

Article

Not peer-reviewed version

Probability is Earned: Information Capacity and the Epistemic Geometry of Inference

[Moriba Kemessia Jah](#) *

Posted Date: 31 March 2026

doi: 10.20944/preprints202603.2319.v1

Keywords: epistemic geometry; information capacity; possibility theory; Possibilistic Cramér–Rao Bound; Jaynesian maximum entropy; Popperian falsification; admissible inference; epistemic validity; orbit determination; observation geometry; verbal probability phrases; epistemic scalarization error; non-Bayesian estimation; uncertainty quantification



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Probability is Earned: Information Capacity and the Epistemic Geometry of Inference

Moriba Kemessia Jah

Black Swan Research Group, GaiaVerse, Ltd., Jah Decision Intelligence Group, Aerospace Engineering & Engineering Mechanics, University of Texas at Austin; moriba@utexas.edu

Abstract

Probability is often treated as the default representation of uncertainty in statistical inference and machine learning. This paper asks a more fundamental question: under what conditions is a probability distribution a valid representation of uncertainty, and what is the information cost of assuming one when those conditions are not met? We show that inference is governed by two joint constraints: maximizing information capacity by preserving the geometric degrees of freedom through which contrast can register, and minimizing false information by asserting nothing the evidence has not forced. These constraints, expressed through Jaynes's principle of maximum entropy and Popper's criterion of falsification, determine the structure of inference without remainder. Bayesian inference emerges not as a competing framework, but as the limiting geometry obtained when epistemic width has contracted sufficiently to justify probabilistic closure. In this sense, probability is not assumed — it is earned. We trace the origin of these ideas through two decades of operational experience in spacecraft navigation, space situational awareness, and orbit determination, where standard probabilistic filters performed well in nominal regimes but failed systematically when uncertainty was driven by genuine ignorance rather than statistical variability. Across problems including debris tracking, attitude estimation, and multi-target inference, the consistent failure mode was premature probabilistic commitment in regimes where observation geometry could not support distinguishability. The central result is that information exists only in the presence of contrast, and that structure destroyed without evidence justification is information permanently lost. We formalize this principle through an epistemic geometry of inference and show that probabilistic representations are valid only when distinguishability, parameterization, and likelihood structure are all earned by the data. When these conditions fail, probabilistic closure incurs a measurable and avoidable information capacity cost.

Keywords: epistemic geometry; information capacity; possibility theory; Possibilistic Cramér–Rao Bound; Jaynesian maximum entropy; Popperian falsification; admissible inference; epistemic validity; orbit determination; observation geometry; verbal probability phrases; epistemic scalarization error; non-Bayesian estimation; uncertainty quantification

1. Introduction

I did not set out to build a non-Bayesian inference framework. I set out to track spacecraft and debris honestly.

What follows is an account of two decades of inference problems that progressively exposed the limits of probabilistic closure — not as a mathematical critique but as an operational reality — and of the framework that eventually emerged from refusing to paper over those limits with convenient assumptions.

1.1. JPL and the Geometry of Information

When I was navigating spacecraft at NASA's Jet Propulsion Laboratory, the question I was asking was not philosophical — it was operational: given the observations I have, from the geometry I have,

what can I *honestly* say about where this spacecraft is? Over the course of my time at JPL, I supported navigation on Mars Global Surveyor, Mars Odyssey, Mars Express (with the European Space Agency), Hayabusa (with the Japan Aerospace Exploration Agency), the Mars Exploration Rovers, and — as interplanetary phase orbit determination lead — the Mars Reconnaissance Orbiter.

The preferred filter at JPL is the Square Root Information Filter (SRIF), developed by Biermann [1], valued for its numerical stability and its clean handling of the information form of the estimation problem. At the US Air Force Research Laboratory, the standard was batch least squares, often called differential correction: a different tradition, same underlying commitment to Gaussian, linearized propagation of uncertainty. I have always been a hybrid practitioner, favoring Kalman-type filters and adapting to the problem at hand, and I was the first to apply Sigma-Point (Unscented Kalman) filters to orbit determination, formalizing that contribution in my PhD dissertation [5].

What all of these filters share — SRIF, differential correction, extended Kalman, unscented Kalman — is the assumption that uncertainty can be represented as a Gaussian, or a sum of Gaussians, or a linearization around a Gaussian. In the nominal regime, with a well-characterized dynamical model and a sensor geometry that is sufficiently rich, this assumption holds and these filters perform extremely well. The problems began when the regime was not nominal.

One of the privileges of supporting a mission through launch is the experience of assembling and interpreting a *launch package*: a document prepared by the navigation team that characterizes the predicted observational environment for the outbound trajectory. Among its many elements are plots of predicted range and range-rate as the spacecraft recedes from Earth. These plots are prescriptive: they show the Deep Space Network (DSN) schedulers exactly *where and when* the most information-rich observations should be acquired to achieve the most accurate and precise determination of the outbound trajectory.

The geometry of information is not uniform in time. As a spacecraft departs, there are windows during which its range-rate is changing most rapidly — windows during which the line-of-sight geometry is evolving fastest relative to the ground station. Observations acquired during these windows are maximally distinguishable from one another: each one constrains the trajectory in a direction that previous observations have not. Outside these windows, the geometry flattens. Observations pile up that are nearly identical, contributing nearly nothing beyond what the previous observation already established. Saturating the DSN schedule with observations during a geometrically flat window is not just wasteful — it is epistemically counterproductive, consuming finite antenna time that could have been allocated where the information actually lives.

This is the operational content of the Fisher information matrix $F(\theta) = G(\theta)^T R^{-1} G(\theta)$, where $G(\theta) = \partial g / \partial \theta$ is the measurement Jacobian. Its eigenvalues peak exactly where the trajectory is changing most rapidly relative to the observer, and collapse where it is nearly static. The experienced navigator learns to read information geometry from predicted observable plots, and to internalize the lesson that resists formalization but demands respect: *more data is not more information — more informative geometry is more information*.

1.2. AFRL and the Onset of Genuine Ignorance

When I moved from spacecraft navigation to Space Situational Awareness at the US Air Force Research Laboratory, the problem was harder in every dimension. Instead of a handful of carefully maintained spacecraft with known physical properties and dedicated tracking assets, I was working with a catalog of thousands of resident space objects — debris, defunct payloads, active satellites — tracked by sensors with heterogeneous geometries, sparse cadences, and no guarantee that the physical forces acting on each object were well-characterized.

This last point proved decisive. In classical orbit determination, gravitational acceleration is modeled to high fidelity. For debris objects, the non-gravitational accelerations — solar radiation pressure, atmospheric drag, thermal re-emission, outgassing — depend on physical properties that are mostly unknown: mass, area, shape, surface reflectivity. I scheduled telescope observations using the Pan-STARRS system based on predicted Fisher information, and the result reinforced what spacecraft

navigation had already taught me. Objects tracked at high-information epochs were where prediction said they would be. Objects tracked during low-information epochs drifted outside the predicted uncertainty. The filter asserted confidence the geometry could not support.

But debris posed a further challenge that JPL had not: the uncertainty was not parametric. It was not that we had the wrong value for a parameter we knew existed. We did not know what we did not know about the forces acting on the object. This is a categorically different epistemic situation, and it is one that probabilistic filters are structurally ill-equipped to represent.

1.3. *A Two-Decade Research Programme and Its Limits*

The work that followed at AFRL and subsequently at the University of Texas at Austin was a sustained attempt to find inference methods adequate to the problem of tracking objects whose physical properties were unknown and whose dynamical models were therefore structurally incomplete. Each approach illuminated the problem more clearly while also revealing its own limits.

Attitude estimation from light curves.

The first challenge was inferring the orientation of debris objects from photometric data — light curves. A tumbling debris object reflects sunlight in patterns that encode its shape and orientation, but extracting that information from noisy, sparse photometric time series is severely ill-posed. The UKF and its variants failed unless the initial guess was already close to the truth. The problem was not filter design — the problem was that the source of uncertainty was genuine ignorance about shape and orientation, not statistical noise around a known model [6].

High area-to-mass ratio objects.

Prof. Thomas Schildknecht at the University of Bern discovered a new class of debris objects exhibiting anomalous orbital dynamics: apparent high area-to-mass ratio (HAMR) objects, whose trajectories deviated substantially from purely ballistic arcs due to enhanced sensitivity to solar radiation pressure. Tracking HAMR objects required confronting the fact that without knowledge of the physical properties governing non-gravitational accelerations, trajectory prediction degrades rapidly. Standard orbit determination methods, applied without a priori information, struggled to recover and maintain tracks [7]. Multiple hypothesis filters provided some improvement [8], but they still relied on the assumption that the uncertainty could be covered by a finite set of Gaussian hypotheses — an assumption that the data repeatedly challenged.

Adaptive Gaussian mixture methods.

Working with then-graduate student Kyle DeMars (now a professor), and building on the Gaussian mixture approach of Prof. Puneet Singla, I mentored the co-development of an adaptive Gaussian mixture model that adjusted the number of mixands online by monitoring differential entropy. We called it AEGIS. It worked remarkably well for initial orbit determination in non-Gaussian regimes [9,10], and it provided a rigorous entropy-based framework for uncertainty propagation in nonlinear systems [11]. But AEGIS was still built on probabilistic closure. It was Gaussian mixtures all the way down, and when the true uncertainty was driven by ignorance rather than statistical dispersion, the mixture model had no principled way to represent what it did not know.

Coupled orbit-attitude dynamics.

Any uncontrolled object in space is tumbling. Working with then-postdoc Carolin Früh (now a professor), one of Prof. Schildknecht's students at the University of Bern, we developed coupled orbit-attitude propagation models for HAMR objects [12–14]. The models were physically more complete, but they also exposed the inference problem more starkly. Six degrees of freedom coupled with unknown physical properties and non-gravitational forces that depended on those properties in nonlinear ways: the posterior distribution over this space was nothing like Gaussian, and Gaussian approximations purchased apparent precision at the cost of real accuracy. Attempts to infer physical properties from

combined astrometric and photometric data mostly failed unless the initial guess was already close [15–17].

Finite set statistics.

The multi-target tracking literature offered a rigorous probabilistic framework for simultaneously tracking an unknown number of objects: Finite Set Statistics (FISST) [4]. Applying FISST to the space catalog problem [18] revealed that while the framework was theoretically elegant, it was computationally near-intractable at catalog scales. The curse of dimensionality, combined with the non-Gaussian uncertainty in object states, made FISST-based approaches difficult to scale.

Joint probabilistic data association.

Joint Probabilistic Data Association (JPDA) proved more tractable and was applied successfully to multi-object space tracking [19]. It worked, but it remained a probabilistic framework applied to a problem where the dominant source of uncertainty — the unknown physical properties of debris objects — was not statistical.

Outer probability measures and possibility functions.

The turning point came when I was introduced to outer probability measures and possibility functions by postdoc Emmanuel Delande (now at CNES) and collaborator Prof. Jérémie Houssineau. For the first time, a framework was available that could represent genuine ignorance — epistemic uncertainty — without forcing it into probabilistic form. Applied to multi-target tracking [20] and initial orbit determination [21,22], the results were encouraging. But the framework still relied on probabilistic machinery for many operations, imposed limitations on the class of possibility distributions that could be used, and could upper-bound uncertainty without lower-bounding it. The approach was a step toward the right territory, but it was not yet the destination.

1.4. Starting from Scratch

After two decades of this — spacecraft navigation, debris tracking, attitude estimation, Kalman filters, sigma-point filters, Gaussian mixtures, FISST, JPDA, outer probability measures — I arrived at a conviction I could no longer defer.

In order to know something, one must measure it.

In order to understand something, one must predict it.

Every filter I had worked with, and every framework I had tried, shared a structural commitment: uncertainty must be represented probabilistically, and the goal of inference is to concentrate that representation as efficiently as possible. When the uncertainty was statistical — when it arose from known noise processes around a known model — this commitment was appropriate and the filters performed well. When the uncertainty was epistemic — when it arose from genuine ignorance about physical properties, dynamical models, or the very identity of the objects being tracked — the commitment imposed false precision. The filters did not fail because they were poorly designed. They failed because they were asked to represent something probabilistic tools cannot honestly represent.

TEAG is the result of starting from scratch and asking: what kind of inference framework satisfies two requirements that two decades of operational experience had made non-negotiable? The first: *maximize information capacity* by preserving the geometric degrees of freedom through which contrast can register. No structure should be collapsed without evidence-justified cause. The second: *minimize false information* by asserting nothing the evidence has not forced. The mathematics that satisfies both requirements simultaneously turned out to be possibilistic rather than probabilistic. That was not the design goal. It was the consequence.

TEAG is not non-Bayesian by intention. It is non-Bayesian by necessity. And as we show in Section 6, Bayesian inference is recovered exactly when the evidence has earned it.

1.5. Paper Organization

Section 2 develops the principle that information exists only in the presence of contrast. Section 3 makes the Jaynesian connection precise. Section 4 makes the Popperian connection precise and introduces the Epistemic Validity Condition — the formal criterion under which a probability distribution is a legitimate representation of uncertainty rather than a projection artifact. Section 5 shows how TEAG holds both principles simultaneously through the PCRB. Section 6 establishes Bayesian inference as the earned limit of TEAG. Section 7 draws implications for machine learning, including a detailed treatment of the CAPphrase verbal probability phrase data as a domain-independent empirical demonstration of the EVC's failure modes. Section 8 concludes.

2. Information Exists in the Presence of Contrast

2.1. Fisher Information and Observation Geometry

Let $\theta \in \mathbb{R}^n$ denote the state of a dynamical system and $y \in \mathbb{R}^m$ an observation through measurement function $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with additive noise covariance R . The Fisher information matrix is

$$\mathbf{F}(\theta) = G(\theta)^\top R^{-1} G(\theta), \quad (1)$$

where $G(\theta) = \partial g / \partial \theta$ is the measurement Jacobian. The eigenvalues of $\mathbf{F}(\theta)$ measure the information available about each principal direction of the state from a single observation at geometry θ .

The key structural fact is that $\mathbf{F}(\theta)$ depends on the observation geometry through $G(\theta)$, not on the number of observations. Repeating an observation at fixed geometry accumulates $k \cdot \mathbf{F}(\theta)$: information scales linearly with count, but the *directions* of maximum information remain fixed. To gain information in a previously unconstrained direction, the geometry must change.

Remark 1 (More data is not more information). *The Fisher information matrix makes precise the distinction that operational practice forces on every working navigator: the information content of an observation is a property of the observation geometry, not the observation count. Scheduling observations where the geometry is richest extracts the maximum information from a finite observation budget. Scheduling observations where the geometry is flat extracts nearly nothing at nontrivial cost.*

2.2. Contrast as the Primitive of Information

The Fisher information matrix captures observation geometry quantitatively, but within a probabilistic framework requiring a known likelihood model and additive Gaussian noise. When these assumptions fail — when the dynamical model is structurally incomplete, when noise is bounded rather than Gaussian, when the hypothesis space does not carry a natural probability measure — Fisher information can be computed formally while misrepresenting the actual information available.

The more primitive concept, underlying Fisher information but not reducible to it, is *contrast*: the degree to which an observation distinguishes one hypothesis from another. An observation is informative to the degree that its outcome differs across the hypothesis space. If every surviving admissible hypothesis predicts approximately the same outcome, the observation is nearly uninformative regardless of what the Fisher information matrix says. This was the lesson from debris tracking: dynamic model uncertainty corrupted the measurement Jacobian, and the formal Fisher information overstated the actual contrast between hypotheses.

Definition 1 (Epistemic contrast). *Let $\pi : \mathcal{H} \rightarrow [0, 1]$ be a normalized possibility distribution and $y_k \in \mathcal{Y}$ an observation. The epistemic contrast generated by y_k is*

$$C_k = \sup_{h \in \mathcal{H}} \min \left(\frac{1}{2} q_k^{(h)}, \pi(h) \right), \quad (2)$$

where $q_k^{(h)} = \|L_e^{-1}(y_k - g(h))\|^2$ is the whitened squared innovation of hypothesis h , and $\Pi_e = L_e L_e^\top$ is the MVEE of the predicted measurement support.

Remark 2 (Epistemic contrast and Choquet information). *The quantity C_k is precisely the aggregate epistemic surprisal $\bar{S}_k$ of TEAG [28] — the Choquet integral of per-hypothesis surprisal with respect to the prior possibility capacity. It is large when the most credible hypotheses are most surprised by the observation, and small when the observation surprises only low-possibility hypotheses. This is the possibilistic analog of Fisher information: it measures contrast among the hypotheses that matter, weighted by their current epistemic standing.*

2.3. The Information Cost of Premature Collapse

The research programme described in Section 1 repeatedly encountered a specific failure mode: a filter would commit to a concentrated posterior before the geometry had forced that commitment, and subsequent high-information observations could not undo the damage. The information was available. The capacity to receive it had been destroyed. This is the operational content of what TEAG formalizes as the cost of premature epistemic collapse.

Proposition 1 (Information capacity and geometric degrees of freedom). *Let $\mathcal{E}_\pi = \int_0^1 \log V_\alpha d\alpha$ denote the possibilistic entropy [28], where $V_\alpha = c_n(\det \Pi_\alpha)^{1/2}$ is the MVEE volume of the α -cut. The maximum admissible entropy reduction per observation is*

$$\Delta \mathcal{E}_{\pi_k}^{\max} = -\frac{n}{2} \log(1 - I_k), \quad (3)$$

where $I_k = 1 - e^{-\bar{S}_k} \in [0, 1)$ is the possibilistic information content [26]. An observation that generates no contrast ($I_k \rightarrow 0$) cannot reduce possibilistic entropy, regardless of how many such observations are accumulated.

Proof. This follows directly from the Possibilistic Cramér–Rao Bound [26]:

$$\mathcal{E}_{\pi_k|k} \geq \mathcal{E}_{\pi_k|k-1} + \frac{n}{2} \log(1 - I_k). \quad (4)$$

The right-hand side vanishes as $I_k \rightarrow 0$, establishing that no entropy reduction is admissible when the observation generates no epistemic contrast. $\square$

3. The Jaynesian Imperative: Maximizing Honest Ignorance

3.1. Maximum Entropy as Epistemic Honesty

Jaynes’s principle of maximum entropy [2] is often presented as a prior selection rule. This understates its scope. Maximum entropy is an epistemological commitment: among all representations consistent with the available evidence, commit to the one that is maximally honest about what you do not know. Any more concentrated representation asserts information the evidence has not provided.

In the operational context of spacecraft navigation this imperative is visceral. Before launch, before any tracking data has been acquired, the navigation team must characterize the initial state uncertainty. A tight prior not justified by evidence resists correction when early observations disagree with it. The filter interprets genuine trajectory errors as measurement noise, and the estimation problem is compromised from the start. The honest prior is as diffuse as actual pre-launch knowledge warrants: no tighter, and no looser.

The AEGIS work with DeMars [11] moved in this direction: monitoring differential entropy online to trigger adaptive expansion of the Gaussian mixture, avoiding premature concentration. But AEGIS was still committed to probabilistic closure. When the uncertainty was epistemic rather than statistical, no number of Gaussian mixands could honestly represent a state of genuine ignorance.

3.2. Possibilistic Entropy as the Geometric Maximum Entropy

When uncertainty is geometric and bounded — when evidence acts by eliminating hypotheses rather than updating probabilities — the appropriate entropy is Boltzmann entropy: the logarithm of the volume of the admissible region.

Definition 2 (Possibilistic entropy [28]). Let π be a normalized possibility distribution over a finite support $\{\chi^{(i)}\}_{i=1}^M \subset \mathbb{R}^n$. The possibilistic entropy is

$$\mathcal{E}_\pi = \int_0^1 \log V_\alpha d\alpha, \quad (5)$$

where $V_\alpha = c_n(\det \Pi_\alpha)^{1/2}$ is the MVEE volume of the α -cut $\mathcal{C}_\alpha = \{i : \pi^{(i)} \geq \alpha\}$.

The possibilistic entropy decomposes into two geometrically distinct terms [26]:

$$\mathcal{E}_\pi = \underbrace{\log V + \kappa_n}_{\text{basin volume term}} + \underbrace{\int_0^1 \log \frac{V_\alpha}{V} d\alpha}_{\text{impossibility gradient term} \leq 0}, \quad (6)$$

where $V = c_n(\det \Pi)^{1/2}$ is the full-support MVEE volume. The basin volume term measures the *extent* of remaining ignorance; the impossibility gradient term measures its *texture*.

The ESPF's expansion step — governed by the asymmetric rate limit $r^+ = 1.15$, permitting rapid support expansion when observations are sparse — is the operational expression of this Jaynesian commitment. When evidence is insufficient to justify concentration, embrace ignorance rather than assert false structure.

Remark 3 (Why Boltzmann, not Shannon). *Possibility theory uses the min rule for conjunction, under which — log is not additive across independent sources. The independence that matters in TEAG is geometric: volumes multiply across independent geometric constraints, and the logarithm converts that product to a sum. Boltzmann entropy respects this geometric independence; Shannon entropy does not [28].*

4. The Popperian Imperative: Minimizing False Information

4.1. Falsification as the Rejection of False Information Capacity

Popper's falsification criterion [3] is the epistemological dual of Jaynes's maximum entropy principle. Where Jaynes says: *commit to nothing beyond what the evidence provides*, Popper says: *reject everything that the evidence refutes*. Together they define the boundaries of legitimate epistemic commitment.

False information capacity is the epistemic cost of asserting certainty the evidence has not warranted. A filter that contracts its posterior below what the observation geometry supports is not merely overconfident — it is generating false information. Every attempt to infer debris physical properties from light curves and angles data [15,17] confronted this directly: unless the initial guess was already close to the truth, the filter would confidently converge to the wrong answer, having expelled the correct hypotheses without evidentiary justification.

The Popperian imperative, operationalized in TEAG, is: eliminate exactly what the evidence eliminates, and nothing more. This is the possibilistic conjunctive update

$$\pi'(h) = \min(\pi(h), \kappa(y | h)), \quad (7)$$

where $\kappa(y | h) \in (0, 1]$ is the compatibility of hypothesis h with observation y . Admissibility is reduced where the evidence is incompatible, and preserved where it is compatible. Crucially, admissibility is *never amplified*: $\pi'(h) \leq \pi(h)$ for all h . This is Popperian monotonicity.

4.2. The Epistemic Validity Condition for Probability

The Popperian imperative, combined with the Jaynesian imperative of Section 3, implies a precise condition under which a probability distribution is a legitimate representation of uncertainty — as opposed to an artifact imposed upon it. We state this formally.

Definition 3 (Epistemic Validity Condition). *A probability distribution P over hypothesis space $\mathcal{H}$ is epistemically admissible for an inference problem if and only if all three of the following conditions hold simultaneously.*

(EVC-1) Justified parameterization. *The mapping $\varphi : \mathcal{U} \rightarrow \mathcal{H}$ from the underlying uncertainty structure $\mathcal{U}$ to the parameter space is bijective and preserves the topology of $\mathcal{U}$. If φ is many-to-one, P is a projection artifact of an unjustified coordinatization choice, not a representation of the uncertainty. Distinct elements of $\mathcal{U}$ that project onto the same element of $\mathcal{H}$ cannot be recovered from any computation performed in $\mathcal{H}$; the information they carried is permanently destroyed at the point of mapping.*

(EVC-2) Grounded likelihoods. *The likelihood $p(y | h)$ corresponds to a physically or observationally justified generative process. A likelihood adopted for computational convenience — Gaussian noise because it is tractable, not because the noise process is Gaussian — does not satisfy this condition. Ungrounded likelihoods produce posteriors that are mathematically well-formed but epistemically invalid: they appear to sharpen under evidence while tracking a model of the noise, not a model of the phenomenon.*

(EVC-3) Earned distinguishability. *The observation geometry generates sufficient epistemic contrast to justify probabilistic commitment. Formally, the Fisher information matrix $\mathbf{F}(\theta)$ must have eigenvalues above a threshold set by the application risk budget across all directions of $\mathcal{H}$ relevant to the decision. When the geometry is flat in a decision-relevant direction — when the evidence cannot distinguish among the hypotheses that matter — a probability distribution that appears to commit to that direction is asserting certainty that the geometry has not provided.*

Remark 4 (When the EVC fails). *When any of EVC-1 through EVC-3 fails, the resulting probability distribution is not an imprecise representation of the truth. It is what we call an epistemic scalarization artifact: a number that looks like a measurement, carries the authority of precision, and propagates through downstream computations as if it were a stable observable — while representing nothing in the underlying uncertainty structure that justified its construction. Two decades of operational experience in orbit determination and space situational awareness provided repeated empirical demonstrations of each failure mode. EVC-1 fails when the mapping from debris physical properties to orbital dynamics is many-to-one and the physical properties are unknown: multiple physically distinct objects project onto indistinguishable trajectories, and averaging over them produces a parameter value that corresponds to no physical object. EVC-2 fails when Gaussian process noise is assumed for non-gravitational accelerations that are bounded, structured, and physically determined by properties we do not know. EVC-3 fails when the observation geometry is flat: when, near apoapsis, successive range-rate measurements are nearly identical and the Fisher information has collapsed in the directions most relevant to orbit determination.*

Remark 5 (The EVC and CAPphrase). *The CAPphrase verbal probability phrase data [23,24] provides a domain-independent empirical demonstration of EVC-1 failure at scale. When the phrase Realistic Possibility is mapped to a probability by 5,174 participants, the mapping φ is many-to-one: different interpretive communities apply distinct, internally coherent but mutually inconsistent functions f_A, f_B, f_C to arrive at genuinely different numeric values. The Bayesian posterior is a projection artifact of pooling across these communities; its MAP value of 60.0% corresponds to no community's actual interpretation. No downstream computation can recover the structure destroyed at intake. TEAG's admissible basin of [25.1, 79.9]% correctly represents what the evidence actually supports: a set of interpretations, not a single number.*

TEAG is the inference framework that operates without assuming the EVC holds. It requires only that hypotheses can be ordered by admissibility, that evidence contracts rather than redistributes, and

that commitment follows from geometric centrality within the surviving basin. When the EVC is satisfied and epistemic width contracts to zero, TEAG recovers probability theory exactly (Theorem 1). When the EVC fails, TEAG provides the honest alternative: the admissible basin, the medioid, and the epistemic width as first-class outputs rather than the false precision of an unearned scalar.

4.3. Two-Stage Falsification and the Tropical Variety

In log-admissibility coordinates $\Phi_{\emptyset}(h) = -\log \pi(h)$, the conjunctive update becomes tropical addition in the max-plus semiring:

$$\tilde{\Phi}_{\emptyset}(h) = \Phi_{\emptyset}(h) \oplus \Phi_S(h) = \max(\Phi_{\emptyset}(h), \Phi_S(h)), \quad (8)$$

where $\Phi_S(h) = \frac{1}{2} \|L_e^{-1}(y - g(h))\|^2$ is the surprisal field. The tropical variety of this update,

$$\mathcal{B}_{\text{active}} = \{h \in \mathcal{H} : \Phi_{\emptyset}(h) = \Phi_S(h)\}, \quad (9)$$

is the *active deformation front*: the exact locus where incoming evidence first matches prior impossibility and begins to deform the posterior field. This is a *necessary* condition for falsification. Sufficient falsification requires exit from the PCRB-admissible basin

$$\mathcal{A}_k = \{h \in \mathcal{H} : \tilde{\Phi}_{\emptyset,k}(h) \leq c_k^*\}, \quad (10)$$

where c_k^* is the equipotential threshold determined by the PCRB at step k . The active deformation front marks where falsification *becomes possible*; $\partial \mathcal{A}_k$ marks where falsification is *complete* [28].

Remark 6 (Non-resurrection and the permanence of false falsification). *The non-resurrection axiom of TEAG (Axiom A3) states that a falsified hypothesis cannot be restored by subsequent evidence. This is epistemically correct when falsification is justified. But it also means that false falsification — expelling a hypothesis without sufficient evidentiary cause — is permanent. The repeated failure to recover debris physical properties from sparse photometric data [6,15] was, in retrospect, a manifestation of this: once the filter had committed to an incorrect orientation or shape, the surviving hypothesis space no longer contained the truth, and no amount of additional data could recover it.*

5. The PCRB: Holding Both Principles Simultaneously

5.1. The Bound as Joint Constraint

The Possibilistic Cramér–Rao Bound bounds the rate at which the admissible basin can contract per observation from below by a quantity determined entirely by the epistemic contrast of the observation:

$$\mathcal{E}_{\pi_k|k} \geq \mathcal{E}_{\pi_k|k-1} + \frac{n}{2} \log(1 - I_k). \quad (11)$$

The right-hand side is non-positive: entropy does not increase under evidence (Popperian monotonicity). But it cannot decrease faster than $-\frac{n}{2} \log(1 - I_k)$: no observation can justify more epistemic contraction than the contrast it generates warrants (Jaynesian honesty).

The PCRB is not a filter-specific result. It is a universal geometric bound on the rate at which information capacity can be legitimately consumed per observation [26]. It operationalizes the navigator’s lesson: the information extractable from an observation is bounded by its geometry, and no estimation algorithm can extract more information than the geometry provides.

5.2. The Asymmetric Rate Limits as Epistemological Commitments

The ESPF implements the PCRB through asymmetric rate limits on admissible support volume: $r^+ = 1.15$ for expansion and $r^- = 0.97$ for contraction [25]. These are not tuning parameters. They are epistemological commitments.

The expansion limit implements the Jaynesian imperative: embracing ignorance when evidence is insufficient to justify concentration is epistemically honest. The contraction limit implements the Popperian imperative: asserting certainty faster than the evidence warrants is false information. The filter is quick to embrace ignorance and slow to assert certainty — not as a design philosophy but as a mathematical consequence of the PCRB.

6. Probability is Earned: Bayesian Inference as the Limit of TEAG

6.1. The Gaussian Collapse Theorem

The relationship between TEAG and Bayesian inference is not one of opposition but of containment. Probability theory is the geometry of TEAG when epistemic width has contracted to zero.

Theorem 1 (Gaussian collapse [27]). *Let $\{\pi_t\}_{t \geq 0}$ be a sequence of normalized, consonant possibility distributions with associated credal sets $\{\mathcal{P}_{\pi_t}\}$. Suppose consonance, credal contraction $\sup_A [\Pi_t(A) - N_t(A)] \rightarrow 0$, L^1 convergence $\pi_t \rightarrow p^*$, and domination hold. Then for every bounded measurable f ,*

$$\int f d\text{Ch}_{\pi_t} \rightarrow \int f(x) p^*(x) d\mu(x) \quad \text{as } t \rightarrow \infty. \quad (12)$$

The ESPF recovers the Kalman filter, and $\mathcal{E}_\pi \rightarrow \frac{1}{2} \log \det \Sigma + \text{const}(n)$.

When epistemic width $W = \int (\pi - n) d\mu$ approaches zero — when the evidence has thoroughly constrained the admissible support — TEAG and Bayesian inference agree. Probability theory is earned by the accumulation of sufficiently constraining evidence. It is the asymptotic geometry of possibilistic inference in the data-rich regime, not its default starting point.

6.2. Convergent Optimality, Not Containment

The ESPF and the Kalman filter are optimal solutions to categorically different problems: the Kalman filter minimizes mean squared error under Gaussian assumptions and a known dynamical model; the ESPF minimizes possibilistic entropy subject to the PCRB under bounded epistemic uncertainty and a structurally uncertain dynamical model. These solutions agree precisely when the world is Gaussian, the model is valid, and epistemic width has contracted to zero. This is *convergent optimality*: two frameworks optimal under different epistemic commitments that coincide when their domains of applicability overlap.

The practitioner who uses the SRIF or differential correction in the nominal regime — where the dynamical model is well-characterized and the uncertainty is genuinely statistical — and the ESPF when the model is structurally uncertain is not switching between competing frameworks. They are applying the appropriate tool to the appropriate epistemic situation, with a precise mathematical characterization of where the transition occurs.

7. Implications for Machine Learning

7.1. The Default Commitment to Probabilistic Closure

The machine learning community's dominant inference paradigm is probabilistic closure: uncertainty is represented as a normalized probability distribution, evidence acts through Bayes' rule, and the posterior is the output. This paradigm is powerful when its assumptions are satisfied. In many real-world settings they fail: observations may be sparse, bounded, or acquired through non-uniformly informative geometries; dynamical models may be structurally incomplete; the hypothesis space may not carry a natural probability measure. In these settings the commitment to probabilistic closure is not neutral — it incurs a false information capacity cost the PCRB makes precise.

7.2. What TEAG Offers

TEAG offers not a replacement for probabilistic inference but a principled criterion for when probabilistic closure is justified. The epistemic width W is a computable diagnostic: when $W < W_{\text{crit}}$, probabilistic inference is appropriate. When $W \geq W_{\text{crit}}$, possibilistic inference is required.

A concrete, large-scale empirical demonstration of this principle comes from a domain far removed from orbital mechanics: the interpretation of verbal probability phrases in natural language.

7.3. Verbal Probability Phrases: Embracing Ignorance Reveals Structure

Adam Kucharski's CAPphrase dataset [23] provides 98,306 absolute numeric probability judgments from 5,174 participants across 19 verbal probability terms — phrases such as *likely*, *realistic possibility*, *almost certain*, and *could happen* — gathered through a large-scale online survey (data and code publicly available at <https://github.com/adamkucharski/CAPphrase>). Each participant assigned a percentage in $[0, 100]$ to each phrase. The dataset therefore provides a direct empirical window into how populations translate linguistic uncertainty into numerical form.

Standard probabilistic analysis of such data asks: what is the central tendency and variance of interpretation for each phrase? Bayesian inference formalizes this by constructing a posterior over a latent true probability value and collapsing to a maximum a posteriori (MAP) point estimate. For phrases whose evidence cloud is genuinely bimodal or multimodal — where different interpretive communities assign systematically different meanings — this collapse is not a summary but a distortion. We call this the *Epistemic Scalarization Error* (ESE): the many-to-one collapse of a structurally contested interpretation space onto a single scalar, together with the fabricated authority that scalar then acquires by virtue of its numeric form.

TEAG applied to this dataset [24] reveals exactly what the Jaynesian and Popperian imperatives predict. For the phrase *Realistic Possibility* — where the population standard deviation is $\sigma = 20.5\%$ and five major institutional bodies (the IPCC, US National Intelligence Council, UK Joint Intelligence Committee, EFSA, and NATO) specify reference ranges that are mutually inconsistent — the Bayesian MAP collapses to a single scalar of 60.0%. That number will travel. It will enter a risk pipeline, a briefing document, a decision threshold. It will look like a measurement. It is not. It is an artifact of averaging across interpretive communities that do not share a common referent for the phrase.

TEAG's admissible basin for *Realistic Possibility* spans $[25.1, 79.9]\%$, with Epistemic Width $W = 54.9\%$. This is not vagueness — it is geometric honesty. The framework correctly identifies that the current evidence does not warrant commitment to a single value. The UK Joint Intelligence Committee's institutional range of $[40, 50]\%$ falls entirely within the TEAG admissible basin but is not recovered by the Bayesian MAP estimate.

Three findings from this analysis bear directly on the thesis of this paper.

Embracing ignorance maximizes information capacity.

The TEAG admissible basin width W tracks population-level epistemic disagreement monotonically and correctly spans the institutional yardstick ranges that MAP estimates systematically miss. For *Highly Likely*, where five institutional bodies themselves disagree (NATO: $[90, 100]\%$; UK: $[80, 90]\%$), the TEAG basin spans $[74.9, 95.0]\%$, encompassing the inter-institutional disagreement. The Bayesian MAP of 90.0% cannot represent this. By refusing premature collapse, TEAG preserves the geometric structure through which the actual disagreement can be seen.

The PCRB bounds how fast certainty can be earned.

The possibilistic information content I_k varies meaningfully across all 19 phrases, distinguishing genuine consensus from epistemic indifference. *About Even* has $\sigma = 3.7\%$ and $W = 7.8\%$: the population is genuinely converged, and the evidence has earned a narrow basin. *Realistic Possibility* has $\sigma = 20.5\%$ and $W = 54.9\%$: the PCRB floor is high because the observation geometry — the distribution of responses across the hypothesis space — generates low contrast among the most credible hypotheses. More respondents cannot fix this. More *informative* evidence geometry could.

May, Might, and Could Happen are epistemically indistinguishable.

The TEAG analysis formally identifies the May/Might/Could Happen cluster as epistemically identical to four decimal places ($\sigma_n = 0.300$, $I_k = 0.484$, $W \approx 51\%$). This is a falsifiable geometric prediction: if Kucharski's independent pairwise comparison data shows high inconsistency rates for any pair drawn from this cluster, the TEAG result is validated. The operative signal for all three phrases is the same one that spacecraft navigation taught: the evidence geometry is too flat here. Gather more — and gather it where the contrast is richest.

The CAPphrase results make the paper's central claim concrete and empirical. Bayesian closure applied to structurally contested linguistic data does not produce imprecise representations of the truth. It produces epistemically invalid representations of scalars that do not exist in the underlying human interpretation space. Embracing ignorance — representing the admissible basin rather than collapsing it — is not a failure of inference. It is the correct response to the evidence geometry, and it is the path that preserves the information capacity to learn more.

7.4. A Design Principle for Information-Aware Inference

In orbit determination, we schedule observations where the trajectory changes fastest — where the Fisher information is maximized. In inference, we should preserve epistemic structure where the hypotheses separate most under evidence — where the epistemic contrast is maximized. TEAG is the framework that makes this precise without collapsing the admissible space prematurely.

Maximize information capacity by preserving geometric degrees of freedom. Minimize false information by eliminating only what the evidence forces. The first principle is Jaynes. The second is Popper. TEAG holds both simultaneously, with Bayesian inference as the geometry they converge to when the evidence has earned it.

8. Conclusions

TEAG is not non-Bayesian. It is epistemically prior to Bayesian inference.

The framework emerged from two decades of operational experience with inference problems that probabilistic closure could not honestly represent: spacecraft filters that asserted confidence they had not earned; debris tracks that drifted outside formally valid uncertainty ellipsoids; light curve inversion problems that failed unless the initial guess was already close; HAMR objects whose trajectories resisted prediction because the physical forces acting on them were genuinely unknown; adaptive Gaussian mixtures that worked beautifully when uncertainty was statistical and struggled when it was epistemic. Every approach illuminated the problem more clearly while also revealing its own limits.

TEAG is what emerged from starting over. Not from dissatisfaction with probability theory as mathematics, but from the repeated operational lesson that probability must be earned by evidence. The Epistemic Validity Condition (Definition 3) makes this precise: a probability distribution is epistemically admissible only when the parameterization is bijective, the likelihoods are grounded in physical process, and the observation geometry has generated sufficient contrast to justify commitment. Two decades of operational inference — spacecraft navigation, debris tracking, attitude estimation, multi-target catalog maintenance — provided repeated empirical demonstrations of each condition failing. The filters did not fail because they were poorly designed. They failed because the conditions under which probability is valid were not met, and no one had a formal criterion for recognizing that.

The empirical validation of this principle comes from a domain as far from orbital mechanics as one can imagine. Applied to Kucharski's CAPphrase dataset [23] — 98,306 probability judgements from 5,174 participants across 19 verbal probability phrases — TEAG demonstrates precisely what the Jaynesian and Popperian imperatives predict [24]. Where the evidence is genuinely contested, as for *Realistic Possibility* ($\sigma = 20.5\%$, five institutional bodies in mutual disagreement), embracing ignorance — preserving the admissible basin $W = 54.9\%$ rather than collapsing to a MAP scalar of 60.0% — is not

a failure of inference. It is the correct epistemic response, and the only response that preserves the information capacity to learn more. Where the evidence is genuinely consensual, as for *About Even* ($\sigma = 3.7\%$, $W = 7.8\%$), the PCRB floor is low and the narrow basin has been earned. The framework does not embrace ignorance indiscriminately — it embraces exactly as much ignorance as the evidence geometry warrants, and no more.

This is the same principle that every spacecraft navigator internalizes through the launch package analysis, and every debris tracker internalizes through the Fisher information geometry of a precious telescope schedule. The field tells you where the information is. Do not assert certainty where the field is flat.

The Possibilistic Cramér–Rao Bound is the joint expression of both the Jaynesian and Popperian imperatives: a universal geometric bound on the rate at which information capacity can be legitimately consumed per observation. The ESPF’s asymmetric rate limits are its operational expression: quick to embrace ignorance, slow to assert certainty.

In order to know something, one must measure it. In order to understand something, one must predict it. TEAG is the framework that makes both honest.

References

1. G.J. Biermann. *Factorization Methods for Discrete Sequential Estimation*. Academic Press, New York, 1977.
2. E.T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–630, 1957.
3. K.R. Popper. *The Logic of Scientific Discovery*. Hutchinson, London, 1959.
4. R. Mahler. *Statistical Multisource-Multitarget Information Fusion*. Artech House, Norwood, MA, 2007.
5. M.K. Jah, M.E. Lisano, G.H. Born, and P. Axelrad. Mars aerobraking spacecraft state estimation by processing inertial measurement unit data. *Journal of Guidance, Control, and Dynamics*, 31(6):1802–1813, 2008. DOI: 10.2514/1.24304..
6. C.J. Wetterer and M. Jah. Attitude estimation from light curves. *Journal of Guidance, Control, and Dynamics*, 32(5):1648–1651, 2009. DOI: 10.2514/1.44254..
7. T. Kelecý and M. Jah. Analysis of high area-to-mass ratio (HAMR) GEO space object orbit determination and prediction performance: Initial strategies to recover and predict HAMR GEO trajectories with no a priori information. *Acta Astronautica*, 69(7–8):551–558, 2011. DOI: 10.1016/j.actaastro.2011.04.019..
8. T. Kelecý, M. Jah, and K. DeMars. Application of a Multiple Hypothesis Filter to near GEO high area-to-mass ratio space objects state estimation. *Acta Astronautica*, 81(2):435–444, 2012. DOI: 10.1016/j.actaastro.2012.08.006..
9. K.J. DeMars, M.K. Jah, and P.W. Schumacher Jr. Initial orbit determination using short-arc angle and angle rate data. *IEEE Transactions on Aerospace and Electronic Systems*, 48(3):2628–2637, 2012. DOI: 10.1109/TAES.2012.6237613..
10. K.J. DeMars and M.K. Jah. Probabilistic initial orbit determination using Gaussian mixture models. *Journal of Guidance, Control, and Dynamics*, 36(5):1324–1335, 2013. DOI: 10.2514/1.59844..
11. K.J. DeMars, R.H. Bishop, and M.K. Jah. Entropy-based approach for uncertainty propagation of non-linear dynamical systems. *Journal of Guidance, Control, and Dynamics*, 36(4):1047–1057, 2013. DOI: 10.2514/1.58987..
12. C. Früh, T.M. Kelecý, and M.K. Jah. Coupled orbit-attitude dynamics of high area-to-mass ratio (HAMR) objects: Influence of solar radiation pressure, Earth’s shadow and the visibility in light curves. *Celestial Mechanics and Dynamical Astronomy*, 117(4):385–404, 2013. DOI: 10.1007/s10569-013-9516-5..
13. C. Früh and M. Jah. Attitude and orbit propagation of high area-to-mass ratio (HAMR) objects using a semi-coupled approach. *Journal of the Astronautical Sciences*, 2013. Published online 9 July 2014.
14. C. Früh and M.K. Jah. Coupled orbit-attitude motion of high area-to-mass ratio (HAMR) objects including efficient self-shadowing. *Acta Astronautica*, 95(1):227–241, 2014. DOI: 10.1016/j.actaastro.2013.11.017..
15. R. Linares, M.K. Jah, J.L. Crassidis, and C.K. Nebelecky. Space object shape characterization and tracking using light curve and angles data. *Journal of Guidance, Control, and Dynamics*, 37(1):13–25, 2014. DOI: 10.2514/1.62986..
16. T. Kelecý, M. Jah, J. Baldwin, and J. Stauch. High area-to-mass ratio object population assessment from data/track association. *Acta Astronautica*, 96(1):166–174, 2014. DOI: 10.1016/j.actaastro.2013.11.037..

17. R. Linares, M.K. Jah, J.L. Crassidis, F.A. Leve, and T. Kelecy. Astrometric and photometric data fusion for inactive space object mass and area estimation. *Acta Astronautica*, 99(1):1–15, 2014. DOI: 10.1016/j.actaastro.2013.10.018..
18. K.J. DeMars, I.I. Hussein, C. Frueh, M.K. Jah, and R.S. Erwin. Multiple-object space surveillance tracking using finite-set statistics. *Journal of Guidance, Control, and Dynamics*, 38(9):1741–1756, 2015. DOI: 10.2514/1.G000987..
19. J. Stauch, T. Bessell, M. Rutten, J. Baldwin, M. Jah, and K. Hill. Joint probabilistic data association and smoothing applied to multiple space object tracking. *Journal of Guidance, Control, and Dynamics*, 2017. DOI: 10.2514/1.G002230..
20. E. Delande, J. Houssineau, J. Franco, C. Frueh, D. Clark, and M. Jah. A new multi-target tracking algorithm for a large number of orbiting objects. *Advances in Space Research*, 64(3):645–667, 2019. DOI: 10.1016/j.asr.2019.04.012..
21. H. Cai, I. Hussein, and M. Jah. Possibilistic admissible region using outer probability measure theory. *Acta Astronautica*, 177:246–257, 2020. DOI: 10.1016/j.actaastro.2020.07.041..
22. H. Cai, J. Houssineau, B.A. Jones, M. Jah, and J. Zhang. Possibility generalized labeled multi-Bernoulli filter for multi-target tracking under epistemic uncertainty. *IEEE Transactions on Aerospace and Electronic Systems*, 2022. DOI: 10.1109/TAES.2022.3200022..
23. A.J. Kucharski. CAPphrase: Comparative and Absolute Probability phrase dataset. Zenodo, 2026. DOI: 10.5281/zenodo.18750055. <https://github.com/adamkucharski/CAPphrase>.
24. M.K. Jah. The Geometry of Linguistic Uncertainty: A Possibilistic Alternative to Bayesian Collapse in Verbal Probability Interpretation. Preprint, GaiaVerse, Ltd., March 2026.
25. M.K. Jah and V. Haslett. The Epistemic Support-Point Filter (ESPF): A bounded possibilistic framework for ordinal state estimation. *arXiv:2508.20806*, 2025.
26. M.K. Jah and V. Haslett. The Epistemic Support-Point Filter: Jaynesian maximum entropy meets Popperian falsification. *arXiv:2603.10065v2*, 2025.
27. M.K. Jah. The Geometry of Knowing: From possibilistic ignorance to probabilistic certainty. Preprint, *arXiv:submit/7362363*, 2026.
28. M.K. Jah. Theory of Epistemic Abductive Geometry (TEAG): A unified theory of admissibility-driven inference across dynamical systems, measure theory, and language. Preprint, doi:10.20944/preprints202603.2010.v2, 2026.
29. M.K. Jah. The Epistemic Support-Point Filter as a Tropical Hamilton–Jacobi System: Wavefront Propagation and Possibilistic Inference. Preprint, doi:10.20944/preprints202603.2110.v1, 2026.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.