Location estimation at variance-gamma logarithmic poles

Scott Schubr
Luminary Oddities

November 2026

Abstract

We determine the sharp behaviour of location estimation for variance-gamma laws at the critical shape, where the density has a logarithmic pole, and settle the rate and limit-distribution questions for the leave-one-out estimator that Nitithumbundit and Chan studied by simulation. In one dimension, the global adjacent-midpoint leave-one-out estimator of Nitithumbundit and Chan converges at rate $(\log n)/n$ to a drifted Brownian argmax, with convergence of all fixed absolute moments; with $I = \pi/\sigma$ its limiting normalized absolute and squared risks are $3/I$ and $26/I^2$. In dimension at least two, shape below one gives a sharp Poisson shift rate, shape above one gives ordinary Gaussian local asymptotics, and shape one gives an exact critical Gaussian experiment at rate $(n \log\log n)^{-1/2}$ in the plane and $(n \log n)^{-1/2}$ in higher dimensions; a truncated score attains the critical minimax squared-risk constants. When shape, skew and scatter are unknown, the joint local experiment at shape one splits into independent critical location and ordinary nuisance Gaussian shifts, and an explicit pilot followed by a centered truncated score preserves the exact critical constant uniformly on compact brackets.

Contents

1 Introduction **2**

2 The leave-one-out location estimator at the variance-gamma logarithmic pole **2**
2.1 Estimator and results . . . 2
2.2 Density estimates and a coarse bound . . . 3
2.3 Fine approximation and increment bounds . . . 4
2.4 Process convergence and maximization . . . 5
2.5 The limiting absolute and squared risks . . . 6

3 Sharp location rates for multivariate variance-gamma laws **7**
3.1 Model and phase diagram . . . 7
3.2 Density and weighted differences . . . 8
3.3 The subcritical Poisson phase . . . 9
3.4 Critical experiments . . . 10
3.5 Attaining the critical information . . . 11
3.6 Compact parameter sets and the minimax constant . . . 11

4 Critical variance-gamma location with unknown shape, skew and scatter **12**
4.1 Model and results . . . 12
4.2 Density bounds near the pole . . . 13
4.3 A moment pilot for the full parameter . . . 13

4.4 The joint local experiment 14
4.5 Estimating location with unknown nuisance 16

1 Introduction

Variance-gamma laws have a logarithmic density singularity at their centre when the shape parameter equals one, and a power singularity below one. Location estimation at such a singularity falls between the Poisson regime of a jump and the regular Gaussian regime. Podgórski and Wallin [2] introduced maximum leave-one-out likelihood for unbounded densities and proved consistency, super-efficiency and a lower bound on the rate for power poles. Nitithumbundit and Chan [4, 3] recorded that the optimal rate and the asymptotic distribution of this estimator had no theoretical treatment and studied both by simulation, including the critical variance-gamma case. Section 2 settles both questions at the logarithmic pole: it gives the exact one-dimensional limit of the leave-one-out estimator at the logarithmic pole. Section 3 determines the sharp rates and minimax constants in every dimension at least two. Section 4 shows that the critical constant is unaffected when shape, skew and scatter must also be estimated. The Brownian argmax constants are classical; see [1].

2 The leave-one-out location estimator at the variance-gamma logarithmic pole

For a symmetric univariate variance-gamma density at shape 1/2, the maximum leave-one-out likelihood location estimator converges at rate $(\log n)/n$ to the maximizer of a two-sided Brownian motion with negative drift proportional to distance from zero. All fixed absolute moments converge after normalization, with explicit limiting absolute and squared risks. This gives an exact limit for the critical case simulated by Nitithumbundit and Chan. Two capped likelihoods, used only in the proof, connect the actual maximizer to the local likelihood process.

2.1 Estimator and results

Let $X_i = \theta_0 + Z_i$, $i = 1, \dots, n$, be independent, with $n \geq 2$, and let the density of Z_i be

$$f_\sigma(x) = \frac{1}{\pi\sigma} K_0(|x|/\sigma), \qquad \sigma > 0.$$

The scale and shape are fixed and known. Write $j(t)$ for a sample index closest to t and let

$$Q_n(t) = \sum_{i \neq j(t)} \log f_\sigma(X_i - t).$$

The data are distinct almost surely. Ties at a midpoint give the same criterion. Define $\widehat{\theta}_n$ to be any measurable global maximizer. It can be obtained by checking the adjacent sample midpoints. Indeed, the integral representation of $K_0(x)$ as a positive mixture of e^{-xu} for $u \geq 1$ gives strict radial decrease and strict log convexity for $x > 0$. On each Voronoi interval of a data point, that point is omitted and the remaining log densities are convex in t. The two unbounded intervals have their maxima at their finite ends by monotonicity. This also proves existence and bounds the maximizer by the sample range.

Let $\mathcal{B}$ be standard two-sided Brownian motion and write $d_n = (\log n)/n$.

Theorem 2.1. *For every θ_0, with $I = \pi/\sigma$,*

$$\frac{\widehat{\theta}_n - \theta_0}{d_n} \implies U_\sigma = \operatorname*{argmax}_{u\in\mathbb{R}} \left\{ \sqrt{I}\, \mathcal{B}(u) - \frac{I}{2}|u| \right\}.$$

For every fixed $a > 0$, $\mathbb{E}|(\widehat{\theta}_n - \theta_0)/d_n|^a \to \mathbb{E}|U_\sigma|^a$. In particular the limiting first absolute moment is $3/I$ and the limiting second moment is $26/I^2$. On every fixed compact interval in u, the centered process $Q_n(\theta_0 + ud_n) - Q_n(\theta_0)$ converges in the space of continuous functions to the Brownian process in the display.

In the notation of Nitithumbundit and Chan [5], this is the model with dimension one, $\nu = 1/2$, $\gamma = 0$ and known $\Sigma = \sigma^2$. Their Section 5.3 maximizes over the adjacent midpoints with the other parameters fixed. Their simulated power-law fit does not determine the logarithmic factor. The theorem also supplies the limiting distribution they examined numerically. Podgórski and Wallin [6] prove super-efficiency for power-law poles under their assumption $-1 < \alpha < 0$. The logarithmic pole has a different normalization. The fixed-parameter logarithmic contact experiment in [9] gives the finite-dimensional likelihood limit here. We prove uniform process control and localization of the unmodified leave-one-out maximizer.

2.2 Density estimates and a coarse bound

Translation and dilation reduce the proof to $\theta_0 = 0$, $\sigma = 1$. Let $f = f_1$, $A = 1/\pi$, and $L = \log n$, so $d_n = L/n$. There is a fixed $0 < x_0 < 1/10$ such that

$$f(x) = A\log(1/|x|) + c_f + o(1), \qquad |f'(x)| \le C/|x| \quad (0 < |x| < x_0).$$

These and the exponential tails follow from the Bessel representations and expansions in [8]. The logarithmic derivative is bounded on $|x| \ge x_0$. Also $|\log f(x)| \le C(1 + |x|)$ there, using the large-argument expansion. In particular $f \in L^2(\mathbb{R})$.

For $0 < r < x_0$, put $f_r(x) = f(|x| \vee r)$, with $g_r = f_r / \int f_r$. This description uses symmetry of f. By integrating the derivative from $|x|$ to r,

$$\|f - g_r\|_1 \le Cr, \qquad \|f^{\otimes n} - g_r^{\otimes n}\|_{\mathrm{TV}} \le Cnr.$$

Writing $\phi_r = \log f_r$, use the centered capped process

$$C_{n,r}(t) = \sum_{i=1}^{n} [\phi_r(X_i - t) - \phi_r(X_i)].$$

At any t with at most one observation within r, the difference between the centered leave-one-out and capped criteria is exactly the change in the capped omitted term, provided the same holds at zero. This identity includes $t = X_i$, where the leave-one-out criterion omits that observation.

We first show that for every fixed $b_0 > 0$ one can choose B such that all maximizers lie within

$$W = BL^2/n \tag{1}$$

with probability $1 - O(n^{-b_0})$. Pick $D > b_0 + 5$ and use a coarse cap $r_g = n^{-D}$. The probability of a pair within $2r_g$ is $O(n^2 r_g)$ by Cauchy–Schwarz and $f \in L^2$. All data are in $[-C_0 L, C_0 L]$ except with probability $O(n^{-b_0})$ for C_0 large. On these two events the deleted capped terms change by at most $C_1 L$ between zero and any t in that deterministic range. Indeed the nearest distance is

at most $2C_0L$, and the largest capped log density is $O(\log L)$. A maximizer there must therefore satisfy $C_{n,r_g}(t) \geq -C_1L$.

Under $g_{r_g}^{\otimes n}$, the exponential half moment of $C_{n,r_g}(t)$ is the n-th power of the affinity of g_{r_g} and its translate. With $H^2(g,h) = \int(\sqrt{g} - \sqrt{h})^2$, it is bounded by $\exp(-nH^2(g_{r_g}, g_{r_g}(\cdot - t))/2)$. For small nonzero t, the squared distance is at least $c|t|/\log(1/|t|)$. By symmetry take $t > 0$ and restrict the integral to $x = tz$, $2 \leq z \leq 3$. Here $f(tz) - f(t(z-1))$ tends uniformly to $A\log((z-1)/z)$ while the squared sum of the square roots is of order $\log(1/t)$. This proves the bound. For $|t|$ bounded away from zero, this distance is bounded away from zero by identifiability, continuity of L^2 translations, and decay of the L^2 correlation of $\sqrt{f}$ at infinity. The triangle inequality for H transfers these lower bounds to g_{r_g}, since $H(f, g_{r_g}) \leq C\sqrt{r_g}$. In particular for $|t| \geq W/2$, $nH^2 \geq cBL$ for sufficiently large n. Markov's inequality on a deterministic mesh of width $n^{-(D+3)}$ in the data range rules out $C_{n,r_g} \geq -C_1L - 1$ at mesh points there except with probability $O(n^{-b_0})$, upon enlarging B. The capped log derivative is bounded globally by C/r_g, so the mesh error for the sum is $o(1)$. Product total variation transfers the union bound from g_{r_g} to f; this costs $O(nr_g)$. Together with the pair and sample range bounds, this proves (1).

2.3 Fine approximation and increment bounds

Fix $b_1 > 0$, choose $K > b_1 + 5$, and use the fine cap $r = n^{-1}L^{-K}$ only on $[-W, W]$. The probability of a pair within $2r$ with both observations in $[-W - 2r, W + 2r]$ is at most

$$Cn^2 r \int_{-W-4r}^{W+4r} f(x)^2\, dx \leq CL^{4-K}.$$

For the first bound apply Cauchy–Schwarz to the density product at each pair displacement. With probability $1 - O(\exp(-cL^3))$ there is an observation in $[-W, W]$, which gives nearest distances at most $2W$ on the entire interval. Hence uniformly on this interval the largest variation of the capped omitted term is at most $\phi_r(0) - \phi_r(2W) = O(\log L/L)$. Consequently $Q_n(t) - Q_n(0)$ and $C_{n,r}(t)$ differ uniformly by $O(\log L/L)$ except with probability $O(L^{-b_1})$. All these observations at true density f can be transferred to g_r at cost $O(L^{-K})$ in probability, and conversely.

Now work under g_r. For $|u| \leq BL$ define

$$Y_u(x) = \phi_r(x - ud_n) - \phi_r(x), \qquad Z_n(u) = \sum_i [Y_u(X_i) - \mathbb{E}_{g_r} Y_u(X_i)].$$

For $|u|, |v| \leq BL$ the following estimates hold with $\epsilon_n = C\log L/L$:

$$\|Y_u - Y_v\|_\infty \leq \epsilon_n, \qquad n\mathbb{E}_{g_r}(Y_u - Y_v)^2 \leq C|u - v|. \tag{2}$$

For $|x| \leq 4W$, use $|\frac{d}{dt}\log f(t)| \leq C/(Lt)$ for $r \leq t \leq 5W$. Since $\log(5W/r) = O(\log L)$ this gives the first bound there. For the second, composition with the lower cap $\log(|x| \vee r) = \max(\log|x|, \log r)$ is a contraction. The squared L^2 shift difference of the uncapped $\log|x|$ is π^2 times the shift length, by the transform of the logarithm, $-\pi/|\xi|$ away from zero, and Plancherel. The reference density g_r is globally at most CL, so the contribution of this region is at most $Cn\eta/L$, where $\eta = d_n|u - v|$.

For $|x| > 4W$ the shifted arguments are comparable to x and outside the cap. The mean value theorem gives the supremum bound $C\eta/(W\log(1/W)) + C\eta$, at most C/L. For the squared integral use the same derivative bound under $g_r(x) \leq C\log(1/|x|)$ near zero. Its value times n is bounded by

$$Cn\eta^2\left[1 + \int_{4W}^{x_0} \frac{dx}{x^2\log(1/x)}\right] \leq Cn\eta^2/(W\log(1/W)) \leq Cn\eta/L.$$

The integral bound follows by splitting at $\sqrt{W}$. This proves (2).

We will truncate the dyadic grids at a polynomially fine power of L. To bound the interpolation error, the number of observations in $[-4W, 4W]$ is $O(L^{b_1+5})$ except with probability $O(L^{-b_1})$, by the expected count $O(L^3)$ and Markov. The individual Lipschitz bound in u there is Cd_n/r. For the rest of the observations the expected sum of their supremum Lipschitz constants is $O(L^2)$: integrate the log derivative times g_r, bounded near zero by $C/|x|$ on $|x| > 4W$, and multiply by $nd_n = L$. Markov gives a bound $O(L^{b_1+4})$ with the same failure probability. The expectation term subtracted in Z_n has deterministic Lipschitz bound of the same polynomial order, using the expected near count. Thus there is a fixed D_1 for which the Lipschitz constant of Z_n on the entire interval is at most L^{D_1} except with probability $O(L^{-b_1})$. Take $J_n = \lceil (D_1+4)\log_2 L \rceil$. Every dyadic grid of level J_n on an interval of length at most $2BL$ then has interpolation error $o(1)$ on this one common event.

2.4 Process convergence and maximization

We give the stochastic bounds with the grid error understood on the common event just described. On an interval of length R, anchored at an endpoint, use dyadic parent and child edges of length at most $R2^{-j+1}$. Bernstein's inequality and (2) show that, except with probability Ce^{-s}, the magnitude of all the level j increments is at most

$$C\big(\sqrt{R2^{-j}(s+2j+1)} + \epsilon_n(s+2j+1)\big).$$

Here the probability bound after summing over all levels uses at most 2^{j+1} edges at a level and includes the top endpoint increment. Summing the thresholds through level J_n bounds the supremum from the anchor by

$$C\sqrt{R}(1+\sqrt{s}) + C\epsilon_n[J_n(1+s) + J_n^2] + o(1). \tag{3}$$

The result holds also with a specified interior anchor, by applying it on the two intervals to its left and right.

On a fixed compact interval, cover by intervals of length at most 2δ, including enough overlap so any pair at distance at most δ is covered. Apply (3) with $s = c + \log(C/\delta)$. After a union bound the failure probability is $O(e^{-c})$ plus the common interpolation failure probability. For fixed δ, c the term with ϵ_n tends to zero. Sending δ to zero gives stochastic equicontinuity of Z_n. The deterministic drift is equicontinuous asymptotically too. Indeed e^{Y_u} is a true one-observation likelihood ratio under g_r. On fixed compacts $\|Y_u\|_\infty = o(1)$, so uniformly there

$$n\mathbb{E}_{g_r} Y_u = -\tfrac{1}{2} n\mathbb{E}_{g_r} Y_u^2 + o(1)$$

by Taylor expansion with (2). The quadratic term is equicontinuous by Cauchy–Schwarz and (2).

For the finite-dimensional limit, put $x = d_n y$. At fixed $y \notin \{0, u\}$, the small-argument expansion gives

$$LY_u(d_n y) \;\longrightarrow\; D_u(y) := \log|y| - \log|y-u|, \qquad \frac{g_r(d_n y)}{L} \;\longrightarrow\; A.$$

On $|x| \le 4W$, the logarithmic contraction used above bounds $|LY_u(d_n y)|$ by $C|D_u(y)|$, and g_r/L is bounded. Thus dominated convergence applies on the whole mark line since $D_u \in L^2$ and $4W/d_n \to \infty$. For $|x| > 4W$ the earlier squared increment bound yields $O(nd_n^2/(W\log(1/W))) = o(1)$ for fixed u. Using $nd_n = L$ and polarization,

$$n\mathbb{E}_{g_r}(Y_u Y_v) \;\longrightarrow\; A\int_{\mathbb{R}} D_u(y) D_v(y)\, dy = \frac{\pi}{2}(|u| + |v| - |u-v|).$$

The product of means times n tends to zero, by the Taylor drift bound. The uniform single-observation envelope tends to zero, so the triangular central limit theorem gives the centered Gaussian vector, and the Taylor drift has limit $-\pi|u|/2$. Combined with tightness and the fine comparison, this proves compact process convergence of the leave-one-out criterion under the true density.

It remains to confine the maximizer to a bounded range in $u = t/d_n$. For $1 \le |u| \le BL$, the same Hellinger comparison as in the coarse bound gives $nH^2(g_r, g_r(\cdot - ud_n)) \ge c|u|$ uniformly. Indeed $nH^2(f, f(\cdot - ud_n)) \ge c'|u|$ by the pole lower bound, and $nH^2(f, g_r) = o(1)$. The log likelihood drift has upper bound $-nH^2$, by $\log x \le 2(\sqrt{x} - 1)$. Thus the drift is at most $-c|u|$ across the range. Apply (3) on the interval from zero to $\min(2R, BL)$, and on its negative, with $s = \gamma R$, where $\gamma > 0$ is sufficiently small. For all sufficiently large R and n the right-hand side is at most $cR/2$: the square-root term has coefficient $O(\sqrt{\gamma})$ at order R, and $\epsilon_n(J_n + J_n^2) = o(1)$ and $\epsilon_n J_n \gamma R = o(R)$. Except with probability $O(e^{-\gamma R})$, the centered capped process therefore cannot overcome the drift on the shell $R \le |u| \le \min(2R, BL)$. The fine comparison changes the criterion by $o(1)$. Sum these bounds over dyadic shells. The single common interpolation event and the fine comparison are paid once. Consequently, for R above a fixed constant and $R \le BL$, the likelihood maximizer has tail bounded by $Ce^{-c_2 R} + O(L^{-b_1}) + O(n^{-b_0})$. The coarse bound already controls the event $|u| > BL$. The probability transfer between the true and capped reference is $O(L^{-K})$, absorbed in these estimates.

The limiting Brownian path with negative linear radial drift tends to minus infinity at both ends and has an almost surely unique global maximizer. On either half-line uniqueness on compact intervals follows from the Brownian maximum property and absolute continuity under addition of linear drift. The two independent half-line suprema have continuous exponential distributions by the Brownian hitting probability with drift. The compact uniform process limit, tightness of all maximizers and the argmax mapping now give the stated limit in law.

Finally the coarse and fine analysis used arbitrary b_0, b_1, while the estimator stayed the same. On the coarse event the scaled maximizer has magnitude at most BL. Integration of the preceding shell tail bound makes any fixed power uniformly integrable there upon choosing b_1 larger than that power. Off the coarse event use the sample range bound $|\widehat{\theta}_n| \le \max_i |X_i|$; exponential tails give $\mathbb{E}(\max_i |X_i|)^{2a} = O(L^{2a})$. Cauchy–Schwarz and the arbitrary n^{-b_0} bound make its contribution to the a-th moment of the scaled maximizer vanish. This proves convergence of all fixed absolute moments. Rescaling the data restores σ, with Brownian information $I = \pi/\sigma$ by Brownian scaling.

2.5 The limiting absolute and squared risks

The canonical Brownian argmax law and the moment 26 go back to Terentyev, as recorded in [7]. We give a short derivation for both risk constants.

By Brownian scaling, $|IU_\sigma|$ is distributed as the absolute maximizing time of $B(t) - |t|/2$ for standard two-sided Brownian motion B. For a one-sided Brownian motion with drift $-a$, $a > 0$, write M, T for the eventual maximum and its time. The first-passage transform to level $m > 0$ is

$$\mathbb{E}[e^{-s\tau_m}; \tau_m < \infty] = e^{-m(a+\sqrt{a^2+2s})}, \qquad s \ge 0.$$

This follows from the exponential Brownian martingale stopped at τ_m and a lower level tending to minus infinity. The case $s = 0$ gives the exponential tail of M. For a grid of levels $m = j\delta$, the strong Markov property at τ_m gives

$$\mathbb{E}[e^{-s\tau_m}; m \le M < m + \delta] = e^{-m(a+\sqrt{a^2+2s})}(1 - e^{-2a\delta}).$$

The first-passage times at lower grid approximations to M converge almost surely to T, by continuity, uniqueness of the maximum and the fact that the path tends to minus infinity. Riemann sums against bounded continuous functions therefore identify the joint transform

$$\mathbb{E}[e^{-sT};\, M \in dm] = 2a\, e^{-m(a+\sqrt{a^2+2s})}\, dm.$$

For the two-sided maximum take $a = 1/2$, multiply by the probability $1-e^{-m}$ that the independent opposite maximum is smaller, and double. The Laplace transform of $|IU_\sigma|$ is thus

$$\frac{2}{(1/2+\sqrt{1/4+2s})(3/2+\sqrt{1/4+2s})}.$$

Its first two moments are 3 and 26, as follows by differentiation at zero. The moment convergence in Theorem 2.1 proves the corresponding absolute and squared risk constants for the leave-one-out estimator.

3 Sharp location rates for multivariate variance-gamma laws

For fixed symmetric variance-gamma noise in dimension $d \geq 2$, location is regular above shape one and has a Poisson local experiment below shape one. At shape one the limit is Gaussian, with a rate involving $\log\log n$ in dimension two and $\log n$ in higher dimensions. We obtain the exact information matrix at both critical rates and construct an estimator attaining its inverse.

3.1 Model and phase diagram

Observe n independent copies of $X = \theta + Z$ for $\theta \in \mathbb{R}^d$, $d \geq 2$, and known shape $\nu > 0$. Normalize the characteristic function of Z to $(1+|t|^2)^{-\nu}$. A fixed known positive definite scatter is included by an invertible linear transformation. Write f for the density of Z, I_d for the identity matrix, and ω_{d-1} for the surface area of the unit sphere in $\mathbb{R}^d$.

Theorem 3.1. *The minimax expected Euclidean error and root mean squared Euclidean error for θ both have order d_n, where*

$$d_n = \begin{cases} n^{-1/(2\nu)}, & 0 < \nu < 1, \\ (n \log\log n)^{-1/2}, & \nu = 1,\ d = 2, \\ (n \log n)^{-1/2}, & \nu = 1,\ d \geq 3, \\ n^{-1/2}, & \nu > 1. \end{cases}$$

For fixed θ_0 and every compact parameter set in u, the local experiments at $\theta_0 + d_n u$ converge in Le Cam distance. If $0 < \nu < 1$, the limit is the Poisson process on $\mathbb{R}^d$ with intensity

$$A_{\nu,d}|x-u|^{2\nu-d}\, dx, \qquad A_{\nu,d} = \frac{\Gamma(d/2-\nu)}{4^\nu \pi^{d/2}\Gamma(\nu)}.$$

If $\nu = 1$, the limit is a Gaussian shift of information

$$\mathcal{I}_d = \begin{cases} \frac{1}{2} I_2, & d = 2, \\ \dfrac{d-2}{2d} I_d, & d \geq 3. \end{cases}$$

There is an estimator with normalized centered limit $N(0, \mathcal{I}_d^{-1})$ and with normalized squared risk tending to $\operatorname{tr}(\mathcal{I}_d^{-1})$ in both critical cases. That trace is also the limiting normalized minimax squared risk. If $\nu > 1$, the Gaussian shift at the root-sample rate uses the ordinary location Fisher information.

In the symmetric parameterization of Nitithumbundit and Chan [5] with positive scatter Σ, the linear map from the normalization above is $(\Sigma/(2\nu))^{1/2}$. Thus at criticality the information in their coordinates is Σ^{-1} in dimension two and $(d-2)\Sigma^{-1}/d$ in higher dimensions, at the corresponding rates. Their one-dimensional numerical study addresses a different local regime, with the logarithmic pole at shape one-half. Its sharp location rate and the actual leave-one-out limit are given in [9] and Section 2. The multivariate results here concern the experiment and attainable location rates with known shape and scatter. The power-pole limits of Ibragimov and Khas'minskii [10] in one-dimensional location and the Poisson models of Dachian [11] are related. Gupta, Lee and Price [12] use smoothed Fisher information for high-probability location bounds for known laws. The planar criticality here uses the annular information of a logarithmic density pole, which accumulates at a double-logarithmic rate.

3.2 Density and weighted differences

Let G be gamma with shape ν and rate one. Conditional on G, Z is centered normal with covariance $2GI_d$. With $r = |x|$ its density is

$$f(r) = \frac{1}{(4\pi)^{d/2}\Gamma(\nu)} \int_0^\infty t^{\nu-d/2-1} e^{-t-r^2/(4t)}\, dt.$$

For $\nu < d/2$, substitution $s = r^2/(4t)$ gives $f(r) \sim A_{\nu,d} r^{2\nu-d}$ and $f'(r) \sim (2\nu-d)A_{\nu,d} r^{2\nu-d-1}$, by differentiating first and using the same substitution. The first two derivative orders are bounded by $Cr^{2\nu-d-j}$, $j = 1, 2$, near zero. For $\nu = d/2$, split the integral at r^2 and at one to get $f(r) \sim a_d \log(1/r)$, where $a_d = 2/[(4\pi)^{d/2}\Gamma(d/2)]$. Differentiation and substitution give $f'(r) \sim -a_d/r$ and $|f''(r)| \le Cr^{-2}$. For $\nu > d/2$, f tends to a positive finite value. The derivative formula then gives the bounds $O(r^{2\nu-d-1})$ for $\nu < d/2+1$, $O(r\log(1/r))$ at equality, and $O(r)$ above it. These bounds follow by splitting the derivative integral at r^2 and at one. The integral is the Bessel expression proportional to $r^{\nu-d/2}K_{\nu-d/2}(r)$, so its tail and differentiated tail estimates follow also from [8]. In particular the tails are exponential and the logarithmic derivative is bounded on radii bounded away from zero.

We record the shift costs in $H^2(f,g) = \int(\sqrt{f} - \sqrt{g})^2$. Let $D = |h|$ be the shift length and fix a small $r_0 > 0$. For a power pole with $0 < \nu \le 1$ and $\nu < d/2$, the part $|x| \le C_0 D$ costs $O(D^{2\nu})$: bound by the sum of the masses of the two translates. For $C_0 D < |x| < r_0$, square-root differentiation and the preceding bounds give a cost at most

$$CD^2 \int_{C_0 D}^{r_0} r^{2\nu-3}\, dr.$$

On an outer ray beyond fixed radius the cost is $O(D^2)$, by the square-integrable derivative of $\sqrt{f}$ there. When $0 < \nu < 1$, these inequalities give $O(D^{2\nu})$ overall and $O(D^{2\nu}R^{2\nu-2}) + O(D^2)$ outside $|x| \le RD$, for $R > C_0$.

For $\nu = 1, d = 2$, the inner ball $|x| \le C_0 D$ has smaller shift cost $O(D^2/\log(1/D))$. To check this, the integral formula, or the K_0 series [8], gives $f(r) = a_2 \log(1/r) + O(1)$ and $a_2 = 1/(2\pi)$. Both densities on this ball are bounded below by $c\log(1/D)$. Set $x = Dy$ and use the square-root identity: the scaled squared density difference is bounded by a constant times one plus the squared logarithms of the distances to the two poles. These logarithms are square integrable on the bounded mark set. From $C_0 D$ to r_0 the derivative bound for $\sqrt{f}$ gives cost

$$CD^2 \int_{C_0 D}^{r_0} \frac{dr}{r\log(1/r)}.$$

These estimates show $O(D^2 \log\log(1/D))$ overall. For $\nu = 1, d \geq 3$ the previous power estimates give $O(D^2 \log(1/D))$. In all cases the bounds in this section can be restricted to a smaller radial annulus, with the corresponding endpoints in the integrals.

For $\nu > 1$, the location Fisher information is finite. Near a power pole the radial integrand is $O(r^{2\nu-3})$; near a logarithmic pole with $d > 2$ it is $O(r^{d-3}/\log(1/r))$. Near a positive finite density at zero, the derivative bounds above are square integrable. There is no boundary term in the weak derivative of $\sqrt{f}$ at zero: its boundary integral over radius r vanishes by the same local density bounds. Thus $\sqrt{f}$ has an L^2 gradient, and translations are differentiable in quadratic mean by the L^2 translation identity. The usual iid local asymptotic normality theorem then gives the positive Fisher limit [13]. It also gives $H^2 = O(D^2)$; the sample mean attains the root-sample bound using the finite second moment.

Multiplying the shift costs by n and taking sufficiently small fixed multiples of d_n gives the two-point lower bounds for expected norm and root mean squared norm in Theorem 3.1. We prove the remaining upper bounds and experiments below.

3.3 The subcritical Poisson phase

Take $0 < \nu < 1$ and analyze the local model at $\theta_0 = 0$. In the mark $x = X/d_n$ the one-draw density multiplied by n converges in L^1 on bounded sets to $A_{\nu,d}|x-u|^{2\nu-d}$, by the near-zero asymptotic and domination by this integrable pole. In a fixed mark ball the binomial count and conditional marks therefore converge in total variation to the corresponding Poisson process. The omitted product squared Hellinger cost relative to $u = 0$ is $O(R^{2\nu-2}) + o(1)$ outside a mark ball of large radius R. The same bound holds for the omitted Poisson intensities by the square-root derivative of the power profile. Regenerate discarded observations from the normalized reference complement conditional on the retained count, and do the same with reference intensities for the Poisson marks. The complement has mass bounded away from zero, so normalization changes the Hellinger bound by at most a fixed factor. Sending R to infinity proves convergence of the finite experiments.

For an estimator, use one portion of the sample for a sample-mean pilot within a small fixed bandwidth of θ. The pilot has exponentially small failure probability at any fixed tolerance because the centered laws have exponential tails. On an independent portion count points within distance h of a proposed center z. For $|z-\theta| \leq \epsilon h$ with fixed $\epsilon < 1/4$, the expected count fraction is at least $ch^{2\nu}$. If $|z-\theta| \geq Bh$, the fraction is at most $C(B-1)^{2\nu-d}h^{2\nu}$ for large B, up to an $O(h^d)$ bound if the trial ball is far from the pole. Choose B large and h small to get a fixed gap between the two normalized probabilities. The probability of any ball is $O(h^{2\nu})$ uniformly, by the radial decrease and the pole estimate. Bernstein's inequality gives $O(\exp(-cnh^{2\nu}))$ failure probability for each candidate to estimate its normalized count fraction to within a fixed limit smaller than the gap.

Start at a small fixed bandwidth and use dyadic scales down to $h_f \asymp n^{-1/(2\nu)}$. At each level consider a deterministic lattice of spacing at most $\epsilon h/\sqrt{d}$ and examine the points within Kh of the previous center, where $K > 2B+2$. Choose one exceeding a threshold strictly between the near and far means, or retain the previous center if none does. Given all previous successes, the candidate set contains a near point and is contained in a deterministic set of bounded cardinality within $O(h)$ of the true location. Thus one may reuse the same data at every level and bound the probability of the first failure by $C\exp(-cnh^{2\nu})$ using an unconditional union bound over that deterministic set. If the first failure is at h, the total remaining movement is $O(h)$, giving error $O(h)$. The sum of squared errors from first failures is $O(\sum_{h \geq h_f} h^2 \exp(-cnh^{2\nu})) = O(h_f^2)$. On pilot failure the error is bounded by the pilot error plus a fixed constant, which contributes exponentially small squared risk by its finite fourth moment and the pilot failure bound. This proves the subcritical upper rate.

3.4 Critical experiments

Take $\nu = 1$, and put $\Lambda_n = \log\log n$ if $d = 2$ and $\Lambda_n = \log n$ if $d \geq 3$. The critical local displacement order is $\delta_n = (n\Lambda_n)^{-1/2}$. Set $\Delta = n^{-1/2}(\log n)^2$ and retain the annulus $\Delta \leq |x| \leq r_0$, with a single censoring symbol for all other values. We first show that this retains the local experiments. If $d \geq 3$, the omitted cost at $|x| < \Delta$ times n is

$$O(n\delta_n^2[1 + \log(\Delta/\delta_n)]) = o(1)$$

by the restricted shift bounds above. For $d = 2$ it is at most

$$Cn\delta_n^2 \left[\frac{1}{\log(1/\delta_n)} + \int_{C_0\delta_n}^{\Delta} \frac{dr}{r\log(1/r)}\right] = o(1).$$

Indeed the integral is the difference of double logarithms with ratio of the inner logarithms tending to one. Constants may depend on the fixed finite set of local parameters. The omitted tail cost is $O(n\delta_n^2) = o(1)$. The discarded region has reference mass bounded below by a fixed positive tail probability. Regenerating its normalized reference values thus proves asymptotic equivalence to censoring.

Let $\ell(x) = -\nabla \log f(x)$ be the location score, and define the matrix

$$J_\Delta = \int_{\Delta \leq |x| \leq r_0} \ell(x)\ell(x)^{\mathsf{T}} f(x)\, dx.$$

The angular average of the radial unit vector's outer product is I_d/d. In dimension two, the density and its radial derivative satisfy

$$f(r) \sim (2\pi)^{-1}\log(1/r), \qquad f'(r) \sim -(2\pi r)^{-1}.$$

They give $J_\Delta \sim \frac{1}{2}\log\log(1/\Delta) I_2$. For $d \geq 3$,

$$f(r) \sim \frac{r^{2-d}}{(d-2)\omega_{d-1}}, \qquad \frac{f'(r)}{f(r)} \sim -\frac{d-2}{r}.$$

The Gamma relation for ω_{d-1} gives the coefficient from $A_{1,d}$. Hence $J_\Delta \sim \frac{d-2}{d}\log(1/\Delta) I_d$. In both cases $J_\Delta/\Lambda_n \to \mathcal{I}_d$.

For clarity the likelihood expansion on the retained experiment is a triangular iid Gaussian expansion. On the annulus, shifted densities differ relatively from the reference by $O(\delta_n/\Delta) = o(1)$. Subtract the linear term $\delta_n u^{\mathsf{T}}\ell(x)$ from the relative difference. The resulting total weighted squared error is at most $Cn\delta_n^4\Delta^{-2} = o(1)$. On the relevant segment the radii are comparable to r. The Hessian of the density is $O(r^{-d})$ in the power case and $O(r^{-2})$ in the planar logarithmic case. Dividing its squared bound by the density and integrating on the annulus gives $O(\Delta^{-2})$ in both. The censoring atom has no linear term since the annulus is centered and the score is odd. Its quadratic remainder cost is bounded by the continuous cost using Cauchy–Schwarz and its positive mass. The exact likelihood increments have mean zero, uniformly vanish, and have limiting n-fold covariance $u^{\mathsf{T}}\mathcal{I}_d u'$. The multivariate triangular central limit theorem gives the linear Gaussian term; squared sums converge in probability to their covariance and the cubic remainder tends to zero by the uniform bound. The limit likelihood vector is that of the stated Gaussian shift. Its coordinates have mean one. Convergence in law and convergence of these means imply finite Le Cam convergence by truncating and quantizing the likelihood vector under the reference, then restoring the reference conditional within each quantization cell. Together with the censoring equivalence this establishes the critical local limits.

3.5 Attaining the critical information

Use a pilot mean $\widetilde{\theta}$ from $n_0 = \lfloor n/\log n \rfloor$ observations and let $n_1 = n - n_0$. Write $e = \theta - \widetilde{\theta}$. The pilot satisfies $\mathbb{E}|e|^4 = O(n_0^{-2})$ and $\Pr(|e| > \Delta/4) \le C\exp(-c(\log n)^3)$, by the exponential moment and a fixed-dimensional coordinate bound on sample means. Put $\ell_\Delta(x) = \ell(x)\mathbf{1}_{\{\Delta \le |x| \le r_0\}}$. Using the remaining sample define

$$\widehat{\theta} = \widetilde{\theta} + J_\Delta^{-1} \frac{1}{n_1} \sum_{i > n_0} \ell_\Delta(X_i - \widetilde{\theta}).$$

For $|e| \le \Delta/4$, its fitted score mean under the true parameter is

$$\int \ell_\Delta(x) f(x - e)\, dx = J_\Delta e + O(|e|^2/\Delta).$$

Oddness gives the zero-order term and the score definition gives the linear one. For the remainder, combine the Hessian estimates with $|\ell(x)| \le C/r$ on the annulus and integrate $C\int_\Delta^{r_0} r^{-2}\, dr$. Then the squared bias of the estimator after normalization by $\sqrt{n\Lambda_n}$ is bounded by

$$C\frac{n\,\mathbb{E}|e|^4}{\Lambda_n \Delta^2} = o(1).$$

The score covariance at $e = 0$ is J_Δ. When $|e|/\Delta \to 0$, it is asymptotic to the same matrix. In fact the reference-relative density error is $O(|e|/\Delta)$ on the score support, using the derivative bounds, and the squared mean is negligible by the preceding expansion. On the whole good pilot event the second moment is bounded by a fixed multiple of J_Δ. The normalized single-observation score envelope is $C/(\Delta\sqrt{n\Lambda_n}) = o(1)$. Conditionally on the pilot, the triangular central limit theorem therefore gives the asserted Gaussian limit and the conditional covariance yields the normalized mean squared risk. The probability of the bad pilot is superpolynomially small. The score is deterministically bounded by C/Δ and the pilot has bounded fourth moments of its centered error, so the bad event contributes $o(1)$ to normalized squared risk. This estimator is translation equivariant. All rate bounds hold uniformly in θ.

3.6 Compact parameter sets and the minimax constant

For compact local parameter sets, choose a finite net. The experiments on the net converge by the preceding proofs, with a Gaussian limit in the regular case as well. For two input parameters u, v with $s = |u - v| \le 1/2$, the product squared Hellinger bounds from Section 2 are $Cs^{2\nu}$ in the Poisson phase, $Cs^2(1 + \log(1/s))$ at criticality, and Cs^2 above it, uniformly for large n. At criticality this uses $D = d_n s$ inside the logarithms. The limits are also uniformly continuous in Hellinger distance, by the shifted intensity bound or the Gaussian affinity. A kernel that compares the restrictions to the net therefore compares every parameter to within the net error plus these two continuity moduli. Letting the mesh tend to zero yields compact-set convergence.

To see the exact critical minimax lower constant, fix a large ball in the local parameter scale and project the local decisions to that ball. Projection cannot increase squared loss for parameters there, and makes the loss bounded. The compact Gaussian limit thus lower bounds the original global minimax risk after normalization by $n\Lambda_n$. In the Gaussian shift, use a centered normal prior of covariance $\tau^2 I_d$ conditioned to the ball of radius R. For estimators valued in that ball, extending the integrated squared loss to the whole prior adds at most $\int_{|u|>R}(R + |u|)^2\, d\Pi_\tau(u)$, tending to zero as $R \to \infty$ for fixed τ. The whole-prior Gaussian Bayes risk is $\mathrm{tr}[(\mathcal{I}_d + \tau^{-2} I_d)^{-1}]$. Let $R \to \infty$, then $\tau \to \infty$. The resulting lower bound equals the upper constant from the risk of the estimator in Section 5.

4 Critical variance-gamma location with unknown shape, skew and scatter

At shape one, multivariate variance-gamma location has a faster rate than the other parameters. We show that the joint local experiment splits into independent Gaussian shifts for location and for shape, skew and scatter. The location normalization is $\sqrt{n \log\log n}$ in two dimensions and $\sqrt{n \log n}$ in higher dimensions. A cumulant pilot and an annular score estimator attain the same limiting location risk as in the experiment with the other parameters known.

4.1 Model and results

Let $d \geq 2$ be fixed. Observe independent copies of

$$X = \theta + \gamma G + \sqrt{G}\, S^{1/2} V, \qquad G \sim \mathrm{Gamma}(\nu, 1), \quad V \sim N(0, I_d),$$

where G, V are independent and S is real symmetric positive definite. Here $\eta = (\nu, \gamma, S)$ is unknown and θ is the location parameter and gives the pole position at shape one. Use ordinary Euclidean coordinates on the vector and independent matrix entries. Fix an interior base point (θ_0, η_0) with $\nu_0 = 1$, and put

$$\Lambda_n = \begin{cases} \log\log n, & d = 2, \\ \log n, & d \geq 3, \end{cases} \qquad \delta_n = (n\Lambda_n)^{-1/2}, \qquad \mathcal{K}(S) = \begin{cases} S^{-1}, & d = 2, \\ \dfrac{d-2}{d} S^{-1}, & d \geq 3. \end{cases}$$

Write f_η for the density with $\theta = 0$, $\ell_\eta = -\nabla_x \log f_\eta$ for the location score away from the pole, and $s_\eta = \nabla_\eta \log f_\eta$ for the ordinary nuisance score. All score identities are understood almost everywhere.

Theorem 4.1. *At $\theta_n = \theta_0 + \delta_n u$, $\eta_n = \eta_0 + h/\sqrt{n}$, the local experiments converge on every compact set of (u, h) in Le Cam distance to independent Gaussian shifts of information $\mathcal{K}(S_0)$ for u and*

$$\mathcal{J}(\eta_0) = \mathbb{E}_{\eta_0}[s_{\eta_0} s_{\eta_0}^{\mathsf{T}}]$$

for h. The nuisance information is finite and positive definite. In particular the location information is the same with the nuisance parameters unknown.

Theorem 4.2. *Fix compact brackets for θ, γ, S with the eigenvalues of S bounded away from zero. For each fixed $H < \infty$, there is an estimator using unknown shape within $|\nu - 1| \leq H/\sqrt{n}$ and unknown γ, S that satisfies*

$$\sqrt{n\Lambda_n}(\widehat{\theta} - \theta) \implies N(0, \mathcal{K}(S_0)^{-1}) \quad \text{along every sequence with } S \to S_0.$$

Its normalized squared risk differs from $\mathrm{tr}(\mathcal{K}(S)^{-1})$ by $o(1)$ uniformly on these brackets. This is the sharp local minimax constant at every interior critical point.

Nitithumbundit and Chan [14, 5] study this normal mixture parameterization and construct likelihood algorithms for the skew multivariate model. Their simulations of the asymptotic location law and rate in [5, Section 5.3] use the univariate symmetric model with the other parameters fixed. Methods of moments for variance-gamma models are classical. Seneta [15] uses moment fitting in the skew univariate model. Fischer, Gaunt and Sarantsev [16] study their properties and modifications in the symmetric case. Here moments produce a preliminary estimate. The critical location information comes from observations near the pole, and the split experiments and attaining estimator show which information survives when the other parameters vary. The known symmetric phase diagram is derived in Section 3.

4.2 Density bounds near the pole

Let $L = (S/2)^{1/2}$, $y = L^{-1}x$, $\beta = L^{-1}\gamma$ and $r = |y|$. Completing the square in the normal density gives

$$f_\eta(x) = \frac{|\det L|^{-1} e^{\beta \cdot y/2}}{(4\pi)^{d/2}\Gamma(\nu)} B(\nu, a, r), \quad B(\nu, a, r) = \int_0^\infty t^{\nu - d/2 - 1} e^{-at - r^2/(4t)}\, dt, \quad a = 1 + |\beta|^2/4. \tag{4}$$

Uniformly on compact brackets for η near $\nu = 1$ and $0 < |x| \le r_0$ for small fixed r_0, this implies

$$|\nabla_x \log f_\eta| \le C/|x|, \qquad \frac{\|D_x^2 f_\eta\|}{f_\eta} \le C/|x|^2, \tag{5}$$

$$|s_\eta| \le C(1 + \log(1/|x|)), \qquad \|D_\eta \ell_\eta\| \le C(1 + \log(1/|x|))/|x|. \tag{6}$$

To check these bounds, inserting t^{-j} in the integral for B costs at most $C_j r^{-2j}$. For $t \ge r^2$ this follows directly. For $t < r^2$ put $z = r^2/(4t)$ and compare the resulting bound to the part of B from $r^2 \le t \le 2r^2$. A further $|\log t|$ costs at most $C(1 + \log(1/r))$ in the same bounds. For $r^2 \le t \le 1$ this is immediate and for $t > 1$ the exponential factor supplies the bound, using a fixed interval away from zero as a lower bound for B. Differentiation in a introduces t, harmless by splitting at one. Differentiating in ν introduces $\log t$. Spatial differentiation uses the $j = 1, 2$ bounds, while differentiation of $L^{-1}x$ in a parameter has size $O(|x|)$. Applying these estimates to (4), including its smooth prefactors, proves (5) and (6). Outside a fixed ball the log spatial derivative is uniformly bounded on the same parameter brackets, by the large-argument asymptotic of the equivalent Bessel kernel. The tails are uniformly exponential, also directly from the normal mixture moment-generating function.

At $\nu = 1$, remove the exponential tilt in (4) and use y coordinates. The radial part has the same leading coefficients and first logarithmic derivative as in the zero-skew density. Indeed the substitution $z = r^2/(4t)$ gives the power coefficient $1/[(d-2)\omega_{d-1}]$ for $d \ge 3$, independent of a. For $d = 2$, the density is $(2\pi)^{-1}\log(1/r) + O(1)$ and its radial derivative is asymptotic to $-(2\pi r)^{-1}$. Splitting at r^2 and one gives the first expression with bounded remainder. Substitution in the differentiated integral gives the second. The tilt tends to one and adds a bounded term to the location score. Its multiplicative error $O(r)$ and the bounded score addition contribute $O(1)$ to the annular information, using integrability of $f_\eta/|x|$ at a critical pole. Thus, for $A_\Delta = \{x : \Delta \le |x| \le r_0\}$ and $\Delta = n^{-1/2}(\log n)^2$, at $\nu = 1$

$$\Lambda_n^{-1} J_\Delta(\eta) \;\longrightarrow\; \mathcal{K}(S), \qquad J_\Delta(\eta) = \int_{A_\Delta} \ell_\eta \ell_\eta^{\mathsf{T}} f_\eta. \tag{7}$$

The convergence is uniform on the brackets. To see the normalization, angular integration in the y coordinates gives $I_2/2$ times $\log\log(1/\Delta)$ in the plane, and $(d-2)I_d/d$ times $\log(1/\Delta)$ in higher dimensions. The radial cutoffs after the linear change lie between fixed multiples of the displayed cutoffs, which changes the leading terms by $o(\Lambda_n)$. Then conjugate by L^{-1} and use $\log(1/\Delta) \sim (\log n)/2$. In particular $f_\eta(1 + \log(1/|x|))/|x|$ is integrable near zero, uniformly at shape one. Multiplication by $|x|^{-\epsilon}$ for a fixed sufficiently small $\epsilon > 0$ preserves this property.

4.3 A moment pilot for the full parameter

A finite moment map is uniformly invertible on the stated brackets. In any coordinate, the marginal characteristic function after removing θ_j factors as

$$(1 - iat)^{-\nu}(1 + ibt)^{-\nu}, \qquad a - b = \gamma_j, \quad ab = S_{jj}/2, \quad a, b > 0.$$

To determine ν from the first coordinate, use its cumulants c_k of orders $k = 2, 3, 4$. They satisfy $c_k = (k-1)!\nu(a^k + (-b)^k)$. Set $\rho = (a^2 + b^2)^{1/2}$, $w = (a-b)/\rho$, $k_1 = c_3/(2c_2)$, and $k_2 = c_4/(6c_2)$. Then

$$\frac{k_1}{\sqrt{k_2}} = \frac{w(3-w^2)}{\sqrt{2}\sqrt{1+2w^2-w^4}}, \qquad k_2 = \frac{\rho^2}{2}(1+2w^2-w^4).$$

The derivative of the first displayed function has strictly positive numerator proportional to $(1-w^2)(3-w^4)$. Since positive a, b stay in compact brackets, w stays a fixed distance from ± 1. Thus the inverse recovers w, ρ, then $\nu = c_2/\rho^2$, with uniform Lipschitz constants.

Given ν, put $B_1 = \mathrm{Cov}(X)/\nu = S + \gamma\gamma^{\mathsf{T}}$. For each coordinate the third cumulant divided by $\nu(B_{1,jj})^{3/2}$ equals $3v_j - v_j^3$ for $v_j = \gamma_j/\sqrt{B_{1,jj}}$. Its derivative is uniformly positive on the compact range away from ± 1. This recovers γ with a Lipschitz inverse, then $S = B_1 - \gamma\gamma^{\mathsf{T}}$ and $\theta = \mathbb{E}X - \nu\gamma$. All these quantities use only the mean, covariance, coordinate third cumulants and the first coordinate's fourth cumulant. They are smooth functions of the requisite raw moments on the relevant compact set.

Fit ν in a fixed compact positive bracket containing one in its interior. Let $n_0 = \lfloor n/a_n \rfloor$ with $a_n = \sqrt{\log\log n}$ for large n. Estimate the raw moments on the pilot sample, truncating their integrands to $|X| \le C_0 \log n$. The omitted expectations are $O(n^{-A_0})$ for any fixed A_0 upon choosing C_0 large enough. Use an approximate minimum-distance fit to the model moments over the compact brackets, to tolerance n^{-1}. The inverse Lipschitz inequalities, bounded variances and fourth moments of the empirical integrands give squared and fourth power pilot errors $O(n_0^{-1})$ and $O(n_0^{-2})$. At error threshold Δ/C for any fixed $C > 0$, the bounded integrand envelope is $O((\log n)^4)$ with $\Delta(\log n)^4 = o(1)$. Bernstein's inequality thus bounds the failure probability by $C\exp(-cn_0\Delta^2)$, which is superpolynomially small. These estimates hold uniformly on the brackets.

The nuisance model at fixed θ is differentiable in quadratic mean. Indeed the complete model (G, X) has square-integrable scores continuous in quadratic mean on compact interior brackets: differentiate the gamma density in shape and the normal conditional density in mean and covariance. Gamma log moments and normal polynomial moments bound the resulting score norms uniformly. Marginalizing gives the usual conditional expectation of the complete score and preserves quadratic mean differentiability [17]. This is s_η from (6). If a linear combination of its coordinates vanished, all the corresponding derivatives of the moment map would vanish, by differentiating under the complete model. The map has a smooth inverse in a neighborhood of the point, by the displayed formulas with nonzero denominators and derivatives. Hence that combination is zero only for the zero direction. This proves positive definiteness of $\mathcal{J}$.

4.4 The joint local experiment

Center at $\theta_0 = 0$. Fix finitely many (u, h) including $(0, 0)$, let $f_h = f_{\eta_0 + h/\sqrt{n}}$ and $f_0 = f_{\eta_0}$, and abbreviate $A = A_\Delta$. First replace $f_h(x - \delta_n u)$ off the annulus A by $f_h(x)$, and normalize the resulting density. The product laws change by $o(1)$ in total variation.

Here are details for the omitted region. At η_0 the planar pole at scale $D = |u|\delta_n$ has inner shift cost $O(D^2/\log(1/D))$ in the ball $|x| \le CD$. After $x = Dy$, the logarithmic density differences there have square integrable dependence on the two pole distances and a square-root denominator bounded below by $c\log(1/D)$. In the power case $d \ge 3$ the inner cost is $O(D^2)$ by the two ball masses. Between CD and Δ the cost is bounded by $CD^2 \int dr/(r\log(1/r))$ in the plane and $CD^2 \int dr/r$ in higher dimensions, by the square-root derivative and the asymptotics of (4). Both omitted costs times n tend to zero: $\log(\Delta/D) = O(\log\log n)$ for fixed nonzero u and the planar difference of inner double logarithms is $o(1)$. These bounds also hold at $\eta_0 + h/\sqrt{n}$ for the discarded inner ball

in the required $o(1/n)$ sense. Indeed the restricted nuisance squared Hellinger cost between f_h and f_0, on a ball of radius 2Δ and likewise after the location shift, is bounded by

$$\frac{C}{n}\int_{|x|\le 3\Delta} f_0(x)(1+\log(1/|x|))^2|x|^{-\epsilon}\,dx = o(1/n)$$

using (6) along the short path. Combine by the triangle inequality for the square-root norms. Outside the fixed outer ball, the shift costs $O(D^2)$ uniformly in the nuisance by tail derivative bounds. Normalization of the spliced density changes the Hellinger bound by at most a fixed multiple, by the reverse triangle inequality for square-root norms. This proves the replacement for the finite experiment.

The replaced density relative to f_h is $(1+k(x))/(1+\mathbb{E}_h k)$, where

$$k(x) = \mathbf{1}_A(x)\left\{\frac{f_h(x-\delta_n u)}{f_h(x)} - 1\right\}, \qquad \mathbb{E}_h k = \int k f_h.$$

By (5), $|k| \le C\delta_n/\Delta = o(1)$. On the annulus the density ratios f_h/f_0 tend uniformly to one by (6). Taylor expansion, (5) and the critical density bounds give

$$n\int \left|k - \delta_n u^{\mathsf{T}}\ell_{\eta_0}\mathbf{1}_A\right|^2 f_0 \;\longrightarrow\; 0.$$

For the location Taylor remainder use a bound $Cn\delta_n^4\Delta^{-2}\log(1/\Delta) = o(1)$ from integrating $f_0/|x|^4$ on the annulus. For the score replacement, (6) bounds the difference by $Cn^{-1/2}(1+\log(1/|x|))/|x|$. The corresponding integral times $n\delta_n^2$ is $O(\delta_n^2(1+\log(1/\Delta))^4) = o(1)$. Also $\mathbb{E}_h k = O(\delta_n)$. The difference between $\mathbb{E}_h k$ and $\mathbb{E}_0 k$, multiplied by n, is $O(\sqrt{n}\,\delta_n) = o(1)$ by (6), the bound $|k(x)| \le C\delta_n/|x|$ on A and the integrability following (7). It follows by expanding the log of the relative likelihood factor that its log contribution under f_0 is

$$\delta_n\sum_{i=1}^n u^{\mathsf{T}}[\ell_{\eta_0}(X_i)\mathbf{1}_A(X_i) - \mathbb{E}_0(\ell_{\eta_0}\mathbf{1}_A)] - \tfrac{1}{2}u^{\mathsf{T}}\mathcal{K}(S_0)u + o_{\Pr}(1).$$

For the quadratic term, the sum of the squares of k concentrates because of the vanishing envelope and bounded total squared mean. The normalization quadratic is $O(n\delta_n^2) = o(1)$. The information limit in (7) applies, since the truncated score mean is uniformly bounded.

Meanwhile the nuisance likelihood for f_h relative to f_0 has the regular $\sqrt{n}$ expansion by quadratic mean differentiability [17]. The two central sequences have asymptotically zero cross covariance:

$$\left\|\sqrt{n}\,\delta_n\mathbb{E}_0[\ell_{\eta_0}\mathbf{1}_A s_{\eta_0}^{\mathsf{T}}]\right\| = O(\Lambda_n^{-1/2}),$$

using (6) and the stated radial integrability. The location rows have envelope $O(\delta_n/\Delta)$, and the nuisance rows meet Lindeberg by their finite second moment. Joint triangular normal convergence thus gives independent Gaussian shifts. The resulting likelihood coordinates have mean one. Truncation and quantization under the reference, followed by regeneration within the quantization cells, give finite Le Cam convergence.

To extend to compact sets, use finite nets and pairwise Hellinger continuity. At fixed location the nuisance product squared Hellinger distance is $O(|h-h'|^2)$, by smooth complete-data quadratic mean bounds uniformly on the brackets. At a fixed nuisance $\eta_0 + h/\sqrt{n}$, the annulus replacement just proved approximates the local locations uniformly for bounded u, h, with negligible error as $n \to \infty$. For uniformity on a bounded set of u, split at $C\delta_n$ with $C > 2\sup|u|$. Within this ball, the bounded mark integrals give $O(\delta_n^2)$ in the power case and $O(\delta_n^2/\log(1/\delta_n))$ in the plane,

uniformly for the allowed shifts. Between $C\delta_n$ and Δ use the same derivative integrals as before with $D^2 \le C^2\delta_n^2$. This gives the uniform $o(1)$ product bound. On the annulus the squared Hellinger cost for two splices at the same nuisance is at most $C|u-u'|^2/n$, since spatial score integration along the segment is bounded by $C\Lambda_n$, using density comparability under the bounded shifts and the critical information integral. Normalization obeys the same bound. The product limits are continuous. The triangle inequality through a nearest net point and contraction under Markov kernels prove compact-set convergence. This proves Theorem 4.1.

4.5 Estimating location with unknown nuisance

Let $\widetilde{\theta}, \widehat{\gamma}, \widehat{S}$ come from the moment pilot. In the score step set $\widehat{\eta} = (1, \widehat{\gamma}, \widehat{S})$; the unknown true shape may differ from one by $H/\sqrt{n}$. Thus the pilot nuisance error remains of order $O(n_0^{-1})$ in squared mean and its fixed-Δ failure probability is still superpolynomially small. Write $\ell_\Delta(x) = \ell_{\widehat{\eta}}(x)\mathbf{1}_A(x)$ and $m_\Delta = \int \ell_\Delta f_{\widehat{\eta}}$. Using the remaining $n_1 = n - n_0$ observations, set

$$\widehat{\theta} = \widetilde{\theta} + J_\Delta(\widehat{\eta})^{-1}\left\{\frac{1}{n_1}\sum_{i>n_0}\ell_\Delta(X_i - \widetilde{\theta}) - m_\Delta\right\}.$$

In the skew model the subtraction of m_Δ is needed. It is uniformly bounded by radial integrability. For $e = \theta - \widetilde{\theta}$, on the good pilot event $|e| \le \Delta/4$ and $|\eta - \widehat{\eta}| \le \Delta/4$, the conditional centered score mean equals

$$J_\Delta(\widehat{\eta})e + O\left(|e|^2\frac{\log(1/\Delta)}{\Delta} + |\eta - \widehat{\eta}|\right).$$

For the location expansion at the fitted nuisance, use the Hessian bound in (5), comparability of densities along the shift, and integrate $f_{\widehat{\eta}}/|x|^3$ on the annulus. This gives the first error bound. For the nuisance discrepancy use (6) on the same shifted annulus and integrate $f_{\widehat{\eta}}(1 + \log(1/|x|))/|x|$. That integral is uniformly bounded, giving the second error. By the pilot moment bounds and (7), the normalized squared bias tends to zero:

$$C\frac{n}{\Lambda_n}\left(n_0^{-2}\frac{\log^2(1/\Delta)}{\Delta^2} + n_0^{-1}\right) = o(1),$$

since $a_n = o(\Lambda_n)$.

On the score support the conditional true density divided by $f_{\widehat{\eta}}$ differs from one by at most a constant times $|e|/\Delta + |\eta - \widehat{\eta}|\log(1/\Delta)$. This tends to zero in probability and stays bounded on the good event. The score means have size $O(1) + o_{\Pr}(1)$ since $J_\Delta e = o_{\Pr}(1)$, so the covariance on the Λ_n scale tends to $\mathcal{K}(S)$. The conditional score rows have normalized envelope $O((\Delta\sqrt{n\Lambda_n})^{-1}) = o(1)$. The conditional triangular central limit theorem gives the stated normal limit, uniformly along convergent scatter sequences. The conditional score second moments are bounded by a fixed multiple of J_Δ on the good event, so the normalized variance converges also in expectation. On the bad event the score is still deterministically bounded by C/Δ, and the pilot lies in its compact brackets. Superpolynomially small probability makes this event's squared risk contribution negligible after normalization. All bounds are uniform. Finally fixing the nuisance gives the Gaussian local minimax lower constant by Theorem 4.1. Project decisions to fixed local balls to make squared loss bounded, then take diffuse normal priors truncated to expanding balls in the limit shift. The Gaussian Bayes risk tends to $\operatorname{tr}(\mathcal{K}(S_0)^{-1})$. The estimator therefore attains the sharp local constant also when the nuisance is unknown.

Statements and Declarations

Funding. This research was self-funded.

Competing interests. The author owns Luminary Oddities and has no other relevant competing interests.

AI assistance. OpenAI Codex assisted with derivations, proof development, source checks and independent internal review of each part. Anthropic Claude edited, consolidated and cross-referenced the parts and checked computations. Scott Schubr directed the research and is responsible for the content.

References

[1] A. Dachian, On limiting likelihood ratio processes of some change-point type statistical models, arXiv:0907.0440. https://arxiv.org/abs/0907.0440.

[2] K. Podgórski and J. Wallin, Maximizing leave-one-out likelihood for the location parameter of unbounded densities, *Ann. Inst. Stat. Math.* (2015). https://doi.org/10.1007/s10463-013-0437-6.

[3] T. Nitithumbundit and J. S. K. Chan, Maximum leave-one-out likelihood method for the location parameter of variance gamma distribution with unbounded density, *J. Stat. Comput. Simul.* (2023). https://doi.org/10.1080/00949655.2023.2202399.

[4] T. Nitithumbundit and J. S. K. Chan, Maximum leave-one-out likelihood estimation for location parameter of unbounded densities, arXiv:1602.01213 (2016). https://arxiv.org/abs/1602.01213.

[5] T. Nitithumbundit and J. S. K. Chan, Maximum leave-one-out likelihood method for the location parameter of variance gamma distribution with unbounded density, J. Stat. Comput. Simul. 93 (2023), 2642–2671. DOI.

[6] K. Podgórski and J. Wallin, Maximizing leave-one-out likelihood for the location parameter of unbounded densities, Ann. Inst. Stat. Math. 67 (2015), 19–38. DOI.

[7] S. Dachian, On limiting likelihood ratio processes of some change-point type statistical models, J. Statist. Plann. Inference 140 (2010), 2682–2692. Author manuscript.

[8] NIST Digital Library of Mathematical Functions, Chapter 10, Section 10.31, Section 10.32, and Section 10.40.

[9] S. Schubr, *Calibrating small noise from exteriors, frontiers, cusps and density singularities*, the part titled *Calibrating a noise endpoint from a logarithmic density singularity*, Luminary Research Archive, 2026.

[10] I. A. Ibragimov and R. Z. Khas'minskii, Asymptotic behavior of statistical estimates of the shift parameter for samples with unbounded density, J. Soviet Math. 16 (1981), 1035–1041. Original paper.

[11] S. Dachian, Estimation of the location of a 0-type or ∞-type singularity by Poisson observations, Statistics 45 (2011), 509–523. Author manuscript.

[12] S. Gupta, J. C. H. Lee and E. Price, High-dimensional location estimation via norm concentration for subgamma vectors, Proc. Mach. Learn. Res. 202 (2023), 12132–12164. Proceedings.

[13] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge University Press (1998), Theorem 7.2.

[14] T. Nitithumbundit and J. S. K. Chan, An ECM algorithm for skewed multivariate variance gamma distribution in normal mean-variance representation. Author manuscript.

[15] E. Seneta, Fitting the variance-gamma model to financial data, J. Appl. Probab. 41A (2004), 177–187. DOI.

[16] A. Fischer, R. E. Gaunt and A. Sarantsev, Modified method of moments for generalized Laplace distributions, Commun. Stat. Simul. Comput. 54 (2025), 2047–2064. Accepted author manuscript.

[17] A. W. van der Vaart, *Asymptotic Statistics*, Cambridge University Press (1998), Chapter 7.