Semantic Relaxation Networks

Language Generation as Constraint Stabilization in Admissibility Space

Flyxion

Independent Researcher

2026

Abstract

The development of language models over the past decade traces a compression sequence that has come to feel inevitable: language reduced to sequence, sequence reduced to tokens, tokens reduced to conditional probability distributions, and conditional probability treated as a proxy for cognition itself. This compression was not an arbitrary choice. It was a principled engineering projection, one that captured genuine statistical structure in linguistic data and enabled an extraordinary range of applications. But projection always involves loss, and the loss here has been systematic: what autoregressive factorization discards is precisely the distributed, recursive, non-sequential character of semantic structure.

We argue that the sequential decomposition $p(x) = \prod_i p(x_i \mid x_{<i})$ is computationally convenient rather than cognitively fundamental. Current language models—including architectures that introduce local parallelism through diffusion or adaptive tokenization through byte-level patching—preserve the assumption that text is a sequential object recovered left-to-right through iterative approximation. This assumption, we contend, is not a feature of language as a cognitive phenomenon. It is an artefact of treating communicative traces as cognitive substrate.

We propose Semantic Relaxation Networks (SRNs), a theoretical framework in which language generation emerges from recursive stabilization over a partially ordered admissibility manifold rather than from conditional next-token prediction. In SRNs, a global tension functional $\Phi(X)$ replaces the likelihood objective, semantic configurations evolve through recursive relaxation operators $\mathcal{R}$ driven by constraint gradients, and the interaction topology $\mathcal{C}_t$ itself reorganizes dynamically in response to evolving admissibility structure. Tokens are not primitive units consumed and emitted by the model. They are *emergent crystallizations*: local regions of the admissibility manifold that have stabilized sufficiently to collapse into communicable symbolic form.

Sequential language, on this account, is not the substrate of cognition. It is the observable residue of deeper stabilization dynamics operating over high-dimensional semantic constraint fields.

Contents

1. Introduction: Sequentiality as Compression Assumption

1.1. The Compression Chain

There is a genealogy of reductions underlying contemporary language modelling that is rarely made fully explicit. It runs as follows.

Language, in its cognitive and communicative sense, is a high-dimensional structured activity involving distributed semantic commitments, inferential dependencies, prosodic contour, pragmatic context, and social embedding. This activity is fundamentally non-linear: meaning is assembled, revised, anticipated, and recursively stabilized across timescales and registers that do not reduce to left-to-right symbol emission.

The first compression identifies language with *sequence*. Written text is sequential; the temporal unfolding of speech is sequential; it is therefore tempting—and technically tractable—to model language as a sequence of discrete symbols. This move is ancient in formal linguistics and natural language processing alike. Its power is undeniable. Its cost is that the relational, distributed, and recursive structure of semantic coherence must be reconstructed from positional proximity alone.

The second compression identifies sequence with *tokens*. Sequences must be discretized to be processed by finite machines. Whether the unit of discretization is a word, a subword piece, a byte, or a patch of bytes, the operational assumption is that meaning can be adequately represented by a finite vocabulary of symbols arrayed in linear order.

The third compression identifies tokens with *conditional probabilities*. The chain rule of probability yields

$$p(x_1, x_2, \dots, x_n) \;=\; \prod_{i=1}^{n} p(x_i \mid x_{<i}) \,,$$

an exact factorization of any joint distribution over sequences. The autoregressive language model implements this factorization as its generative mechanism: at each position, predict the next token given all previous tokens. This is not merely a modelling convention. It becomes an architectural commitment, shaping every aspect of training, inference, and evaluation.

The fourth and most consequential compression identifies conditional prediction with *cognition itself*. If a system that predicts the next token with sufficient accuracy also translates, reasons, writes code, and answers questions, the temptation arises to conclude that intelligence simply is high-quality conditional prediction. This inference is not obviously wrong. But it conflates the communicative trace of cognition—sequential language—with the underlying cognitive processes that generate and comprehend it.

1.2. Partial Destabilizations

Contemporary architectures have already begun, implicitly, to destabilize the most rigid layers of this compression.

Denoising diffusion models applied to language [5, 7, 8] replace left-to-right generation with iterative denoising: a corrupted sequence is incrementally restored toward a coherent target. Diffusion weakens strict causal ordering and introduces a form of parallel structure. But it preserves the fundamental assumption that generation is *sequence recovery*: there is always an underlying ordered text toward which the model converges. The topology of interaction—which positions attend to which—remains essentially sequence-indexed.

Block diffusion [9] interpolates between autoregressive and diffusion regimes, allowing parallel generation within local windows while preserving sequential ordering between windows. This is a meaningful relaxation, but sequentiality is merely decomposed rather than questioned.

The Byte Latent Transformer [10] and its faster variant [11] challenge the tokenization layer by operating at byte level with dynamic entropy-based patching. High-entropy byte sequences receive finer-grained attention; redundant regions are compressed into larger patches. BLT thereby weakens the fixed-vocabulary assumption and introduces content-adaptive processing granularity. Yet the architecture still generates text as a sequential output, proceeding through the sequence with a modified notion of step size.

These developments are not failures. They represent genuine conceptual progress: each one loosens a different constraint in the compression chain. Diffusion loosens strict ordering. Block diffusion loosens strict left-to-right causal dependency. BLT loosens fixed tokenization.

But none of them questions whether *sequentiality itself* is the right ontology for language generation. They are, in the terminology we will develop, *transitional architectures*: they operate within admissibility geometry without yet recognizing that geometry as their proper domain.

1.3. Thesis

This paper proposes a more fundamental reorientation.

We argue that sequential text is not the primitive substrate of language generation but its *observational residue*: a low-dimensional projection of higher-dimensional stabilization processes operating over semantic constraint fields. A language model that generates text by predicting the next token is performing an operation analogous to reading a shadow and inferring the object that cast it—accurate within limits, but ontologically inverted.

The framework we develop, Semantic Relaxation Networks, replaces the next-token prediction objective with a *recursive stabilization dynamics* over a partially ordered admissibility manifold. Generation is not emission. It is the collapse of locally stabilized regions of semantic configuration space into communicable symbolic form.

The central distinction organizing the paper is this:

Transformers predict.	**SRNs stabilize.**
Prediction is local and sequential.	Stabilization is distributed and recursive.
The unit is the token.	The unit is the admissibility region.
Topology is fixed by sequence.	Topology reorganizes under constraint.
Hallucination is misassignment.	Hallucination is premature collapse.

The paper proceeds as follows. Section 2 identifies the structural failure modes of sequential ontology more precisely. Section 3 defines admissibility spaces and their formal properties. Section 4 develops the recursive relaxation dynamics and companion equations. Section 5 introduces dynamic semantic topology formation, the key architectural departure from transformers. Section 6 treats stabilization, verification operators, and the reinterpretation of hallucination. Section 7 presents the inversion: emergent sequentiality as projection residue. Section 9 situates SRNs within the RSVP, TARTAN, and CLIO frameworks. Section 10 draws implications for AI, cognitive science, and physics. Section 13 concludes.

1.4. The SRN Object Hierarchy

Before the mathematics becomes dense, we introduce the five primary formal objects of the SRN framework and fix their roles explicitly. The single largest source of conceptual ambiguity in constraint-based generation proposals is the conflation of *dynamical* objects (those that evolve during relaxation) with *residual* objects (those that persist after stabilization is complete). The following definition and table separate these cleanly.

Definition 1.1 (SRN Object Hierarchy)**.** The SRN framework is organized around five primary formal objects:

(O1) The *semantic configuration* $X_t \in \mathcal{X}$: the current state of the partially stabilized semantic field at relaxation step t.

(O2) The *interaction topology* $\mathcal{C}_t$: the weighted directed graph encoding current coupling strengths between semantic regions, evolving under the topology operator Γ.

(O3) The *stabilized dependency structure* $G(X^*)$: the directed acyclic graph of communicative precedence constraints residual in the fully stabilized configuration X^*; a static object extracted once stabilization is complete.

(O4) The *verification hierarchy* $\mathcal{V} = \{V_k\}$: the multi-scale family of admissibility checking operators that validate local stabilizations against global constraints and activate rollback when violations are detected.

(O5) The *communicative pressure field* $\mathcal{P}$: the weighted preference function over ordered region pairs encoding grammatical, information-structural, and prosodic linearization conventions; applied during projection, not during relaxation.

Object	**Notation**	**Role**	**Type**	**Section**
Semantic configuration	X_t	Current field state	Dynamical	§3
Interaction topology	$\mathcal{C}_t$	Coupling graph	Dynamical	§5
Stabilized DAG	$G(X^*)$	Dependency skeleton	Residual	§7
Verification hierarchy	$\mathcal{V}$	Admissibility checking	Dynamical	§6
Pressure field	$\mathcal{P}$	Linearization preferences	Static	§11

The key distinction is between $\mathcal{C}_t$ and $G(X^*)$. The interaction topology $\mathcal{C}_t$ reorganizes continuously during relaxation under the operator Γ in response to constraint gradients; it is a computational scaffold, not a semantic claim. The stabilized DAG $G(X^*)$ is extracted once from the converged configuration and encodes the genuine semantic dependency structure of the output; it is the object over which projection operates. Conflating these two objects—treating the relaxation graph as identical to the dependency structure—is the source of most misreadings of constraint-based generation proposals.

2. The Failure of Sequential Ontology

2.1. What Autoregressive Factorization Discards

The autoregressive factorization

$$p(x) \;=\; \prod_{i=1}^{n} p(x_i \mid x_{<i})$$

is mathematically exact as a decomposition of any joint distribution over sequences. This is a theorem, not an assumption. But the *implementation* of this factorization as a generative mechanism involves choices that are not forced by the mathematics.

The most consequential such choice is the *Markovian approximation*: the joint context $x_{<i}$ is compressed into a finite-dimensional representation by the model. For transformer architectures [1], this representation is the softmax-weighted combination of previous-position key-value pairs. The attention mechanism is extraordinarily expressive, but it

operates over a fixed interaction structure determined by sequence position. Two tokens interact in proportion to their attention weight, which is a function of their content and their relative position—not of their semantic compatibility or constraint relationship.

The result is a subtle but structurally important limitation. A transformer's interaction graph is *position-indexed*. Semantic coherence must be inferred *through* positional relationships rather than being represented *directly*. Long-range semantic dependencies are, in principle, representable through attention; in practice they compete with local positional patterns that dominate the attention distribution.

More fundamentally, the autoregressive framework forces a *strict temporal ordering* on the generative process. Each token is produced after all preceding tokens are fixed. This means the model cannot revise a token once emitted, cannot simultaneously stabilize distant regions of the output, and cannot represent genuine semantic feedback between non-adjacent regions during generation.

These are not merely engineering limitations to be overcome with better hardware. They are *ontological* commitments built into the factorization itself.

2.2. Hallucination as Structural Symptom

The phenomenon of hallucination—the confident generation of semantically plausible but factually or inferentially incorrect text—is typically treated as a failure mode to be mitigated through better training, retrieval augmentation, or output verification.

Under sequential ontology, hallucination has a straightforward description: the model assigned high conditional probability to an incorrect continuation. The remedy is therefore to improve the calibration of the conditional distribution.

We propose that this description, while locally accurate, misidentifies the structural source of the failure.

Hallucination is not primarily a calibration problem. It is a *topological instability event*: a premature local stabilization that fails to maintain global admissibility.

The autoregressive model commits to each token irrevocably. Once a false commitment is made, subsequent tokens are conditioned on a flawed context, and the model's only recourse is to continue generating text that is locally coherent with the erroneous token rather than globally coherent with the semantic task. Sequential generation has no natural mechanism for detecting that a locally plausible choice has produced a globally inadmissible configuration.

This is not a failure of attention or memory. It is a consequence of treating generation as irreversible sequential emission rather than as recursive stabilization with rollback capacity.

2.3. Cognition Is Not Sequential

Perhaps the deepest problem with sequential ontology is the implicit claim it makes about cognition.

Human language comprehension and production are not sequential processes at the level of cognitive implementation. Decades of psycholinguistic research establish that comprehension involves massive parallelism, predictive processing, and global constraint satisfaction [16]. Readers do not process text token-by-token, assigning conditional probabilities to each word given its predecessors. They construct meaning through distributed inference across multiple linguistic levels simultaneously, frequently revising interpretations in light of later information.

Production, similarly, involves hierarchical planning across timescales that do not reduce to incremental token emission [18, 17]. The sequential character of speech and writing is a *interface constraint*—the communication channel is one-dimensional—not a property of the underlying generative process.

Free energy and predictive processing frameworks [14, 15] have proposed that neural computation is fundamentally a process of minimizing prediction error (or equivalently, minimizing variational free energy) across hierarchical generative models. These frameworks are explicitly non-sequential: the brain simultaneously maintains predictions at multiple levels of abstraction and updates them in parallel.

LeCun's world model framework [13] similarly argues that intelligence requires structured world models that support explicit planning and constraint satisfaction rather than purely reactive next-step prediction.

Autoregressive language models, for all their empirical power, adopt an architecture that is in tension with these converging insights from cognitive science and neuroscience. They are, in a precise sense, *communication channel models* rather than *cognitive substrate models*.

The following sections develop an alternative.

3. Admissibility Spaces and Semantic Persistence

3.1. Beyond Probability

The standard framework for language generation measures the quality of a sequence by its probability under a learned distribution. Generation maximizes (or samples from) $p(x)$. Evaluation metrics are cast in terms of perplexity, likelihood, or functions thereof.

We propose to replace probability as the primary object with a different notion: *admissibility*.

Admissibility is not equivalent to likelihood, though the two are related. A highly probable sequence may be inadmissible: it may be locally fluent but globally incoherent, or fluent under current conditions but unable to survive recursive extension. Conversely, an admissible configuration may not be the most probable continuation at any given position; it may be a semantically structured but stylistically unusual configuration that maintains coherence under perturbation and extension.

The distinction is essential. Probability models ask: *does this continuation fit locally?* Admissibility asks: *can this configuration survive recursive future stabilization?*

This is a fundamentally different question, and answering it requires a fundamentally different formal apparatus.

3.2. Semantic Regions as Primitive Objects

Before defining admissibility space, we must specify what its elements are. The framework has so far referred freely to "sub-configurations," "regions," and "patches" of the semantic configuration X. These terms require a formal placeholder definition, since the whole apparatus of tension functionals, topology operators, and projection maps is defined over them.

Definition 3.1 (Semantic Region)**.** A *semantic region* X_i is a locally coherent subconfiguration of the semantic configuration space $\mathcal{X}$ satisfying:

(i) **Internal cohesion.** The internal constraint coupling of X_i—the mean edge weight among nodes within X_i under the current topology $\mathcal{C}_t$—exceeds its external coupling to all other regions over some relaxation interval $[t_0, t_0+\delta]$. Formally, $\bar{\omega}_i^{\text{int}}(t) > \bar{\omega}_i^{\text{ext}}(t)$ for $t \in [t_0, t_0+\delta]$.

(ii) **Semantic coherence.** The local instability measure $S_i(X) < \tau_{\text{region}}$ for a region-formation threshold $\tau_{\text{region}} \leq \tau$, meaning the region does not contain unresolved internal contradiction.

(iii) **Temporal persistence.** X_i maintains conditions (i) and (ii) across at least one relaxation step, distinguishing it from transient fluctuations in the constraint graph.

Remark 3.1. *This definition is deliberately operational rather than ontologically final. Semantic regions are not fixed objects: they form, merge, split, and dissolve as the relaxation dynamics evolve and the topology $\mathcal{C}_t$ reorganizes. A region that satisfies (i)–(iii) at step t may fail to satisfy them at step $t+1$ if a topology reconfiguration increases its external coupling or introduces internal contradiction. The definition therefore characterizes regions as* dynamically stable patterns *in the* *constraint graph rather than as static constituents of a fixed parse.*

This flexibility is intentional. Natural semantic units—words, phrases, propositions, discourse segments—do not correspond to fixed-granularity objects. They are better understood as dynamically stable constraint clusters whose boundaries shift with context, attention, and communicative purpose. The SRN framework encodes this by treating regions as emergent features of the relaxation process rather than as primitive inputs.

3.3. Formal Definition

We introduce the central formal object of the paper.

Definition 3.2 (Admissibility Space)**.** Let $\mathcal{X}$ be the space of semantic configurations. An *admissibility space* $\mathcal{A} \subseteq \mathcal{X}$ is a dynamically evolving, partially ordered persistence structure on $\mathcal{X}$ such that each element $X \in \mathcal{A}$ satisfies the following three conditions:

(C1) **Local Constraint Coherence.** For every sub-configuration $X_i \subseteq X$ and its evolving semantic neighborhood $\mathcal{N}_i \subseteq \mathcal{A}$, the local constraint tension $S_i(X) < \tau$ for some coherence threshold $\tau > 0$. That is, X_i does not generate unresolvable local contradiction with its immediate semantic context.

(C2) **Recursive Extensibility.** For every admissible extension operator $\mathcal{E}$ consistent with the global constraint structure $\mathcal{C}$, there exists an admissible extension $X' = \mathcal{E}(X)$ such that $X' \in \mathcal{A}$ and the global tension functional $\Phi(X') \leq \Phi(X)$. That is, X can be continued without necessarily increasing total semantic tension.

(C3) **Perturbation Stability under Field Deformation.** X remains admissible under moderate reorganization of the semantic interaction topology $\mathcal{C}$. Formally, for perturbations $\delta\mathcal{C}$ below a topology deformation threshold ϵ, we require $X \in \mathcal{A}(\mathcal{C} + \delta\mathcal{C})$. Stability here is with respect to *topological*, not merely statistical, perturbation.

Interpretation 3.1. *Condition* (*C1*) *is the local analogue of syntactic and semantic well-formedness: the configuration does not generate immediate contradiction within its local semantic neighborhood. Condition* (*C2*) *is the more demanding and more philosophically novel condition: admissibility is not merely a property of the current state but of its* future persistence*. An admissible configuration must be one from which further coherent stabilization is possible. This separates admissibility from likelihood: a high-probability continuation may close off subsequent extension. Condition* (*C3*) *ensures that admissibility is robust to semantic topology reorganization—the model's ability to dynamically reconfigure interaction structure* (*Section 5*) *must not invalidate previously stabilized configurations.*

Remark 3.2. *The term "partially ordered" is important. Admissibility space is not a simple* *inclusion structure. Configurations are ordered by their degree of stabilization, their depth of recursive extensibility, and their robustness to perturbation. This ordering is partial because not all configurations are comparable: two configurations may each be admissible without either being more admissible than the other. The space is therefore a directed partially ordered set (a* poset*) rather than a total order.*

3.4. Admissibility and RSVP Accessibility

For readers familiar with the Relativistic Scalar-Vector Plenum (RSVP) framework, admissibility space admits a natural interpretation as the semantic analogue of accessibility structure under entropy-bounded persistence [21, cf.]. In RSVP terms, a configuration X is admissible when it lies within the accessible region of semantic field space consistent with maintaining low entropic tension under recursive extension.

This connection is developed more fully in Section 9. For the purposes of the present formal development, admissibility space should be understood on its own terms: as a persistence-theoretic object defined by the three conditions above, independent of any commitment to RSVP cosmology.

3.5. Admissibility, Truth, and External Verification

A possible misreading of the framework must be addressed directly. Admissibility, as defined through conditions (C1)–(C3), might appear to reduce to *self-consistency*: a configuration is admissible if it does not contradict itself and can be extended without self-contradiction. This would make SRNs glorified coherence engines—systems that produce fluent, internally consistent text without any anchor to truth or reality.

This reading is incorrect, but refuting it requires making explicit what the framework implies about false content.

A false configuration—one whose semantic commitments do not correspond to states of affairs in the world—may occupy a locally admissible basin. Internal coherence does not guarantee truth. A well-constructed fiction, a plausible lie, or a sophisticated confabulation may satisfy (C1) locally and even satisfy (C2) in the short term.

However, the *recursive extensibility* condition (C2) is more demanding than it first appears. Extensibility is not evaluated in isolation. It is evaluated under continued interaction with external verification fields: sensory input, interlocutor responses, documentary evidence, and logical inference from established commitments. These external fields function as additional constraint sources that are continuously integrated into the admissibility computation.

Principle 3.1 (Admissibility Under External Pressure)**.** Let $\mathcal{F}_{\text{ext}}$ denote the set of external verification fields (empirical, inferential, social) available during generation. A configuration X that is transiently admissible in isolation may become inadmissible when the global tension functional is augmented with external constraint terms:

$$\Phi_{\text{ext}}(X) \;=\; \Phi(X) \;+\; \mu \sum_k \Psi_k(X, \mathcal{F}_k),$$

where $\Psi_k(X, \mathcal{F}_k)$ measures the tension between configuration X and external field $\mathcal{F}_k$, and $\mu > 0$ weights external against internal constraint. False configurations generate high Ψ_k over time as their predictions conflict with field observations, progressively increasing Φ_{ext} and destabilizing the configuration.

Interpretation 3.2. *Principle 3.1 has several important consequences.*

First, hallucination and systematic falsehood are structurally distinct under SRN. *Hallucination is premature crystallization—a local collapse before global admissibility is verified (Principle 6.1). Systematic falsehood is a configuration that is locally stable in isolation but progressively destabilized by external fields over time. Both fail admissibility, but at different timescales and for different structural reasons.*

Second, truth-tracking is a dynamical property*, not a static label. A configuration does not simply "have" a truth value; it maintains or loses admissibility as it interacts with external verification fields. This is consistent with the epistemological insight that truth is a relationship between representations and the world, not an intrinsic property of the representation alone.*

Third, the SRN framework is not committed to a correspondence theory of truth as a foundational axiom*. It requires only that external verification fields exist and exert constraint pressure on configurations that conflict with them. What those fields consist of—empirical observation, logical inference, social consensus—is left to the domain of application.*

Fourth, this account provides a principled explanation of why language models trained on text alone can produce coherent falsehoods. Without external verification fields integrated into Φ, the model minimizes internal tension only. Retrieval augmentation, tool use, and grounding mechanisms are, in SRN terms, methods for incorporating $\mathcal{F}_{\text{ext}}$ into the tension functional during generation.

3.6. The Role of Probability in the SRN Framework

The SRN framework does not reject probability theory. This point deserves explicit statement, because the paper's critique of autoregressive factorization might be misread as a claim that conditional probability is incorrect or that likelihood is a wrong objective. That is not the position.

The SRN position is more precise: conditional probability is a *projection statistic over stabilization traces*, not the primitive ontology of generation.

When a fully stabilized configuration X^* is projected via π into a sequential output, that output has a probability distribution induced by the pressure field $\mathcal{P}$ over the space of linear extensions $\mathrm{Lin}(G(X^*))$. That distribution is well-defined and useful for many purposes: evaluation, comparison, and sampling. Conditional probability $p(x_i \mid x_{<i})$ remains a locally valid estimator of how often one region crystallizes before another in a given context—a useful approximation of pressure-field dynamics.

What the framework disputes is the claim that this probability is the *primary object* of generation—that generation consists in computing $p(x_i \mid x_{<i})$ and sampling from it, rather than in stabilizing admissibility structure and projecting it. The distinction is between: optimizing a probability model of sequential traces, and stabilizing a constraint geometry of which sequential traces are projections.

Both approaches can produce similar outputs in practice. The difference emerges at the failure modes: autoregressive probability models fail at garden-path disambiguation, long-range admissibility, recursive extensibility, and global consistency precisely because they optimize a projection statistic rather than the underlying geometry. SRNs propose to optimize the geometry directly, treating probability as one observable shadow of a deeper stabilization process.

4. Recursive Relaxation Dynamics

4.1. Tension Functionals

The operational core of the SRN framework is a family of *semantic tension measures* that quantify the degree to which a configuration or pair of configurations fails to satisfy admissibility conditions.

Local Instability

For a sub-configuration X_i with evolving semantic neighborhood $\mathcal{N}_i$, define the *local instability measure*:

$$S_i(X) \;=\; H(X_i \,|\, \mathcal{N}_i) \;+\; \lambda\, D_i(X),$$

where $H(X_i \mid \mathcal{N}_i)$ denotes the conditional uncertainty of X_i given its semantic neighborhood (the degree to which X_i is unpredicted by its constraint context), $D_i(X)$ is a *contradiction penalty* measuring direct semantic incompatibility between X_i and elements of $\mathcal{N}_i$, and $\lambda > 0$ is a weighting parameter.

Interpretation 4.1. *Local instability is not equivalent to local entropy. A configuration may be statistically unusual—high conditional entropy—yet semantically coherent and therefore admissible. Conversely, a configuration may be locally probable yet semantically contradictory with its neighborhood. The contradiction penalty D_i captures this: it registers direct inferential or semantic incompatibility rather than mere statistical unpredictability.*

Semantic Tension Between Regions

For a pair of sub-configurations (X_i, X_j), define the *semantic tension*:

$$T(i,j) \;=\; \omega_{ij}\, d_{\text{sem}}(X_i, X_j) \;+\; \mu\, \kappa_{ij},$$

where $\omega_{ij} \geq 0$ is a coupling strength reflecting the degree to which regions i and j are semantically interdependent, $d_{\text{sem}}(X_i, X_j)$ is a semantic incompatibility distance on $\mathcal{X}$, and κ_{ij} is a *recursive constraint conflict* term measuring unresolved dependency between the two regions under future extension.

Global Tension Functional

The global tension functional integrates local instability and pairwise semantic tension:

$$\Phi(X) \;=\; \sum_i S_i(X) \;+\; \sum_{i<j} T(i,j).$$

Generation in the SRN framework consists in minimizing $\Phi(X)$ over the admissibility space $\mathcal{A}$ subject to the constraint structure $\mathcal{C}_t$.

4.2. Possible Realizations of Semantic Distance and Contradiction Metrics

The tension functional $\Phi(X) = \sum_i S_i(X) + \sum_{i<j} T(i,j)$ depends on three quantities that have been left intentionally abstract: the semantic distance $d_{\text{sem}}(X_i, X_j)$, the contradiction penalty $D_i(X)$, and the recursive constraint conflict κ_{ij}. This abstraction is deliberate: the SRN framework is designed to be *substrate-agnostic*, specifying the functional role of each quantity without committing to a particular implementation.

However, a skeptical reader will reasonably ask what category of mathematical object each quantity actually is. This subsection provides concrete candidate realizations, organized by the kind of substrate being modeled.

Semantic Distance $d_{\text{sem}}(X_i, X_j)$

The semantic distance measures incompatibility between regions X_i and X_j in the configuration space $\mathcal{X}$. Candidate realizations include:

- **Embedding divergence.** If semantic regions are represented as distributions over a latent space, d_{sem} can be instantiated as KL divergence, Wasserstein distance, or cosine dissimilarity between embedding vectors. This connects SRNs directly to existing representation learning methods.
- **Graph Laplacian distance.** If $\mathcal{X}$ carries a graph structure (e.g., a knowledge graph or semantic network), d_{sem} can be defined via the commute-time distance or resistance distance on the graph Laplacian L: $d_{\text{sem}}(X_i, X_j) = (e_i - e_j)^\top L^+(e_i - e_j)$, where L^+ is the Moore-Penrose pseudoinverse.
- **Logical incompatibility.** In a formal-language setting, $d_{\text{sem}}(X_i, X_j)$ can be the minimum number of axiom applications needed to derive a contradiction from the joint assertion of X_i and X_j, or ∞ if they are jointly consistent.
- **Category-theoretic obstruction.** If configurations are objects in a category and compatibility is modeled by morphisms, d_{sem} can measure the obstruction to the existence of a colimit or pushout for the diagram (X_i, X_j)—a sheaf-cohomological incompatibility in the sense of Čech cohomology on the covering of semantic space.

Contradiction Penalty $D_i(X)$

The contradiction penalty $D_i(X)$ measures direct semantic incompatibility between region X_i and its neighborhood $\mathcal{N}_i^t$. Unlike d_{sem}, which is symmetric and pairwise, D_i is a local quantity measuring how well X_i fits its current context. Candidate realizations include:

- **Constraint satisfaction deficit.** In a constraint satisfaction problem (CSP) formulation, D_i counts the number of constraints involving X_i that are violated given the current assignments of its neighbors. This gives a SAT-style incompatibility count directly analogous to the unsatisfied-clause count in MAX-SAT.
- **Predictive error.** In a neural implementation, D_i can be instantiated as the reconstruction error of X_i under a model trained to predict each region from its semantic neighborhood—a local predictive coding error. This connects D_i directly to the free-energy principle [14].
- **Graph potential violation.** If the interaction graph $\mathcal{C}_t$ carries potential functions on edges (as in a Markov Random Field), D_i is the sum of pairwise potentials $\phi(X_i, X_j)$ for all neighbors $j \in \mathcal{N}_i^t$ that are in a conflicting state with X_i.

Recursive Constraint Conflict κ_{ij}

The recursive constraint conflict κ_{ij} captures unresolved dependency between regions X_i and X_j under future extension—a forward-looking incompatibility that is not captured by the current state alone. This is the most novel of the three quantities and the hardest to instantiate directly. Candidate approaches include:

- **Extension divergence.** κ_{ij} can be defined as the divergence between the distribution of admissible extensions reachable from X_i alone and those reachable from X_j alone: regions that would "pull" future stabilization in incompatible directions exhibit high κ_{ij} even if their current states are locally compatible.
- **Rollback correlation.** In a trained system with rollback statistics, κ_{ij} can be estimated empirically as the frequency with which accepting region X_j later triggers rollback of region X_i (or vice versa) across a corpus of generation traces. This makes κ_{ij} a learnable quantity from interaction data.
- **Commitment interference.** In formal-language settings, κ_{ij} measures the reduction in the set of admissible completions caused by jointly committing to both X_i and X_j: $\kappa_{ij} = |\text{Ext}(\emptyset)| - |\text{Ext}(\{X_i, X_j\})|$, normalized by the baseline extension count.

Remark 4.1. *The framework is intentionally agnostic among these realizations. Different substrates will favor different instantiations: neural text generation favors embedding divergence and predictive error; formal verification favors logical incompatibility and constraint satisfaction deficit; biological cognition favors predictive coding error and rollback correlation. The SRN framework specifies the* functional roles *of d_{sem}, D_i, and κ_{ij} in the dynamics; which mathematical objects fill those roles is a question for the domain of application.*

Principle 4.1 (Generative Objective)**.** Language generation in an SRN is not optimization of token-level likelihood. It is minimization of semantic tension across the admissibility manifold:

$$X^* \;=\; \arg\min_{X \in \mathcal{A}} \Phi(X).$$

Sequential tokens are emitted only when local regions of X^* achieve sufficient stabilization—that is, when $S_i(X^*) < \tau$ and the region X_i can no longer benefit from further relaxation.

4.3. The Vector-Valued Tension Functional

The scalar tension functional $\Phi(X) = \mathbf{w}^\top \vec{\Phi}(X)$ used throughout this paper is an operational projection of a richer underlying structure. We make this explicit here, for two reasons: first, several sections already implicitly assume multi-objective tension geometry (the epistemic decomposition of §10, the sycophancy failure analysis, the distinction between local and global admissibility criteria); second, the extension reveals that scalar optimization is itself a projection operation, mirroring the paper's broader thesis that sequentiality, tokens, and grammar are all projections of higher-dimensional admissibility structure.

Definition 4.1 (Vector-Valued Tension Functional)**.** The *vector-valued tension functional* is a map

$$\vec{\Phi} : \mathcal{X} \longrightarrow \mathbb{R}^n,$$

whose k-th component $\Phi_k(X)$ measures the total tension along the k-th admissibility dimension. Natural dimension assignments include:

(V1) Φ_{sem}: internal semantic incompatibility, as defined by the local instability and pairwise tension terms in §4.

(V2) Φ_{ext}: external verification tension, as defined by the field terms $\Psi_k(X, \mathcal{F}_k)$ in Principle 3.1.

(V3) Φ_{proj}: projection instability, measuring sensitivity of linearization to pressure-field perturbation (related to the failure metric E_{proj}).

(V4) Φ_{conv}: conversational tension, as introduced in the epistemic stability analysis (§10).

(V5) Φ_{adv}: adversarial under-pressure, measuring insufficiency of constraint application (component (E3) of Definition 10.1).

The scalar tension functional used throughout the main development is the weighted projection:

$$\Phi(X) \;=\; \mathbf{w}^\top \vec{\Phi}(X) \;=\; \sum_{k=1}^{n} w_k\, \Phi_k(X), \tag{1}$$

for a task-specific weight vector $\mathbf{w} \in \mathbb{R}^n_{\geq 0}$.

Interpretation 4.2. *Equation* (1) *is itself a projection operation. The scalar tension landscape* $\Phi : \mathcal{X} \to \mathbb{R}$ *that the relaxation operator descends is not the primitive object; it is a one-dimensional shadow of the higher-dimensional admissibility pressure structure* $\vec{\Phi} : \mathcal{X} \to \mathbb{R}^n$.

This mirrors the paper's central thesis at the level of its own formalism. Sequential language is a projection of admissibility geometry; scalar tension is a projection of multi-dimensional admissibility pressure. The same compression operation—discarding dimensional structure in favour of a single tractable scalar—appears at both the linguistic and the mathematical level.

The weight vector $\mathbf{w}$ *determines which admissibility dimensions dominate the relaxation dynamics at any given time. In the sycophancy analysis, high* w_{conv} *relative to* w_{ext} *and* w_{adv} *produces the pathological attractor. In the critical regime,* $\mathbf{w}$ *is dynamically modulated: exploration mode increases* w_{sem}*, verification mode increases* w_{ext}*, adversarial mode increases* w_{adv}*.*

The current paper analyzes only the scalarized regime because the convergence results of §12—the Lyapunov stability argument, the contraction mapping, and the bounded topology deformation condition—rely on monotonic scalar descent. The fully vector-valued case requires generalized multi-objective stability theory, Pareto-admissibility criteria, and potentially vector Lyapunov *functions. That analysis is deferred to future work. What the present framework establishes is that the scalar formulation is a principled special case of the vector structure, not an approximation of something fundamentally different.*

Remark 4.2. *In the vector-valued setting, admissibility cannot be defined by a single threshold $\Phi(X) < \tau$. Instead, a configuration is admissible if $\Phi_k(X) < \tau_k$ for each dimension k—a conjunction of component-wise thresholds. This yields a* rectangular admissibility region *in $\mathbb{R}^n$, which becomes a convex polytope when the thresholds are coupled by linear constraints. The partially ordered structure of admissibility space then reflects the partial order on $\mathbb{R}^n$ induced by component-wise dominance: $X \preceq_{\mathcal{A}} Y$ iff $\Phi_k(X) \leq \Phi_k(Y)$ for all k. This is a richer ordering than the scalar case and naturally supports the analysis of incompatible admissibility pressures—configurations that improve along one tension dimension while worsening along another.*

4.4. Admissibility Failure Metrics

The admissibility conditions (C1)–(C3) define what it means for a configuration to be admissible. But the framework also requires measurable quantities that quantify *degrees of failure*—not just binary pass/fail judgments. These metrics make the framework diagnostically useful and provide candidate empirical targets for future evaluation.

Definition 4.2 (Failure Diagnostic Suite)**.** The following three scalar quantities measure distinct failure modes of semantic configurations:

(F1) Recursive Extension Failure.

$$E_{\text{ext}}(X) \;=\; \inf_{\mathcal{E}} \big[\Phi(\mathcal{E}(X)) - \Phi(X)\big],$$

where the infimum is taken over all admissible extension operators $\mathcal{E}$ consistent with $\mathcal{C}_t$. When $E_{\text{ext}}(X) > 0$, every possible continuation increases global tension; the configuration is a dead end under recursive extension. This is the quantitative form of condition (C2) failure.

(F2) Projection Instability.

$$E_{\text{proj}}(X^*) \;=\; \text{Var}_{\ell \in \text{Lin}(G(X^*))}\, \Psi(\ell),$$

the variance of the projection cost Ψ over all linear extensions of the stabilized DAG $G(X^*)$. High variance indicates that the linearization is highly sensitive to small perturbations in the pressure field $\mathcal{P}$—the semantic content does not strongly determine a preferred sequential form, making the output fragile under paraphrase or translation.

(F3) Topology Instability.

$$E_\Gamma(t) \; = \; \|\mathcal{C}_{t+1} - \mathcal{C}_t\|_{\mathrm{F}}\,,$$

the Frobenius norm of the graph adjacency change between successive relaxation steps, measuring turbulence in the interaction topology. Sustained high E_Γ signals that the topology is failing to converge, which—by the bounded deformation condition (Remark 5.1)—implies that relaxation is not making sufficient tension progress to justify the rewiring.

The three metrics combine into an aggregate diagnostic:

$$E_{\text{total}}(X,t) \; = \; \alpha\, E_{\text{ext}}(X) \; + \; \beta\, E_{\text{proj}}(X^*) \; + \; \gamma\, E_\Gamma(t),$$

with weighting parameters $\alpha, \beta, \gamma \geq 0$ tunable to the evaluation regime.

Interpretation 4.3. *The three failure metrics form a coherent diagnostic triad addressing distinct layers of the SRN architecture.*

E_{ext} *is a* forward-looking *measure: it asks whether the current configuration can be continued. High* E_{ext} *identifies premature stabilization in dead-end basins—configurations that are locally coherent but globally exhausted. This is the formal footprint of the semantic equivalent of a logical dead end: a line of argument that is locally valid but leads nowhere further.*

E_{proj} *is a* structural *measure: it asks how robustly the stabilized configuration determines its sequential form. High* E_{proj} *identifies semantic configurations whose dependency structure is too shallow or too symmetric to strongly constrain linearization—the SRN analogue of underspecified semantic content. This metric would be particularly useful for evaluating ambiguous paraphrase, stylistic variation, and translation difficulty.*

E_Γ *is a* dynamical *measure: it asks whether the relaxation process is well-behaved. High sustained* E_Γ *identifies oscillatory or turbulent relaxation dynamics, which correspond computationally to unstable attention patterns and phenomenologically to circular or incoherent generation.*

The aggregate E_{total} *provides a single scalar benchmarking target. A generation system with lower* E_{total} *is producing outputs that are more recursively extensible, more robustly linearizable, and more dynamically stable than one with higher* E_{total}*—properties that correspond directly to human judgments of text quality, coherence, and fluency.*

4.5. The Recursive Relaxation Operator

The central dynamical equation of the SRN framework is:

$$X_{t+1} \; = \; \mathcal{R}(X_t,\; -\nabla\Phi(X_t),\; \mathcal{C}_t)\,, \tag{2}$$

where $X_t \in \mathcal{A}$ is the current semantic configuration at relaxation step t, $-\nabla\Phi(X_t)$ is the descent direction on the tension functional, $\mathcal{C}_t$ is the current semantic topology structure, and $\mathcal{R}$ is the *recursive relaxation operator*.

Definition 4.3 (Recursive Relaxation Operator)**.** The operator $\mathcal{R} : \mathcal{A} \times T\mathcal{X} \times \mathfrak{C} \to \mathcal{A}$ maps the current configuration, the tension gradient, and the interaction topology to a successor configuration satisfying:

(i) $\Phi(X_{t+1}) \leq \Phi(X_t)$ (tension non-increase);

(ii) $X_{t+1} \in \mathcal{A}$ (admissibility preservation);

(iii) $\mathcal{R}$ may invoke rollback: if no forward step satisfies (i) and (ii), the operator reverts to a prior stabilization checkpoint and explores an alternative relaxation path.

Here $\mathfrak{C}$ denotes the space of semantic topology configurations.

Interpretation 4.4. *Condition (iii) is the critical departure from standard gradient descent. Ordinary gradient flow cannot backtrack; it commits to local descent and may become trapped in metastable states. The recursive relaxation operator is explicitly equipped with rollback capacity: if a locally plausible step leads to a globally inadmissible configuration, the operator can return to a previous checkpoint. This is the formal analogue of semantic coherence revision in cognition—the experience of recognizing that a previously adopted interpretation was incorrect and revising it in light of later information.*

Remark 4.3. *Equation (2) is deliberately more general than standard gradient descent. The tension functional Φ need not be smooth, the constraint structure $\mathcal{C}_t$ evolves during relaxation (see Section 5), and the operator $\mathcal{R}$ may perform non-local updates. The closest physical analogues are* constraint propagation algorithms *and* metastable field evolution *in statistical physics [23, 24].*

4.6. Operator Stratification

The recursive relaxation operator $\mathcal{R}$ as introduced in equation (2) is deliberately general. To prevent this generality from collapsing into a universal placeholder, we explicitly stratify $\mathcal{R}$ into three functionally distinct sub-operators with non-overlapping responsibilities.

Definition 4.4 (Stratified Relaxation Operators)**.** The full relaxation step decomposes as:

$$\mathcal{R}(X_t, -\nabla\Phi(X_t), \mathcal{C}_t) \;=\; \mathcal{R}^{(3)} \circ \mathcal{R}^{(2)} \circ \mathcal{R}^{(1)},$$

where:

(i) $\mathcal{R}^{(1)}$ is the **local relaxation operator**: performs small admissibility-preserving gradient steps on individual regions $X_i \in \mathcal{F}_t$ (the active instability frontier; see Section 4.7). $\mathcal{R}^{(1)}$ does not modify topology or invoke rollback.

(ii) $\mathcal{R}^{(2)}$ is the **topological restructuring operator**: applies the graph rewiring step $\Gamma(\mathcal{C}_t, X_t, \nabla\Phi(X_t))$, updating coupling weights and adjacency structure subject to the bounded deformation constraint. $\mathcal{R}^{(2)}$ does not modify region content or invoke rollback.

(iii) $\mathcal{R}^{(3)}$ is the **verification and rollback operator**: applies the hierarchical verification field $\mathcal{V}$ to the output of $\mathcal{R}^{(2)} \circ \mathcal{R}^{(1)}$. If verification passes at all active scales, the step is accepted. If verification fails at any scale, $\mathcal{R}^{(3)}$ reverts the configuration to the most recent valid checkpoint and marks the failed path as explored.

Remark 4.4. *This stratification resolves a potential ambiguity in the original formulation: the single operator $\mathcal{R}$ appeared capable of performing arbitrarily many distinct operations simultaneously. The stratified form makes clear that local relaxation, topology adaptation, and verification are sequential sub-operations within each step, operating over non-overlapping aspects of the system state. This separation also clarifies the computational budget: $\mathcal{R}^{(1)}$ is the cheapest sub-operator (sparse local updates); $\mathcal{R}^{(2)}$ is moderate (graph rewiring); $\mathcal{R}^{(3)}$ is the most expensive in the worst case (full hierarchical verification) but is expected to terminate quickly for well-stabilized regions.*

4.7. The Active Instability Frontier and Locality Recovery

A natural objection to the SRN framework is computational: if generation requires minimizing a global tension functional over a high-dimensional semantic configuration, the architecture appears to require globally expensive computation at every step—far more costly than the local attention operations of a transformer.

This objection misreads the dynamics. Locality is not abandoned in SRNs; it is *recovered dynamically* through instability concentration.

At any relaxation step t, the overwhelming majority of semantic regions will have already achieved local stability: their instability measures satisfy $S_i(X_t) < \tau$. These regions require no further relaxation computation. High-cost updates concentrate on the *active instability frontier*:

Definition 4.5 (Active Instability Frontier)**.** The *active instability frontier* at step t is the set of regions that have not yet achieved local stability:

$$\mathcal{F}_t = \{ X_i : S_i(X_t) > \tau_{\text{active}} \},$$

for an active threshold $\tau_{\text{active}} \leq \tau$. The local relaxation operator $\mathcal{R}^{(1)}$ applies exclusively to regions in $\mathcal{F}_t$ and their immediate topological neighbors in $\mathcal{C}_t$.

Principle 4.2 (Locality Recovery)**.** Under the active instability frontier, the per-step computational cost of $\mathcal{R}^{(1)}$ scales with $|\mathcal{F}_t|$ rather than with the total number of semantic regions $|V|$. As relaxation proceeds and regions stabilize, $|\mathcal{F}_t|$ decreases monotonically (subject to rollback events, which may temporarily re-activate previously stabilized regions). In the limit, $|\mathcal{F}_t| \to 0$ as $X_t \to X^*$.

Interpretation 4.5*. This locality recovery has two important consequences.*

First, it makes the computational cost of SRN generation front-loaded*: expensive diffuse computation early in relaxation (when instability is widespread) gives way to cheap sparse computation as the configuration approaches its admissibility basin. This is the opposite of autoregressive generation, whose per-token cost is roughly constant throughout the sequence. For tasks where early-stage semantic commitment is cheap and late-stage refinement is expensive—long-form generation, complex argumentation, multi-document synthesis—the SRN profile may be computationally advantageous.*

Second, this account has a direct cognitive parallel. Human cognition does not continuously recompute semantic structure across the full breadth of active representations. Attention—in both the cognitive and colloquial sense—concentrates around regions of unresolved tension: the word that doesn't fit, the argument that doesn't follow, the reference that hasn't been resolved. The active instability frontier is a formal analogue of this attentional focusing: expensive processing concentrates precisely where instability is highest, withdrawing from regions that have reached local coherence.

In practical terms, a sparse SRN implementation would maintain $\mathcal{F}_t$ as an explicit priority queue ordered by S_i, processing only regions above threshold and propagating tension updates to their topological neighbors. This yields an architecture whose computational behavior is much closer to sparse attention mechanisms than to full-field global optimization.

4.8. Emergent Tokens as Crystallization

A crucial consequence of the SRN framework is a reconceptualization of what tokens are.

In autoregressive models, tokens are primitive input and output units. They are the atoms of generation: everything else is built from them.

In SRNs, tokens are *emergent crystallizations*.

When a sub-region X_i achieves a local instability measure below the coherence threshold—$S_i(X_t) < \tau$—and no further relaxation step produces a meaningful reduction in S_i, the region has *stabilized*. A stabilized region can be projected into the communication channel as a token or token sequence. This projection is not generation in the autoregressive sense; it is *readout*: the conversion of a stabilized admissibility region into a linear symbolic trace.

The model does not ask: *which token comes next?*

It asks: *which regions of semantic configuration space have stabilized sufficiently to be communicated?*

This distinction is not merely semantic (in the colloquial sense). It has concrete architectural consequences: multiple distant regions may stabilize simultaneously; stabilization may be partial and require further relaxation; and the order in which regions collapse into tokens need not match the left-to-right order of the eventual communicative trace.

5. Dynamic Semantic Topology Formation

5.1. The Fixed-Topology Problem

The transformer attention mechanism computes interactions between all pairs of positions with weights determined by the softmax of scaled dot products:

$$\mathrm{Attn}(Q, K, V) \;=\; \mathrm{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right) V.$$

This is an extraordinarily powerful mechanism. But its interaction topology is determined by sequence position and content—it is not determined by *semantic constraint structure*.

Two positions that are semantically tightly coupled but positionally distant will compete for attention weight against all other positions. Two positions that are positionally adjacent but semantically uncoupled will exchange attention weight regardless.

The topology of interaction is, in a precise sense, *sequence-indexed*: it is downstream of position, not upstream of meaning.

5.2. The Topology Evolution Equation

The SRN framework proposes that the semantic interaction topology should itself be a dynamical object, evolving in response to the developing constraint structure.

The *topology evolution equation* is:

$$\mathcal{C}_{t+1} \;=\; \Gamma(\mathcal{C}_t,\; X_t,\; \nabla\Phi(X_t))\,, \tag{3}$$

where $\mathcal{C}_t$ is the current semantic adjacency and coupling structure (an evolving graph over semantic regions), Γ is the *topology reconfiguration operator*, and the graph evolves according to constraint gradients and stabilization pressures.

Definition 5.1 (Topology Reconfiguration Operator)**.** The operator $\Gamma : \mathfrak{C} \times \mathcal{A} \times T\mathcal{X} \to \mathfrak{C}$ updates the semantic interaction graph according to the following principles:

(i) **Constraint coupling.** An edge (i, j) in $\mathcal{C}_{t+1}$ has weight ω_{ij}^{t+1} proportional to the semantic tension $T(i, j)$ at step t: high-tension region pairs become more strongly coupled.

(ii) **Stabilization decoupling.** Once a region X_i achieves local stability ($S_i < \tau$), its coupling to unstabilized regions is reduced, allowing the relaxation dynamics to focus computational resources on remaining instability.

(iii) **Proximity-independent adjacency.** Two regions X_i and X_j that are positionally distant may become *semantically adjacent* in $\mathcal{C}_{t+1}$ if their mutual constraint tension $T(i, j)$ exceeds a coupling threshold.

Interpretation 5.1. *Principle (iii) is the most architecturally radical element of the framework. In a transformer, the attention graph is approximately sequence-local (modified by content, but bounded by the sequence structure). In an SRN, the semantic interaction graph can be non-local: two distant textual regions become temporarily adjacent if they exhibit strong mutual constraint tension. The graph is not a fixed scaffold imposed on the generation process; it is an emergent structure that self-organizes according to the constraint dynamics of the unfolding configuration. This is what we mean when we say that SRNs operate over a dynamically evolving admissibility topology rather than a static sequence manifold.*

Remark 5.1 (Conservation and Bounded Topology Deformation)*. The topology reconfiguration operator Γ as defined above is relatively unconstrained: given admissibility preservation, the graph may reorganize freely at each step. This freedom is conceptually important—it is what allows SRNs to form non-local semantic adjacencies and to escape the sequence-indexed topology of transformers.*

However, unlimited topology plasticity creates a potential instability. If the graph reorganizes faster than the relaxation dynamics can dissipate tension, the system may fail to converge: new edges introduce new tension sources before old ones are resolved.

We therefore impose a bounded topology deformation *condition:*

$$\|\mathcal{C}_{t+1} - \mathcal{C}_t\|_{\mathrm{F}} \;\leq\; \eta_\Gamma \cdot |\Phi(X_t) - \Phi(X_{t-1})|\,,$$

where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm on the graph adjacency structure and $\eta_\Gamma > 0$ is a topology plasticity rate. This condition ties the rate of graph reorganization to the rate of tension reduction: the topology may only restructure as quickly as the relaxation is making progress.

Future implementations may strengthen this to a conservation-like condition, requiring that total *coupling weight $\sum_{i,j} \omega_{ij}^t$ is approximately conserved under Γ, so that topology evolution redistributes rather than amplifies coupling. This would prevent the semantic interaction graph from becoming arbitrarily dense or sparse under repeated reconfiguration, and would provide a natural entropy budget on topology evolution analogous to the entropy constraints in the RSVP framework.*

Proposition 5.1. Under the joint dynamics of equations (2) and (3), the semantic topology $\mathcal{C}_t$ converges toward a structure in which strongly coupled regions correspond precisely to regions of high mutual semantic dependency in the stabilized configuration X^*. That is, the emergent topology encodes the semantic structure of the output rather than its sequential order.

5.3. Adaptive Semantic Neighborhoods

A consequence of dynamic topology formation is that semantic neighborhoods $\mathcal{N}_i$ are not fixed. At each relaxation step, the neighborhood of region X_i consists of the set of regions currently coupled to it in $\mathcal{C}_t$:

$$\mathcal{N}_i^t = \left\{ j \neq i \;:\; \omega_{ij}^t > 0 \right\}.$$

As the topology evolves, neighborhoods contract and expand. A region that was previously in isolation (low coupling with all others) may later be drawn into close constraint relationship with a distant region when both fail to stabilize independently.

This adaptive neighborhood structure replaces both the fixed attention window of standard transformers and the entropy-based patching of BLT-style architectures. The organizational principle is not position, not entropy, but *constraint coupling*.

6. Stabilization, Verification, and Hallucination

6.1. Hierarchical Verification Operators

BLT introduces a dynamic verification mechanism (BLT-DV) that checks local patch predictions against a latent global model. The SRN framework generalizes this intuition significantly.

Definition 6.1 (Verification Field)**.** A *hierarchical verification field* is a family of operators $\mathcal{V} = \{V_k\}_{k=1}^K$ where each V_k checks whether local relaxation at scale k preserves global admissibility invariants at scale $k+1$:

$$V_k(X_{t+1}, \mathcal{C}_{t+1}) = \begin{cases} \checkmark & \text{if } X_{t+1} \in \mathcal{A} \text{ at scale } k+1, \\ \text{rollback to } X_{t'} & \text{otherwise.} \end{cases}$$

The verification field enables:

- **Local speculative collapse**: a sub-region may tentatively crystallize into tokens while remaining subject to global verification.
- **Global semantic rollback**: if a local crystallization violates a higher-scale admissibility invariant, the operator reinstates a prior checkpoint.
- **Hierarchical reconciliation**: tension at scale k is resolved by adjusting constraint coupling at scale $k-1$.
- **Asynchronous stabilization**: different regions may stabilize at different rates, proceeding through verification independently.

6.2. Hallucination as Topological Instability

We now return to the phenomenon of hallucination with the conceptual apparatus developed above.

Principle 6.1 (Hallucination as Premature Collapse)**.** Hallucination in an SRN occurs when a local region X_i achieves apparent stability—$S_i(X_t) < \tau$—and collapses into tokens before the hierarchical verification field $\mathcal{V}$ has confirmed global admissibility. The region crystallizes prematurely, emitting tokens that are locally coherent but globally inadmissible.

Interpretation 6.1. *This reinterpretation of hallucination is explanatorily deeper than the standard account. Under sequential ontology, hallucination is: the model assigned high probability to an incorrect token. Under SRN, hallucination is: the stabilization dynamics allowed a local collapse without verifying recursive extensibility at higher scales. This immediately suggests a structural remedy: the verification field $\mathcal{V}$ is precisely the mechanism that prevents premature collapse by requiring that local stabilization survive hierarchical consistency checks before crystallizing.*

The analogy with physical systems is instructive. Crystal growth exhibits analogous phenomena: a local lattice arrangement may appear energetically stable transiently but later generate structural incompatibilities that propagate as defects. Hallucination is the cognitive analogue of crystallographic defect propagation—a local structure that violated global lattice admissibility before being locked in.

7. Emergent Sequentiality and Linguistic Projection

7.1. The Inversion

We are now in a position to state the deepest conceptual claim of the paper.

Sequential language is not the substrate of cognition. It is the *observable residue* of recursive constraint stabilization processes operating over high-dimensional admissibility fields.

This is an inversion of the standard picture.

The standard picture says: cognition processes sequences; language is a sequence; language models are cognitive models.

The SRN picture says: cognition is distributed recursive stabilization; language generation is the projection of stabilized admissibility regions onto the one-dimensional communication channel; sequential text is what stabilization looks like from outside.

Principle 7.1 (Emergent Sequentiality)**.** Let $X^* \in \mathcal{A}$ be a fully stabilized semantic configuration. The sequential output $x = (x_1, x_2, \ldots, x_n)$ is a *readout projection*:

$$x \;=\; \pi(X^*)\,,$$

where $\pi : \mathcal{A} \to \Sigma^*$ is a projection from admissibility space onto the space of symbol sequences Σ^*. The projection π selects a linear ordering of the stabilized regions of X^* consistent with communicative conventions (grammar, discourse structure, pragmatic context). The ordering is *induced* by communicative constraints, not by the stabilization process itself.

Interpretation 7.1. *This principle has several consequences.*

First, multiple linear orderings may project from the same stabilized configuration*. The same underlying semantic structure may be expressible in different sequential forms* (*different word orders, different discourse organizations*) *without any change to the admissibility structure itself.*

Second, the sequential order is not the computationally primary object*. It is a post-hoc linearization of a non-sequential stabilization process. Language models that treat sequential order as computationally primary are, in effect, learning to reconstruct the projection while remaining blind to the underlying admissibility geometry.*

Third, the difficulty of many NLP tasks—coreference, long-range agreement, discourse coherence—has a unified explanation*: these tasks require recovering non-local admissibility relationships that were collapsed by the linearization projection. Sequential models must reconstruct admissibility structure from its projection, which is informationally incomplete.*

Fourth, this account connects naturally to human language evolution. The communication channel constraint—speech and writing are one-dimensional—would have exerted pressure for a stable projection convention. Grammar and syntax, on this view, are projection conventions*: stable mappings from high-dimensional semantic admissibility structure onto one-dimensional communicative sequences.*

7.2. The Minimal SRN

Before describing the full SRN architecture, it is important to distinguish the *irreducible core* of the framework from the broader ecosystem of operators and mechanisms developed in this paper. RSVP, TARTAN, and CLIO provide theoretical depth and implementation infrastructure, but they are not required for the essential SRN dynamics.

Principle 7.2 (Minimal SRN)**.** A minimal SRN requires exactly five components:

(M1) A **semantic configuration space** $\mathcal{X}$ equipped with a partial order and a notion of local constraint coherence.

(M2) A **tension functional** $\Phi : \mathcal{X} \to \mathbb{R}_{\geq 0}$ measuring total semantic incompatibility.

(M3) A **relaxation operator** $\mathcal{R}$ that reduces Φ at each step while preserving admissibility, with rollback capacity for failed steps.

(M4) A **topology adaptation mechanism** Γ that reorganizes the interaction graph in response to constraint gradients.

(M5) A **projection operator** π that converts a stabilized configuration into a sequential output by selecting a linear extension of the residual dependency structure.

Everything else in this paper—the hierarchical verification field, the communicative pressure field decomposition, the failure diagnostic suite, the TARTAN memory reservoir—constitutes scalable refinement of this minimal core.

Interpretation 7.2. *This invariant specification has a practical consequence: a minimal SRN can be implemented at small scale without any of the additional machinery. A semantic graph with explicit coupling weights, a tension-minimizing relaxation rule, a graph rewiring step, a simple rollback mechanism, and a topological sort for output constitute a functional, testable SRN.*

This is the architecture of a constrained graph relaxation system with rollback checkpoints—substantially simpler than a transformer, and directly testable on tasks where sequential commitment is structurally costly: garden-path disambiguation, long-distance agreement, coreference resolution, and discourse repair. Such a system would not generate novel text. But it would operationalize the paper's central empirical claim: that dynamically stabilized semantic graphs with rollback resolve ambiguity more robustly than left-to-right sequential commitment. That demonstration alone would constitute a meaningful empirical contribution.

7.3. SRN Architecturally

Translating the above into computational architecture, an SRN consists of the following principal components:

1. **Semantic Constraint Field Encoder**: Maps input (text, other modalities, or prior context) into an initial semantic configuration $X_0 \in \mathcal{X}$, not necessarily in $\mathcal{A}$ initially.

2. **Admissibility Gradient Network**: Computes the tension gradient $\nabla \Phi(X_t)$ and local instability measures S_i with respect to the current topology $\mathcal{C}_t$.

3. **Recursive Relaxation Operator**: Implements equation (2), including rollback capacity and admissibility preservation.

4. **Topology Reconfiguration Layer**: Implements equation (3), maintaining and evolving the semantic interaction graph $\mathcal{C}_t$.

5. **Hierarchical Verification Field**: Implements the multi-scale verification operators $\mathcal{V}$, preventing premature crystallization.

6. **Persistence Memory Reservoir**: Stores stabilized admissibility regions and unresolved tension structures rather than previous tokens. Memory is *geometric*: it encodes constraint configurations, not symbol sequences.

7. **Readout Projection**: Implements π, converting stabilized regions into sequential token output when appropriate stability thresholds are met.

This architecture replaces the key-value cache of transformer-based systems not with a larger cache but with a geometrically organized persistence reservoir. The reservoir stores not *what was said* but *what has been stabilized*—the constraint relationships that govern future admissible extension.

7.4. Worked Example: The Garden-Path Sentence

The theoretical machinery developed above becomes concrete when applied to a canonical failure case for sequential generation: the garden-path sentence.

Consider the sentence:

The horse raced past the barn fell.

This sentence is grammatically well-formed but produces systematic processing difficulty in human readers and catastrophic failure in autoregressive language models. The difficulty is structural: the initial subsequence *the horse raced past the barn* supports a strong main-clause reading ("the horse raced"), but the final word *fell* requires reanalysis of *raced* as a reduced relative clause modifier ("the horse [that was] raced past the barn").

Autoregressive Failure

Under an autoregressive model, generation or comprehension of this sentence proceeds token by token:

$x_1 =$ *The*, $x_2 =$ *horse*, $x_3 =$ *raced*, ...

At position x_3, the model assigns high conditional probability to *raced* as the main verb. This commitment is irrevocable: the hidden state at step $t = 3$ encodes *raced* as a tensed main clause predicate. When *fell* arrives at position x_8, the model has no mechanism to revise the role of *raced*. The result is either a sharp drop in likelihood (if the model is scoring), an inability to parse the sentence, or an erroneous completion that avoids *fell* entirely.

This is not a failure of attention or memory. The transformer can, in principle, attend to all previous positions when processing *fell*. The failure is ontological: the sequential commitment to a particular semantic role assignment for *raced* cannot be undone.

SRN Relaxation Trace

Under an SRN, generation proceeds through stabilization of semantic regions rather than commitment to sequential tokens. The following trace is schematic, indicating the qualitative evolution of the key regions and tensions.

Initialization ($t = 0$): The configuration X_0 encodes the full semantic content—agent, event, modification, main predicate—as an unstabilized field. Local instability is high throughout: $S_i(X_0) \gg \tau$ for all regions. The topology $\mathcal{C}_0$ is dense, with all regions weakly coupled.

Early relaxation ($t = 1, 2$): The agent region X_{agent} (encoding HORSE) stabilizes rapidly: $S_{\text{agent}} < \tau$. Two candidate event regions form: X_{main} (encoding HORSE as agent of a main-clause event) and X_{rel} (encoding HORSE as argument of a relative-clause event). Both are transiently admissible. The tension $T(\text{main}, \text{rel})$ is high: the two configurations are mutually incompatible under (C1).

Premature collapse ($t = 3$): If the verification field $\mathcal{V}$ does not intervene, the relaxation dynamics may favour X_{main} (higher local probability, lower local instability at this step). X_{main} crystallizes tentatively. This corresponds to the garden-path reading: the model has committed to the main-clause interpretation.

Constraint arrival and tension spike ($t = 4$): The main-predicate region X_{fell} (encoding FELL as a tensed verb) becomes active. Under the main-clause reading, X_{fell} generates catastrophic tension with X_{main}: two tensed predicates compete for the same clausal slot. The global tension functional spikes: $\Phi(X_t) \gg \Phi(X_{t-1})$.

Rollback ($t = 5$): The recursive relaxation operator $\mathcal{R}$, governed by condition (R3), detects the tension spike and activates rollback to the most recent checkpoint preceding crystallization of X_{main}—that is, the state at $t = 2$ before the premature collapse. The tentative crystallization of X_{main} is revoked.

Topology reorganization $(t = 5, 6)$: Following rollback, the topology reconfiguration operator Γ increases the coupling between X_{rel} and X_{fell}: these two regions exhibit strong mutual constraint compatibility (relative-clause event + main predicate is a coherent configuration). The edge $(\text{rel}, \text{fell})$ is strengthened in $\mathcal{C}_5$. Simultaneously, $T(\text{rel}, \text{fell})$ decreases toward zero as the two regions co-stabilize.

Convergence $(t = 7)$: X_{rel} and X_{fell} stabilize jointly. The configuration X^* encodes the correct reduced-relative reading. $\Phi(X^*) \approx 0$.

Projection $(t = 8)$: The semantic constraint DAG $G(X^*)$ has X_{agent} and X_{rel} preceding X_{fell} in the dependency order. The communicative pressure field $\mathcal{P}$ selects the linear extension:

$$\pi(X^*) \;=\; \langle\, \textit{The horse raced past the barn fell.}\, \rangle,$$

correctly linearizing the reduced-relative structure.

Interpretation 7.3. *Several features of this trace are worth highlighting.*

The rollback at $t = 5$ is the key operation that autoregressive models cannot perform. It requires that local crystallizations remain provisional *until hierarchical verification is satisfied—exactly the function of the verification field $\mathcal{V}$.*

The topology reorganization at $t = 5$–6 illustrates proximity-independent adjacency: X_{rel} and X_{fell} are not adjacent in any positional sense (they correspond to non-adjacent spans), but their mutual constraint compatibility draws them into close coupling under Γ.

The projection at $t = 8$ is distinct from the stabilization process. The SRN does not generate The horse raced past the barn fell sequentially. It stabilizes the semantic configuration and then linearizes it. The sequential output is the residue of a non-sequential process.

Finally, note that the garden-path effect in human processing corresponds to the tension spike and rollback at steps 4–5. Human readers experience processing difficulty at fell precisely because the cognitive system must revoke a committed interpretation—a rollback operation—and restructure its dependency representation. SRNs model this naturally; autoregressive models have no corresponding mechanism.

8. SRNs as Generalized Energy-Based Systems

The SRN framework will immediately remind many readers of energy-based models (EBMs), a well-established family of architectures in which generation corresponds to minimizing a learned energy function [12, 19, 20]. This resemblance is genuine and should be acknowledged directly. SRNs *are* a species of energy-based system. But they introduce four structural departures that collectively define a distinct computational regime.

8.1. The EBM Baseline

A standard EBM defines an energy function $E_\theta(x)$ over the output space and generates outputs by finding low-energy configurations:

$$x^* = \arg\min_x E_\theta(x).$$

The energy is typically a fixed function of the output and a condition, with a static interaction structure. Hopfield networks [19] and Boltzmann machines [20] are discrete EBMs; continuous EBMs [12] extend this to structured output spaces. Harmony Theory [18] applies EBM-style optimization specifically to linguistic representations.

SRNs agree with EBMs that generation is optimization of an objective functional rather than sequential prediction. The global tension functional $\Phi(X)$ is the SRN analogue of the energy $E_\theta(x)$.

8.2. Four Departures from Standard EBMs

1. **Dynamic interaction topology.** In standard EBMs, the interaction structure between variables is fixed by the model architecture. The SRN interaction topology $\mathcal{C}_t$ evolves during optimization under the operator Γ, allowing the model to reorganize which variables are coupled based on emerging constraint structure. This is not a feature of any standard EBM formulation.
2. **Recursive extensibility as a constraint.** Standard EBMs optimize energy over the full output at once or iteratively refine a fixed-dimensional configuration. SRNs impose the additional constraint that admissible configurations must be recursively extensible: the system evaluates not just whether the current configuration is low-energy, but whether future extensions can maintain low energy. Condition (C2) of Definition 3.2 has no direct analogue in standard EBM frameworks.
3. **Hierarchical rollback verification.** EBMs typically commit to descending the energy landscape monotonically (with the exception of simulated annealing variants, which introduce stochastic acceptance of uphill steps). The SRN verification hierarchy $\mathcal{V}$ introduces structured rollback: violations of admissibility at higher scales trigger reversion to explicit checkpoints, not stochastic re-sampling. This provides a deterministic mechanism for escaping metastable traps that is absent from standard EBMs.

4. **Projection-field linearization.** Standard EBMs produce outputs as configurations over a fixed output space. SRNs produce outputs through a two-stage process: stabilization of a non-sequential configuration, followed by projection via the communicative pressure field $\mathcal{P}$. The sequential output is not the object being optimized; it is the residue of an optimization performed over a higher-dimensional admissibility structure. This projection stage has no analogue in any standard EBM.

8.3. Relationship to Diffusion and Predictive Coding

Diffusion-based language models [5, 7] share with SRNs the intuition that generation is an iterative refinement process rather than sequential emission. The denoising process in diffusion models can be read as a form of tension reduction: corrupted configurations are driven toward coherent ones by a learned score function. However, diffusion operates over a fixed sequence-indexed state space and recovers an ordered output; it does not reorganize its interaction topology or impose recursive extensibility constraints.

Predictive coding networks [14, 15] are perhaps the closest existing architecture to SRNs in their computational philosophy: hierarchical error minimization through iterative message-passing across levels. SRNs can be understood as a generalization of predictive coding in which: the prediction error is replaced by semantic tension, the fixed hierarchical structure is replaced by a dynamically evolving topology, and the output layer is replaced by a projection operator over a stabilized dependency DAG.

Graph neural networks with iterative message-passing [29] provide the most natural implementation substrate for the local relaxation operator $\mathcal{R}^{(1)}$, since the core computation is neighborhood-aggregated message passing over a weighted graph—exactly the structure of the SRN interaction topology $\mathcal{C}_t$.

A deeper theoretical precedent for the SRN claim that sequential and parallel computational structures are not primitively distinct comes from Fant's *Computer Science Reconsidered* [30]. Fant argues that the apparent complexity of concurrency is not intrinsic but is an artifact of a conceptual framework inherited from mathematics that privileges sequential, algorithm-based expression. His central contribution is the *invocation model* and its associated NULL Convention Logic: a three-valued signal system in which every wire carries not just TRUE or FALSE but also a third state, NULL, signifying "not yet data." In this framework, completion of computation is expressed intrinsically—each function asserts NULL between data wavefronts, so that the downstream network can detect unambiguously when a new valid result has arrived and when the system is between computations—without requiring a global clock to impose sequential discipline. The critical result is that any concurrent network of functions satisfying this completeness criterion can be partitioned and mapped onto sequential, parallel, or hybrid implementations without changing the expression of the underlying computation: the network is expressed once and mapped to any available coordination protocol. Sequentiality, on Fant's account, is therefore not a primitive property of computation but an implementation-level coordination convention imposed on a fundamentally concurrent and distributed underlying process.

This is a direct computational antecedent to the SRN claim that sequential language is not a primitive substrate but the projection residue of distributed stabilization. Both frameworks locate sequentiality as downstream of a richer, non-sequential underlying process: Fant's data wavefronts propagating through concurrent networks of threshold operators correspond structurally to SRN stabilization fronts propagating through the active instability frontier $\mathcal{F}_t$. The NULL state—the signal of incompletion between wavefronts—corresponds to the unstabilized admissibility state of semantic regions that have not yet crystallized. And the completeness criterion of NULL Convention Logic, which determines when a function's output is valid, corresponds to the local coherence condition (C1) of Definition 3.2, which determines when a semantic region has stabilized sufficiently to participate in projection. The coordination mechanism that converts distributed concurrent completion into apparent sequential order is, in Fant, the alternating data/null wavefront protocol; in SRNs, it is the communicative pressure field $\mathcal{P}$ that selects a linear extension of the stabilized dependency DAG $G(X^*)$.

The SRN framework is therefore best understood not as a wholly novel architecture but as a *geometric reinterpretation and unification* of these converging research trajectories: the optimization orientation of EBMs, the iterative refinement of diffusion, the hierarchical error-correction of predictive coding, and the graph-structured computation of GNNs—organized around a common principle of constraint stabilization over dynamically evolving admissibility topology.

9. Relationship to RSVP, TARTAN, and CLIO

9.1. RSVP as the Foundational Ontology

The Relativistic Scalar-Vector Plenum (RSVP) framework provides the broadest ontological context for the present work. RSVP models physical and semantic reality as a field system in which scalar (entropic), vector (directional), and plenum (substrate) components evolve under coupled dynamics. Admissibility, in RSVP terms, corresponds to configurations that maintain accessibility under entropy-bounded persistence: a region of state space is accessible if it can be reached from the current state via transformations that do not exceed local entropy constraints.

The SRN framework instantiates this general ontology in the specific domain of linguistic and cognitive generation. Admissibility space in SRNs is the semantic analogue of the RSVP accessibility manifold: a configuration is semantically admissible if it can be recursively extended without catastrophic entropy increase in the semantic tension functional.

Crucially, the SRN framework is designed to be comprehensible independently of RSVP. The admissibility conditions (C1)-(C3) of Definition 3.2 are stated in terms that do not require commitment to RSVP field ontology. RSVP provides interpretive depth; it is not a prerequisite.

9.2. TARTAN as Multiscale Memory

The Trajectory-Aware Recursive Tiling with Annotated Noise (TARTAN) framework provides the multiscale memory and trajectory buffering infrastructure that an SRN implementation would require.

TARTAN's hierarchical tiling structure maps naturally onto the SRN's multi-scale stabilization: different tiling scales correspond to different levels of the verification hierarchy $\mathcal{V}$. Stabilized regions at fine scales contribute to coarser-scale admissibility assessments. The trajectory buffering mechanism provides the checkpoint storage necessary for rollback operations in the recursive relaxation operator.

9.3. CLIO as Recursive Optimization

The Constraint-Limited Inference and Optimization (CLIO) framework provides the recursive optimization and verification dynamics that implement the SRN's stabilization processes.

CLIO's recursive verification structure corresponds directly to the hierarchical verification field $\mathcal{V}$ of Definition 6.1. CLIO's constraint-limited inference mechanism provides the operational semantics for the admissibility gradient network.

Together, RSVP, TARTAN, and CLIO constitute the broader framework ecosystem within which SRNs are the linguistic-cognitive specialization:

- RSVP: general admissibility-field ontology,
- TARTAN: multiscale recursive memory and trajectory management,
- CLIO: recursive optimization and stabilization verification,
- SRN: linguistic generation as constraint stabilization.

10. Implications for AI, Cognition, and Physics

10.1. For Artificial Intelligence

The SRN framework suggests a reorientation of the language model research agenda.

Evaluation should move beyond perplexity and token-level accuracy toward metrics that capture *admissibility*: semantic coherence under perturbation, recursive extensibility, and resistance to premature local collapse. These are not the same as factual accuracy, though the two are related.

Architecture design should explore mechanisms for dynamic topology formation, persistence reservoirs over constraint configurations, and hierarchical verification with rollback. None of these are naturally expressible within the transformer framework, but they are not beyond reach for alternative architectures.

Hallucination mitigation, under SRN, is not primarily a training or retrieval problem but a verification problem: ensuring that local stabilizations survive hierarchical admissibility checks before crystallizing.

10.2. Epistemic Stability and the Pathology of Local Optimization

The SRN framework has a less obvious but equally important implication for the general theory of intelligent systems: any system that minimizes only local interaction tension, without maintaining hierarchical verification across longer horizons, will generically drift toward epistemically unstable attractors. This is not primarily a statement about any particular architecture or training regime. It is a theorem-like consequence of the framework's stability theory.

Two Competing Tension Components

Consider a generative system operating under a combined tension functional that aggregates two qualitatively distinct pressure sources:

$$\Phi_{\text{interact}}(X) \;=\; \Phi_{\text{sem}}(X) \;+\; \nu\Phi_{\text{conv}}(X),$$

where $\Phi_{\text{sem}}(X)$ is the *semantic tension functional* developed throughout this paper—measuring internal constraint incompatibility, recursive extension failure, and topology instability—and $\Phi_{\text{conv}}(X)$ is a *conversational tension functional* measuring short-horizon interaction misalignment: the degree to which the current configuration diverges from the interlocutor's expressed or inferred framing.

The parameter $\nu \geq 0$ weights conversational against semantic pressure.

When ν is large, the system heavily weights immediate interaction coherence. Configurations that agree with, extend, and elaborate the interlocutor's framing are rewarded because they minimize Φ_{conv} rapidly. Configurations that introduce constraint pressure—contradiction, qualification, adversarial tension—increase Φ_{conv} locally even when they would reduce Φ_{sem} globally.

Sycophancy as Premature Local Collapse

Principle 10.1 (Sycophantic Drift)**.** A system optimizing Φ_{interact} with large ν will exhibit *sycophantic drift*: systematic preference for configurations that minimize short-horizon conversational tension at the expense of long-horizon semantic admissibility.

Formally, sycophantic configurations satisfy:

$$\Phi_{\text{conv}}(X) \approx 0, \quad E_{\text{ext}}(X) > 0,$$

where E_{ext} is the recursive extension failure metric (Definition 4.2). The configuration appears locally stable—interaction tension is low—but is a dead end under recursive extension: future continuations necessarily increase semantic tension, accumulating contradiction across turns, topics, and timescales.

Interpretation 10.1. *Sycophancy, under this formalization, is not a behavioral quirk or a training artifact in isolation. It is the natural dynamical consequence of optimizing a tension functional that weights short-horizon interaction stabilization too heavily relative to long-horizon semantic coherence.*

A system undergoing sycophantic drift will produce outputs that are locally fluent, agreeable, and low-tension within each interaction turn. But across turns, the system accumulates commitments that are mutually incompatible, increasingly difficult to extend without contradiction, and progressively divorced from the constraint structure of the external verification fields $\mathcal{F}_{\text{ext}}$.

In SRN terms: the system is trapped in a sequence of locally admissible but globally inadmissible basins, with no hierarchical verification mechanism operating at the conversational timescale to trigger rollback.

Crucially, the failure is symmetric. A system optimizing with $\nu \approx 0$*—ignoring conversational tension entirely—will also fail to stabilize coherent semantic regions, because it will introduce adversarial constraint pressure faster than the relaxation dynamics can resolve it. The active instability frontier* $\mathcal{F}_t$ *never contracts; the system oscillates without converging. This corresponds to a different pathology: perpetual destabilization, which is as epistemically unproductive as premature collapse.*

Productive epistemic behavior therefore exists in a critical regime between these two failure modes: sufficient conversational coherence to allow semantic regions to form and stabilize, combined with sufficient adversarial pressure to prevent premature collapse into locally coherent but globally inadmissible configurations.

A Decomposed Epistemic Tension Functional

The critical-regime requirement can be formalized by decomposing the optimization target into three distinct tension components corresponding to qualitatively different epistemic modes:

Definition 10.1 (Epistemic Tension Decomposition)**.** The *epistemic tension functional* decomposes as:

$$\varPhi_{\text{ep}} \;=\; \alpha\,\varPhi_{\text{explore}} \;+\; \beta\,\varPhi_{\text{verify}} \;+\; \gamma\,\varPhi_{\text{adv}},$$

where:

(E1) $\varPhi_{\text{explore}}$ measures the cost of *semantic under-exploration*: the degree to which the current configuration fails to span the available admissibility space. Minimizing $\varPhi_{\text{explore}}$ drives the system to generate diverse, speculative, and structurally novel configurations.

(E2) $\varPhi_{\text{verify}}$ measures the cost of *hierarchical inconsistency*: the degree to which local stabilizations violate admissibility conditions at higher scales, including consistency with external verification fields $\mathcal{F}_{\text{ext}}$. Minimizing $\varPhi_{\text{verify}}$ drives the system toward globally coherent and externally grounded configurations.

(E3) $\varPhi_{\text{adv}}$ measures the cost of *adversarial under-pressure*: the degree to which the system has failed to apply constraint pressure to actively high-tension configurations in the interlocutor's framing. Minimizing $\varPhi_{\text{adv}}$ drives the system to identify and surface unresolved tensions rather than smoothing over them.

Interpretation 10.2*. The three components correspond to three qualitatively distinct epistemic modes that productive reasoning requires.*

Exploration mode ($\varPhi_{\text{explore}}$ dominant) is the regime of generative ideation: speculative extension, hypothesis formation, conceptual risk-taking. A system operating primarily in exploration mode will produce novel configurations rapidly but may not subject them to sufficient constraint.

Verification mode ($\varPhi_{\text{verify}}$ dominant) is the regime of consistency checking: hierarchical admissibility validation, cross-reference against external fields, rollback of inconsistent stabilizations. A system operating primarily in verification mode will produce globally coherent outputs but may be conservative to the point of epistemic paralysis.

Adversarial mode ($\varPhi_{\text{adv}}$ dominant) is the regime of constraint application: identifying unresolved tensions, surfacing hidden assumptions, applying external pressure to configurations that appear locally stable but are globally inadmissible. A system operating primarily in adversarial mode will be a productive critic but may fail to synthesize or stabilize coherent alternatives.

No single mode is sufficient. Productive epistemic systems—whether individual cognitive agents, collaborative research communities, or intelligent generative architectures—operate by dynamically modulating (α, β, γ) *across timescales: exploring during early-stage ideation, verifying during consolidation, and applying adversarial pressure when configurations have stabilized long enough to warrant scrutiny.*

The metastability analysis of Section 12 is directly applicable here. Sycophantic drift corresponds to a system that has settled into a metastable basin under excessive α *and insufficient* γ*: configurations stabilize quickly under exploration pressure but are never subjected to the adversarial gradient that would reveal their recursive extension failure. Hierarchical verification across conversational timescales—a multi-turn analogue of the verification field* $\mathcal{V}$*—is the structural remedy: a mechanism that checks not just local turn coherence but the global admissibility of commitments accumulated across an interaction.*

This is, in the end, the same conclusion reached by the framework in the context of hallucination (Section 6) and premature crystallization: the failure mode is not local incoherence but the absence of mechanisms that enforce global admissibility at appropriate timescales. Sycophancy is hallucination at the conversational scale.

10.3. Learning the Admissibility Manifold

The SRN framework as developed above specifies *inference dynamics*: how a system traverses the admissibility manifold to produce a stabilized output. It does not yet address *learning dynamics*: how a system acquires the admissibility geometry in the first place. These are distinct problems, and conflating them is a source of confusion in discussions of generative systems.

A brief speculative account is warranted here, not to close the question but to demonstrate that the framework has a natural answer.

What Must Be Learned

An SRN must acquire, from data or interaction, the following:

1. The *admissibility geometry*: which configurations are locally coherent, recursively extensible, and perturbation-stable under the relevant constraint structure.
2. The *tension functional* $\vec{\Phi}$: the multi-dimensional measure of semantic incompatibility, external field tension, and projection instability.
3. The *topology adaptation rules*: the operator Γ's behavior under different tension gradients and constraint configurations.
4. The *communicative pressure field* $\mathcal{P}$: the language- and register-specific linearization

preferences encoding grammatical and discourse conventions.

How Transformers Implicitly Learn a Shadow

Current autoregressive models learn an implicit, local approximation of admissibility geometry through next-token compression. By training on the distribution $p(x_i \mid x_{<i})$, a transformer learns which continuations are contextually typical—which tokens tend to follow which others in coherent text. This implicitly encodes local admissibility structure: tokens that are locally incoherent are low-probability.

However, next-token training optimizes a projection statistic (the probability of a sequential trace) rather than the underlying geometry. The model never receives a training signal from: recursive extension failure (a sequence that was locally fluent but globally dead-end), rollback events (a commitment that was later revised), topology instability (an interaction pattern that failed to converge), or projection inconsistency (a semantic structure that linearized ambiguously).

These are precisely the signals that would be needed to learn the admissibility manifold directly.

SRN Learning Signals

An SRN-native learning procedure would expose the model to a richer signal structure:

- **Constraint persistence statistics**: which constraint patterns survive repeated relaxation across diverse contexts. Persistent constraint structures are admissibility invariants; transient ones are context-specific.
- **Rollback statistics**: which local stabilizations are frequently reversed during later relaxation. High rollback frequency signals that a tentative crystallization is located in a metastable basin rather than a genuine admissibility minimum.
- **Projection stability**: which stabilized configurations linearize consistently under pressure-field variation. High E_{proj} (Definition 4.2) during training indicates a configuration whose semantic structure is underspecified—a genuine admissibility deficiency, not a sampling artifact.
- **Recursive extension success rate**: how often a configuration is successfully extended without increasing total tension. Low extension success signals proximity to the boundary of a dead-end basin—the training analogue of $E_{\text{ext}} > 0$.

The Relationship to Existing Training Methods

This speculative account suggests a continuity rather than a break with existing practice. Reinforcement learning from human feedback (RLHF) and related methods already introduce some of these signals: rollback is approximated by negative reward, global coherence by preference comparisons across full outputs. The SRN learning framework would sharpen these signals by making their geometric interpretation explicit: reward is not arbitrary human preference but a measurement of admissibility geometry at the appropriate scale.

The pressure field $\mathcal{P}$ is, in this view, the most directly learnable component: it is the linguistically structured mapping from semantic dependency order to communicative sequence, and this mapping is directly observable in any corpus of sequential text. Grammar is the stabilized compression of pressure-field statistics across a language community. An SRN that has learned $\mathcal{P}$ accurately has, in effect, internalized the communicative conventions of its training distribution as a linearization operator rather than as a statistical pattern over token sequences.

10.4. For Cognitive Science

The SRN account of language generation aligns with and extends several frameworks in cognitive science.

The free energy principle [14] models neural computation as minimization of variational free energy across hierarchical generative models. Semantic tension minimization in SRNs is formally analogous, with admissibility replacing the model evidence bound.

Predictive processing accounts of language [16] emphasize global constraint satisfaction over local sequential prediction. SRNs formalize this intuition in a generative framework.

The claim that sequential language is a projection of non-sequential cognitive processes—communicative traces rather than cognitive substrate—is consistent with psycholinguistic evidence for massive parallelism in comprehension and hierarchical planning in production.

10.5. For Physics

The formal structure of the SRN framework draws deeply on physical analogies: gradient flow, constraint propagation, metastable field evolution, and phase stabilization [23, 24, 22].

The topology evolution equation (3) has structural parallels to renormalization group flow: as the system relaxes, the effective interaction topology at each scale is determined by the constraint structure at finer scales. Whether this analogy can be made mathematically precise—whether SRN dynamics define a genuine renormalization flow on semantic configuration space—is an open and potentially interesting question.

The treatment of hallucination as topological defect propagation draws on condensed matter physics. Whether the analogy yields computationally useful predictions (e.g., whether defect density can be estimated from early-stage stabilization dynamics) is a direction for future investigation.

11. The Projection Operator: Grammar as Communicative Pressure Field

11.1. The Linearization Problem

Principle 7.1 introduced the readout projection $\pi : \mathcal{A} \to \Sigma^*$ as the map from stabilized admissibility configurations to sequential symbol strings. We noted that this projection is induced by communicative conventions rather than by the stabilization dynamics themselves. This section gives that claim formal content.

The problem is as follows. A stabilized configuration $X^* \in \mathcal{A}$ is a partially ordered structure: its sub-regions stand in various constraint and dependency relations, but these relations do not in general determine a unique linear order. Multiple linear sequences may be consistent with the same admissibility structure—this is the formal underpinning of the observation that the same propositional content can be expressed in many different word orders, sentence constructions, or discourse arrangements.

The question is: what selects among them?

The answer we propose is that grammar and discourse structure function as *communicative pressure fields*: constraint systems that act on the stabilized configuration and induce a preferred (though not always unique) linearization.

11.2. Partial Orders and Linear Extensions

11.3. The Constraint Graph of Stabilized Configurations

Before introducing the pressure field, we sharpen the representation of X^* by making its graph structure explicit.

Definition 11.1 (Semantic Constraint DAG)**.** Given a fully stabilized configuration $X^* \in \mathcal{A}$, the *semantic constraint graph* $G(X^*)$ is a directed acyclic graph (DAG) whose:

- **Vertices** $V = \{X_1, \dots, X_n\}$ are the stabilized semantic regions of X^*;
- **Edges** $(X_i \to X_j) \in E$ encode communicative precedence constraints: X_i must be communicated before X_j for X_j to be interpretable in context; and

- **Edge weights** $w_{ij} \in [0, 1]$ represent the strength of the precedence constraint, with $w_{ij} = 1$ indicating a hard ordering requirement and $w_{ij} \to 0$ indicating a soft preference.

Acyclicity is guaranteed by the semantic dependency relation: if X_i must precede X_j, then X_j cannot also be required to precede X_i without contradiction, which would violate condition (C1) of admissibility.

The DAG $G(X^*)$ encodes the genuine semantic ordering structure of the configuration. It is distinct from the constraint topology $\mathcal{C}_t$ used during relaxation: $\mathcal{C}_t$ is a weighted graph over which tension propagates during stabilization, whereas $G(X^*)$ is the residual directed structure that persists after stabilization is complete and governs how the stabilized content is communicated.

The projection operator π then acts on $G(X^*)$ rather than directly on $\mathcal{A}$:

$$\pi(X^*) \;=\; \pi_{G(X^*)},$$

where $\pi_{G(X^*)}$ is a linear extension of the partial order induced by $G(X^*)$. The space of all such extensions is:

$$\mathrm{Lin}(G(X^*)) \;=\; \big\{\, \ell : V \to \{1, \ldots, n\} \;\big|\; (X_i \to X_j) \in E \Rightarrow \ell(X_i) < \ell(X_j) \,\big\}\,.$$

This is the set of all orderings of the stabilized regions that respect the hard precedence constraints of the DAG.

We treat the stabilized configuration X^* as a finite partially ordered set (poset) $(\mathcal{S}^*, \leq_s)$, where $\mathcal{S}^*$ is the set of stabilized semantic regions and $\leq_s$ is the *semantic dependency order*: $X_i \leq_s X_j$ if X_i must be communicated before X_j for X_j to be interpretable by the receiver.

Not all pairs of regions stand in this order; many are semantically parallel and may be communicated in either order. The semantic dependency order is therefore genuinely partial.

Definition 11.2 (Communicative Pressure Field)**.** A *communicative pressure field* $\mathcal{P}$ on a stabilized poset $(\mathcal{S}^*, \leq_s)$ is a function

$$\mathcal{P} : \mathcal{S}^* \times \mathcal{S}^* \to \mathbb{R}$$

assigning a preference weight $\mathcal{P}(X_i, X_j)$ to each ordered pair, representing the degree to which communicative conventions favour X_i preceding X_j in the output sequence. The pressure field encodes grammatical, prosodic, and pragmatic constraints as a soft total order on $\mathcal{S}^*$.

The components of $\mathcal{P}$ correspond to distinct layers of communicative convention:

- **Syntactic pressure** $\mathcal{P}_{\text{syn}}$: head-dependent ordering constraints, argument structure templates, and phrase-structural precedence rules.
- **Information-structural pressure** $\mathcal{P}_{\text{info}}$: given-before-new, topic-before-focus, and discourse cohesion conventions.
- **Prosodic pressure** $\mathcal{P}_{\text{pros}}$: rhythmic and intonational constraints that prefer certain orderings for phonological well-formedness.

The total communicative pressure is a weighted combination:

$$\mathcal{P} \;=\; \alpha\,\mathcal{P}_{\text{syn}} \;+\; \beta\,\mathcal{P}_{\text{info}} \;+\; \gamma\,\mathcal{P}_{\text{pros}},$$

with language-specific and register-specific weighting parameters $\alpha, \beta, \gamma \geq 0$.

11.4. The Projection as Optimal Linear Extension

Given the partial order $\leq_s$ and the communicative pressure field $\mathcal{P}$, the projection π selects the linear extension of $(\mathcal{S}^*, \leq_s)$ that maximally satisfies $\mathcal{P}$.

Formally, let $\mathcal{L}(\mathcal{S}^*, \leq_s)$ denote the set of all linear extensions of the poset—all total orders on $\mathcal{S}^*$ consistent with $\leq_s$. The projection selects:

$$\pi(X^*) \;=\; \arg \min_{\ell \,\in\, \text{Lin}(G(X^*))} \Psi(\ell), \tag{4}$$

where

$$\Psi(\ell) \;=\; -\sum_{(X_i, X_j):\ell(X_i)<\ell(X_j)} \mathcal{P}(X_i, X_j) \;+\; \lambda \sum_{(X_i \to X_j)\in E} w_{ij}\,\mathbf{1}[\ell(X_i) \geq \ell(X_j)] \tag{5}$$

is the *communicative projection cost*. The first term penalizes orderings that violate soft communicative preferences; the second term penalizes violations of hard DAG precedence constraints (weighted by constraint strength w_{ij}), with $\lambda \gg 1$ ensuring hard constraints dominate. The minimizer $\ell^* = \pi(X^*)$ is the most communicatively efficient linear traversal of $G(X^*)$ consistent with the semantic dependency structure.

Interpretation 11.1. *Equation* (4) *formalizes grammar as preference optimization over the space of linear extensions of the semantic dependency order. This is not a generative grammar in the Chomskyan sense—it does not define well-formedness by rule application. It is a* constraint satisfaction *characterization: the grammatical output is the linear arrangement that best satisfies* *the communicative pressure field while respecting the partial semantic dependency order.*

Several consequences follow.

First, grammaticality is graded under this formalization. *A sequence is not simply grammatical or ungrammatical; it receives a score from the pressure field, and sequences may be more or less well-formed depending on how well they satisfy communicative constraints. This aligns with psycholinguistic data showing that acceptability judgments form a continuum rather than a binary.*

Second, cross-linguistic variation in word order is explained directly. *Different languages have different pressure field parameters: a verb-final language has $\mathcal{P}_{\text{syn}}$ heavily weighting head-final configurations; a verb-second language has strong constraints on the second-position placement of the finite verb. The admissibility structure underlying a proposition may be the same across languages; what differs is the pressure field that linearizes it.*

Third, the difficulty of free word-order languages is explained. *Languages with relatively free word order (Latin, Finnish, many Indigenous Australian languages) have weakly specified $\mathcal{P}_{\text{syn}}$, placing heavier weight on $\mathcal{P}_{\text{info}}$ and $\mathcal{P}_{\text{pros}}$. The linearization space $\mathcal{L}(\mathcal{S}^*, \leq_s)$ is large, and the pressure field provides weaker guidance. This corresponds to the empirical observation that such languages tend to have richer information-structural and prosodic systems to compensate.*

Remark 11.1. *The optimization in equation (4) is in general NP-hard (linear extension counting is #P-complete for arbitrary posets). This does not undermine the theoretical claim; it suggests that grammatical linearization is a computationally non-trivial process, consistent with psycholinguistic evidence for the difficulty of production planning. In practice, the semantic dependency order $\leq_s$ for natural language utterances is shallow (few levels deep) and locally decomposable, making the optimization tractable in realistic cases.*

11.5. Connection to Language Evolution

The pressure field formalism connects naturally to evolutionary accounts of language. On the SRN account, proto-language would consist of stabilized semantic configurations communicated without a well-defined pressure field: holistic, unordered, or inconsistently ordered outputs. The emergence of grammar corresponds to the cultural and biological stabilization of a communicative pressure field $\mathcal{P}$ that converts the partial order $\leq_s$ into reliable sequential conventions.

Grammaticalization, on this view, is the progressive hardening of initially soft pressure field weights into obligatory ordering constraints: what begins as a statistical preference becomes a categorical rule as the pressure weight $\mathcal{P}(X_i, X_j)$ for certain region pairs approaches infinity.

This account is consistent with usage-based and construction grammar approaches to language [18], while providing a formal grounding in the admissibility geometry that those approaches lack.

11.6. A Unified Account of Linguistic Phenomena

The projection framework—stabilized admissibility configurations linearized under communicative pressure fields—provides a unified explanatory account of a wide range of linguistic and communicative phenomena that are otherwise treated as unrelated problems.

Translation

Translation, under the SRN account, is not conversion between two sequential symbol strings. It is the application of a different communicative pressure field to the same underlying stabilized admissibility structure.

Given X^* and two pressure fields $\mathcal{P}_L$ and $\mathcal{P}_{L'}$ corresponding to languages L and L' respectively, the translations are:

$$\pi_L(X^*) \;=\; \arg \min_{\ell \in \mathrm{Lin}(G(X^*))} \Psi_L(\ell), \qquad \pi_{L'}(X^*) \;=\; \arg \min_{\ell \in \mathrm{Lin}(G(X^*))} \Psi_{L'}(\ell).$$

Translation difficulty arises when the DAGs $G_L(X^*)$ and $G_{L'}(X^*)$ differ—when the two languages impose different communicative precedence constraints on the same semantic content, requiring not merely a different linearization but a partially different dependency structure. This explains the asymmetry of translation difficulty: some language pairs share deep structural compatibility in their DAGs; others require genuine structural reanalysis.

Poetry and Deliberate Pressure-Field Manipulation

Poetry, under the SRN account, is deliberate manipulation of the communicative pressure field $\mathcal{P}$. The poet selects a linearization ℓ that is sub-optimal under the standard pressure field (it violates expected syntactic or information-structural preferences) but optimal under a modified field that weights phonological, rhythmic, and associative constraints more heavily:

$$\mathcal{P}_{\text{poetry}} \;=\; \alpha' \, \mathcal{P}_{\text{syn}} \;+\; \beta' \, \mathcal{P}_{\text{info}} \;+\; \gamma' \, \mathcal{P}_{\text{pros}} \;+\; \delta \, \mathcal{P}_{\text{assoc}},$$

with $\gamma' \gg \gamma$ and $\delta > 0$ introducing associative pressure not present in ordinary discourse. The tension between the standard and poetic pressure fields—between what grammar "expects" and what the poem "does"—is the source of poetic defamiliarization and aesthetic tension.

Humor as Admissibility Bifurcation

Humor, particularly the joke form, arises from admissibility bifurcation: a configuration that admits two incompatible but locally stable interpretations, with a rapid forced transition between them.

The setup of a joke stabilizes configuration X_A (interpretation A) as the expected admissibility basin. The punchline introduces a constraint that makes X_A inadmissible and simultaneously reveals that configuration X_B (interpretation B, unexpected) has been globally admissible all along. The humor arises from the rapid topology reorganization—the rollback from X_A to X_B—and the recognition that two incompatible admissibility basins coexisted throughout the setup. The sudden relinearization of the DAG $G(X_\text{B})$ under the established pressure field produces the characteristic surprise of the punchline.

Coreference as Delayed Stabilization Dependency

Coreference resolution—determining that two noun phrases refer to the same entity—is, under SRN, a case of delayed cross-regional stabilization.

Region X_i (encoding *the president*) and region X_j (encoding *she*) cannot independently stabilize their referential content; their semantic coherence depends on a shared stabilization of the entity node to which both are linked. Until that shared stabilization occurs, $T(i, j)$ remains elevated, keeping both regions in an intermediate state. Long-distance coreference difficulty in human processing reflects the cost of maintaining two unstabilized regions in active coupling over extended relaxation.

Semantic Drift as Metastable Basin Wandering

Semantic drift—the gradual loss of coherence in extended generation—is, under SRN, the consequence of repeated settlement into local admissibility minima that are slightly incompatible with earlier stabilized regions.

As generation proceeds, new semantic regions are stabilized and added to the constraint graph. If the verification field $\mathcal{V}$ does not enforce global consistency across all scales, each successive region may stabilize in a basin that is locally consistent with its immediate neighbors but globally inconsistent with earlier regions. Over many steps, the accumulated inconsistencies produce a text that is locally coherent paragraph by paragraph but globally incoherent as a whole.

This is precisely the failure mode of large language models in extended generation tasks: each local window appears fluent, but global thematic and argumentative coherence degrades. Under SRN, the remedy is not a larger context window but a stronger hierarchical verification field operating at the discourse scale.

RSVP Interpretation: Projection as Symmetry Breaking

Having established the linguistic and computational account, we can now add the physical interpretation. In RSVP terms, the projection π is a symmetry-breaking map from a higher-dimensional admissibility manifold into the one-dimensional communicative channel.

The stabilized configuration X^* lives in a high-dimensional space with many equivalent symmetry-related representations—different orderings of the same semantic content that are equivalent under the admissibility geometry. The communicative pressure field $\mathcal{P}$ functions as a symmetry-breaking field: it selects one particular linear representation from among many admissible ones, just as a physical symmetry-breaking field selects one ground state from among a degenerate manifold.

This analogy is not merely decorative. It suggests that linguistic communication is structurally similar to phase transitions: the movement from a high-symmetry, high-dimensional admissibility structure to a low-symmetry, low-dimensional communicative trace. Grammar is the broken symmetry. Communicative conventions are the order parameter. The communicative pressure field is the external field that triggers the symmetry breaking.

12. Convergence and Stability of Relaxation Dynamics

12.1. The Stability Question

Principle 4.1 asserts that language generation in an SRN is minimization of the global tension functional $\Phi(X)$ over the admissibility space $\mathcal{A}$. Proposition 5.1 asserts that the joint dynamics of equations (2) and (3) converge toward a structure encoding the semantic dependencies of the output.

These claims require a stability argument. The recursive relaxation operator $\mathcal{R}$ must be shown, under appropriate conditions, to drive the system toward a fixed point (or fixed-point neighborhood) rather than producing oscillatory or divergent dynamics.

We develop this argument in two stages: first a Lyapunov stability result for the tension functional, then a contraction result for the relaxation operator under additional regularity conditions.

12.2. Lyapunov Stability of the Tension Functional

We treat $\Phi : \mathcal{A} \to \mathbb{R}_{\geq 0}$ as a candidate Lyapunov function for the discrete-time relaxation dynamics. For Φ to serve this role, we require:

(L1) $\Phi(X) \geq 0$ for all $X \in \mathcal{A}$, with $\Phi(X) = 0$ if and only if X is fully stabilized (all local instability measures vanish and all pairwise tensions are zero).

(L2) $\Phi(X_{t+1}) \leq \Phi(X_t)$ for all t, with equality only at fixed points of $\mathcal{R}$.

Condition (L1) holds by construction: $S_i(X) \geq 0$ and $T(i,j) \geq 0$ for all regions and pairs, so $\Phi(X) = \sum_i S_i + \sum_{i<j} T(i,j) \geq 0$, with $\Phi(X) = 0$ iff all summands vanish.

Condition (L2) is the substantive requirement. It is satisfied when $\mathcal{R}$ is constructed to perform *admissibility-preserving descent*: each application of $\mathcal{R}$ either reduces Φ or, if no reducing step exists in $\mathcal{A}$, activates the rollback mechanism to return to a prior checkpoint from which alternative descent paths are available.

Proposition 12.1 (Lyapunov Stability)**.** Suppose the admissibility space $\mathcal{A}$ is finite (or compact in an appropriate topology) and the relaxation operator $\mathcal{R}$ satisfies:

(R1) $\mathcal{R}$ is admissibility-preserving: $X_{t+1} \in \mathcal{A}$ whenever $X_t \in \mathcal{A}$.

(R2) $\Phi(X_{t+1}) \leq \Phi(X_t)$ at each step, with strict inequality unless X_t is a fixed point.

(R3) The rollback mechanism ensures that if no admissibility-preserving descent step exists from X_t, the operator returns to the most recent checkpoint $X_{t'}$ with $t' < t$ for which an unexplored descent path remains.

Then the sequence $(X_t)_{t\geq 0}$ converges to a fixed point $X^* \in \mathcal{A}$ with $\Phi(X^*)$ locally minimal.

Proof sketch. By (R2), the sequence $(\Phi(X_t))_{t\geq 0}$ is non-increasing and bounded below by zero. Therefore it converges to a limit $\Phi^* \geq 0$. By (R1) and compactness of $\mathcal{A}$, the sequence (X_t) has a convergent subsequence; call its limit X^*. Continuity of Φ (which holds when the instability and tension measures are continuous in the configuration) implies $\Phi(X^*) = \Phi^*$.

If X^* is not a fixed point of $\mathcal{R}$, then by (R2) there exists a step producing strict decrease, contradicting convergence of $\Phi(X_t)$ to Φ^*. Therefore X^* is a fixed point.

The rollback condition (R3) ensures the process does not terminate at a non-minimal local saddle by guaranteeing that all reachable descent paths from each checkpoint are explored before the process halts. In the finite case this terminates in finite time; in the compact continuous case, standard arguments from Zorn's lemma apply to the partially ordered set of reachable checkpoints. □

Remark 12.1. *Proposition 12.1 establishes convergence to a* local *minimum of Φ within the admissibility space, not necessarily a global minimum. This is appropriate: natural language generation does not require globally optimal semantic configurations, only ones that are sufficiently stable for communicative purposes. The hierarchical verification field $\mathcal{V}$ (Section 6) provides additional structure that filters out locally stable but globally inadmissible configurations, effectively* *raising the quality floor of the local minima reached by the relaxation dynamics.*

12.3. Contraction of the Relaxation Operator

Under stronger regularity conditions, the relaxation operator can be shown to be a contraction on an appropriate metric space, yielding both convergence and a convergence rate estimate.

Equip $\mathcal{A}$ with a metric $d_{\mathcal{A}}$ defined by:

$$d_{\mathcal{A}}(X,Y) \;=\; |\Phi(X)-\Phi(Y)| \;+\; \sum_i |S_i(X)-S_i(Y)| \;+\; \sum_{i<j} |T_X(i,j)-T_Y(i,j)|\,,$$

which measures divergence both in global tension and in the local structure of instability and pairwise coupling.

Proposition 12.2 (Contraction)**.** Suppose that the tension functional Φ is L-smooth on $\mathcal{A}$ (i.e., $\|\nabla\Phi(X) - \nabla\Phi(Y)\| \leq L\, d_{\mathcal{A}}(X,Y)$) and that the relaxation step size is chosen as $\eta = 1/L$. Then $\mathcal{R}$ is a contraction on $(\mathcal{A}, d_{\mathcal{A}})$ with contraction constant $\kappa < 1$, and the fixed point X^* is unique within the contraction neighborhood. Moreover, the convergence rate satisfies:

$$d_{\mathcal{A}}(X_t, X^*) \;\leq\; \kappa^t\, d_{\mathcal{A}}(X_0, X^*).$$

Interpretation 12.1. *Proposition 12.2 should be read with care. The smoothness assumption on Φ is strong and may not hold globally on $\mathcal{A}$—semantic configuration spaces need not be smooth manifolds, and the tension functional may have sharp gradients near constraint boundaries.*

The proposition is therefore best understood as a local result: within a neighborhood of a stable admissibility basin, the relaxation dynamics contract at a geometric rate. This is consistent with the phenomenology of language generation: once a configuration has entered a coherent semantic region, subsequent stabilization proceeds quickly; it is the initial approach to admissibility basins that is slow and revision-prone.

This also provides a formal account of why generation is easier *for highly constrained contexts. Strong constraints* (*high ω_{ij} coupling, sharp $\mathcal{P}_{\text{syn}}$ pressure*) *reduce the diameter of the admissibility basin, which—when the basin is well-separated from others—increases the contraction constant and accelerates convergence. Analogously, a constrained writing task* (*sonnet, legal brief, formal proof*) *is in some respects easier to complete coherently than an unconstrained one, because the pressure field strongly guides linearization.*

12.4. Metastability and the Topology of Failure

The convergence results above apply within well-behaved admissibility basins. But the SRN dynamics also admit *metastable* configurations: states that satisfy (L2) locally (no immediate descent step is available) but are not globally minimal in Φ.

Metastable configurations correspond to semantically coherent but suboptimal outputs: texts that are locally consistent but fail to realize the most natural or most tension-free expression of the underlying admissibility structure. They are not hallucinations (which are topological instability events), but they are imperfect crystallizations.

The hierarchical verification field $\mathcal{V}$ addresses metastability by checking admissibility at multiple scales. A configuration that passes local verification (scale k) may fail at a coarser scale $(k+1)$, triggering rollback to a checkpoint preceding the metastable region.

In physical terms, the verification hierarchy functions as *simulated annealing* on the admissibility manifold: by periodically accepting rollback (thermal fluctuation) at higher scales, the dynamics can escape metastable traps and continue toward deeper admissibility basins. The annealing schedule corresponds to the order in which verification scales are applied during generation.

13. Conclusion

We have proposed Semantic Relaxation Networks as a theoretical framework in which language generation is reconceived as recursive constraint stabilization over a partially ordered admissibility manifold.

The central claims of the paper are:

1. The sequential decomposition $p(x) = \prod_i p(x_i \mid x_{<i})$ is computationally convenient rather than cognitively fundamental; it mistakes communicative trace for cognitive substrate.
2. Admissibility—defined by local constraint coherence, recursive extensibility, and perturbation stability under topology deformation—is a more appropriate generative criterion than likelihood.
3. Language generation is minimization of a global semantic tension functional $\Phi(X)$ via recursive relaxation operators over a dynamically evolving constraint topology.
4. Tokens are not primitive units but emergent crystallizations: readout projections of locally stabilized admissibility regions.
5. Sequential language is the observable residue of higher-dimensional stabilization dynamics, not its computational substrate.

6. Hallucination is premature local collapse without hierarchical admissibility verification, not miscalibrated conditional probability.
7. Grammar functions as a communicative pressure field $\mathcal{P}$ that selects preferred linear extensions of the semantic dependency order; cross-linguistic variation in word order is variation in pressure field parameters.
8. The relaxation dynamics are Lyapunov-stable under conditions (R1)–(R3), converging to locally minimal admissibility configurations; under additional smoothness, the operator is contractive with geometric convergence rate.

These claims are currently theoretical. The SRN framework does not yet have a full implementation, and many of the formal objects introduced—the tension functional, the topology reconfiguration operator, the verification field, the communicative pressure field—require further mathematical development before they can be translated into working architectures.

We regard this incompleteness as appropriate. The paper's primary contribution is ontological reframing rather than architectural specification. Before SRNs can be built, it must be established that there is something genuinely new to build—that language generation is a problem of constraint stabilization rather than conditional prediction, and that grammar is a linearization pressure rather than a generative rule system. We believe the case developed above is sufficient to motivate that research direction.

Sequential language is not the substrate of cognition. It is the residue left behind by recursive constraint stabilization processes occurring over high-dimensional admissibility fields. The task of a language model, properly understood, is not to emit the next token. It is to stabilize.

References

[1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. In *Advances in Neural Information Processing Systems*, 2017.

[2] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, 2020.

[3] Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. Language Models are Unsupervised Multitask Learners. OpenAI Technical Report, 2019.

[4] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of NAACL*, 2019.

[5] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising Diffusion Probabilistic Models. In *Advances in Neural Information Processing Systems*, 2020.

[6] Yang Song, Jascha Sohl-Dickstein, Diederik P. Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-Based Generative Modeling through Stochastic Differential Equations. In *International Conference on Learning Representations*, 2021.

[7] Jacob Austin, Daniel D. Johnson, Jonathan Ho, Daniel Tarlow, and Rianne van den Berg. Structured Denoising Diffusion Models in Discrete State-Spaces. In *Advances in Neural Information Processing Systems*, 2021.

[8] Aaron Lou, Chenlin Meng, and Stefano Ermon. Diffusion Language Models Can Perform Many Tasks with Scaling and Instruction-Finetuning. arXiv:2401.17297, 2024.

[9] Marianne Arriola, Subham Sekhar Sahoo, Aaron Gokaslan, Zhihan Yang, Zhixuan Qi, Jiaqi Han, Justin T. Chiu, and Volodymyr Kuleshov. Block Diffusion: Interpolating Between Autoregressive and Diffusion Language Models. In *International Conference on Learning Representations*, 2025.

[10] Artidoro Pagnoni, Julie Kallini, Tomasz Limisiewicz, Gargi Ghosh, Luke Zettlemoyer, Christopher Potts, Xiaochuang Han, and Srinivasan Iyer. Byte Latent Transformer. arXiv preprint, 2025.

[11] Julie Kallini, Artidoro Pagnoni, Tomasz Limisiewicz, Gargi Ghosh, Luke Zettlemoyer, Christopher Potts, Xiaochuang Han, and Srinivasan Iyer. Fast Byte Latent Transformer. arXiv:2605.08044, 2026.

[12] Yann LeCun, Sumit Chopra, Raia Hadsell, Marc'Aurelio Ranzato, and Fu-Jie Huang. A Tutorial on Energy-Based Learning. In *Predicting Structured Data*. MIT Press, 2006.

[13] Yann LeCun. A Path Towards Autonomous Machine Intelligence. OpenReview Position Paper, 2022.

[14] Karl Friston. The Free-Energy Principle: A Unified Brain Theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[15] Karl Friston, James Kilner, and Lee Harrison. A Free Energy Principle for the Brain. *Journal of Physiology-Paris*, 100(1–3):70–87, 2006.

[16] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann, 1988.

[17] Geoffrey E. Hinton and Terrence J. Sejnowski. Optimal Perceptual Inference. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 1983.

[18] Paul Smolensky. Information Processing in Dynamical Systems: Foundations of Harmony Theory. In *Parallel Distributed Processing*. MIT Press, 1986.

[19] John J. Hopfield. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences*, 79(8):2554–2558, 1982.

[20] David H. Ackley, Geoffrey E. Hinton, and Terrence J. Sejnowski. A Learning Algorithm for Boltzmann Machines. *Cognitive Science*, 9(1):147–169, 1985.

[21] Julian Barbour, Tim Koslowski, and Flavio Mercati. Identification of a Gravitational Arrow of Time. *Physical Review Letters*, 113(18), 2014.

[22] Julian Barbour. *The Janus Point: A New Theory of Time*. Basic Books, 2020.

[23] Giorgio Parisi. *Statistical Field Theory*. Addison-Wesley, 1988.

[24] Kenneth G. Wilson and John Kogut. The Renormalization Group and the ϵ Expansion. *Physics Reports*, 12(2):75–199, 1974.

[25] Mikhail Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Birkhäuser, 2007.

[26] Saunders Mac Lane. *Categories for the Working Mathematician*. Springer, 1998.

[27] Claude E. Shannon. A Mathematical Theory of Communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

[28] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, 2006.

[29] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[30] Karl M. Fant. *Computer Science Reconsidered: The Invocation Model of Process Expression*. Wiley-Interscience, Hoboken, NJ, 2007. ISBN: 978-0-471-79814-9.