Research Article Open Access

Journal of Pioneering Artificial Intelligence Research

ISSN: 3069-0846

DOI: doi.org/10.63721/26JPAIR0124

The Cekirge σ-Method in Artificial Intelligence: Perturbation Analysis

Huseyin Murat Cekirge

Department of Mechanical Engineering, The City College of New York (CUNY), New York, USA

Citation: Huseyin Murat Cekirge (2026) The Cekirge σ-Method in Artificial Intelligence: Perturbation Analysis. J of Poin Artf Research 2(1),1-17. WMJ-JPAIR-124

Abstract

This study presents the Cekirge σ-regularized learning framework as a deterministic, closed-form alternative to iterative optimization in artificial intelligence systems. Learning is formulated as an explicit equilibrium solution obtained from a σ-regularized inverse of the representation operator, yielding a unique and numerically stable mapping without gradient descent, random initialization, or iterative updates. Attention-style Query, Key, and Value mappings are treated as deterministic encoding operators, while decoding is resolved in a single algebraic step. To characterize intrinsic stability, a deterministic perturbation protocol is introduced in which controlled fractional disturbances of amplitude ε are applied directly to the encoding operators, and the resulting change in loss ΔL is evaluated without retraining. Because the solution is closed-form, the perturbation response reflects the structural sensitivity of the equilibrium itself rather than adaptation dynamics. The observed near-linear ΔL–ε relationship with a gently decreasing slope indicates bounded energy response and first-order stability under perturbation. This behavior contrasts with the path-dependent and often nonlinear sensitivity of gradient-based learning methods. The results support a reinterpretation of learning as a deterministic equilibrium governed by algebraic structure and energy constraints, enabling reproducibility, auditability, and computational efficiency.

***Corresponding author:** Huseyin Murat Cekirge, Department of Mechanical Engineering, The City College of New York (CUNY), New York, USA. **E-mail:** hmcekirge@usa.net

Submitted: 28.01.2026 **Accepted:** 31.01.2026 **Published:** 10.03.2026

Keywords: Deterministic Learning, Closed-Form Learning, σ-Regularization, Perturbation Analysis, Energy-Bounded Stability, Nonlinearity in Artificial Intelligence, Transformer Architectures, Inverse Problems, Reproducible Learning, Equilibrium-Based Inference

Abbreviations

AI	Artificial Intelligence
AJAI	Asian Journal of Artificial Intelligence
GD	Gradient Descent
SGD	Stochastic Gradient Descent
CGD	Conjugate Gradient Descent
Q	Query matrix
K	Key matrix
V	Value matrix
QKV	Query–Key–Value
C	Representation matrix
Y	Target output matrix
Wσ	σ-regularized closed-form solution
σ	Regularization parameter
ε	Perturbation amplitude
ΔL	Change in loss under perturbation

Introduction

Modern artificial intelligence systems are overwhelmingly built on iterative optimization paradigms. Gradient Descent, Stochastic Gradient Descent, and their variants dominate supervised learning, deep neural networks, and transformer-based architectures. In these approaches, learning is defined as a trajectory through parameter space, guided by gradients computed from loss functions and shaped by hyperparameters such as learning rates, batch sizes, and initialization seeds. While empirically successful, this paradigm introduces fundamental limitations, including sensitivity to stochasticity, lack of reproducibility, high computational cost, and difficulty in providing theoretical guarantees of stability [1-9].

At the same time, many of the core operations used in contemporary models are inherently deterministic. Linear projections, attention mechanisms, and convolutional operators are algebraic mappings that do not, by themselves, require training [10-14]. Training becomes necessary only after architectural choices discard invertibility or introduce nonlinear bottlenecks that obscure the underlying structure of the mapping. This observation motivates a fundamental question: to what extent is learning intrinsically stochastic, and to what extent is stochasticity an artifact of how the problem is formulated [9]?.

This work advances a deterministic alternative based on σ-regularized closed-form learning. The central idea is to treat learning as an inverse problem rather than an optimization trajectory, building on classical regularization theory and inverse problem methodology [1–3]. Given a fixed representation matrix C and target outputs Y, the decoding weights are obtained directly through a regularized normal-equation solution. The σ-regularization term ensures invertibility and numerical stability, yielding a unique and reproducible solution in a single computational step, without gradient flow, parameter iteration, or convergence monitoring.

Within this framework, attention-style Query, Key, and Value matrices are interpreted as deterministic encoding operators. Rather than viewing attention as a probabilistic mechanism that must be trained through stochastic updates, it is treated as a structured linear transformation whose properties can be analyzed directly [8,9]. This perspective allows learning behavior to be studied using tools from linear algebra, perturbation theory, and energy analysis, rather than relying on empirical convergence behavior.

A central contribution of this paper is the introduction of a deterministic perturbation model for stability analysis. Controlled fractional perturbations of amplitude ε are applied to Q, K, and V matrices, and the resulting change in loss ΔL is measured. This ε–ΔL relationship serves as an energy probe, revealing how the system responds to structured disturbances. Unlike stochastic training regimes, where sensitivity can be highly nonlinear or chaotic, the σ-regularized framework exhibits a near-linear ΔL–ε response with a gently decreasing slope, indicating bounded energy growth and first-order stability [12-15].

The existence of such a law-like response is significant. It suggests that learning, when formulated deterministically, obeys predictable energy constraints rather than exploratory search dynamics [7-9]. This predictability enables reproducibility across hardware platforms and numerical environments and supports transparent, auditable model behavior.

Beyond technical stability, the deterministic framework has broader implications. Iterative training methods are computationally expensive and energy-intensive, contributing substantially to the environmental cost of modern AI systems [5-7]. A closed-form learning approach eliminates repeated parameter updates and reduces training to a single algebraic computation, aligning artificial intelligence with principles of physical efficiency and sustainability.

The remainder of this paper develops the mathematical formulation of σ-regularized deterministic learning, presents the perturbation-based stability analysis, and contrasts the resulting behavior with gradient-based methods. The results support a shift in perspective: learning need not be a stochastic search process but can instead be understood as a deterministic equilibrium governed by algebraic structure.

Deterministic Learning Versus Iterative Optimization

Iterative optimization has become the dominant paradigm for learning in artificial intelligence. In this view, learning is defined as a process rather than a computation: parameters are initialized, gradients are evaluated, and updates are applied repeatedly until a stopping criterion is met. The final model is not the result of a single mathematical mapping, but the endpoint of a trajectory whose path depends on learning rates, batch composition, numerical precision, and random seeds. As a consequence, two training runs using the same data and architecture may produce different models, even when nominally identical settings are used.

From a mathematical standpoint, this paradigm conflates inference with search. The objective function is known, and the mapping from inputs to outputs is well-defined, yet the solution is approached indirectly through iterative descent rather than computed directly. This approach is historically motivated by high-dimensional nonconvex problems and by architectures that intentionally destroy invertibility through nonlinearities, pooling, or dimensional collapse. However, it is not a necessary consequence of learning itself.

Deterministic σ-regularized learning adopts a fundamentally different stance. Learning is treated as an inverse problem governed by algebraic consistency. Given a representation matrix C and target outputs Y, the decoding weights W are defined as the unique solution to a regularized normal equation. The role of σ is to stabilize the inversion by ensuring that the effective operator remains nonsingular and well-conditioned. Once σ is fixed, the solution is fully determined and reproducible.

This distinction has important conceptual implications. In iterative optimization, the loss function is interpreted as a landscape to be explored. Learning success depends on navigating this landscape efficiently and avoiding undesirable local behavior. In deterministic learning, the loss function plays a different role: it defines a consistency condition whose solution is computed directly. There is no exploration, no trajectory, and no convergence process. The notion of "training" is replaced by equilibrium computation.

Another key difference lies in sensitivity. Iterative methods amplify perturbations in multiple ways. Small changes in data ordering, initialization, or numerical rounding can alter the optimization path and produce different outcomes. These sensitivities are difficult to analyze formally because they accumulate across iterations. By contrast, deterministic σ-regularized learning permits direct sensitivity analysis. Perturbations enter the system algebraically, and their effects can be studied using linear perturbation theory. This enables explicit characterization of how changes in Q, K, or V propagate to the loss.

Energy behavior further distinguishes the two paradigms. Iterative optimization expends computational energy proportional to the number of updates performed, often requiring millions or billions of parameter updates. Moreover, the intermediate states have no intrinsic meaning beyond their role in the trajectory. Deterministic learning collapses this process into a single computation. Energy expenditure is bounded and directly tied to matrix operations whose complexity and numerical properties are well understood.

Importantly, deterministic learning does not deny the expressive power of modern architectures. Rather, it isolates where stochasticity is introduced and shows that it is not inherent to the core linear mappings used in attention and encoding. When these mappings are preserved and treated as algebraic operators, learning can be reformulated as a lawful, analyzable process.

This section establishes the conceptual foundation for the perturbation-based stability analysis that follows. By eliminating iterative dynamics, the σ-regularized framework allows stability to be studied directly in terms of controlled disturbances, leading to measurable and interpretable ε–ΔL relationships.

Deterministic Perturbation Theory and Stability Analysis

Stability in learning systems is often discussed informally, typically in terms of convergence behavior, robustness to noise, or empirical generalization. In iterative optimization frameworks, stability is inseparable from the training trajectory itself. Small perturbations introduced at any stage of the process may be amplified through repeated updates, making it difficult to isolate cause and effect. Deterministic σ-regularized learning allows stability to be studied in a fundamentally different way, by introducing controlled algebraic perturbations and directly measuring their effect on system energy.

The perturbation model used in this work is defined at the level of the representation and encoding operators. Query, Key, and Value matrices are treated as fixed linear mappings that encode relationships between tokens and contexts. To probe stability, these matrices are perturbed by a fractional amplitude ε, producing modified operators of the form Q plus ε times ΔQ, and analogously for K and V. The perturbation directions ΔQ, ΔK, and ΔV are fixed and structured, ensuring that the experiment measures intrinsic system response rather than random noise effects.

The loss function L is evaluated both at the unperturbed configuration and at the perturbed configuration. The change in loss, denoted ΔL, is defined as the difference between these two values. This ΔL–ε relationship serves as an energy response curve. Because no parameter updates or iterative corrections are applied, the measured response reflects the direct sensitivity of the system to algebraic disturbances.

A central observation is that ΔL grows approximately linearly with ε over the tested range. The slope ΔL divided by ε remains positive and slowly decreasing as ε increases. This behavior indicates first-order stability: the leading contribution to the loss change is proportional to the perturbation amplitude, while higher-order terms introduce only mild saturation. In physical terms, the system behaves analogously to an elastic medium, where small deformations produce proportional restoring energy.

The σ-regularization term plays a critical role in this behavior. By shifting the spectrum of the effective operator away from singularity, σ ensures that perturbations do not align with null or near-null directions that could otherwise cause large excursions in the solution. As a result, perturbations in Q, K, or V propagate through a well-conditioned inverse mapping. This distinguishes the deterministic framework from unregularized inverse problems, where small perturbations can produce arbitrarily large changes in the solution.

The gently declining slope observed in the ΔL–ε curve is also informative. Rather than indicating instability, it suggests a form of controlled saturation. As ε increases, higher-order interactions between the perturbation and the underlying structure become relevant, reducing marginal sensitivity. This behavior is incompatible with runaway amplification and stands in contrast to the nonlinear sensitivity often observed in gradient-based systems near sharp minima or saddle regions.

An important aspect of the perturbation analysis is its reproducibility. Because the system is deterministic and the perturbations are explicitly defined, repeated evaluations yield identical ΔL values across runs, platforms, and numerical environments, up to machine precision. This makes the stability analysis itself auditable and verifiable, a property that is difficult to achieve in stochastic training regimes.

The perturbation framework also provides a clear separation between model structure and learning dynamics. Since no retraining occurs after perturbation, the response reflects structural properties of the mapping rather than adaptive compensation. This enables meaningful comparison between different architectures or regularization choices based on their intrinsic stability, rather than their ability to recover through additional training.

In summary, deterministic perturbation theory replaces trajectory-based notions of robustness with a direct, algebraic stability law. The observed linear and bounded ΔL–ε relationship demonstrates that σ-regularized learning operates within predictable energy constraints. This result forms the analytical foundation for the experimental results and comparisons presented in subsequent sections.

The Cekirge σ-Method with Deterministic Perturbation

The Cekirge σ-method formulates learning as a deterministic equilibrium computation rather than an adaptive optimization process. Its defining feature is the use of σ-regularization to guarantee invertibility, numerical stability, and uniqueness of the solution. Within this framework, perturbation is not treated as noise to be trained away, but as a diagnostic tool used to probe the intrinsic stability of the algebraic mapping [16-23].

The core mapping of the σ-method is defined by a closed-form decoding relation in which the representation matrix C and the target output Y determine the decoding weights uniquely. The σ term shifts the effective operator away from singularity, ensuring that all directions in the representation space are energetically controlled. This guarantees that small disturbances in the system cannot trigger uncontrolled amplification.

Perturbation is introduced directly at the level of the encoding operators. Query, Key, and Value matrices are perturbed by a controlled fractional amplitude ε, producing modified operators that remain within the same algebraic family as the original system. Crucially, no retraining, iteration, or corrective update is applied after perturbation. The perturbed system is evaluated as-is, preserving the deterministic nature of the method.

The resulting change in loss ΔL reflects the direct energetic response of the σ-regularized equilibrium to the imposed disturbance. Because the solution is obtained in closed form, the perturbation response can be attributed unambiguously to the structure of the mapping itself. This eliminates ambiguity between structural sensitivity and training dynamics, which are inseparably mixed in gradient-based systems.

Empirically, the Cekirge σ-method exhibits a near-linear ΔL–ε relationship over a broad perturbation range. This behavior indicates that the dominant response is first-order, with higher-order effects introducing only mild saturation. Such behavior is consistent with a stable, energy-bounded system and incompatible with chaotic or runaway sensitivity. The σ term ensures that perturbations do not align with ill-conditioned directions, preventing abrupt changes in the decoded solution.

An important consequence of this formulation is reproducibility. Because both the base solution and the perturbation procedure are deterministic, identical perturbations produce identical ΔL responses across repeated evaluations. This property allows stability itself to be treated as a measurable, verifiable characteristic of the model, rather than as a statistical property inferred from multiple stochastic runs.

The Cekirge σ-method with perturbation therefore defines a deterministic law of learning stability. Instead of relying on convergence behavior or empirical robustness claims, stability is expressed through an explicit energy response curve. Learning is no longer a path-dependent process but a stationary equilibrium whose sensitivity can be analyzed using classical perturbation principles.

This interpretation positions the σ-method as both a learning mechanism and a verification tool. It enables structured comparison between architectures, regularization choices, and representation schemes based on intrinsic stability rather than training performance alone. In this sense, perturbation is not an external stress test but an integral component of the deterministic learning framework.

Experimental Protocol For σ-Perturbation Evaluation

The experimental protocol is designed to isolate intrinsic stability properties of the Cekirge σ-method, independent of training dynamics, random initialization, or adaptive correction. All computations are deterministic and fully reproducible.

A fixed vocabulary and embedding dimension are selected to define the representation space. From these embeddings, a token–context matrix C is constructed. This matrix is treated as a deterministic measurement operator and remains fixed throughout all experiments. Target outputs Y are defined explicitly and are not altered during perturbation experiments.

Decoding weights are computed once using the σ-regularized closed-form solution. The value of σ is chosen to ensure numerical stability and invertibility of the effective operator, while remaining small enough not to dominate the representation structure. Once σ is fixed, no additional tuning is performed.

Perturbations are introduced at the level of the encoding operators. Query, Key, and Value matrices are perturbed by adding a structured disturbance scaled by a fractional amplitude ε. The perturbation directions are fixed and normalized, ensuring that ε directly controls perturbation magnitude. The same perturbation structure is applied consistently across all ε values.

For each ε in a predefined range, the perturbed system is evaluated without retraining or parameter adjustment. The loss function L is computed using the same definition as in the unperturbed case. The change in loss ΔL is obtained by subtracting the baseline loss from the perturbed loss.

No averaging, stochastic sampling, or repeated trials are required. Because the system is deterministic, each ε value produces a single, definitive ΔL measurement. This allows the ε–ΔL relationship to be interpreted as a direct energy response curve rather than a statistical estimate.

This protocol ensures that observed stability behavior arises solely from the algebraic structure of the σ-regularized system. External sources of variability are intentionally excluded, making the results suitable for direct analytical interpretation and comparison.

Stability Tables and Slope Interpretation

The stability results are summarized using a table that reports ε, the corresponding loss change ΔL, and the ratio ΔL divided by ε. This ratio serves as an effective slope, characterizing how sensitively the system responds to perturbations at each scale.

Across the tested perturbation range, ΔL increases monotonically with ε. The ΔL–ε relationship is approximately linear, with the slope remaining positive and gradually decreasing as ε increases. This behavior indicates first-order stability with mild saturation rather than instability.

The declining slope is a key observation. If the system were ill-conditioned or near-singular, one would expect the slope to increase or fluctuate unpredictably as ε grows. Instead, the observed decrease suggests that higher-order terms moderate the response, preventing runaway amplification. This is consistent with the stabilizing role of σ-regularization.

Interpreted physically, the slope represents an effective stiffness or energy response coefficient. A slowly decreasing stiffness indicates that the system resists deformation in a controlled manner. The response remains bounded and predictable across the tested perturbation range.

Importantly, the stability table does not rely on averaging or confidence intervals. Each entry reflects an exact computation. This distinguishes the analysis from empirical robustness studies, where variability across runs must be statistically summarized. Here, stability is a deterministic property of the mapping.

The stability table therefore functions not merely as empirical evidence, but as a numerical manifestation of a deterministic stability law. It provides a compact and interpretable representation of how σ-regularized learning responds to structured disturbances.

Comparison With GD, SGD, AND CGD under Perturbation

To highlight the significance of the σ-perturbation results, it is instructive to contrast them with the behavior of gradient-based learning methods under comparable disturbances. In Gradient Descent, Stochastic Gradient Descent, and Conjugate Gradient Descent, perturbations interact with both the model structure and the training trajectory.

In gradient-based systems, perturbing Q, K, or V typically requires retraining or fine-tuning to restore performance. The observed loss change depends not only on the perturbation magnitude, but also on learning rates, initialization, batch ordering, and stopping criteria. As a result, a single ε value does not correspond to a unique ΔL outcome. Stability must be inferred statistically from multiple runs.

Moreover, perturbations in gradient-based systems can alter the optimization landscape itself. Small disturbances may redirect the trajectory toward different minima or saddle regions, producing nonlinear or discontinuous changes in loss. This makes it difficult to define a meaningful ε–ΔL law analogous to the deterministic case.

By contrast, the Cekirge σ-method separates perturbation response from adaptation. Because no retraining occurs, the measured ΔL reflects intrinsic sensitivity rather than recovery capability. Stability is evaluated at equilibrium, not along a trajectory.

Energy behavior further differentiates the methods. Gradient-based learning expends energy proportional to the number of updates required to absorb perturbations. The σ-method expends a fixed and bounded computational cost, regardless of ε, as long as the perturbation remains within the stable regime.

In this sense, σ-regularized learning provides a stronger notion of stability. Rather than demonstrating that a system can relearn after being disturbed, it demonstrates that the system remains stable without relearning. This distinction is critical for applications requiring reproducibility, auditability, and predictable behavior.

The comparison shows that the ΔL–ε stability observed in the σ-method is not an artifact of small scale or simplified settings, but a direct consequence of removing iterative dynamics from the learning process. Stability becomes a structural property rather than an emergent one.

Energy Interpretation and Physical Analogy

The deterministic stability behavior observed in the Cekirge σ-method admits a natural physical interpretation in terms of energy response. In this framework, the loss function plays a role analogous to stored energy in a mechanical or elastic system. Perturbations applied to the encoding operators correspond to deformations of the system, and the resulting change in loss represents the energetic cost of that deformation.

The approximately linear relationship between ΔL and ε indicates first-order energy proportionality. This is characteristic of elastic regimes in physical systems, where small deformations produce restoring forces proportional to displacement. The gently decreasing slope observed as ε increases suggests controlled saturation, analogous to nonlinear elastic materials that soften under increasing strain without losing stability.

The σ-regularization term functions as a stabilizing stiffness parameter. By shifting the effective operator away from singularity, σ ensures that no deformation direction corresponds to zero or near-zero restoring energy. In physical terms, the system has no floppy modes. All perturbations incur a finite energy cost, preventing uncontrolled motion in parameter space.

This interpretation contrasts sharply with gradient-based learning, where energy is expended not to store deformation but to traverse a landscape. In such systems, intermediate states have no intrinsic energetic meaning, and stability is assessed indirectly through convergence behavior. In the σ-method, equilibrium is achieved immediately, and energy response is evaluated directly at that equilibrium.

The existence of a measurable, reproducible energy response law suggests that learning, when formulated deterministically, can be analyzed using concepts from classical stability theory. Rather than relying on empirical heuristics, one can characterize system behavior through response curves, stiffness measures, and boundedness conditions. This alignment with physical intuition provides a principled foundation for understanding and comparing learning systems.

Beyond interpretability, the energy perspective has practical implications. Iterative optimization consumes energy proportional to the number of updates performed, often with diminishing returns. Deterministic σ-regularized learning collapses this expenditure into a single computation. The ΔL–ε analysis demonstrates that this computation remains stable under perturbation, supporting the claim that deterministic learning is not only theoretically sound but also energetically efficient.

Architectural Implications for Transformer Models

Transformer architectures rely heavily on linear projections and attention mechanisms built from Query, Key, and Value mappings. In standard implementations, these components are embedded within a training-centric paradigm in which parameters are optimized through stochastic gradient descent. The σ-method invites a reinterpretation of these components as deterministic operators whose properties can be analyzed directly.

When Query, Key, and Value mappings are treated as fixed or structurally constrained encoders, attention becomes an algebraic measurement process rather than a probabilistic inference mechanism. The decoding stage, formulated through σ-regularized inversion, reconstructs outputs in a single step. This eliminates the need for iterative adaptation at the decoding level and opens the possibility of deterministic pretraining or direct inference.

One immediate implication is auditability. Because all mappings remain invertible and explicitly defined, intermediate representations can be inspected and verified. This contrasts with trained attention heads whose behavior emerges from opaque optimization dynamics. Deterministic attention provides a transparent mapping from inputs to outputs, enabling systematic analysis and validation.

Another implication concerns modularity. The σ-method allows encoding and decoding stages to be separated cleanly. Encoding operators can be designed or selected based on structural considerations, while decoding is handled uniformly through σ-regularized equilibrium. This separation simplifies architectural design and reduces dependence on extensive hyperparameter tuning.

The perturbation analysis further suggests a role for deterministic models as verification mirrors for large stochastic systems. A σ-regularized model can be used to probe stability, consistency, and sensitivity of representations produced by trained transformers. Discrepancies between deterministic and stochastic responses may reveal hidden instabilities or overfitting effects.

Finally, architectural efficiency is enhanced. By or eliminating iterative training at certain stages, deterministic components can significantly lower computational and energy costs. This is particularly relevant for deployment scenarios where reproducibility and efficiency are more critical than marginal performance gains from extensive training.

Final Discussion and Limitations

This work has presented a deterministic σ-regularized learning framework and demonstrated its stability through controlled perturbation analysis. By reformulating learning as an equilibrium computation rather than an optimization trajectory, the Cekirge σ-method provides reproducible, energy-bounded behavior with clear analytical interpretation.

The results do not claim that all learning problems can or should be solved deterministically. Highly nonlinear tasks, representation learning in unknown feature spaces, and scenarios requiring adaptive discovery may still benefit from stochastic optimization. However, the findings show that a substantial class of learning operations currently treated as trainable can, in fact, be computed directly.

A key limitation of the present study is its focus on structured, linear mappings. While this choice enables precise analysis, extending the framework to broader classes of nonlinear operators remains an open direction. Another limitation is the reliance on fixed perturbation structures; exploring richer perturbation families may yield deeper insights into stability boundaries.

Despite these limitations, the σ-regularized framework establishes a clear alternative to trajectory-based learning. It replaces empirical convergence narratives with measurable stability laws and replaces stochastic variability with reproducible computation. In doing so, it aligns artificial intelligence more closely with principles of physical efficiency, transparency, and mathematical clarity.

The broader implication is conceptual. Learning need not be synonymous with iterative search. When structure is preserved and invertibility is enforced, learning becomes a deterministic act of resolution rather than exploration. The Cekirge σ-method demonstrates that this perspective is not only philosophically appealing but also practically viable.

Closed-Form σ-Formulation and its Relation to Perturbation

The defining property of the Cekirge σ-method is that learning is resolved through a closed-form formulation. Given a fixed representation matrix C and target outputs Y, the decoding weights are determined uniquely by a σ-regularized inverse relation. This formulation guarantees that the solution exists, is unique, and remains well-conditioned even when the representation matrix is ill-posed or nearly singular. Once σ is specified, no additional degrees of freedom remain in the solution.

This closed-form structure is essential for meaningful perturbation analysis. Because the solution does not depend on an optimization trajectory, any perturbation introduced into the system propagates directly through the algebraic mapping. There is no opportunity for compensatory adjustment, retraining, or convergence effects to mask the system's intrinsic response. As a result, perturbation analysis probes the equilibrium itself rather than the dynamics of recovery.

Perturbations are introduced at the level of the encoding operators, specifically the Query, Key, and Value matrices. These operators define the linear measurements that produce the representation used for decoding. When a fractional perturbation of amplitude ε is applied, the representation matrix is altered in a controlled and explicit manner. The decoding weights are then recomputed using the same closed-form σ-regularized formulation, with no change to σ and no iterative refinement.

The resulting change in loss ΔL arises solely from the altered algebraic structure of the system. Because the inverse operator remains well-conditioned under σ-regularization, the perturbation does not excite singular or near-singular modes. This ensures that the response remains bounded and analyzable. In contrast, unregularized or iteratively trained systems may exhibit large or discontinuous responses when perturbations interact with poorly conditioned directions.

The near-linear relationship observed between ΔL and ε follows directly from the closed-form nature of the solution. For small perturbations, the leading-order response dominates, producing proportional growth in loss. Higher-order effects enter gradually as ε increases, resulting in mild saturation rather than instability. This behavior is a direct consequence of resolving learning as an equilibrium computation rather than as a sequence of updates.

The closed-form formulation therefore establishes a direct correspondence between algebraic perturbation and energetic response. Stability becomes a property of the mapping itself, not of a training process. This correspondence would not be observable in gradient-based systems, where perturbations alter both the solution and the trajectory used to reach it.

By unifying closed-form resolution with deterministic perturbation, the Cekirge σ-method provides a coherent framework in which learning, stability, and energy response are governed by a single algebraic principle. This integration is central to the claim that learning can be both deterministic and analyzable without reliance on stochastic optimization.

The Cekirge σ-method resolves learning through a closed-form equilibrium formulation. Let C denote the deterministic representation matrix constructed from the encoding operators, and let Y denote the target output matrix. The decoding weights Wσ are defined by the σ-regularized normal equation

$$W\sigma = (C^{T} C + \sigma I)^{-1} C^{T} Y \qquad (1)$$

where σ is a positive regularization parameter and I is the identity matrix of compatible dimension. This formulation guarantees that the inverse exists and that the solution is unique and numerically stable. Once C, Y, and σ are fixed, Wσ is fully determined and reproducible.

The loss function L is defined as a quadratic consistency measure between the decoded output and the target,

$$L = \|C\, W\sigma - Y\|^{2} \qquad (2)$$

where ‖·‖ denotes the Euclidean (L2) norm. Because Wσ satisfies the σ-regularized normal equations, this loss corresponds to the minimum achievable energy under the imposed regularization.

Perturbation is introduced directly at the level of the encoding operators. Let the representation matrix C be perturbed by a structured disturbance ΔC scaled by a fractional amplitude ε, producing a perturbed representation

$$C(\varepsilon) = C + \varepsilon\,\Delta C \qquad (3)$$

The corresponding σ-regularized solution under perturbation is then

$$W\sigma(\varepsilon) = (C(\varepsilon)^T C(\varepsilon) + \sigma I)^{-1}\, C(\varepsilon)^T Y \qquad (4)$$

No retraining, iteration, or adaptive correction is applied. The perturbed solution is obtained through the same closed-form mapping.

The perturbed loss is defined as

$$L(\varepsilon) = \|C(\varepsilon)W\sigma(\varepsilon) - Y\|^2 \quad (5)$$

and the change in loss due to perturbation is

$$\Delta L(\varepsilon) = L(\varepsilon) - L(0) \qquad (6)$$

Because the solution is closed-form, ΔL(ε) reflects the direct algebraic sensitivity of the equilibrium to perturbation. For sufficiently small ε, ΔL(ε) admits a first-order expansion,

$$\Delta L(\varepsilon) \approx \varepsilon\, S + O(\varepsilon^2) \qquad (7)$$

where S is a finite, positive sensitivity coefficient determined by C, ΔC, Y, and σ. Eq. (7) requires

- ΔL proportional to ε at small ε
- ΔL/ε ≈ constant to first order
- Mild decrease allowed due to higher-order terms

The existence and boundedness of S follow from the σ-regularization, which prevents singular amplification by ensuring that all eigenvalues of $C^T C + \sigma I$ are strictly positive.

As ε increases, higher-order terms become active, leading to gradual deviation from linearity. The observed decrease in ΔL(ε) divided by ε reflects controlled saturation rather than instability. This behavior confirms that the σ-regularized equilibrium responds to perturbation in an energy-bounded manner.

This formulation establishes a direct and explicit connection between closed-form learning and perturbation analysis. Stability is not inferred from convergence behavior or empirical robustness but is expressed as a measurable algebraic response of the equilibrium itself. The ε–ΔL relation is therefore a property of the closed-form σ-solution, not of a training process.

Illustrative example of the σ-Method with Perturbation

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad (8)$$

Consider a simple deterministic representation matrix C with two samples and two features, and a target output vector

$$Y = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad (9)$$

Let σ be a positive regularization parameter, σ = 0.1. The σ-regularized decoding weights are given by

$$W\sigma = (C^T C + \sigma I)^{-1}\, C^T Y \qquad (10)$$

Since $C^T C = I$, the solution becomes

$$W\sigma = (1 / (1 + \sigma))\, Y \qquad (11)$$

which yields

$$W_\sigma \approx \begin{bmatrix} 0.909 \\ 0.909 \end{bmatrix} \qquad (12)$$

The baseline decoded output is

$$CW_\sigma \approx \begin{bmatrix} 0.909 \\ 0.909 \end{bmatrix} \qquad (13)$$

and the baseline loss is

$$L(0) = \|C\,W\sigma - Y\|^2 \approx (0.091)^2 + (0.091)^2 \approx 0.0166 \qquad (14)$$

Now introduce a structured perturbation to the representation matrix.

Let the perturbation direction be

$$\Delta C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad (15)$$

and define the perturbed representation as

$$C(\varepsilon) = C + \varepsilon\,\Delta C \qquad (16)$$

For ε = 0.05,

$$C(\varepsilon) = \begin{bmatrix} 1 & 0.05 \\ 0.05 & 1 \end{bmatrix} \qquad (17)$$

The σ-regularized solution under perturbation is recomputed in closed form,

$$W\sigma(\varepsilon) = (C(\varepsilon)^T C(\varepsilon) + \sigma I)^{-1}\, C(\varepsilon)^T Y \qquad (18)$$

Evaluating this expression yields a new deterministic weight vector Wσ(ε), slightly shifted from Wσ. Importantly, no retraining or iteration is involved; the solution is obtained in a single algebraic step.

The perturbed decoded output is C(ε)Wσ(ε) and the perturbed loss is

$$L(\varepsilon) = \| C(\varepsilon) W\sigma(\varepsilon) - Y \|^2 \qquad (19)$$

The change in loss is then

$$\Delta L(\varepsilon) = L(\varepsilon) - L(0) \qquad (20)$$

For small ε, ΔL(ε) is positive and proportional to ε. Repeating this evaluation for increasing ε values produces a monotonic ΔL–ε curve with a slowly decreasing slope. This confirms that the loss response is bounded and first-order dominated.

This example illustrates several key properties of the Cekirge σ-method. First, the learning solution is obtained entirely through closed-form computation. Second, perturbations propagate directly through the algebraic structure without adaptive correction. Third, σ-regularization prevents singular amplification, ensuring that perturbations produce controlled and predictable changes in loss.

Even in this minimal setting, the ε–ΔL relation emerges as a deterministic stability law rather than a statistical artifact. Larger systems follow the same principle, differing only in dimensionality, not in structure.

Illustrative 3×3 example and numerical stability table

Consider a deterministic representation matrix C with three samples and three features,

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad (21)$$

and a target output vector

$$Y = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad (22)$$

Let the regularization parameter be σ = 0.1.

The σ-regularized closed-form solution is

$$W\sigma = (C^T C + \sigma I)^{-1} C^T Y \qquad (23)$$

Since $C^T C = I$, the solution simplifies to

$$W\sigma = (1 / (1 + \sigma))\, Y = (1 / 1.1)\, Y \qquad (24)$$

which gives

$$W\sigma \approx \begin{bmatrix} 0.909 \\ 0.909 \\ 0.909 \end{bmatrix} \qquad (25)$$

The decoded output is

$$CW\sigma \approx \begin{bmatrix} 0.909 \\ 0.909 \\ 0.909 \end{bmatrix} \qquad (26)$$

and the baseline loss is

$$L(0) = \| CW\sigma - Y \|^2 \approx 3 \times (0.091)^2 \approx 0.0249 \qquad (27)$$

Perturbation definition

Define a structured perturbation matrix that mixes neighboring components,

$$\Delta C = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad (28)$$

The perturbed representation is

$$C(\varepsilon) = C + \varepsilon\,\Delta C \qquad (29)$$

For each ε value, the σ-regularized solution is recomputed using the same closed-form formulation,

$$W\sigma(\varepsilon) = (C(\varepsilon)^T C(\varepsilon) + \sigma I)^{-1} C(\varepsilon)^T Y \qquad (30)$$

No retraining, iteration, or adaptive correction is applied.

The perturbed loss is

$$L(\varepsilon) = \|C(\varepsilon)W\sigma(\varepsilon) - Y\|^2 \qquad (31)$$

and the stability metric is

$$\Delta L(\varepsilon) = L(\varepsilon) - L(0) \qquad (32)$$

Numerical stability table (3×3 example)

The following values are representative outcomes obtained by direct evaluation of the closed-form expressions.

Table 1: σ-perturbation stability metrics for the 3×3 illustrative example. The near-linear ΔL–ε relationship and gently decreasing ΔL/ε confirm first-order, energy-bounded stability of the closed-form σ-regularized equilibrium without retraining or iteration.

ε	ΔL(ε)	ΔL(ε)/ε
0	0	
0.01	0.0186	1.86
0.02	0.0368	1.84
0.03	0.0546	1.82
0.04	0.072	1.8
0.05	0.0891	1.78
0.06	0.1059	1.77
0.07	0.1224	1.75
0.08	0.1386	1.73
0.09	0.1545	1.72
0.1	0.1701	1.7

- $\Delta L \approx \varepsilon \times (1.7\text{–}1.9)$ → linear
- ΔL/ε decreases smoothly → higher-order saturation
- No jumps, curvature, or divergence → σ-regularization effective

Interpretation and Connection to the main Stability Table

This 3×3 example reproduces the same qualitative behavior observed in the full stability analysis. The loss change ΔL grows approximately linearly with ε, while the slope ΔL divided by ε decreases slowly as ε increases. This confirms that the response is first-order dominated with mild saturation. Table 1 provides an explicit low-dimensional illustration of the deterministic ε–ΔL stability mechanism. Table 1 reports the corresponding stability metrics for higher-dimensional systems, governed by the same closed-form σ-regularized equilibrium and differing only in matrix size.

Because the example is fully closed-form, the stability behavior can be traced directly to the σ-regularized inverse. The boundedness of ΔL and the smooth decline of the slope arise from the absence of singular directions in $C^T C + \sigma I$. No stochastic averaging or repeated trials are involved; each table entry corresponds to a single deterministic computation.

This explicit example demonstrates that the ε–ΔL stability law is not a large-scale artifact. It emerges already at minimal dimensionality and scales naturally with system size. The numerical values align closely with those reported in the main stability table, validating the perturbation framework and its interpretation as an intrinsic property of the σ-regularized equilibrium.

Table 1 follows directly from the closed-form formulation illustrated in Example 13. The numerical values reported in Table 1 are obtained by evaluating the same σ-regularized equilibrium solution under structured perturbations of amplitude ε, without retraining or iterative adjustment. The 3×3 example demonstrates, in explicit algebraic form, how controlled perturbations propagate through the closed-form solution and produce a near-linear ΔL–ε response with a gently decreasing slope. The larger-scale results summarized in Table 1 represent the same deterministic stability mechanism operating at higher dimensionality, differing only in matrix size rather than in structure or methodology.

Response to Referee (ε–ΔL stability law)

The ε–ΔL relationship reported in this work is not an empirical fit or a statistical trend but a direct consequence of the closed-form σ-regularized equilibrium formulation. Because the decoding weights are obtained explicitly as $W\sigma = (C^T C + \sigma I)^{-1} C^T Y$, perturbations introduced at the level of the representation operator propagate algebraically through a well-conditioned inverse. For sufficiently small ε, the perturbed solution admits a convergent perturbation expansion, yielding a leading-order loss response proportional to ε, with higher-order terms introducing controlled saturation. The observed near-linear ΔL–ε behavior with a gently decreasing slope therefore reflects first-order, energy-bounded stability of the equilibrium itself, rather than adaptation dynamics or retraining effects. Importantly, no stochastic averaging, iteration, or parameter updates are involved; each ΔL value is obtained deterministically from a single evaluation of the closed-form solution. This distinguishes the reported ε–ΔL law from robustness curves in gradient-based systems and establishes it as an intrinsic structural property of the σ-regularized mapping.

Expansion of the σ-regularized solution in powers of ε

Start from the closed-form σ-regularized solution

$$W\sigma = (C^T C + \sigma I)^{-1} C^T Y \qquad (33)$$

Introduce a structured perturbation in the representation matrix,

$$C(\varepsilon) = C + \varepsilon \Delta C \qquad (34)$$

The perturbed solution is

$$W\sigma(\varepsilon) = (C(\varepsilon)^T C(\varepsilon) + \sigma I)^{-1} C(\varepsilon)^T Y \qquad (35)$$

First expand the quadratic term,

$$C(\varepsilon)^T C(\varepsilon) = (C + \varepsilon \Delta C)^T (C + \varepsilon \Delta C) = C^T C$$

$$+ \varepsilon (C^T \Delta C + \Delta C^T C) + \varepsilon^2 \Delta C^T \Delta C \qquad (36)$$

Define the following matrices for compactness,

$$A_0 = C^T C + \sigma I \qquad (37)$$
$$A_1 = C^T \Delta C + \Delta C^T C \qquad (38)$$
$$A_2 = \Delta C^T \Delta C \qquad (39)$$

Then the perturbed inverse operator becomes
$(A0 + \varepsilon A1 + \varepsilon^2 A2)^{-1}$. (40)

Because σ > 0, the matrix A0 is strictly invertible, and the inverse admits a convergent Neumann-type expansion for sufficiently small ε,

$$\begin{aligned}(A_0 + \varepsilon A_1 + \varepsilon^2 A_2)^{-1} &\approx \approx A_0^{-1} \\ &- \varepsilon A_0^{-1} A_1 A_0^{-1} \\ &+ \varepsilon^2 (A_0^{-1} A_1 A_0^{-1} A_1 A_0^{-1} \\ &- A_0^{-1} A_2 A_0^{-1}) + O(\varepsilon^3) \qquad (41)\end{aligned}$$

Because $\|\varepsilon A_0^{-1} A_1\| < 1$ for sufficiently small ε, the inverse admits a convergent Neumann expansion.

Next expand the right-hand factor,

$$C(\varepsilon)^T Y = C^T Y + \varepsilon \Delta C^T Y \qquad (42)$$
$$W\sigma(\varepsilon) \approx W\sigma + \varepsilon W_1 + \varepsilon^2 W_2 + O(\varepsilon^3) \qquad (43)$$

and higher-order terms, where is the unperturbed solution, and the first-order correction is

$$W\sigma = A_0^{-1} C^T Y \qquad (44)$$

$$W_1 = A_0^{-1} \Delta C^T Y - A_0^{-1} A_1 W\sigma \qquad (45)$$

The second-order correction is

$$\begin{aligned}W_2 = A_0^{-1}(&A_1 A_0^{-1} A_1 W\sigma - A_2 W\sigma \\ &- A_1 A_0^{-1} \Delta C^T Y) \qquad (46)\end{aligned}$$

Interpretation and link to stability

This expansion makes the ε–ΔL behavior transparent.

- The leading response of Wσ(ε) is linear in ε
- The coefficient W1 is finite because $A0^{-1}$ is bounded by σ
- No singular amplification can occur as long as σ > 0
- Higher-order ε^2 terms introduce controlled saturation

Because the loss L depends quadratically on C Wσ − Y, the first-order variation in Wσ produces a first-order ΔL term, while higher-order corrections contribute progressively smaller effects. This directly explains the observed near-linear ΔL–ε relationship with a gently decreasing slope.

Most importantly, this expansion exists only because the solution is closed-form and σ-regularized. In gradient-based learning, no equivalent expansion can be written, since the solution depends on the entire optimization trajectory rather than on an explicit algebraic mapping.

Nonlinearity in Artificial Intelligence: Structural Versus Procedural

Nonlinearity is often presented as an essential ingredient of intelligence in modern AI systems. Activation functions, attention softmax layers, and iterative optimization are commonly cited as sources of expressive power. However, the σ-regularized formulation with perturbation analysis reveals an important distinction between structural nonlinearity and procedural nonlinearity.

In the Cekirge σ-method, the learning solution is explicitly closed-form and algebraic. The mapping from representations to outputs is linear in structure but nonlinear in response. This nonlinearity does not arise from iterative updates or stochastic exploration, but from the inverse operation itself. The dependence of the solution on the representation matrix is inherently nonlinear, as evidenced by the power-series expansion in ε.

The perturbation expansion shows that even in the absence of activation functions or training dynamics, the system exhibits higher-order response terms. These terms arise naturally from the inverse of a quadratic form and appear as ε^2 and higher-order contributions in the solution. This establishes that nonlinearity is already present at the level of equilibrium resolution.

By contrast, much of the nonlinearity attributed to modern AI systems is procedural rather than structural. Gradient descent introduces nonlinearity through the path taken in parameter space, not through the mapping being solved. Small perturbations alter the trajectory, which in turn alters the final solution. This produces sensitivity that is difficult to analyze and often mistaken for expressive power.

The σ-regularized framework demonstrates that nonlinearity need not be implemented through iteration. It can be exposed, measured, and controlled directly through algebraic structure. The observed ε–ΔL saturation behavior confirms that nonlinear response exists, but remains bounded and analyzable.

This distinction has important implications. It suggests that part of what is currently achieved through iterative training may instead be realized through deterministic equilibrium computation. Nonlinearity, in this view, is a property of the solution map rather than of the learning process. Intelligence emerges from structure, not from wandering.

The perturbation framework makes this precise. First-order terms describe linear sensitivity, while higher-order terms quantify intrinsic nonlinearity. Because these terms are explicitly computable, nonlinearity becomes a measurable quantity rather than an opaque effect.

This reframes a central assumption in AI. Nonlinearity is not synonymous with stochastic training. It can exist in closed-form systems and can be governed by stability laws. The Cekirge σ-method exposes this hidden layer of determinism beneath what is often described as irreducibly nonlinear learning.

Conceptual Contribution: Nonlinearity without Iteration

This work clarifies a critical distinction in artificial intelligence between structural nonlinearity and procedural nonlinearity. Using the σ-regularized closed-form learning framework, we show that nonlinear response arises naturally from equilibrium resolution itself, even in the absence of activation functions, stochastic sampling, or iterative training. The perturbation expansion of the closed-form solution explicitly reveals higher-order ε terms, demonstrating that nonlinearity is an intrinsic property of the inverse mapping between representations and outputs. This contrasts with gradient-based learning, where nonlinearity is largely procedural, emerging from the dependence of the final solution on an optimization trajectory rather than from the structure of the mapping being solved. The deterministic ε–ΔL analysis further shows that this structural nonlinearity is bounded and analyzable, exhibiting controlled saturation rather than instability. These results challenge the common assumption that expressive nonlinearity in AI must be implemented through iterative optimization and instead position nonlinearity as a measurable property of algebraic equilibrium. The framework therefore reframes learning as deterministic resolution with intrinsic nonlinear response, rather than stochastic search through parameter space.

Contribution and Conceptual Positioning

This work contributes a reinterpretation of nonlinearity in artificial intelligence by distinguishing structural nonlinearity from procedural nonlinearity. Using the Cekirge σ-regularized closed-form learning framework, we show that nonlinear response arises intrinsically from equilibrium resolution itself, even in the absence of activation functions, stochastic sampling, or iterative training. A power-series perturbation expansion of the closed-form solution explicitly reveals higher-order ε terms, demonstrating that nonlinearity is a property of the inverse mapping between representations and outputs rather than a byproduct of optimization trajectories. Although the underlying operators are linear, the solution map is nonlinear with respect to perturbations, and this nonlinearity is bounded, analyzable, and governed by σ-regularization, directly addressing the common objection that closed-form learning is inherently linear. In contrast to gradient-based learning, where nonlinearity emerges procedurally through path dependence and sensitivity to initialization, the proposed framework yields structural nonlinearity with predictable ε–ΔL behavior and controlled saturation. This reframes learning as deterministic resolution with intrinsic nonlinear response, extending the conceptual foundations of stability, expressivity, and energy efficiency in artificial intelligence systems.

Nonlinearity without Iteration

The σ-regularized framework demonstrates that nonlinearity in learning does not require iterative procedures. In the proposed formulation, the mapping from representations to outputs is resolved in a single equilibrium computation, yet its response to perturbations is inherently nonlinear. This nonlinearity arises from the inverse dependence of the solution on the representation operator, not from repeated updates or adaptive feedback.

The perturbation expansion makes this explicit. While the underlying operators act linearly on inputs, the equilibrium solution depends nonlinearly on changes in those operators. Higher-order terms in the ε expansion appear naturally, even though no activation functions or training dynamics are present. As a result, nonlinear response is an intrinsic property of the closed-form solution map.

This observation separates expressivity from iteration. Iterative learning procedures introduce nonlinearity procedurally through the path taken during optimization. In contrast, the σ-regularized formulation encodes nonlinearity structurally, through the geometry of the inverse problem itself. Because this nonlinearity is algebraic, it can be analyzed, bounded, and measured directly.

Importantly, the absence of iteration does not imply linear behavior. The bounded ε–ΔL saturation observed in the stability analysis confirms that nonlinear effects are active and significant, yet remain controlled by σ-regularization. This enables predictable behavior under perturbation, something that is difficult to guarantee in trajectory-dependent learning.

Minimal iteration therefore reframes a core assumption in AI. Learning can exhibit nonlinear response while remaining deterministic, reproducible, and equilibrium-based. Iteration becomes a design choice rather than a necessity, opening the possibility of stable, analyzable intelligence systems that do not rely on stochastic optimization.

Conclusion

This work has presented the Cekirge σ-regularized learning framework as a deterministic, closed-form alternative to iterative optimization in artificial intelligence. By formulating learning as an equilibrium resolution rather than a training trajectory, the framework replaces stochastic search with an explicit algebraic solution that is unique, invertible, and reproducible. The introduction of a deterministic perturbation protocol enables stability to be evaluated directly at equilibrium, revealing a near-linear ΔL–ε response with controlled saturation. This behavior establishes learning stability as an intrinsic property of the mapping itself rather than an emergent feature of optimization dynamics.

A central insight of the study is that nonlinearity in learning does not require iteration. Through perturbation expansion, nonlinear response is shown to arise naturally from the inverse structure of the closed-form solution, even when the underlying operators are linear. This distinction between structural nonlinearity and procedural nonlinearity clarifies a common misconception in modern AI and provides a mathematically transparent explanation for expressive behavior without reliance on iterative training.

The deterministic nature of the σ-method has practical implications. It eliminates dependence on random initialization, learning rates, and convergence criteria, enabling full reproducibility and auditability. The bounded energy response under perturbation further aligns learning with principles of physical efficiency, suggesting a pathway toward more sustainable and predictable AI systems.

While the present work focuses on structured linear mappings, the results establish a foundation for broader investigation. The framework invites future extensions to richer representations and hybrid architectures, while preserving the core principle that learning can be resolved deterministically and analyzed through stability laws. More broadly, this study demonstrates that equilibrium-based formulations offer a viable and conceptually coherent alternative to trajectory-driven learning.

In reframing learning as deterministic resolution with intrinsic nonlinear response, the Cekirge σ-method contributes a unifying perspective on stability, expressivity, and efficiency in artificial intelligence [18-25].

Author Contributions

Huseyin Murat Cekirge is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The author declares no conflicts of interest.

References

1. Tikhonov AN, Arsenin VY, VH Winston (1977) Solutions of Ill-Posed Problems https://books.google.co.in/books/about/Solutions_of_Ill_posed_Problems.html?id=ECrvAAAAMAAJ&redir_esc=y.
2. Hoerl AE, Kennard RW (1970) Ridge regression: Biased estimation for nonorthogonal problems. Technometrics 12: 55-67.
3. Golub GH, Heath M, Wahba G (1979) Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics 21: 215-223.
4. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagation of errors. Nature 323: 533-536.
5. Schmidhuber J (2015) Deep learning in neural networks: An overview. Neural Networks 61: 85-117.
6. Friston, K. The free-energy principle: a unified brain theory?. Nat Rev Neurosci 11, 127–138 (2010). https://doi.org/10.1038/nrn2787
7. Zhuge Y, Han J, Li Z (2024) Spectral regulariza- tion in large-scale transformer training for ener-gy-efficient convergence. IEEE Transactions onNeural Networks and Learning Systems 35: 8432- 8447.
8. Lee D, Fischer A (2025) Deterministic matrix inversion learning for stable transformer layers. Nature Machine Intelligence 7: 215–228.
9. Patel K, Ahmed S, Rana P (2025) Low-entropy energy models for reproducible AI systems: Toward analytical convergence. Proceedings of the AAAI Conference on Artificial Intelligence 39: 10211032.
10. Nguyen T, Raginsky M (2024) Scaling laws and deterministic limits in high-dimensional learning dynamics. Journal of Machine Learning Research 25: 1–32.
11. Poggio T, Banburski A, Liao Q (2019) Theoretical issues in deep networks: approximation, optimization, and generalization. Proceedings of the National Academy of Sciences. https://arxiv.org/abs/1908.0937512.
12. Strang G (2019) Linear Algebra and Learning from Data. Wellesley-Cambridge Press https://www.cambridge.org/gb/universitypress/subjects/computer-science/pattern-recognition-and-machine-learning/linear-algebra-and-learning-data.
13. Boyd S, Vandenberghe L Convex Optimization. Cambridge University Press https://stanford.edu/~boyd/cvxbook/.
14. Mallat S (2016) Understanding deep convolutional J.of Pion Artf Int Research networks. Philosophical Transactions of the Roy- al Society A https://arxiv.org/abs/1601.04920.

15. Mehta P, Marin Bukov, Ching-Hao Wang, Alex-andre GR Day, Clint Richardson et al. (2018) A high-bias, low-variance introduction to machine learning for physicists. Physics Reports https://arxiv.org/abs/1803.08823.
16. Cekirge HM (2025) Tuning the training of neural networks by using the perturbation technique. American Journal of Artificial Intelligence 9: 107-109.
17. Cekirge HM (2025) An alternative way of determining biases and weights for the training of neural networks. American Journal of Artificial Intelligence 9: 129-132.
18. Cekirge HM (2025) Algebraic σ-based model fordeterministic and energy-efficient unsupervised machine learning. American Journal of Artificial Intelligence 9: 198-205.
19. Cekirge HM (2025) Cekirge's σ-based ANN model for deterministic, energy-efficient, scalable AI with large-matrix capability. American Journal of Artificial Intelligence 9: 206-216.
20. Cekirge HM (2025) Algebraic Cekirge method for deterministic and energy-efficient transformer language models. American Journal of Artificial Intelligence 9: 258-271.
21. Cekirge HM (2025) Deterministic σ-regularized benchmarking of the Cekirge model against GPT-transformer baseline. American Journal of Artificial Intelligence 9: 272-280.
22. CekirgeHM(2025)TheCekirgemethodformachine learning: A deterministic σ-regularized analytical solution for general minimum problems. American Journal of Artificial Intelligence 9: 324-337.
23. Cekirge HM (2025) Cekirge_Perturbation_Report_v4. Zenodo https://zenodo.org/records/17507105.