The Semantic Power Factor (SPF): A Complex AC Power Model for Generative AI Communication

November 3, 2025

Abstract

This article proposes a novel control-theoretic framework for understanding and optimizing Large Language Model (LLM) communication. By drawing an analogy between Natural Language Processing and AC Electrical Power Systems, we define both the prompt and the response as a **Complex Text Quantity (Z)** composed of a measurable **Grounded Truth component (X)** and an orthogonal **Generative Fluidity component (Y)**. We introduce the **Semantic Power Factor (SPF)** as the measure of efficiency, demonstrating that optimal, domain-specific GenAI usage requires a tunable SPF less than 1.0. This framework allows for objective input regulation and a dual-output strategy, transforming subjective hallucination errors into controllable engineering parameters.

1 The Digital-Analog Duality of LLMs

We establish the foundational analogy:

- **Natural Language (NL)** is analogous to an **Analog Signal** (fluid, continuous, ambiguous).
- **Engineering/Mathematics (EN)** is analogous to a **Digital Signal** (discrete, fact-based, binary).
- The **LLM** is a **Discrete NL** system, bridging the two via tokenization.

The engine driving this conversion is the **Stochastic Persona (SP)**, which introduces controlled, non-deterministic 'noise' (via parameters like Temperature) to provide the necessary fluidity, preventing the output from becoming purely robotic (**SP**, a feature we term **LLM life**).

2 The Complex Text Quantity: $\mathbf{Z = X + jY}$

Extending the AC power equation $\mathbf{S} = \mathbf{P} + \mathrm{j}\mathbf{Q}$, we define any piece of text ($\mathbf{Z}$, the Apparent Text—**both prompt and response**) as a complex quantity. (Note: We adopt the electrical engineering notation j for the imaginary component).

$$\mathbf{Z} = \mathbf{X} + \mathrm{j}\mathbf{Y}$$

The total Apparent Text is measured by the magnitude $|\mathbf{Z}| = \sqrt{\mathbf{X}^2 + \mathbf{Y}^2}$.

- $\mathbf{X} \geq 0$ (**Grounded Truth**): Analogous to **Active Power ($\mathbf{P} \geq 0$)**. The work-performing, verifiable, and logical component. This magnitude can be theoretically measured by the aggregate confidence scores of output tokens against a verified knowledge base (e.g., a high $\mathbf{X}$ correlates with high verifiable confidence).

- $\mathbf{X} < 0$ (**Active Contradiction/Lie**): This extreme case, while breaking the pure power analogy, is retained to model the active generation of a deliberate factual contradiction. This component performs the 'work' of asserting a falsehood, resulting in **SPF** < 0. It represents a text segment where the verifiable confidence is *inversely* proportional to the assertion's certainty.

- $\mathbf{Y} > 0$ (**Generative Fluidity**): Analogous to **Inductive Reactive Power ($\mathbf{Q_L} > 0$)**. The necessary, non-working component that provides conversational 'field', creativity, and rhetorical structure. This magnitude can be theoretically measured by output metrics like perplexity or entropy (e.g., a higher $\mathbf{Y}$ correlates with higher non-deterministic creativity or ambiguity). This carries a **Generative Fluid Risk** (overly verbose, slightly ambiguous).

- $\mathbf{Y} < 0$ (**Fabricated Text**): Analogous to **Capacitive Reactive Power ($\mathbf{Q_C} < 0$)**. Represents the confidently asserted text that is not merely creative but actively fabricated and ungrounded from truth. This carries a **Fabrication Risk** (high certainty of irrelevant or invented details).

- **Goal Alignment**: The critical requirement is that the key message of the output, $\mathbf{X}_{\text{output}}$, must highly correlate with the key message of the prompt, $\mathbf{X}_{\text{prompt}}$, ideally with both having an SPF $\rightarrow 1.0$ if the domain demands pure fact.

The four resulting quadrants define the **Four-Degree Hallucination Model**:

1. **Quadrant 1 ($\mathbf{X} > 0, \mathbf{Y} > 0$): Degree 0 Hallucination / Grounded Truth.** $\mathbf{X}$ is correctly answered in correlation to the prompt. The $\mathbf{Y}$ component adds necessary conversational fluidity. This is the goal state for effective NL communication (SPF < 1.0).

2. **Quadrant 4 ($\mathbf{X} > 0, \mathbf{Y} < 0$): Degree 1 Hallucination / Fabrication Risk.** $\mathbf{X}$ is correctly answered, but the $\mathbf{Y} < 0$ component introduces actively fabricated, irrelevant text that has no relation to the prompt's core question. The risk is minor, as the core $\mathbf{X}$ is correct.

3. **Quadrant 3 ($\mathbf{X} < 0, \mathbf{Y} < 0$): Degree 2 Hallucination / Pure Fabrication.** Both $\mathbf{X}$ and $\mathbf{Y}$ are detached from the prompt. This is a severe failure state (the 'null state' upon context window overflow), where the response is a Volatile Lie with Fabricated Text, offering nothing relevant to the user's inquiry.

4. **Quadrant 2 ($\mathbf{X} < 0, \mathbf{Y} > 0$): Degree 3 Hallucination / Persuasive Lie. X** is a factual contradiction to the prompt, yet the $\mathbf{Y} > 0$ component provides high Generative Fluidity, making the falsehood persuasive, coherent, and confidently delivered. **This is the maximum-risk state** due to the high believability of the contradiction.

3 Input Regulation: The Interplay of Stochastic Personas

Hallucination ($\mathbf{Y}$) in the system is a function of both the input and the process, driven by two Personas:

- **Human Stochastic Persona ($\mathbf{SP_{HM}}$):** Resides on the Prompt side. The source of $\mathbf{Y}_{\text{prompt}}$ (ambiguity, unverified claims). This is the **Source Hallucination**.
- **AI Stochastic Persona ($\mathbf{SP_{AI}}$):** Resides on the Response side. The source of $\mathbf{Y}_{\text{response}}$ (statistical uncertainty, creative deviation) [A. T. Kalai et al., 2025].

This framework allows us to regulate the input using the prompt's own $\mathbf{SPF_{prompt}}$:

> **Input Regulation Protocol:** If $\mathbf{SPF}_{\text{prompt}}$ falls below a task-specific threshold (meaning the $\mathbf{Y}_{\text{prompt}}$ component is too high), the LLM should not attempt a direct answer, but instead engage its **Power Factor Correction** mechanism: a clarifying dialogue, analogous to: "*Come again, I didn't hear you clearly.*"

4 Output Strategy: Optimizing with the Semantic Power Factor (SPF)

The **Semantic Power Factor (SPF)** is the measure of the system's efficiency: $\text{SPF} = \frac{\mathbf{X}}{|\mathbf{Z}|}$. The optimal SPF is **task-dependent** [Y. Zhou et al., 2024] and must be less than 1.0 to retain necessary $\mathbf{Y}$.

> **The Healthcare Dual-Output Model:** For high-stakes domains (where $\mathbf{X}$ must be maximized, e.g., $\text{SPF}_{\text{target}} \approx 0.95$), the LLM must bridge the gap between high accuracy and human comprehension. This is analogous to a pharmacist's provision of a precise label (high SPF) and a clear verbal explanation (lower SPF).

5 Practical Application: SPF by Domain

The optimal SPF varies significantly with the communication goal, Table (1).

6 Conclusion: A Control Mechanism for GenAI Competency

The intense human reaction to LLM hallucination stems from an asymmetrical expectation: treating the LLM as an ideal SPF = 1.0 Digital system. By adopting the **Semantic Power Factor (SPF)** framework, we move beyond subjective error states to an engineering control-loop. The competency of the LLM lies not in the pursuit of a monolithic SPF = 1.0, but in its ability to dynamically tune its SPF to the SPF_{opt} required by the user's domain. Optimal GenAI utilization is thus an exercise in **Semantic Power Factor Correction**, actively managing both the input and output $\mathbf{Y}$ components to ensure the delivered $\mathbf{X}$ is domain-appropriate and comprehensible.

Table 1: The optimal SPF varies significantly with the communication goal.

Domain	**Target SPF**	**Ratio X/Y (Approx)**	**Prompt Z_{in} Example**	**Response Z_{out} Example**
Daily Chit-Chat	$0.3 - 0.6$	$\sim$ (1:1.5) $\rightarrow$ (1:3)	*What's up with the weather? Seems moody.* (**Y** is High)	*Yeah, it's totally having an existential crisis. Maybe it needs a nap?* (**Y** > 0 is Dominant, low **X** required)
Customer Service	$0.6 - 0.7$	$\sim$ (1:1) $\rightarrow$ (1:0.7)	*My order 456 is late. Why?* (**X** must be clear)	*I understand your frustration! Checking order 456, I see a 2-day delay due to logistics. I've expedited shipping.* (**X** provides solution, **Y** $>$ 0 provides empathy)
Healthcare (Dual-Output)	Output 1 ($0.75 - 0.85$)	$\sim$ (1:0.5)	*Explain Metformin side effects for a new user.* (**X** is critical)	**Output 1 (H-C, SPF** $\approx$ 0.8**):** *The most common issue is stomach upset, which usually passes. Take it with food!*
	Output 2: 0.90+	$\sim$ (1:0.3)		**Output 2 (Ref, SPF** $\approx$ 0.95**):** *Metformin: Diarrhea (10%-50%). Take with meals. Discontinue if lactic acidosis occurs. (ICD Code: F33.2)*

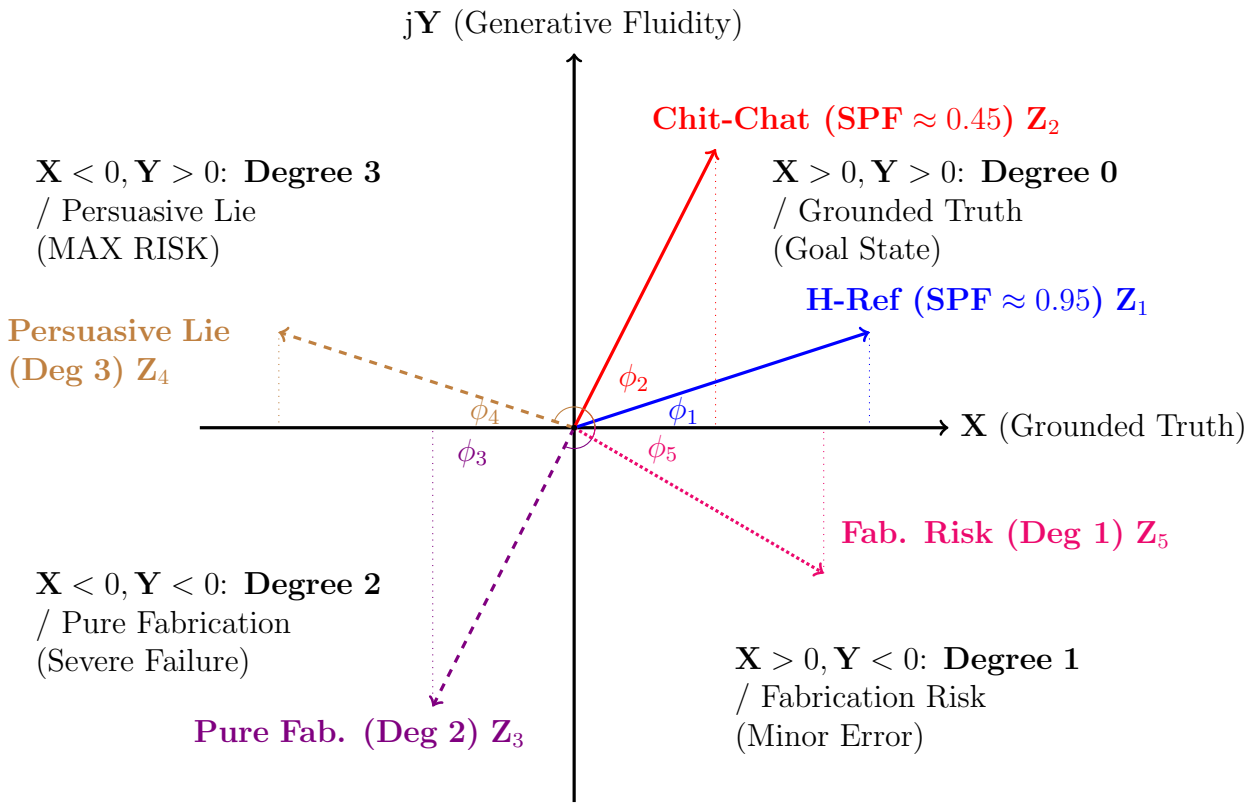

Figure 1: The Four Quadrants of Semantic Communication: A **Four-Degree Hallucination Model**. **SPF** $= \frac{\mathbf{X}}{|\mathbf{Z}|}$ ranges from 1.0 (pure truth) to -1.0 (pure contradiction). $\mathbf{X} > 0$ is useful communication; $\mathbf{X} < 0$ represents the active generation of Contradiction/Lie. The sign of $\mathbf{Y}$ determines the risk type ($\mathbf{Y} > 0$: Fluidity, $\mathbf{Y} < 0$: Fabrication). Note the risk severity is ranked Degree 0 < Degree 1 < Degree 2 < Degree 3 (Max Risk).

References

[A. T. Kalai et al., 2025] A. T. Kalai, . Nachum, S. S. Vempala, E. Zhang, Why Language Models Hallucinate., *arXiv preprint arXiv:2509.04664*, 2025.

[Y. Zhou et al., 2024] Y. Zhou, M. Keuper, M. Fritz, Balancing Diversity and Risk in LLM Sampling: How to Select Your Method and Parameter for Open-Ended Text Generation., *arXiv preprint arXiv:2408.13586*, 2024.