Gaussian Halfspace Testing, Distributional Robustness, and the Lognormal Pullback

Jonathan R. Landers

1 Context

A recent preprint of Chen, De, Huang, Nadimpalli, Servedio, and Yang studies relative-error testing of halfspaces in Gaussian space [1]. The underlying object is an unknown Boolean function $f\colon \mathbb{R}^n \to \{0,1\}$, and the target class is the family of halfspaces

$$\mathcal{H} := \left\{ h_{w,\theta}(x) = \mathbf{1}[w^\top x \geq \theta] : w \in \mathbb{R}^n,\ \theta \in \mathbb{R} \right\}.$$

Here $\mathbf{1}[E]$ denotes the indicator of an event E, so $h_{w,\theta}$ is just a linear threshold classifier. The ambient input law is the standard Gaussian distribution $G := N(0, I_n)$, where I_n is the $n \times n$ identity matrix. Their distance notion is the relative-error quantity

$$d_G^{\mathrm{rel}}(f,h) := \frac{G(f \triangle h)}{G(f=1)},$$

where $f \triangle h$ is the disagreement set and $G(A)$ denotes the Gaussian probability of an event A.

The qualitative message is elegant: Gaussian geometry makes halfspaces unusually testable. In contrast with the Boolean-cube setting, where a logarithmic lower bound is known, the Gaussian setting admits a tester whose query complexity is sublinear in the dimension [1]. That is the starting point of the present note.

2 A robustness question

Suppose a Gaussian tester $\mathcal{A}_G$ uses $q_G(n)$ queries in dimension n. We ask: *what survives if the ambient law is not exactly Gaussian, but is close to Gaussian after an appropriate change of variables?*

Let D be another distribution on the input space, and let T be a measurable transform from the input space into $\mathbb{R}^n$ that places D into Gaussian coordinates. Write $\widetilde{D} := T_\# D$ for the pushforward of D by T, meaning that $\widetilde{D}(B) = D(T^{-1}(B))$ for measurable sets $B \subseteq \mathbb{R}^n$. Define the total variation distance

$$\Delta := \|\widetilde{D} - G\|_{\mathrm{TV}} := \sup_A |\widetilde{D}(A) - G(A)|,$$

where the supremum ranges over measurable events A. The point of Δ is that it measures how different the two sampling laws can look to any event, and therefore to any fixed randomized tester. If we run the Gaussian tester in T-coordinates, each query sees samples from $\widetilde{D}$ instead of G, so the natural first-order penalty is proportional to $q_G(n)\Delta$.

3 Main transfer statement

Theorem 1 (Distributional robustness of a Gaussian tester). *Let $\mathcal{A}_G$ be a tester for the halfspace class $\mathcal{H}$ under the Gaussian law $G = N(0, I_n)$. Assume that in dimension n the tester uses $q_G(n)$ oracle queries and has completeness c and soundness s, with gap $\gamma := c - s > 0$. Let D be another input distribution, let T be a measurable transform, and set $\widetilde{D} := T_\# D$ and $\Delta := \|\widetilde{D} - G\|_{\mathrm{TV}}$. Run the same tester in T-coordinates. Then its acceptance gap under D is at least*

$$\gamma_D \;\geq\; \gamma - 2\, q_G(n)\, \Delta.$$

In particular, whenever $q_G(n)\, \Delta < \gamma/2$, the Gaussian tester remains valid under D with a positive constant gap.

Proof. Fix the dimension n and abbreviate $q := q_G(n)$. The tester induces a random transcript consisting of its internal randomness, its queried points, and the corresponding oracle answers. When the tester is run in T-coordinates, the only change in law is that the queried samples are drawn from $\widetilde{D}$ rather than G.

By the definition of total variation distance, there exists a coupling of one sample from $\widetilde{D}$ and one sample from G that disagrees with probability at most Δ. Coupling the q queried samples independently and applying the union bound, the entire q-sample portion of the transcript disagrees with probability at most $q\Delta$. Hence the total variation distance between the transcript law under $\widetilde{D}$ and the transcript law under G is at most $q\Delta$.

Acceptance is itself an event in transcript space, so the difference in acceptance probabilities is at most $q\Delta$. Therefore the completeness can drop by at most $q\Delta$ and the soundness can rise by at most $q\Delta$, yielding

$$\gamma_D \;\geq\; (c - q\Delta) - (s + q\Delta) \;=\; \gamma - 2q\Delta.$$

The criterion $q\Delta < \gamma/2$ follows immediately. □

The theorem is modest on purpose. It does not say that the optimal query complexity under D is a continuous function of Δ. It says something more immediate and operational: a tester already known to work in Gaussian space continues to work under nearby laws, provided $q_G(n)\Delta$ is small enough. In compact form:

$$\text{Gaussian theorem gives } q_G(n) \implies \Delta \text{ controls gap deterioration}$$
$$\implies q_G(n)\Delta \ll 1 \text{ implies that } D \text{ inherits the Gaussian result.}$$

Exact equality of asymptotic complexity occurs in the special case $\Delta = 0$.

4 Exact lognormal corollary

The lognormal case is the cleanest possible illustration because the mismatch parameter vanishes exactly.

Corollary 2 (Exact lognormal pullback). *Let $Y \sim \mathrm{LogNormal}(\mu, \Sigma)$, where $\mu \in \mathbb{R}^n$ and Σ is positive definite, and define $T(y) := \Sigma^{-1/2}(\log y - \mu)$ with the logarithm taken coordinatewise. Then $T(Y) \sim N(0, I_n)$, so $\|T_\# D - G\|_{\mathrm{TV}} = 0$. Consequently, the Gaussian halfspace-testing result transfers exactly to the class of log-halfspaces*

$$\mathcal{H}_{\log} := \big\{ y \mapsto \mathbf{1}[a^\top \log y \geq b] : a \in \mathbb{R}^n,\; b \in \mathbb{R} \big\}.$$

Proof. If Y is lognormal with parameters (μ, Σ), then by definition the random vector $\log Y$ is Gaussian with law $N(\mu, \Sigma)$. Whitening by $\Sigma^{-1/2}$ centers and rescales this law to $N(0, I_n)$. Thus the pushed-forward law is exactly G, so $\Delta = 0$. A log-halfspace in the original y-coordinates becomes an ordinary linear halfspace in the transformed coordinates. $\square$

This is the pleasant endpoint of the flow. The Gaussian theorem is not merely approximated in the lognormal setting; after the right coordinate change it is literally the same geometric problem.

5 Multivariate-t perturbation bound

The lognormal corollary, while satisfying, sits at $\Delta = 0$ and therefore does not exercise the robustness bound. To see the bound at work, we need a distribution family where Δ is nonzero and explicitly computable.

Corollary 3 (Multivariate-t perturbation bound)**.** *Let $Y \sim t_\nu(0, I_n)$ be the standard multivariate t-distribution with $\nu > n$ degrees of freedom. Then*

$$\Delta(n, \nu) \;:=\; \|t_\nu(0, I_n) - N(0, I_n)\|_{\mathrm{TV}} \;=\; C_n \cdot \frac{n}{\nu} + O\!\left(\frac{n^2}{\nu^2}\right),$$

where $C_n \to 1/4$ as $\nu/n \to \infty$. Consequently:

(i) ***Gap robustness.*** *The Gaussian halfspace tester retains a positive acceptance gap under t_ν provided $q_G(n) \cdot n/\nu$ is small, i.e. $\nu \gg n \cdot q_G(n)$.*

(ii) ***Complexity overhead.*** *Under the same condition with room to spare, Theorem 5 gives*

$$q_{t_\nu}(n) \;\le\; q_G(n) + O\!\left(\frac{q_G(n)^2 \cdot n}{\nu}\right).$$

In particular, if $q_G(n) = n^\alpha$ for $\alpha < 1$, stability holds whenever $\nu \gg n^{2\alpha+1}$.

Proof. Both $t_\nu(0, I_n)$ and $N(0, I_n)$ are spherically symmetric, so the total variation distance reduces to an integral over the radial densities. The Gaussian radial density is the chi density with n degrees of freedom, and the t_ν radial density is proportional to $r^{n-1}(1 + r^2/\nu)^{-(n+\nu)/2}$.

Write R_G for a chi-distributed radius and R_T for the t_ν radius. For large ν, the t_ν density converges to the Gaussian density, and a Taylor expansion of the log-density ratio in powers of $1/\nu$ yields a leading-order mismatch proportional to n/ν. The constant C_n is determined by the integrated absolute deviation of the radial densities; numerical evaluation across a range of dimensions gives $C_n \approx 0.24$–0.25 for $\nu \gg n$, converging to $1/4$.

The two consequences follow by substituting $\Delta \approx n/(4\nu)$ into Theorem 1 and Theorem 5, respectively. $\square$

Remark 4. *The multivariate-t corollary complements the lognormal one in a precise sense. Corollary 2 shows that the framework can be exact: when the coordinate change eliminates all mismatch, the Gaussian result transfers for free. Corollary 3 shows that the framework degrades gracefully: when the coordinate change leaves residual mismatch, the cost is controlled and explicit. The trade-off surface has a natural interpretation—the heavier the tail, the more queries are spent fighting the tail rather than testing the halfspace, and the rate at which this happens is governed by the ratio n/ν.*

6 A stronger next theorem: statement and proof sketch

The natural next step is stronger than the transfer theorem above. One would like to tie the distance in *query complexity itself* to the distance between distributions.

Theorem 5 (Sketch: complexity stability under distributional perturbation)**.** *Assume that, for each dimension n, there exists a Gaussian halfspace tester with optimal query complexity $q_G(n)$ and constant gap γ, and assume moreover that near-optimal testers can be boosted by independent repetition without changing the class being tested. Then for distributions D satisfying $\Delta := \|T_{\#}D - G\|_{\mathrm{TV}} \ll 1/q_G(n)$, one expects a quantitative bound of the form*

$$q_D(n) \;\le\; q_G(n) + C\, q_G(n)^2\, \Delta,$$

for some absolute constant C, where $q_D(n)$ denotes the query complexity under D in the same transformed coordinates.

Proof sketch. Run the Gaussian-optimal tester under D. By the robustness theorem, its acceptance gap decreases from γ to roughly $\gamma - 2q_G(n)\Delta$. If $q_G(n)\Delta$ is small, the gap remains positive but weaker than before. Standard amplification by repetition restores a fixed target gap by increasing the number of repetitions by a factor of the form $1 + O(q_G(n)\Delta)$. Multiplying this factor by the original query cost $q_G(n)$ gives a total complexity of the form

$$q_G(n)\bigl(1 + O(q_G(n)\Delta)\bigr) \;=\; q_G(n) + O(q_G(n)^2\, \Delta).$$

This argument is only a sketch because it suppresses constants, assumes a uniform boosting mechanism, and does not prove that the amplified tester is itself optimal under D. But it captures the right next principle: once the tester is stable, one may hope that the *complexity law* is stable as well. □

The point, then, is not that we already know $|q_D(n) - q_G(n)| \lesssim \Delta$. We do not. The present note proves only the first rung: a known Gaussian tester survives under nearby distributions. The sketch theorem identifies the next rung: turn that survival statement into a quantitative perturbation law for query complexity itself.

References

[1] X. Chen, A. De, Y. Huang, S. Nadimpalli, R. A. Servedio, and T. Yang, "Sublinear-query relative-error testing of halfspaces," 2026. Available: `https://arxiv.org/abs/2604.01557`.