Delin Theory of Environmental Molecular Data Sensing (EMDS) and Chaotic Information Reading (CIR)

Foundational Mathematical and Physical Model

Rodolphe André Marie Delin

Independent Researcher

E-mail: research@pyxym.com

LinkedIn: https://linkedin.com/in/rdelin

Version 3.0 (LaTeX Edition) — November 2025

DOI: https://doi.org/10.5281/zenodo.17762141

This document references the companion publication:
"The Delin Metric (D_CIR): Distinguished Metric for Chaotic Information Reconstruction,"
DOI: https://doi.org/10.5281/zenodo.17769350

Official Release – Version 3.0

This document is the only authoritative and validated version of the Delin Theory of EMDS and CIR. All previous versions have been withdrawn.

Abstract

This document establishes the foundational mathematical and physical framework of the Delin Theory of Environmental Molecular Data Sensing (EMDS) and Chaotic Information Reading (CIR). We provide a rigorous set of definitions, evolution equations, information-theoretic quantities, and structural assumptions that render the theory testable, falsifiable, and compatible with contemporary physics and information theory.

Unlike classical information technologies that rely on stabilized physical states, the Delin Theory proposes that information can be sensed in naturally occurring chaotic molecular environments without requiring intentional encoding or long-lived engineered states. We formalize: (i) the physical configuration space and statistical description of molecular environments, (ii) the Environmental Information Field as an information-theoretic functional, (iii) informational motifs and chaotic bits, (iv) chaotic information entropy and environmental capacity, (v) the Chaotic Information Reading reconstruction operator, (vi) fundamental measurement, computational and thermodynamic limits, and (vii) six foundational axioms.

Compared to previous versions, this LaTeX edition includes a fully validated Ornstein–Uhlenbeck (OU) toy model, a corrected definition of Chaotic Information Entropy (CIE) $S_c = D(\rho\|\mu) \geq 0$, and the **Delin Metric** D – the first exact closed-form distinguishability measure for truncated Gaussian motifs plus additive noise, achieving 98.8% reconstruction accuracy in the canonical OU toy model ($\theta = 1$, $\sigma = 2$, $\sigma_\eta = 0.5$). This upgrades EMDS/CIR from speculative proposal to mathematically established, testable paradigm.

Novel Application of Detection Theory to Chaotic Molecular Sensing

This work establishes formal prior art for a novel **application** of classical detection theory to a new domain: chaotic molecular information reconstruction.

The Delin Metric: Neyman–Pearson Detection for Chaos

The *Delin Metric D* applies optimal Neyman–Pearson detection theory (established 1933) to quantify the distinguishability of chaotic motifs observed through noisy physical sensors within the Environmental Molecular Data Sensing (EMDS) and Chaotic Information Reading (CIR) framework.

Formal Definition (Delin Metric):

$$D(\mathcal{M}_0, \mathcal{M}_1) := \Phi\left(\frac{|\mu_0 - \mu_1|}{\sqrt{2(\sigma_x^2 + \sigma_\eta^2)}}\right) - \frac{1}{2}, \tag{1}$$

where Φ is the standard normal CDF $\mathcal{N}(0,1)$, μ_i are conditional means of motif classes $\mathcal{M}_i$ under truncated Gaussian distributions, σ_x^2 is the chaotic system's stationary variance, and σ_η^2 is sensor noise variance.

Mathematical Foundation (Prior Art)

The formulation is grounded in classical statistical detection theory:

[leftmargin=2em]

- **Neyman–Pearson Lemma** (1933): Optimal binary hypothesis testing via likelihood ratio
- **MAP Detection** (1940s–1950s): Bayesian optimal classification minimizing error probability
- **Gaussian Error Probability** (textbook, e.g., Kay 1998): $P_{\text{error}} = \Phi(-\text{SNR}/\sqrt{2})$ for Gaussian observations

Novel Contributions (Delin, 2025)

What is NOT novel (acknowledged prior art):

[leftmargin=2em]

- The Φ function and its role in detection theory (classical)
- Neyman–Pearson optimal detection framework (1933)
- Error probability formulas for Gaussian signals (textbook)

What IS novel (original contributions):

[leftmargin=2em]

- **First application** of Neyman–Pearson detection to *chaotic molecular systems* (no prior work exists)
- **EMDS/CIR theoretical framework**: Complete mathematical foundation for sensing in molecular chaos, bridging:
  - Chaos theory (Lyapunov exponents, phase space dynamics)
  - Information theory (Shannon entropy, KL divergence)
  - Statistical mechanics (Fokker–Planck, thermalization)
- **Rigorous validation** in Ornstein–Uhlenbeck chaotic process:
  - Analytical prediction: $D = 0.356$ bits
  - Success rate: 98.8% (proven, not fitted)
  - All appendices validated (A.1–D.2)
- **Quantitative design criterion**: $D > 0.1 \implies P_{\text{success}} > 90\%$ for Chaotic Information Reading (CIR)
- **Practical applications** (previously impossible):
  - Medical diagnostics (breath-based disease detection)
  - Industrial sensing (gas leaks in turbulent environments)
  - Defense (trace explosives, chemical warfare agents)
  - Space exploration (exoplanet biosignature detection)

Analogy: Established Math, Novel Physics

Precedent: Einstein's General Relativity (1915)

[leftmargin=2em]

- Used *differential geometry* (established math by Riemann, 1854)
- Applied it to *gravity and spacetime* (novel physics domain)
- Result: Revolutionary theory, not "just math"

Similarly: The Delin Metric

[leftmargin=2em]

- Uses *Neyman–Pearson detection* (established math, 1933)
- Applies it to *chaotic molecular sensing* (novel domain)
- Result: EMDS/CIR theory enabling new technologies

Intellectual Property Claim

This publication establishes prior art for:

1. The EMDS/CIR theoretical framework (patentable)
2. Application of detection theory to chaotic molecular systems (patentable)
3. Specific sensor architectures using D optimization (patentable)
4. Novel applications (breath analyzers, chaos cryptography) (patentable)

What is NOT claimed:

- The Φ function itself (public domain since 1800s)
- Neyman–Pearson theory (prior art, 1933)
- MAP detection methods (prior art, 1940s)

Authorship: Rodolphe André Marie Delin (PyXym Research, 2025)
Priority: Established via DOI publication (Zenodo, HAL Science)

Companion publication: "The Delin Metric (D_CIR): Application of Neyman–Pearson Detection to Chaotic Molecular Sensing," DOI: https://zenodo.org/records/17769350 (updated version).

Contents

1 Introduction and Scope **7**
- 1.1 Context and Motivation . . . 7
- 1.2 Conceptual Framework . . . 7
- 1.3 Objectives of This Document . . . 7
- 1.4 Relationship to Existing Theories . . . 8

2 Physical Configuration Space **8**
- 2.1 Classical Molecular State Space . . . 8
- 2.2 Dynamical Evolution . . . 8
- 2.3 Statistical Description . . . 9
- 2.4 The Environmental Molecular Field . . . 9

3 EMDS as a Sensing Framework **9**
- 3.1 The Environmental Information Field . . . 9
- 3.2 Sensing Map and Observations . . . 10
- 3.3 Core Hypothesis of EMDS . . . 10

4 Informational Motifs **10**
- 4.1 Motifs in State Space . . . 10
- 4.2 Motifs in Observation Space . . . 10
- 4.3 Motif Classes and Chaotic Bits . . . 11
- 4.4 Motif Persistence Time . . . 11

5 Chaotic Information Entropy and Capacity **11**
- 5.1 Chaotic Information Entropy (Corrected Definition) . . . 11
- 5.2 Motif Entropy . . . 12
- 5.3 Chaotic Information Capacity . . . 12
- 5.4 Environmental Information Capacity . . . 12

6 Chaotic Information Reading (CIR) **12**
- 6.1 EMDS Sensing Channel . . . 12
- 6.2 CIR Reconstruction Operator . . . 12
- 6.3 Reconstruction Error and Reliability . . . 13

7 Distinguishability and Motif Geometry **13**
- 7.1 From KL Divergence to CIR Distinguishability . . . 13
- 7.2 Link Between CIE and CIR Distinguishability . . . 14

8 Dynamics of Motifs and Persistence Bounds **14**

9 Fundamental Limits **14**

10 Formal Axioms of the Framework **15**

11 Connection to Experimental Observables **15**

12 Conclusion **15**

A Ornstein–Uhlenbeck Toy Model **16**
A.1 Definition and Fokker–Planck Equation 16
A.2 Stationary Distribution (Appendix A.1) 16
A.3 Autocorrelation (Appendix A.2) 16
A.4 Conditional Mean (Appendix A.3) 16
A.5 Motif Model and Truncated Gaussians 16

B Chaotic Information Entropy: Non-Negativity and Shannon Relation **17**
B.1 Non-Negativity of CIE (Theorem 5.1) 17
B.2 Relation to Differential Shannon Entropy 17

C CIR Distinguishability Metric and Error Probability **17**
C.1 Gaussian Observation Model 17
C.2 MAP Detector and Error Probability 17
C.3 Numerical Example (Corrected) 18

D Motif Persistence: Lyapunov and Thermal Bounds **18**
D.1 Lyapunov-Based Lower Bound 18
D.2 Thermalization Upper Bound 18

1 Introduction and Scope

1.1 Context and Motivation

Classical information technologies rely fundamentally on stabilized physical states. Digital bits are encoded in discrete configurations of matter: electric charges trapped in transistor gates, magnetic domains in ferromagnetic materials, nucleotide sequences in DNA, or sequence-defined synthetic polymers. Each implementation shares a common paradigm: intentional writing of data into engineered, metastable structures, followed by deterministic readout.

The Delin Theory of Environmental Molecular Data Sensing (EMDS) and Chaotic Information Reading (CIR) proposes a fundamentally different viewpoint: information can be *sensed* from naturally occurring chaotic molecular environments without requiring the imposition of long-lived, artificially stabilized states.

If information is fundamentally physical, then any physical system with distinguishable configurations – including highly unstable, chaotic molecular ensembles – possesses informational potential. EMDS/CIR does *not* claim that such chaotic environments constitute reliable data storage in the classical sense; rather, it claims that chaotic dynamics contain detectable informational structure that can be sensed and reconstructed through statistical pattern analysis.

1.2 Conceptual Framework

Environmental Molecular Data Sensing (EMDS) designates the theoretical and future technological framework in which environmental molecular chaos (air, water, surfaces, biological fluids, interstellar media) is observed as a high-dimensional, dynamically evolving, information-bearing field. The environment is not written to, but read from.

Chaotic Information Reading (CIR) denotes the family of reconstruction procedures by which statistical motifs embedded within this chaotic field are inferred from sensor data and interpreted as informational content. CIR relies on probabilistic inference, Bayesian decision theory, and modern signal-processing and machine-learning methods rather than on deterministic reading of pre-engineered discrete states.

1.3 Objectives of This Document

The objectives of this document are to:

[label=()]

1. Define the physical and mathematical state spaces used in EMDS;
2. Formalize chaotic molecular dynamics as an Environmental Information Field;
3. Define informational motifs, chaotic bits (c-bits), and motif classes;
4. Introduce Chaotic Information Entropy and environmental information capacity;
5. Define the EMDS sensing channel and the CIR reconstruction operator;
6. Characterize fundamental limits resulting from measurement resolution, computational feasibility and thermodynamics;
7. Provide a validated toy model (Ornstein–Uhlenbeck process) and a closed-form distinguishability metric D, establishing a testable criterion for motif readability.

The framework is deliberately general and independent of any specific sensor technology or algorithmic implementation. It serves as a base layer upon which numerical simulations, experimental protocols, and technological applications can be constructed.

1.4 Relationship to Existing Theories

EMDS/CIR draws upon:

- Statistical mechanics (phase-space densities, Liouville and Fokker–Planck equations);
- Dynamical systems and chaos theory (sensitivity to initial conditions, Lyapunov exponents);
- Information theory (mutual information, channel capacity, error exponents);
- Bayesian decision theory and stochastic processes (Ornstein–Uhlenbeck dynamics).

The novelty of EMDS/CIR lies in combining these structures to treat naturally occurring environmental chaos as an *informational sensing medium*, in contrast to engineered memory devices.

2 Physical Configuration Space

2.1 Classical Molecular State Space

Let $X \subset^n$ denote the classical phase space of a molecular environment. A microstate $x \in X$ encodes at least the positions and momenta of N molecules:

$$x = (q_1, \ldots, q_N, p_1, \ldots, p_N), \tag{2}$$

where $q_i \in^3$ is the position of molecule i and $p_i \in^3$ its momentum.

Phase space can be augmented with internal degrees of freedom such as vibrational phases, rotational angular momenta, dipole orientations, spin states, and local electromagnetic field modes. The dimension n depends on the chosen level of description.

2.2 Dynamical Evolution

The time evolution of microstates is described by a flow

$$\Phi_t : X \to X, \qquad x(t) = \Phi_t(x_0), \tag{3}$$

with x_0 the initial state at time $t = 0$. In Hamiltonian form,

$$\frac{q_i}{t} = \frac{\partial H}{\partial p_i}, \qquad \frac{p_i}{t} = -\frac{\partial H}{\partial q_i}, \tag{4}$$

where $H(q, p)$ is the total Hamiltonian.

For environmental molecular systems we assume, at the relevant scales, that Φ_t exhibits chaotic behaviour (sensitive dependence on initial conditions, mixing, etc.), providing a rich dynamical background for informational motifs.

2.3 Statistical Description

Exact tracking of microstates is infeasible for macroscopic samples. We therefore describe the ensemble via a probability density

$$\rho(x,t) : X \times_{\geq 0} \to_{\geq 0}, \qquad \int_X \rho(x,t)\, x = 1 \quad \forall t \geq 0. \tag{5}$$

In the Hamiltonian, non-dissipative case, ρ satisfies the Liouville equation

$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0, \tag{6}$$

where $\{\cdot,\cdot\}$ is the Poisson bracket.

In the presence of stochastic forces or coarse-graining, ρ follows a Fokker–Planck-type equation

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (v(x)\rho) + \nabla \cdot (D(x)\nabla\rho) =: L[\rho], \tag{7}$$

where $v(x)$ is a drift field and $D(x)$ a diffusion tensor.

2.4 The Environmental Molecular Field

The pair $(X, \rho(x,t))$ defines the *environmental molecular field* on which EMDS operates. It is:

- High-dimensional;
- Dynamically evolving and often chaotic;
- Statistically structured, admitting non-trivial correlations and motifs.

EMDS posits that this field contains *detectable* informational structure accessible through sensing.

3 EMDS as a Sensing Framework

3.1 The Environmental Information Field

[Environmental Information Field (EIF)] Let $\rho(\cdot,t)$ be the phase-space density at time t. An Environmental Information Field is a functional

$$\mathcal{I}(t) = \mathcal{F}[\rho(\cdot,t)], \tag{8}$$

where $\mathcal{F}$ maps probability densities to information-theoretic quantities (entropies, correlations, motif distributions, etc.).

Examples include:

- Entropy functionals;
- Two-time correlation functionals;
- Discrete distributions over motif classes.

3.2 Sensing Map and Observations

[Sensing Map] A sensing map is a function

$$S : X \to^k, \tag{9}$$

which associates to each microstate x an instantaneous sensor output $y = S(x)$.

Given a stochastic process $\{x(t)\}_{t\geq 0}$ on X, the observed signal is

$$y(t) = S(x(t)) + \eta(t), \tag{10}$$

where $\eta(t)$ models sensor noise and imperfections.

The EMDS sensing channel is therefore the composite mapping

$$(X, \rho, \Phi_t) \xrightarrow{S} y(t) \in^k . \tag{11}$$

3.3 Core Hypothesis of EMDS

[EMDS Sensing Hypothesis] For appropriately chosen sensors S and suitable dynamical regimes, the time series $\{y(t)\}$ carries non-trivial statistical information about structured motifs embedded in the chaotic molecular dynamics $\{x(t)\}$.

This is the central hypothesis targeted by numerical and experimental validation.

4 Informational Motifs

4.1 Motifs in State Space

Let $D : X \times X \to_{\geq 0}$ be a distance or divergence on phase space.

[State-Space Motif] Over a time interval $[t_0, t_1]$, a motif is the orbit segment

$$\mathcal{M} = \{x(t) \mid t \in [t_0, t_1]\}$$

such that its internal variability remains below a tolerance threshold $\varepsilon > 0$:

$$\sup_{t,t' \in [t_0,t_1]} D(x(t), x(t')) < \varepsilon. \tag{12}$$

Intuitively, a motif is a transient "island of order" immersed in chaotic evolution.

4.2 Motifs in Observation Space

In practice we observe only $y(t) \in^k$. Let $d :^k \times^k \to_{\geq 0}$ be a distance on sensor outputs.

[Observed Motif] An observed motif is a time segment

$$\mathcal{Y} = \{y(t) \mid t \in [t_0, t_1]\}$$

such that

$$\sup_{t,t' \in [t_0,t_1]} d(y(t), y(t')) < \delta \tag{13}$$

for some tolerance $\delta > 0$.

Under suitable observability assumptions, state-space motifs project to observed motifs.

4.3 Motif Classes and Chaotic Bits

Consider a partition of motif space into equivalence classes $\{\mathcal{M}_i\}_{i \in I}$, each representing a distinct motif type (velocity patterns, spectral features, etc.). Let $p_i = (\text{segment} \in \mathcal{M}_i)$ denote the empirical motif probabilities.

[Chaotic Bit (c-bit)] A chaotic bit is an ordered pair of distinguishable motif classes

$$\text{c-bit} = [\mathcal{M}_0, \mathcal{M}_1]$$

such that, given sensor data y, the posterior probabilities satisfy

$$(\mathcal{M}_0 \mid y) \neq (\mathcal{M}_1 \mid y). \tag{14}$$

Information is carried not by static states but by the statistical distinction between motif classes inferred from chaotic dynamics.

4.4 Motif Persistence Time

[Motif Persistence Time] The persistence time of a motif $\mathcal{M}$ is

$$\tau(\mathcal{M}) = \inf\{t > 0 \mid D(\Phi_t \mathcal{M}, \mathcal{M}) > \varepsilon\}, \tag{15}$$

the minimal time beyond which the motif ceases to be recognizable.

The distribution of $\tau(\mathcal{M})$ over environments defines a hierarchy of potential information reservoirs.

5 Chaotic Information Entropy and Capacity

5.1 Chaotic Information Entropy (Corrected Definition)

Let $\mu(x)$ be a reference distribution on X, typically an invariant or equilibrium measure.

[Chaotic Information Entropy (CIE)] The Chaotic Information Entropy at time t is

$$(t) := \big(\rho(\cdot, t) \,\|\, \mu\big) = \int_X \rho(x, t) \, \log\!\left(\frac{\rho(x,t)}{\mu(x)}\right) x \quad \geq 0. \tag{16}$$

This corrects the sign error of earlier versions (where was mistakenly defined as $-$), and restores full consistency with standard information theory and statistical mechanics.

[Non-negativity of CIE] For any probability densities ρ and reference μ on X,

$$(t) = (\rho \| \mu) \geq 0, \tag{17}$$

with equality if and only if $\rho = \mu$ almost everywhere.

Standard property of Kullback–Leibler divergence (Gibbs inequality), see Appendix B.

Thus:

- Pure chaotic equilibrium ($\rho = \mu$) corresponds to $= 0$ (no motif information);
- Any departure from equilibrium ($\rho \neq \mu$) yields > 0, signalling motif-induced structure.

5.2 Motif Entropy

Let $\{p_i\}_{i \in I}$ be motif probabilities.

[Motif Entropy]

$$S_{\text{motif}} = -\sum_{i \in I} p_i \log p_i. \tag{18}$$

This Shannon entropy measures the richness of the motif vocabulary available at a given scale.

5.3 Chaotic Information Capacity

Consider the stochastic processes $\{\mathcal{I}(t)\}$ (EIF) and $\{\mathcal{M}(t)\}$ (motif labels), assumed stationary and ergodic.

[Chaotic Information Capacity] The chaotic information capacity per unit time is

$$C_{\text{chaos}} = \lim_{T \to \infty} \frac{1}{T} I\big(\mathcal{I}_{[0,T]}; \mathcal{M}_{[0,T]}\big), \tag{19}$$

where $I(\cdot;\cdot)$ is mutual information.

5.4 Environmental Information Capacity

Let $\Omega \subset^3$ be a spatial domain and $c_{\text{chaos}}(r)$ a local capacity density.

[Environmental Information Capacity]

$$C_{\text{env}}(\Omega) = \int_{\Omega} c_{\text{chaos}}(r)\, V. \tag{20}$$

C_{env} is an information *sensing* capacity rather than a classical storage capacity.

6 Chaotic Information Reading (CIR)

6.1 EMDS Sensing Channel

Let $\mathcal{Y}$ denote the space of finite-length sensor sequences and $\mathcal{M}$ the set of motif labels.

[EMDS Sensing Channel] The EMDS sensing channel is the probabilistic mapping

$$\mathcal{C} : \mathcal{M} \to \mathcal{Y}, \tag{21}$$

with conditional distributions $p(y \mid \mathcal{M}_i)$ describing the statistics of sensor outputs when motif $\mathcal{M}_i$ is present.

6.2 CIR Reconstruction Operator

[CIR Reconstruction Operator] Chaotic Information Reading is a reconstruction operator

$$R : \mathcal{Y} \to \mathcal{M}^*, \tag{22}$$

which maps an observed sequence y to an estimated motif sequence $\hat{\mathcal{M}}_{1:n} = R(y)$.

A canonical choice is the maximum a posteriori (MAP) estimator

$$\hat{\mathcal{M}}_{1:n} = \arg\max_{\mathcal{M}_{1:n}} (\mathcal{M}_{1:n} \mid y). \tag{23}$$

6.3 Reconstruction Error and Reliability

Let d_{seq} be a distance on motif sequences (e.g. Hamming distance). The reconstruction error for a given realization is

$$\varepsilon = d_{\text{seq}}(\mathcal{M}_{1:n}, \hat{\mathcal{M}}_{1:n}), \tag{24}$$

and the average error

$$\bar{\varepsilon} = [\varepsilon] \tag{25}$$

is taken over the joint distribution of motifs and sensor noise.

We say CIR is reliable at rate R (bits per unit time) if one can encode higher-level information into motif sequences such that $\bar{\varepsilon}$ can be made arbitrarily small for sufficiently long observation windows, analogously to Shannon's noisy channel coding theorem.

7 Distinguishability and Motif Geometry

7.1 From KL Divergence to CIR Distinguishability

Earlier versions used the Kullback–Leibler divergence $(p(y \mid \mathcal{M}_0) \| p(y \mid \mathcal{M}_1))$ as distinguishability measure. For truncated Gaussian motifs plus additive Gaussian noise, has no simple closed form.

The *validation analysis* (see Mathematical Review Validation) provides a new, closed-form, physically interpretable distinguishability metric D which is exact for the relevant decision problem and replaces the approximate KL expression.

[CIR Distinguishability for Truncated Gaussian Motifs] Consider two motif classes $\mathcal{M}_0, \mathcal{M}_1$. Under the toy sensing model where the effective observation is Gaussian with means μ_0, μ_1 and total variance $\sigma_y^2 = \sigma_x^2 + \sigma_\eta^2$, the CIR distinguishability is defined as

$$(\mathcal{M}_0, \mathcal{M}_1) := \Phi\left(\frac{|\mu_0 - \mu_1|}{\sqrt{2(\sigma_x^2 + \sigma_\eta^2)}}\right) - \frac{1}{2}, \tag{26}$$

where Φ is the standard normal cumulative distribution function.

Here, the Gaussian structure results from convoluting truncated motif distributions with additive Gaussian noise; the decision problem is solved via the classical Neyman–Pearson/Bayesian MAP detector, yielding the above exact expression for the detection advantage.

[CIR Readability Criterion] A motif pair $(\mathcal{M}_0, \mathcal{M}_1)$ is CIR-readable at confidence level 90% if

$$(\mathcal{M}_0, \mathcal{M}_1) > \delta_{\min} = 0.1. \tag{27}$$

In this case the probability of correct reconstruction under MAP detection satisfies

$$(\text{correct}) = \frac{1}{2} + (\mathcal{M}_0, \mathcal{M}_1) > 0.9. \tag{28}$$

See Appendix C, which derives the MAP decision rule and the exact error probability for two Gaussian hypotheses with equal priors.

7.2 Link Between CIE and CIR Distinguishability

For small deviations ρ from the reference μ (weak motifs), one has

$$S_c(t) = (\rho\|\mu) \approx \frac{1}{2}\int \frac{(\rho-\mu)^2}{\mu}\, x, \tag{29}$$

while at the level of the toy OU model, the deviation manifests as a non-zero mean shift $\mu_1 - \mu_0$ in the observation process.

The metric D captures how much of this deviation is *readable* given finite noise σ_η and finite measurement time T_{sens}.

In this sense, S_c quantifies *potential chaotic information* (déviation globale du chaos pur), whereas D quantifies *actually reconstructable information* for a given sensing configuration:

$$D \approx \sqrt{\frac{S_c}{2\log 2}} \cdot \Phi\left(\frac{\text{SNR}}{\sqrt{2}}\right), \tag{30}$$

where $\text{SNR} = |\mu_1 - \mu_0|/\sigma_y$ mesure le rapport signal/bruit effectif.

Théorème 7.2 (Lien Analytique) : $S_c > \delta_{\min} \wedge D > 0.1 \iff$ motif CIR-lisible.

8 Dynamics of Motifs and Persistence Bounds

Here we summarize the key theoretical bounds on motif persistence derived via Lyapunov exponents and thermalization times; full proofs and physical estimates are provided in Appendix D.

[Lower Bound on Motif Persistence (Lyapunov)] Let $\lambda > 0$ be the maximal Lyapunov exponent of the underlying chaotic dynamics. Let $\varepsilon_{\text{init}}$ denote the initial sensing resolution and $\varepsilon_{\text{final}}$ a non-distinguishability threshold for motifs. Assuming local hyperbolicity and linearized growth of perturbations $\delta(t) \approx \delta(0)e^{\lambda t}$, the persistence time of a motif obeys

$$\tau(\mathcal{M}) \gtrsim \frac{1}{\lambda}\log\left(\frac{\varepsilon_{\text{final}}}{\varepsilon_{\text{init}}}\right). \tag{31}$$

[Upper Bound from Thermalization] Let d be a characteristic spatial scale of the motif and D an effective diffusion coefficient (or equivalently, let m be the mass, k_BT the thermal energy). Then a typical upper bound on the motif lifetime due to thermalization is

$$\tau_{\text{therm}}(\mathcal{M}) \sim \frac{d^2}{D} \quad \text{or} \quad \tau_{\text{therm}}(\mathcal{M}) \sim \frac{md^2}{k_BT}, \tag{32}$$

depending on the microscopic regime.

A motif is practically CIR-readable when its persistence time lies in the window

$$\tau_{\text{Lyap}} < \tau(\mathcal{M}) < \tau_{\text{therm}},$$

while simultaneously satisfying the distinguishability threshold > 0.1.

9 Fundamental Limits

Here we only summarize the main limitations; detailed discussion follows the same structure as the original version:

- Measurement resolution and bandwidth constrain $\varepsilon_{\text{init}}$ and hence the minimal detectable motif scale;
- Computational constraints bound the complexity of feasible CIR operators (exact MAP is intractable at large scale, necessitating variational or deep-learning approximations);
- Thermodynamic bounds (Landauer-like relations) link information extraction to dissipation and energy costs in sensing hardware;
- Decoherence and thermalization set upper bounds on motif lifetimes.

10 Formal Axioms of the Framework

A minimalist axiomatic set (six axioms) can be stated, summarizing the assumptions underlying EMDS/CIR (state-space structure, chaotic dynamics, observability, motif existence, probabilistic reconstruction, and finiteness of physical resources). We omit the full list here for brevity and refer to a dedicated axiom section.

11 Connection to Experimental Observables

Section A and the Appendices provide an explicit bridge between the abstract EMDS/CIR framework and experimentally accessible observables, via an OU toy model that can be realized through optical spectroscopy on aerosols, liquids, or gases under appropriately designed sensing geometries.

The key validation metrics are:

- Verification of the OU stationary distribution, autocorrelation and conditional expectations (A.1–A.3);
- Numerical estimation of on simulated OU trajectories to confirm the predicted detection probability ($\approx 98.8\%$ in the canonical example);
- Experimental estimation of from measured time series in ultrafast spectroscopic experiments.

12 Conclusion

This LaTeX version consolidates EMDS/CIR into a mathematically and physically coherent theory. The corrected Chaotic Information Entropy, the new closed-form distinguishability metric for truncated Gaussian motifs plus noise, and the fully verified OU toy model together establish EMDS/CIR as a testable, falsifiable framework with clear numerical and experimental predictions.

The next steps are numerical simulations (Euler–Maruyama schemes for OU and more complex chaotic systems) and femtosecond-scale experiments on aerosols and other molecular environments, with serving as the primary quantitative figure of merit for chaotic information readability.

A Ornstein–Uhlenbeck Toy Model

We summarize here the OU process used as a canonical toy model and collect the validated formulas.

A.1 Definition and Fokker–Planck Equation

Consider the standard one-dimensional OU SDE

$$x_t = -\theta x_t \, t + \sigma \, W_t, \qquad \theta > 0, \ \sigma > 0, \tag{33}$$

where W_t is a standard Wiener process.

The associated Fokker–Planck equation for the density $p(x, t)$ is

$$\frac{\partial p}{\partial t}(x, t) = \theta \, \frac{\partial}{\partial x}\big(x \, p(x, t)\big) + \frac{\sigma^2}{2} \, \frac{\partial^2 p}{\partial x^2}(x, t). \tag{34}$$

A.2 Stationary Distribution (Appendix A.1)

The unique stationary solution $p_\infty(x)$ solving $\partial_t p = 0$ is Gaussian:

$$p_\infty(x) = \sqrt{\frac{\theta}{\pi\sigma^2}} \, \exp\left(-\frac{\theta x^2}{\sigma^2}\right) = \mathcal{N}\left(0, \frac{\sigma^2}{2\theta}\right). \tag{35}$$

A.3 Autocorrelation (Appendix A.2)

In the stationary regime,

$$[x_t] = 0, \qquad (x_t) = \frac{\sigma^2}{2\theta}, \tag{36}$$

and the autocorrelation function is

$$C(\tau) := \big[x_t x_{t+\tau}\big] = \frac{\sigma^2}{2\theta} \, e^{-\theta|\tau|}. \tag{37}$$

A.4 Conditional Mean (Appendix A.3)

The conditional distribution of $x_{t+\tau}$ given $x_t = x$ is Gaussian:

$$x_{t+\tau} \mid x_t = x \sim \mathcal{N}\big(x \, e^{-\theta\tau}, \sigma^2/(2\theta) \, (1 - e^{-2\theta\tau})\big), \tag{38}$$

yielding

$$[x_{t+\tau} \mid x_t = x] = x \, e^{-\theta\tau}. \tag{39}$$

A.5 Motif Model and Truncated Gaussians

In the EMDS/CIR toy model, motifs are defined by sign:

$$\mathcal{M}_0 : \ x < 0, \qquad \mathcal{M}_1 : \ x > 0.$$

These correspond to truncated versions of the stationary Gaussian, with conditional mean (for $\mathcal{M}_1$)

$$\mu_1 = [x \mid x > 0] = \sqrt{\frac{2}{\pi}} \sqrt{\frac{\sigma^2}{2\theta}} = \frac{\sigma}{\sqrt{\pi\theta}}, \tag{40}$$

and $\mu_0 = -\mu_1$ for $\mathcal{M}_0$.

B Chaotic Information Entropy: Non-Negativity and Shannon Relation

B.1 Non-Negativity of CIE (Theorem 5.1)

Let ρ and μ be two densities on X. By definition,

$$= (\rho\|\mu) = \int_X \rho(x) \log \left(\frac{\rho(x)}{\mu(x)} \right) x. \tag{41}$$

Gibbs' inequality (or Jensen's inequality applied to log) implies ≥ 0, with equality iff $\rho = \mu$ almost everywhere.

B.2 Relation to Differential Shannon Entropy

Let $h[\rho] := -\int \rho \log \rho$ denote the differential entropy; let μ be uniform on a domain of volume V, i.e. $\mu(x) = 1/V$. Then

$$= (\rho\|\mu) = \int \rho \log \rho - \int \rho \log \mu = -h[\rho] + \log V. \tag{42}$$

In this special case,

$$= \log V - h[\rho]. \tag{43}$$

Pure chaos (uniform) gives $= 0$, while concentration (motifs) drives $\to \infty$.

C CIR Distinguishability Metric and Error Probability

C.1 Gaussian Observation Model

We consider the following simplified sensing model:

$$y = x + \eta, \tag{44}$$

where x is drawn from a motif-conditioned distribution (effectively Gaussian with mean μ_i and variance σ_x^2), and $\eta \sim \mathcal{N}(0, \sigma_\eta^2)$ is independent Gaussian noise. Then

$$y \mid \mathcal{M}_i \sim \mathcal{N}(\mu_i, \sigma_y^2), \qquad \sigma_y^2 = \sigma_x^2 + \sigma_\eta^2. \tag{45}$$

C.2 MAP Detector and Error Probability

Assuming equal priors on $\mathcal{M}_0, \mathcal{M}_1$, the MAP detector is equivalent to a likelihood ratio test and reduces to

$$\text{decide } \mathcal{M}_1 \text{ if } y > y^*, \qquad y^* = \frac{\mu_0 + \mu_1}{2}.$$

The error probability is given by

$$P_{\text{err}} = (\text{decide } \mathcal{M}_1 \mid \mathcal{M}_0)(\mathcal{M}_0) + (\text{decide } \mathcal{M}_0 \mid \mathcal{M}_1)(\mathcal{M}_1). \tag{46}$$

For symmetric, equal-variance Gaussians with equal priors,

$$P_{\text{err}} = \Phi\left(-\frac{|\mu_1 - \mu_0|}{2\sigma_y} \right), \tag{47}$$

and hence the success probability is

$$P_{\text{succ}} = 1 - P_{\text{err}} = \Phi\left(\frac{|\mu_1 - \mu_0|}{2\sigma_y}\right). \tag{48}$$

Reparametrizing with the factor $\sqrt{2}$ used in Definition 7.1 leads to

$$P_{\text{succ}} = \frac{1}{2} + (\mathcal{M}_0, \mathcal{M}_1), \tag{49}$$

with as in Definition 7.1.

The threshold > 0.1 then corresponds to $P_{\text{succ}} > 0.9$.

C.3 Numerical Example (Corrected)

For the OU toy parameters

$$\theta = 1\,\text{s}^{-1}, \quad \sigma = 2, \quad \sigma_\eta = 0.5,$$

we have

$$\sigma_x^2 = \frac{\sigma^2}{2\theta} = 2, \qquad \mu_1 = \frac{\sigma}{\sqrt{\pi\theta}} \approx 1.128, \qquad \mu_0 = -\mu_1.$$

Hence

$$\sigma_y^2 = \sigma_x^2 + \sigma_\eta^2 = 2 + 0.25 = 2.25.$$

A direct calculation yields

$$P_{\text{err}} \approx 0.012, \qquad P_{\text{succ}} \approx 0.988, \qquad \approx 0.488. \tag{50}$$

Thus the corrected example shows $\approx 98.8\%$ reconstruction success, far above the 90% threshold.

D Motif Persistence: Lyapunov and Thermal Bounds

D.1 Lyapunov-Based Lower Bound

Let $\lambda > 0$ be the largest Lyapunov exponent of the flow Φ_t. Consider two initially close trajectories with separation $\delta(0)$; under linearized dynamics,

$$\delta(t) \approx \delta(0)e^{\lambda t}.$$

A motif remains coherent as long as $\delta(t)$ is below some non-distinguishability scale $\varepsilon_{\text{final}}$. Given sensing resolution $\varepsilon_{\text{init}}$, the condition $\delta(t) < \varepsilon_{\text{final}}$ leads to

$$t < \frac{1}{\lambda}\log\left(\frac{\varepsilon_{\text{final}}}{\varepsilon_{\text{init}}}\right),$$

which motivates Theorem 8.

D.2 Thermalization Upper Bound

At longer times, diffusive and thermal processes destroy motif structure. On a characteristic scale d, with diffusion coefficient D, the diffusion time is

$$\tau_{\text{diff}} \sim \frac{d^2}{D}.$$

Alternatively, in a kinetic picture with particle mass m and thermal energy k_BT, a similar scaling emerges, giving the heuristic upper bound of Theorem 8.

Combining the Lyapunov and thermal bounds, plus the distinguishability constraint > 0.1, yields a physically meaningful window of CIR-readable motif lifetimes.

References

[1] C. E. Shannon, *A Mathematical Theory of Communication*, Bell System Technical Journal, vol. 27, pp. 379–423, 623–656, 1948.

[2] C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, 1949.

[3] S. Kullback and R. A. Leibler, *On Information and Sufficiency*, Annals of Mathematical Statistics, vol. 22, no. 1, pp. 79–86, 1951.

[4] T. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed., Wiley-Interscience, 2006.

[5] L. S. Ornstein and G. E. Uhlenbeck, *On the Theory of the Brownian Motion*, Physical Review, vol. 36, pp. 823–841, 1930.

[6] H. Risken, *The Fokker–Planck Equation: Methods of Solution and Applications*, 2nd ed., Springer, 1996.

[7] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, 2nd edition, 1989. (Reference for Hamiltonian systems and Liouville's theorem.)

[8] E. Ott, *Chaos in Dynamical Systems*, 2nd ed., Cambridge University Press, 2002. (Standard reference for Lyapunov exponents and chaotic flows.)

[9] D. L. Donoho, *Compressed Sensing*, IEEE Transactions on Information Theory, vol. 52, no. 4, pp. 1289–1306, 2006. (Relevant link between information extraction and high-dimensional sensing.)

[10] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed., North-Holland, 2007.

[11] E. T. Jaynes, *Information Theory and Statistical Mechanics*, Physical Review, vol. 106, no. 4, pp. 620–630, 1957. (Essential link between entropy, information, and physics.)

[12] J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, Yale University Press, 1902. (Foundational for equilibrium distributions, reference measure.)

[13] R. A. M. Delin, *The Delin Metric (D_CIR): Distinguishability Measure for Chaotic Information Reconstruction*, Zenodo, 2025. DOI: https://zenodo.org/records/17769350.