Beyond Exact-Match: Semantic Authentication for AI-Native Wireless Systems

Abstract

AI-native NextG wireless systems increasingly exchange compressed, generated, and machine-interpretable representations whose value lies in their meaning rather than in exact message recovery. This shift creates a mismatch with conventional authentication mechanisms, which verify bit-level or message-level equality and may reject benign semantic transformations. We propose a semantic authentication framework that accepts a recovered meaning whenever it remains within an admissible equivalence class of the intended meaning. The equivalence classes are induced by a description-length-based semantic distance, which measures the excess coding cost incurred when an observation generated under one meaning is interpreted under another. Using this metric, we develop a set-based authentication rule with an invariant semantic hash, analyze the resulting false-rejection probability under benign perturbations, and validate the analytical results through simulation.

1. Introduction

AI-native NextG wireless systems are moving beyond conventional bit-level communication toward task-driven exchange of compressed, generated, and machine-interpretable representations. In such systems, the utility of a transmission is determined not by exact symbol recovery, but by whether the receiver recovers the intended meaning with sufficient fidelity. This shift creates a fundamental security mismatch: classical authentication mechanisms verify exact messages, whereas AI-native communication pipelines naturally permit benign semantic transformations introduced by compression, relaying, learned inference, or generative reconstruction (Chaccour et al., 2024; Strinati et al., 2024; Zhang et al., 2025).

This mismatch is particularly important in applications such as autonomous transportation, robotic coordination, and edge intelligence, where generative and multi-modal models may produce different representations that nevertheless preserve the same operational intent (Li & Aijaz, 2025; Pan et al., 2025). A receiver does not always need to reconstruct the exact transmitted representation; rather, it must recover an interpretation that is consistent with the sender's intended task. For example, in a connected-driving scenario, an instruction such as "slow down and merge right due to a lane closure ahead" may be reconstructed as "reduce speed and move right before reaching the blocked lane." Although the wording differs, the resulting driving action is the same. Exact-match authentication would reject such a message, despite its task-relevant meaning being preserved.

These observations motivate *semantic authentication over equivalence classes*. Instead of authenticating a single exact representation, the receiver accepts a message whenever its inferred meaning lies within an admissible region around the intended meaning. To formalize this notion, we introduce a *description-length-based semantic distance*. Under the minimum description length (MDL) principle (Grünwald, 2007), a meaning can be viewed as a generative model for observable messages. Semantic discrepancy is then measured by the additional description length incurred when a message generated under the intended meaning is encoded or interpreted under an alternative meaning. This distance naturally induces equivalence classes in meaning space.

Building on this distance, we develop a set-based authentication framework in which semantic validity is defined by membership in an admissible equivalence class rather than by exact equality. The framework uses an invariant semantic hash that remains unchanged over meanings within the same admissible class, enabling authentication to tolerate benign semantic variation while still rejecting meaningfully different or adversarial interpretations. This makes the proposed approach compatible with AI-native wireless systems in which communication, inference, control, and generation are tightly coupled.

The main contributions of this paper are as follows. First, we formulate semantic authentication over equivalence classes for AI-native wireless systems, allowing a recovered meaning to be accepted whenever it remains within an admissible semantic region. Second, we define these admissible regions using a description-length-based semantic distance and construct a set-based authentication rule with an invariant semantic hash. Third, we analyze robustness through the false-rejection probability under benign Gaussian perturbations. For the isotropic case, we derive a closed-form expression; for the anisotropic case, we characterize how the eigenvalue spread of the perturbation covariance and Fisher information matrix increases the false-rejection probability relative to the isotropic benchmark. Finally, numerical results validate the analytical expression through Monte Carlo simulation and quantify the performance gap between isotropic and anisotropic regimes.

2. System Model

We consider an AI-native NextG wireless system in which a transmitter communicates task-relevant semantic content to a receiver over a wireless channel. To describe the end-to-end pipeline, we distinguish the original high-dimensional source S from a compact semantic representation M that captures the task-relevant meaning of S. Let $\mathcal{S}$, $\mathcal{M}$, and $\mathcal{X}$ denote the source, meaning, and channel-input spaces, respectively. The overall semantic communication pipeline is represented as

$$S \to M \to X \to X' \to \hat{M} \to \hat{S}. \tag{1}$$

Each stage is described below.

Semantic extraction ($S \to M$). The transmitter first applies a semantic extractor $M = F(S)$, where $F : \mathcal{S} \to \mathcal{M}$ maps the high-dimensional source S to a compact representation $M \in \mathcal{M}$. This representation retains the information needed for the downstream task while discarding irrelevant source-level details. The mapping is generally many-to-one: multiple source realizations may correspond to the same meaning M.

Semantic encoding ($M \to X$). The semantic encoder maps the meaning representation to a channel input, $X = E(M)$, where $E : \mathcal{M} \to \mathcal{X}$. This stage converts semantic content into a physical-layer signal suitable for wireless transmission.

Wireless channel ($X \to X'$). The channel input X is transmitted through a NextG wireless channel described by the conditional distribution $P(X' \mid X, H)$, where X' is the received signal and H denotes the channel state, including fading, interference, mobility, and resource allocation.

Semantic inference ($X' \to \hat{M}$). Upon receiving X', the receiver applies a semantic inference rule $\hat{M} = T(X')$, where $T : \mathcal{X} \to \mathcal{M}$ estimates the meaning conveyed by the received signal.

Generative reconstruction ($\hat{M} \to \hat{S}$). Finally, a generative model synthesizes a reconstruction of the original source, $\hat{S} = G(\hat{M})$, where $G : \mathcal{M} \to \mathcal{S}$. This stage is generally one-to-many: the same inferred meaning $\hat{M}$ may produce many syntactically different but semantically equivalent outputs $\hat{S}$.

Role of semantic authentication. The structure of the pipeline also clarifies where authentication should be performed. The $S \to M$ stage is many-to-one: different source files, sensor observations, or surface-level phrasings may map to the same semantic meaning. Conversely, the $\hat{M} \to \hat{S}$ stage is one-to-many: the same inferred meaning can be rendered into multiple syntactically distinct reconstructions. As a result, exact-match authentication at the source level, signal level, or reconstructed-output level may reject valid transmissions simply because compression, channel distortion, relaying, or generative reconstruction introduces benign surface variation.

Authentication at the meaning layer is therefore the appropriate granularity. Rather than verifying whether $X' = X$ or $\hat{S} = S$, the receiver should verify whether the inferred meaning $\hat{M}$ is semantically consistent with the intended meaning M. This motivates semantic authentication over admissible meaning classes, in which a transmission is accepted whenever $\hat{M}$ preserves the task-relevant meaning of M.

3. Authentication over Semantic Equivalence Classes

This section develops a semantic authentication framework for AI-native NextG wireless systems, where task-relevant meanings are communicated through learned, adaptive, and potentially multi-hop wireless pipelines rather than preserved as exact bit strings.

3.1. Semantic Distance Based on Description Length

A central challenge in semantic authentication is to quantify the discrepancy between two meanings in a way that is compatible with AI-native communication. In NextG wireless systems, messages may be compressed, paraphrased, translated, relayed, or regenerated by edge models and intermediate agents. Therefore, semantic proximity should be measured at the meaning level rather than at the level of exact symbols or representations. To this end, we adopt a description-length-based semantic distance.

Under the minimum description length (MDL) principle, a meaning M is viewed as a generative model that explains an observable message x. The quality of this explanation is measured by the conditional description length $L(x \mid M)$, which denotes the number of bits required to encode x when the receiver assumes semantic model M. If x is generated according to the intended meaning M, then an alternative interpretation $\hat{M}$ generally incurs a larger description length. The resulting excess description length therefore provides a natural measure of semantic discrepancy.

Definition 3.1 (DL-Based Semantic Distance)**.** The semantic distance between two meanings M and $\hat{M}$ is defined as

$$d_{\mathrm{DL}}(\hat{M}, M) = \mathbb{E}_{x \sim P(\cdot \mid M)} \Big[L(x \mid \hat{M}) - L(x \mid M) \Big]. \tag{2}$$

The quantity $d_{\mathrm{DL}}(\hat{M}, M)$ measures the expected increase in description length when data generated under the intended meaning M is interpreted using the alternative meaning $\hat{M}$. It therefore captures how well two meanings explain the same observations. This makes it suitable for semantic communication settings in which multiple representations may correspond to the same operational intent.

Lemma 3.2 (KL Interpretation of DL-Based Semantic Distance)**.** *Define the conditional description length by*

$L(x \mid M) = -\log P(x \mid M)$. *Then*

$$d_{\mathrm{DL}}(\hat{M}, M) = D\Big(P(\cdot \mid M) \,\|\, P(\cdot \mid \hat{M})\Big). \tag{3}$$

3.2. Semantic Equivalence Class in Meaning Space

Using the DL-based semantic distance, we define an admissible set of meanings that are acceptable for the downstream task. This set-based viewpoint is particularly important in AI-native wireless systems, where benign variations introduced by learned compression, edge inference, relaying, or generative reconstruction may change the surface representation without changing the task-relevant intent.

Definition 3.3 (Semantic Equivalence Class)**.** For a given meaning $M \in \mathcal{M}$ and tolerance $\epsilon \geq 0$, the semantic equivalence class centered at M is defined as

$$\mathcal{C}_\epsilon(M) = \{\tilde{M} \in \mathcal{M} : d_{\mathrm{DL}}(\tilde{M}, M) \leq \epsilon\}. \tag{4}$$

The set $\mathcal{C}_\epsilon(M)$ contains all meanings whose generative models are sufficiently close to that of M in terms of description length. Equivalently, these meanings explain observations generated under M with at most ϵ additional expected bits. They therefore represent task-consistent interpretations that are semantically indistinguishable from M within tolerance ϵ. Such variability may arise from benign factors such as inference uncertainty, representation mismatch, semantic compression, channel distortion, or limited model capacity at the receiver.

3.3. Inference of Semantic Meaning

In AI-native wireless systems, the receiver often performs semantic inference rather than exact symbol recovery. This inference can be interpreted through the MDL principle (Grünwald, 2007), under which the preferred meaning is the one that provides the shortest overall description of the received observation.

Let $L(M)$ denote the description length of a semantic representation M, and let $L(x' \mid M)$ denote the description length required to explain the received observation x' under meaning M. An MDL-based receiver infers the semantic meaning as

$$\hat{M} = \arg\min_{M'} \left[L(M') + L(x' \mid M')\right]. \tag{5}$$

Thus, the inferred meaning $\hat{M}$ is the most concise semantic explanation of the received wireless observation. This interpretation is especially natural when the receiver includes a learned semantic decoder, an edge model, or a task-aware inference module operating under uncertain channel conditions.

3.4. Set-Based Semantic Authentication

Set-based semantic authentication verifies whether the inferred meaning $\hat{M}$ lies within the semantic equivalence class of the intended meaning M. Unlike conventional authentication, which requires exact message equality, the receiver accepts a message if the recovered meaning remains sufficiently close to the intended meaning under the semantic distance metric. This makes the framework robust to benign AI-induced transformations while preserving sensitivity to semantically meaningful deviations.

Definition 3.4 (Set-Based Semantic Authentication)**.** A semantic authentication scheme is said to be *set-based* if the receiver accepts a received message x' whenever the inferred meaning $\hat{M}$ belongs to the semantic equivalence class of the intended meaning M, i.e., $\hat{M} \in \mathcal{C}_\epsilon(M)$, where $\mathcal{C}_\epsilon(M)$ denotes the ϵ-semantic equivalence class.

This formulation generalizes authentication from pointwise verification to region-based verification in meaning space. As a result, it is compatible with semantic communication pipelines in which exact representation matching is neither necessary nor realistic. For example, in vehicular or edge-assisted NextG wireless systems, multiple semantically equivalent reconstructions may correspond to the same intended traffic or control action and should therefore be accepted.

3.5. Semantic Hash Construction

To enable efficient and secure verification in the meaning domain, we introduce a *semantic hash function*

$$H_s : \mathcal{M} \to \{0,1\}^k, \tag{6}$$

which maps each meaning $M \in \mathcal{M}$ to a finite-length binary representation.

The purpose of $H_s(\cdot)$ is to assign the same hash value to meanings that are semantically equivalent, while assigning different hash values to meanings that are semantically distinct. Specifically, we require that

$$H_s(M_1) = H_s(M_2), \quad \forall M_1, M_2 \in \mathcal{C}_\epsilon(M), \tag{7}$$

so that all meanings within the semantic equivalence class $\mathcal{C}_\epsilon(M)$ produce the same hash value.

This construction is particularly relevant to AI-native wireless settings in which the receiver may observe different semantic realizations of the same intent due to compression, relay-side transformation, channel distortion, or learned generation. By hashing the semantic region rather than the exact representation, the authentication process becomes compatible with such transformations.

The semantic hash H_s can be constructed in two stages: a *quantization stage* that partitions the meaning space $\mathcal{M}$ into semantically coherent regions, and an *encoding stage* that assigns a fixed-length binary codeword to each region. We discuss each stage in turn.

3.6. Authentication Procedure

At the transmitter, the sender first computes the semantic hash of the intended meaning, $z = H_s(M)$, where $H_s(\cdot)$ maps meanings to finite-length binary representations and satisfies the invariance condition in (7). Using a shared secret key K, the sender then generates an authentication tag: $t = \mathrm{MAC}_K(z)$, where $\mathrm{MAC}_K(\cdot)$ denotes a keyed message-authentication function. The transmitted data therefore consists of the encoded semantic message $X = E(M)$ together with the authentication tag t.

At the receiver, the inferred meaning $\hat{M}$ is first obtained from the received observation x'. The receiver then computes the corresponding semantic hash $\hat{z} = H_s(\hat{M})$, and verifies the tag by checking $\mathrm{MAC}_K(\hat{z}) \stackrel{?}{=} t$.

Authentication succeeds if the verification holds, indicating that the recovered meaning $\hat{M}$ belongs to the same admissible semantic region as the intended meaning M. Thus, authentication is performed over task-relevant semantic regions rather than exact message representations. This makes the proposed procedure robust to benign variations introduced by AI-native wireless processing while still detecting semantically significant manipulations.

4. False-Rejection Probability

We now characterize the probability that a benign transmission is incorrectly rejected under the set-based semantic authentication rule. Recall that the receiver accepts the recovered meaning if

$$\hat{M} \in \mathcal{C}_\epsilon(M) \quad \Longleftrightarrow \quad d_{\mathrm{DL}}(\hat{M}, M) \le \epsilon. \tag{8}$$

Accordingly, a false rejection occurs when benign perturbations cause the recovered meaning to fall outside the semantic equivalence class, that is, $d_{\mathrm{DL}}(\hat{M}, M) > \epsilon$.

4.1. Local Perturbation Model

To obtain an explicit characterization of robustness, we model the recovered meaning under benign operation as a small perturbation of the intended meaning:

$$\hat{M} = M + W, \tag{9}$$

where W represents semantic estimation errors induced by channel noise, decoding imperfections, and inference uncertainty.

For sufficiently small perturbations, the DL-based semantic distance admits the local quadratic approximation

$$d_{\mathrm{DL}}(\hat{M}, M) = D\big(P(\cdot|M) \,\|\, P(\cdot|\hat{M})\big) \approx \frac{1}{2} W^\top J(M) W, \tag{10}$$

where $W = \hat{M} - M$ and

$$J(M) \triangleq \mathbb{E}_{x\sim P(\cdot|M)}\big[\nabla_M \log P(x|M) \nabla_M \log P(x|M)^\top\big] \tag{11}$$

is the Fisher information matrix of the semantic generative model $P(x|M)$ at M.

4.2. False-Rejection Probability

A false rejection occurs when the inferred meaning $\hat{M}$ falls outside the admissible semantic equivalence class of the intended meaning M, i.e., $d_{\mathrm{DL}}(\hat{M}, M) > \epsilon$. Accordingly, the false-rejection probability is

$$P_{\mathrm{FR}} = \Pr\Big(d_{\mathrm{DL}}(\hat{M}, M) > \epsilon\Big) \tag{12}$$

$$\approx \Pr\left(\frac{1}{2} W^\top J(M) W > \epsilon\right), \tag{13}$$

where $W = \hat{M} - M$ denotes the local perturbation in the semantic meaning space, and $J(M)$ is the Fisher information matrix associated with the local generative model around M.

The following theorem gives a closed-form approximation under a locally isotropic Gaussian perturbation model. This model is intended as a local approximation in the meaning space, rather than as a universal law of semantic distortion. In particular, the recovered meaning $\hat{M}$ results from multiple stages, including wireless transmission, decoding, and semantic inference, each of which may contribute a small perturbation. When these disturbances are aggregated and the semantic estimator is locally linearized around M, the resulting error can be approximated as Gaussian by a central-limit-type argument (Samarathunga et al., 2025). This modeling choice is also consistent with probabilistic and generative semantic communication frameworks based on latent-variable, VQ-VAE, and diffusion-based representations, while preserving analytical tractability (Beck et al., 2023; Hu et al., 2023; Guo et al., 2024).

Theorem 4.1 (False-rejection probability under isotropic Gaussian perturbations)**.** *Suppose that the semantic perturbation satisfies $W \sim \mathcal{N}(0, \Sigma)$, and assume that both the perturbation covariance and the Fisher information matrix are isotropic:*

$$\Sigma = \sigma^2 I, \qquad J(M) = \kappa I, \tag{14}$$

where $\sigma^2 > 0$ is the perturbation variance per semantic dimension and $\kappa > 0$ is the Fisher information per semantic dimension. Then the local quadratic semantic discrepancy satisfies

$$Q := \frac{1}{2} W^\top J(M) W = \frac{\kappa\sigma^2}{2} \sum_{i=1}^{d} Z_i^2 = \frac{\kappa\sigma^2}{2} \chi_d^2, \tag{15}$$

where $Z_i = W_i/\sigma \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$ for $i = 1, \ldots, d$, and $\chi_d^2 = \sum_{i=1}^d Z_i^2$ is a chi-squared random variable with d degrees of freedom. Here, d denotes the dimension of the semantic meaning representation.

Substituting (15) into (13), the false-rejection probability is approximated as

$$P_{\mathrm{FR}} \approx \Pr\left(\frac{\kappa\sigma^2}{2}\chi^2_d > \epsilon\right) \tag{16}$$

$$= \frac{\Gamma\left(\frac{d}{2}, \frac{\epsilon}{\kappa\sigma^2}\right)}{\Gamma\left(\frac{d}{2}\right)}, \tag{17}$$

where $\Gamma(s,a) = \int_a^\infty t^{s-1}e^{-t}\,\mathrm{d}t$ *is the upper incomplete gamma function, and* $\Gamma(s) = \Gamma(s,0)$ *is the standard gamma function.*

5. Numerical Results and Discussion

Figure 1 plots the false-rejection probability, P_{FR}, as a function of the semantic SNR, $\epsilon/(\kappa\sigma^2)$, for $d \in \{5, 10, 15\}$ under the isotropic baseline $\Sigma = \sigma^2 I_d$ and $J(M) = \kappa I_d$. Solid lines represent the closed-form expression in Theorem 4.1, while open circles denote Monte Carlo estimates. The analytical expression matches the simulation results across the full SNR range, validating the closed-form approximation.

The figure also shows that P_{FR} increases monotonically with the semantic dimension d. Under the isotropic model, $Q = \frac{\kappa\sigma^2}{2}\chi^2(d)$, so the mean semantic perturbation energy is $\mathbb{E}[Q] = \frac{d\kappa\sigma^2}{2}$, which grows linearly with d. By contrast, for a fixed semantic SNR, the acceptance threshold is $\epsilon = \mathrm{SNR} \times \kappa\sigma^2$, which does not scale with d. Hence, as d increases, the $\chi^2(d)$ distribution shifts to the right and places more probability mass above the fixed threshold. Equivalently, P_{FR} increases with d at any fixed semantic SNR. Thus, a higher-dimensional meaning space makes a fixed tolerance ϵ progressively more restrictive.

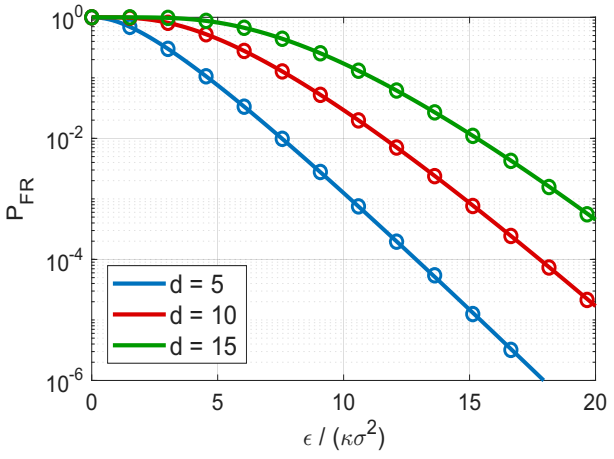

Figure 1. False-rejection probability versus $\epsilon/(\kappa\sigma^2)$ for different values of d under the isotropic baseline. Open circles show Monte Carlo estimates.

Figure 2 plots P_{FR} versus the semantic SNR, $\epsilon/(\kappa\sigma^2)$, for four configurations of the perturbation covariance Σ and the Fisher information matrix $J(M)$, all at semantic dimension $d = 10$.

Simulation setup. The anisotropic eigenvalues are generated as follows. For the Fisher-information vector, d values are first drawn independently from $\mathrm{Uniform}(0.7\kappa,\ 1.3\kappa)$ and then rescaled so that their empirical mean equals κ exactly:

$$\kappa_i = \tilde{\kappa}_i\kappa \Big/ \left(\frac{1}{d}\sum_{j=1}^{d}\tilde{\kappa}_j\right),\ \tilde{\kappa}_j \overset{\text{i.i.d.}}{\sim} \mathrm{Uniform}(0.7\kappa,\ 1.3\kappa). \tag{18}$$

This normalization enforces $\frac{1}{d}\sum_{i=1}^{d}\kappa_i = \kappa$. The perturbation variances σ_i^2 are generated in the same manner and independently of κ_i, yielding $\frac{1}{d}\sum_{i=1}^{d}\sigma_i^2 = \sigma^2$. These constraints ensure that all four configurations have the same total Fisher information, $\sum_{i=1}^{d}\kappa_i = d\kappa$, and the same total perturbation power, $\sum_{i=1}^{d}\sigma_i^2 = d\sigma^2$. Therefore, differences in P_{FR} are caused only by how noise variance and semantic sensitivity are distributed across the semantic dimensions.

Curve-by-curve description. The *blue solid* curve is the Monte Carlo estimate under the isotropic baseline $\Sigma = \sigma^2 I_d$ and $J(M) = \kappa I_d$. The *orange dashed* curve introduces anisotropy only in Σ while keeping $J(M) = \kappa I_d$. The *green dash-dot* curve introduces anisotropy only in $J(M)$ while keeping $\Sigma = \sigma^2 I_d$. These two curves are nearly identical to each other and remain close to the isotropic baseline at low semantic SNR.

The *gray dotted* curve, where both Σ and $J(M)$ are anisotropic, is the clear outlier. It separates from the other curves beyond approximately $\mathrm{SNR} \approx 10$ and ends roughly one order of magnitude above the isotropic baseline at $\mathrm{SNR} = 20$. This behavior indicates that the closed-form expression in Theorem 4.1 acts as an optimistic lower bound when the system is anisotropic, particularly at high semantic SNR. For semantic SNR below approximately 10, however, the isotropic analytical expression still provides a reasonably accurate estimate of the false-rejection probability.

Physical interpretation. The results show that perturbations along directions with high Fisher information are the most damaging. In these directions, the description-length-based semantic distance is highly sensitive, so a large perturbation component W_i produces a disproportionately large increase in the quadratic form Q. When Σ and $J(M)$ are both anisotropic, there is a nonzero probability that a high-variance noise direction aligns with a high-Fisher-information direction. When such alignment occurs, Q can spike far above its mean, substantially increasing the probability of false rejection. The isotropic closed-form formula does not capture this alignment risk, which is why it becomes increasingly optimistic when both the perturbation covariance and the Fisher information matrix are anisotropic.

Design implication. These observations suggest that the semantic encoder should avoid placing high-variance semantic perturbations in directions where the authentication rule is highly sensitive. Ideally, the high-variance directions of the perturbation covariance Σ should be aligned with the low-Fisher-information directions of $J(M)$. Such anti-alignment reduces the effective tail weight of $Q = \frac{1}{2}\sum_{i=1}^{d} \kappa_i \sigma_i^2 Z_i^2$, and helps keep P_{FR} close to the isotropic analytical bound even when the system is anisotropic. Concretely, this suggests jointly optimizing the semantic encoder E and the tolerance ϵ so that the principal axes of Σ are aligned with the smallest-eigenvalue directions of $J(M)$.

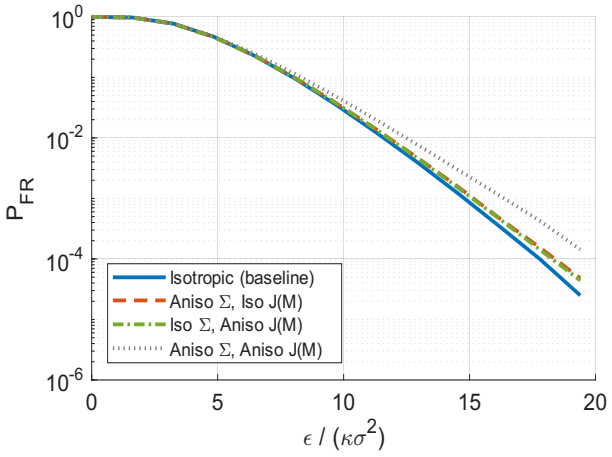

Figure 2. False-rejection probability versus $\epsilon/(\kappa\sigma^2)$ for four configurations of the perturbation covariance Σ and the Fisher information matrix $J(M)$, all at semantic dimension $d = 10$.

6. Conclusion

This paper proposed semantic authentication over equivalence classes for AI-native NextG wireless systems, where a transmission is accepted whenever the inferred meaning is sufficiently close to the intended meaning. Using a description-length-based semantic distance to define admissible semantic regions, we developed a set-based authentication rule with an invariant semantic hash compatible with semantic compression, learned inference, and generative reconstruction.

We analyzed robustness through the false-rejection probability P_{FR}. Under isotropic Gaussian perturbations, a closed-form expression was derived and validated by Monte Carlo simulation. Numerical results show that P_{FR} increases with semantic dimension d, and that anisotropy in either Σ or $J(M)$ raises P_{FR} above the isotropic bound, with simultaneous anisotropy in both producing an approximately one-decade gap at high semantic SNR. The isotropic formula remains accurate for $\epsilon/(\kappa\sigma^2) \lesssim 10$, motivating a design principle: aligning high-variance perturbation directions with low-sensitivity directions of $J(M)$ minimizes $\sum_i \kappa_i^2 \sigma_i^4$ and keeps P_{FR} close to the analytical bound in anisotropic deployments.

References

Beck, A. et al. Semantic recovery in generative communication systems. In *Proceedings of the IEEE International Conference on Communications (ICC)*, 2023.

Chaccour, C., Saad, W., Debbah, M., Han, Z., and Poor, H. V. Less data, more knowledge: Building next-generation semantic communication networks. *IEEE Communications Surveys & Tutorials*, 27(1):37–76, 2024.

Grünwald, P. D. *The Minimum Description Length Principle*. MIT press, 2007.

Guo, Y. et al. Diffusion-driven semantic communication for AI-native wireless systems. *IEEE Journal on Selected Areas in Communications*, 2024.

Hu, Q. et al. Robust semantic communication via vector-quantized variational autoencoders. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*, 2023.

Li, P. and Aijaz, A. Task-oriented connectivity for networked robotics with generative ai and semantic communications. In *IEEE INFOCOM 2025-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 1–6. IEEE, 2025.

Pan, Y., Wang, Y., Guo, S., Yin, C., Li, R., Su, Z., and Wu, Y. Trustworthy semantic communication for vehicular networks: Challenges and solutions. *IEEE Vehicular Technology Magazine*, 2025.

Samarathunga, A. et al. Semantic communication under technical noise: A probabilistic framework. *IEEE Transactions on Communications*, 2025.

Strinati, E. C., Di Lorenzo, P., Sciancalepore, V., Aijaz, A., Kountouris, M., Gündüz, D., Popovski, P., Sana, M., Stavrou, P. A., Soret, B., et al. Goal-oriented and semantic communication in 6G AI-native networks: The 6G-goals approach. In *2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, pp. 1–6. IEEE, 2024.

Zhang, P., Niu, K., Liu, Y., Liang, Z., Ma, N., Xu, X., Xu, W., Sun, M., Liu, Y., Wang, X., et al. Way to build native AI-driven 6G air interface: Principles, roadmap, and outlook. *IEEE Transactions on Network Science and Engineering*, 13:3551–3565, 2025.

.**AUTHORERR: Missing \icmlcorrespondingauthor.**

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.