla \psi(\theta_0))(v).$$

Comments. van der Vaart (1998, Theorem 8.11) states the lower bound (4.1) for so-called bowl-shaped loss functions (not just quadratic forms), under the $\mathbb{L}^2(\mu)$ -differentiability of \mathcal{P} at θ_0 (only, not on a neighborhood thereof) in all directions (while Theorem 16 considered canonical directions only), and for ψ differentiable at θ_0 (in sharp contrast with the continuity and boundedness assumptions on ψ and $\mathcal{I}_{\mathcal{P}}$ in Theorem 16). That Theorem 4.1 may only deal with quadratic forms is unsurprising, given the quadratic nature of the van Trees inequality; but it came to us as a surprise that the results of Section 3, and thus Theorem 4.1, hold for differentiability assumed only along canonical directions. Takatsu and Kuchibhotla (2024) explore further local (asymptotic, or not) minimax bounds based, among others, on the van Trees inequality, and relax the regularity conditions both on the model and on the target function ψ .

REMARK 17. The non-singularity of $\mathcal{I}_{\mathcal{P}}(\theta_0)$ in Assumption 15 is actually not required to get a meaningful LAM bound from the van Trees inequality. We consider, for instance, the case of $\psi(\theta) = \theta$ and only assume that $\mathcal{I}_{\mathcal{P}}$ is continuous at θ_0 : the $I(\theta_0, c, n)$ still converge to $\mathcal{I}_{\mathcal{P}}(\theta)$, which may however be singular. Now, the proof above reveals that if $U \in \mathbb{R}^p$ is in the kernel of $\mathcal{I}_{\mathcal{P}}(\theta)$, then the LAM lower bound in (4.1) with $\ell(v) = (U^\top v)^2$ equals $+\infty$. Conversely, still under the continuity assumption of $\mathcal{I}_{\mathcal{P}}$ at θ_0 , if there exists an estimator having a finite local asymptotic maximum in quadratic risk $\ell(v) = \|v\|^2$, as in

the left-hand side of (4.1), then $\mathcal{I}_{\mathcal{P}}(\theta_0)$ is non singular. This can be used to get a simple proof of the non singularity of the efficient Fisher information in semiparametric estimation problems: such an argument has been used in Gassiat, Rousseau and Vernet (2018, for multidimensional mixtures) and in Moss and Rousseau (2024, for hidden Markov models) by applying a preliminary version of the proof of Theorem 16.

APPENDIX: PROOF OF LEMMA 13

We consider the following version of Stokes' theorem, where we use again the notation of Definition 8. Lemma 13 follows from it by considering the set $\mathcal{S} = \{g \neq 0\}$ and the product $\varphi = fg$, which is nice as absolute continuity in the classical sense is itself stable by products (Titchmarsh, 1939, page 375, §12.11). By continuity of g , the set \mathcal{S} is open and g vanishes at its boundary, while f (because it is nice) is such that $f(\cdot, u_{-i})$ is locally bounded for almost all $u_{-i} \in D_{-i}$.

LEMMA 18. *Let $D \subseteq \mathbb{R}^d$ be an open domain. Fix a nice function $\varphi : \Theta \rightarrow \mathbb{R}$ and $i \in \{1, \dots, d\}$ such that*

$$\int_D |\varphi| dm < +\infty \quad \text{and} \quad \int_D |\partial_i \varphi| dm < +\infty,$$

and such that $\varphi(u)$ tends to 0 as u approaches any boundary point of D with finite norm along the i -th canonical direction. Consider an open subset \mathcal{S} such that for almost all $u_{-i} \in D_{-i}$, one has $\varphi(u_i, u_{-i}) \rightarrow 0$ as u_i approaches a boundary point of \mathcal{S} located in the interior of $D(u_{-i})$. Then,

$$\int_{D \cap \mathcal{S}} \partial_i \varphi dm = 0.$$

PROOF. We introduce $G = D \cap \mathcal{S}$. By Fubini's theorem, it suffices to show that for almost all $u_{-i} \in G_{-i}$,

$$\int_{G(u_{-i})} \partial_i \varphi(u_i, u_{-i}) du_i = 0.$$

Now almost all $u_{-i} \in D_{-i}$ are such that the following holds: as φ is nice on D , $\varphi(\cdot, u_{-i})$ is absolutely continuous on the open domains $D(u_{-i})$ and $G(u_{-i})$; by Fubini's theorem,

$$\int_{D(u_{-i})} |\varphi(u_i, u_{-i})| du_i < +\infty$$

and

$$\int_{D(u_{-i})} |\partial_i \varphi(u_i, u_{-i})| du_i < +\infty;$$

by the \mathcal{S} boundary assumption, $\varphi(u_i, u_{-i}) \rightarrow 0$ as u_i approaches a boundary point of \mathcal{S} located in the interior of $D(u_{-i})$. We consider such a point $u_{-i} \in G_{-i}$ and mimic the one-dimensional arguments located in the second part

of Section 2.3.3. Namely, we write $G(u_{-i})$ as an (at most) countable disjoint union of open intervals,

$$G(u_{-i}) = \bigsqcup_{n \geq 1} (a_n(u_{-i}), b_n(u_{-i})),$$

where $a_n(u_{-i}) \in \mathbb{R} \cup \{-\infty\}$ and $b_n(u_{-i}) \in \mathbb{R} \cup \{+\infty\}$. By absolute continuity in the classical sense, for all $n \geq 1$, for all real numbers $a > a_n(u_{-i})$ and $b < b_n(u_{-i})$,

$$\int_a^b \partial_i \varphi(u_i, u_{-i}) du_i = \varphi(b, u_{-i}) - \varphi(a, u_{-i}).$$

The boundary of G is included in union of the boundaries of D and \mathcal{S} . The D and \mathcal{S} boundary assumptions on φ ensure $\varphi(a, u_{-i}) \rightarrow 0$ and $\varphi(b, u_{-i}) \rightarrow 0$ as $a \rightarrow a_n(u_{-i})$ and $b \rightarrow b_n(u_{-i})$, except maybe in the cases where $a_n(u_{-i}) = -\infty$ or $b_n(u_{-i}) = +\infty$. In the latter cases, we use that by integrability of $\varphi(\cdot, u_{-i})$ over $D(u_{-i})$, the liminf of this function must be null and let $a \rightarrow a_n(u_{-i})$ or $b \rightarrow b_n(u_{-i})$ in a careful way. In all cases,

$$\int_{a_n(u_{-i})}^{b_n(u_{-i})} \partial_i \varphi(u_i, u_{-i}) du_i = 0$$

and may sum the obtained equalities over $n \geq 1$, by dominated convergence, to get the equality claimed at the beginning of this proof. \square

ACKNOWLEDGMENTS

The authors would like to thank David Pollard for suggesting to study the van Trees inequality under the angle of a Cramér–Rao bound for a location model, and for following and encouraging this work since 2001, when he delivered a series of lectures during the statistics semester at Institut Henri Poincaré, Paris.

FUNDING

Elisabeth Gassiat was supported by Institut Universitaire de France and by ANR grants ANR-21-CE23-0035-02 and ANR-23-CE40-0018-02.

REFERENCES

- BICKEL, P. J., KLAASSEN, C. A. J., RITOV, Y. and WELLNER, J. A. (1993). *Efficient and Adaptive Estimation for Semiparametric Models*. Johns Hopkins University Press.
- BOBROVSKY, B. Z., MAYER-WOLF, E. and ZAKAI, M. (1987). Some classes of global Cramér–Rao bounds. *The Annals of Statistics* **15** 1421–1438.
- EVANS, L. C. and GARIÉPY, R. F. (1992). *Measure Theory and Fine Properties of Functions*. *Studies in Advanced Mathematics*. CRC Press.
- GASSIAT, E., ROUSSEAU, J. and VERNET, E. (2018). Efficient semi-parametric estimation and model selection for multidimensional mixtures. *Electronic Journal of Statistics* **12** 703–740.
- GILL, R. and LEVIT, B. (1995). Applications of the van Trees inequality: a Bayesian Cramér–Rao bound. *Bernoulli* **1** 59–79.
- HÁJEK, J. (1970). A characterization of limiting distributions of regular estimates. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* **14** 323–330.
- HÁJEK, J. (1972). Local asymptotic minimax and admissibility in estimation. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability* **1** 175–194.
- IBRAGIMOV, I. A. and HAS’MINSKII, R. Z. (1981). *Statistical Estimation: Asymptotic Theory*. Springer.
- JUPP, P. E. (2010). A van Trees inequality for estimators on manifolds. *Journal of Multivariate Analysis* **101** 1814–1825.
- LEHMANN, E. L. and ROMANO, J. P. (2005). *Testing Statistical Hypotheses*, third ed. Springer.
- LENSTRA, A. J. (2005). Cramér–Rao revisited. *Bernoulli* **11** 263–282.
- LETAC, G. (2022). The randomization by Wishart laws and the Fisher information. Preprint, arXiv:2211.14137.
- MOSS, D. and ROUSSEAU, J. (2024). Efficient Bayesian estimation and use of cut posterior in semiparametric hidden Markov models. *Electronic Journal of Statistics* **18** 1815–1886.
- POLLARD, D. (2001; 2005). Asymptotia (book in progress), chapter on Hellinger differentiability. Lecture notes from the Paris 2001 statistics semester at IHP, with a final edit in 2005; available at <http://www.stat.yale.edu/~pollard/Courses/607.spring05/handouts/DQM.pdf>.
- TAKATSU, K. and KUCHIBHOTLA, A. K. (2024). Generalized van Trees inequality: Local minimax bounds for non-smooth functionals and irregular statistical models. Preprint, arXiv:2405.06437.
- TITCHMARSH, E. C. (1939). *The Theory of Functions*, second ed. Oxford University Press.
- VAN DER VAART, A. W. (1998). *Asymptotic Statistics*. Cambridge University Press.
- VAN TREES, H. L. (1968). *Detection, Estimation and Modulation Theory*. Wiley & Sons.