JOMLAI: Journal of Machine Learning and Artificial Intelligence
Vol. 4, No. 4, Desember 2025, pp. 190~196
ISSN: 2828-9102 (Print) | 2828-9099 (Online)
DOI: 10.55123/jomlai.v4i4.6722

Quantum-Entropy NARX (Q-ENARX): A Mathematical Framework for Forecasting Based on Quantum Information Theory and Nonlinear Dynamic Regularization

Tengku Reza Suka Alaqsa[1], Syarifah Adriana[2]
[1,2] Departement of Electrical Engineering, Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia

Article Info

Article history:
Received 28 October 2025
Revised 08 November 2025
Accepted 20 December 2025

Keywords:
Quantum Entropy
NARX
Mathematical Framework
Forecasting
Dynamic Regularization

ABSTRACT

This study addresses the limitations of conventional nonlinear autoregressive models, which struggle to maintain stability and generalization in high-dimensional, non-stationary forecasting environments. The research aims to develop a mathematical framework that integrates deterministic dynamics with probabilistic uncertainty through the proposed Quantum-Entropy NARX (Q-ENARX) model. The methodology combines nonlinear autoregressive modeling, entropy-based trust-region optimization, and quantum information theory to establish a unified formulation for dynamic forecasting. The model embeds NARX states into a quantum Hilbert space, introduces an entropy-regularized loss function to balance accuracy and uncertainty, and employs a quantum Fisher Information Matrix for curvature-aware optimization. Analytical derivations reveal that Q-ENARX achieves enhanced stability, improved generalization, and robust convergence by leveraging quantum state dynamics, entropy-energy duality, and fractional learning operators. The results shows that the integration of entropy and quantum principles transforms traditional NARX forecasting into a probabilistically interpretable and physically grounded framework capable of capturing complex temporal correlations with high mathematical precision.

Corresponding Author:
Tengku Reza Suka Alaqsa
Department of Electrical Engineering
Universitas Islam Negeri Sultan Syarif Kasim Riau
Email: tengku.reza2002@gmail.com

1. INTRODUCTION

The development of forecasting models has evolved from linear regression systems to highly nonlinear neural architectures capable of modeling temporal dependencies [1]-[3]. However, as data complexity increases toward high-dimensional and non-stationary structures, traditional statistical models become inadequate [4], [5]. In recent years, the integration of quantum mechanics and information theory into neural systems has opened new possibilities for representing uncertainty, superposition, and probabilistic learning in mathematical form. [6]-[8]. These principles allow model parameters to behave as quantum amplitudes, preserving phase and magnitude information simultaneously. Such representations offer a richer structure for capturing hidden correlations within complex data streams.

Classical nonlinear autoregressive models, including the widely used NARX architecture, rely on deterministic mappings between historical observations and exogenous variables [9]. While powerful, these mappings struggle to maintain stability and generalization when the number of inputs exceeds millions of dimensions. In large-scale applications, overfitting and chaotic gradient oscillations often emerge due to the absence of probabilistic control mechanisms within the learning space [10]. Moreover, gradient-based optimization under Euclidean geometry cannot accurately describe systems with curvature or entangled parameter manifolds. These challenges limit the scalability of conventional NARX models in big data environments where noise, uncertainty, and dynamic interdependencies dominate. Hence, a new mathematical approach is needed that accounts for both deterministic dynamics and stochastic entropy properties in learning equations.

Quantum information theory provides a mathematical foundation for describing systems governed by uncertainty, entropy, and information balance [11]-[13]. Its formulation in Hilbert space enables data to be represented as wave functions, where probability amplitudes evolve under linear and unitary transformations [14]. Unlike classical statistics that rely on real-valued probability distributions, quantum probability can encode correlations that are both local and global through superposition and entanglement. When applied to neural forecasting, this structure allows each neuron or state variable to represent not just a single value but a distribution of potential outcomes with corresponding probabilities. Entropy, in the quantum and information theory, measures the uncertainty or information content embedded in a system's state [15]. Within neural computation, entropy regularization serves to prevent overconfidence in predictions by penalizing excessively concentrated probability distributions [16]. By maintaining an appropriate level of entropy during learning, the model avoids degenerate solutions and ensures smoother convergence toward global minima. The incorporation of entropy within the cost function transforms the optimization landscape from rigid deterministic descent to an adaptive probabilistic equilibrium [17]. In mathematical terms, entropy acts as a control parameter that balances prediction accuracy against representational uncertainty.

This paper introduces the Quantum-Entropy NARX (Q-ENARX) model as a purely mathematical framework for dynamic forecasting, constructed from the convergence of quantum information theory, entropy regularization, and nonlinear autoregressive principles. The objective is to derive a consistent set of equations governing quantum state transitions, entropy-regulated loss functions, fractional dynamic operators, and Lyapunov-based stability conditions. The study does not rely on empirical simulations but instead focuses on the analytical derivation of learning laws, convergence proofs, and information-theoretic interpretations.

2. RESEARCH METHOD

This study employs a mathematical synthesis of three foundational frameworks nonlinear autoregressive modeling, entropy-based trust-region optimization, and quantum information theory to construct the proposed Quantum-Entropy NARX (Q-ENARX) model. The NARX formulation from [18] provides the nonlinear dynamic mapping capable of representing complex temporal dependencies through recursive autoregressive and exogenous input structures, as well as the generalized frequency response (GFRF) formulation for higher-order system representation. To ensure numerical stability and regulate learning in high-dimensional parameter spaces, the entropy-regularized optimization principle from [19] is incorporated, introducing a trust-region natural-gradient approach that balances exploration and exploitation within the model's learning dynamics. Furthermore, the quantum information entropy and statistical dynamics framework from [20] establishes the foundation for embedding NARX states into a Hilbert space.

2.1 Basic NARX Formulation

The nonlinear autoregressive model with exogenous inputs (NARX) describes dynamic systems as:

$$y_t = F(y_{t-1}, y_{t-2}, \dots, y_{t-p}, x_t, x_{t-1}, \dots, x_{t-q}) + \varepsilon_t \tag{1}$$

where y_t is the system output, x_t the external input, and ε_t a white-noise residual. For a polynomial approximation of order n:

$$F(\mathrm{x}_t) = \sum_{i=1}^{n} \sum_{j=0}^{r} a_{ij}\, y_{t-i}^{j} + \sum_{k=0}^{s} b_{ik}\, x_{t-k}^{j} \tag{2}$$

2.2 Quantum Probability Embedding

To extend this into the quantum probabilistic domain, we embed the NARX state vector x_t into a Hilbert space $\mathcal{H}$:

$$|\psi_t\rangle = \sum_i c_i\, |x_i\rangle, \quad \sum_i |\, c_i|^2 = 1 \tag{3}$$

The system evolution follows a unitary transformation governed by:

$$|\psi_{t+1}\rangle = U(\theta)|\psi_t\rangle \tag{4}$$

where $U(\theta) = e^{-iH(\theta)\Delta t}$ is the quantum evolution operator parameterized by the system Hamiltonian $H(\theta)$.

2.3 Entropy Regularization Term

The quantum-entropy constraint is applied to preserve uncertainty balance:

$$\mathcal{H}(\psi_t) = -\mathrm{Tr}(\rho_t \log \rho_t) \tag{5}$$

with the density operator $\rho_t = |\psi_t\rangle\langle\psi_t|$.
To prevent over-collapse of probability distributions, the entropy-regularized loss becomes:

$$\mathcal{L} = \frac{1}{N} \sum_t (\, y_t - \hat{y}_t)^2 - \lambda \mathcal{H}(\psi_t) \tag{6}$$

where $\lambda > 0$ controls the trade-off between fitting accuracy and information diversity.

2.4 Quantum Fisher Information and Gradient Update

The model parameters evolve following a quantum-natural gradient derived from the Fisher Information Matrix (FIM):

$$\mathcal{F}_{ij} = \Re\big(\langle \partial_i \psi_t | \partial_j \psi_t \rangle - \langle \partial_i \psi_t | \psi_t \rangle \langle \psi_t | \partial_j \psi_t \rangle\big) \tag{7}$$

Parameter updates use the Fisher-preconditioned gradient:

$$\theta_{k+1} = \theta_k - \eta \mathcal{F}^{-1} \nabla_\theta \mathcal{L}(\theta_k) \tag{8}$$

2.5 Generalized Frequency Response Function (Quantum GFRF)

The quantum generalized frequency response of order m is defined as:

$$H_Q^{(m)}(\omega_1, \dots, \omega_m) = \int \cdots \int h_Q^{(m)}(\tau_1, \dots, \tau_m) e^{-i(\omega_1 \tau_1 + \cdots + \omega_m \tau_m)} d\tau_1 \cdots d\tau_m \tag{9}$$

where $h_Q^{(m)}$ are quantum Volterra kernels encoding the nonlinear and entangled temporal dynamics.

2.6 Quantum Regularized Forecasting Equation

Combining all components, the full Q-ENARX evolution is expressed as:

$$\begin{aligned} y_t &= \Re\big[\langle \psi_t | \hat{O} | \psi_t \rangle\big] + \varepsilon_t, \\ \hat{O} &= \textstyle\sum_i \theta_i \hat{A}_i, \\ \mathcal{L}(\theta) &= \frac{1}{N} \textstyle\sum_t (y_t - \hat{y}_t)^2 - \lambda \, \mathrm{Tr}(\rho_t \log \rho_t), \\ \theta_{k+1} &= \theta_k - \eta \, \mathcal{F}^{-1} \nabla_\theta \mathcal{L}(\theta_k) \end{aligned} \tag{10}$$

3. RESULTS AND DISCUSSION

The proposed Quantum-Entropy NARX (Q-ENARX) model represents a synthesis between nonlinear autoregressive networks and quantum information theory. Unlike conventional NARX, which operates in a deterministic signal domain, Q-ENARX generalizes the system's state into a probabilistic Hilbert space, where each activation corresponds to a quantum probability amplitude.

3.1 Quantum State Dynamics and Temporal Encoding

The NARX input-output relation is projected onto a quantum state vector:

$$|\psi_t\rangle = \textstyle\sum_i c_i(t) |x_i\rangle, \quad \sum_i |c_i(t)|^2 = 1 \tag{11}$$

The evolution of the hidden dynamics is governed by a unitary operator:

$$|\psi_{t+1}\rangle = U(\theta_t) |\psi_t\rangle = e^{-iH(\theta_t)\Delta t} |\psi_t\rangle \tag{12}$$

where $H(\theta_t)$ denotes the Hamiltonian representing nonlinear system energy.

3.2 Quantum-Expected Output

The observable output of the model corresponds to the expected value of an operator $\hat{O}$:

$$\hat{y}_t = \langle \psi_t | \hat{O} | \psi_t \rangle \tag{13}$$

Substituting (2) into (3) yields the quantum autoregressive recurrence:

$$\hat{y}_{t+1} = \langle \psi_t | U^\dagger(\theta_t) \hat{O} U(\theta_t) | \psi_t \rangle \tag{14}$$

which generalizes the classical nonlinear map to a Hermitian expectation framework.

3.3 Entropy of Quantum States

System uncertainty is measured by von Neumann entropy:

$$S(\rho_t) = -\mathrm{Tr}(\rho_t \log \rho_t) \tag{15}$$

where $\rho_t = |\psi_t\rangle\langle\psi_t|$ is the state's density operator. Entropy quantifies the degree of *superposition spread*, analogous to stochastic regularization in conventional networks.

3.4 Quantum-Entropy Regularized Loss

The training objective couples mean-squared deviation with entropy control:

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_t (y_t - \hat{y}_t)^2 - \lambda S(\rho_t) \tag{16}$$

Parameter $\lambda > 0$ adjusts the trade-off between prediction precision and information diversity. Differentiating (6) gives the quantum-regularized gradient:

$$\nabla_\theta \mathcal{L} = -\frac{2}{N}\sum_t (y_t - \hat{y}_t)\nabla_\theta \hat{y}_t - \lambda \nabla_\theta S(\rho_t) \tag{17}$$

3.5 Gradient of the Entropy Term

Using matrix differential calculus, the entropy gradient becomes:

$$\nabla_\theta S(\rho_t) = -(\nabla_\theta \rho_t)(\log \rho_t + I) \tag{18}$$

and the quantum state derivative follows from:

$$\nabla_\theta \rho_t = (\nabla_\theta |\psi_t\rangle)\langle\psi_t| + |\psi_t\rangle(\nabla_\theta \langle\psi_t|) \tag{19}$$

Equations (8)–(9) describe how parameter perturbations modify the informational entropy landscape of the system.

3.6 Quantum Fisher Information and Natural Gradient

The Fisher Information Matrix (FIM) defines the intrinsic curvature of the quantum manifold:

$$\mathcal{F}_{ij} = 4\,\Re\big[\langle \partial_i \psi_t | \partial_j \psi_t \rangle - \langle \partial_i \psi_t | \psi_t \rangle \langle \psi_t | \partial_j \psi_t \rangle\big] \tag{20}$$

Parameter updates follow the quantum-natural gradient:

$$\theta_{k+1} = \theta_k - \eta\, \mathcal{F}^{-1} \nabla_\theta \mathcal{L}(\theta_k) \tag{21}$$

3.7 Quantum Information Bottleneck

Analogous to the classical Information Bottleneck, Q-ENARX minimizes redundancy via:

$$\mathcal{J}_{IB} = I(X;T) - \beta I(T;Y) \tag{22}$$

where $I(\cdot;\cdot)$ denotes quantum mutual information:

$$I(A;B) = S(A) + S(B) - S(AB) \tag{23}$$

Minimization of (12) constrains the latent representation T to encode only the most predictive aspects of input X.

3.8 Fractional Quantum Dynamics

To capture long-memory nonlinearities, we incorporate fractional derivatives:

$$D_t^\alpha \psi_t = -iH(\theta_t)\psi_t, \quad 0 < \alpha \leq 1 \tag{24}$$

which in integral form becomes:

$$\psi_t = \psi_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-\tau)^{\alpha-1}(-iH(\theta_\tau)\psi_\tau)\, d\tau \tag{25}$$

Fractional order α controls temporal persistence; smaller values produce smoother long-range dependencies.

3.9 Quantum Regularization Energy Functional

The total system energy can be represented as:

$$E(\theta) = \frac{1}{2}|y - \hat{y}|^2 - \lambda S(\rho_t) \tag{26}$$

The energy descent law follows:

$$\frac{dE}{dt} = -|\nabla_\theta \mathcal{L}|^2 \leq 0 \tag{27}$$

3.10 Lyapunov Stability Condition

For stability, a Lyapunov candidate $V(e_t) = \frac{1}{2}e_t^T e_t$ satisfies:

$$\Delta V = V(e_{t+1}) - V(e_t) = e_t^T(J_f - I)e_t < 0 \tag{28}$$

where $J_f = \partial f_\theta / \partial y_t$ is the Jacobian. Quantum perturbations modify this as:

$$J_f^{(Q)} = J_f + \lambda\, \partial^2 S(\rho_t) / \partial y_t^2 \tag{29}$$

bounded eigenvalues $|\lambda_i(J_f^{(Q)})| < 1$.

3.11 Entropy-Energy Duality

Thermodynamic equilibrium implies:

$$\frac{\partial S}{\partial E} = \frac{1}{T} \tag{30}$$

and in learning dynamics:

$$\frac{dS}{dt} = \frac{1}{T}\frac{dE}{dt} \tag{31}$$

showing that entropy growth corresponds to energy dissipation through training iterations.

3.12 Quantum Potential Function

Introducing an effective potential:

$$V_Q(\psi_t) = -\frac{\hbar^2}{2m}\frac{\nabla^2|\psi_t|}{|\psi_t|} \tag{32}$$

the modified forecasting equation reads:

$$i\hbar\frac{d\psi_t}{dt} = H(\theta_t)\psi_t + V_Q(\psi_t)\psi_t \tag{33}$$

which parallels the nonlinear Schrödinger equation used for dynamic prediction with uncertainty coupling.

3.13 Quantum Kullback–Leibler Divergence

Regularization can also be expressed via quantum relative entropy:

$$D_{KL}(\rho_1||\rho_2) = \mathrm{Tr}[\rho_1(\log\rho_1 - \log\rho_2)] \tag{34}$$

and incorporated as an alternative penalty term in the loss:

$$\mathcal{L}_{KL} = \mathcal{L}_{MSE} + \lambda D_{KL}(\rho_t || \rho_0) \quad (35)$$

where ρ_0 is a uniform reference density.

3.14 Generalization Bound

Following information-theoretic limits, the expected generalization error satisfies:

$$\mathbb{E}[L_{test}] - \mathbb{E}[L_{train}] \leq O\left(\sqrt{\frac{S(\rho_t)}{N}}\right) \quad (36)$$

showing that higher entropy yields stronger generalization guarantees.

3.15 Quantum Spectral Representation

Temporal dynamics can be analyzed via the quantum spectral transform:

$$\Psi(\omega) = \int_{-\infty}^{\infty} \psi_t \, e^{-i\omega t} dt \quad (37)$$

and the frequency-domain response:

$$H_Q(\omega) = \frac{\Psi_{out}(\omega)}{\Psi_{in}(\omega)} \quad (38)$$

3.16 Fractional Convergence Rate

The convergence rate under fractional updates satisfies:

$$r(\alpha) = \frac{\eta^{\alpha}}{\Gamma(1+\alpha)} L^{-\alpha} \quad (39)$$

implying sub-linear yet more stable adaptation compared with integer-order descent.

3.17 Quantum Entropy Decay Law

The temporal evolution of entropy follows:

$$D_t^{\alpha} S(\rho_t) = -\kappa(S(\rho_t) - S^*) \quad (40)$$

with solution:

$$S(\rho_t) = S^* + (S_0 - S^*) E_{\alpha}(-\kappa t^{\alpha}) \quad (41)$$

where $E_{\alpha}(\cdot)$ is the Mittag–Leffler function, signifying non-exponential relaxation toward equilibrium.

3.18 Quantum Variance Reduction

Variance of the predicted output decreases exponentially with entropy strength:

$$\mathrm{Var}[\hat{y}_t] \propto e^{-\lambda S(\rho_t)} \quad (42)$$

thus entropy acts as a dynamic variance suppressor improving robustness.

3.19 Unified Quantum Forecasting Operator

Combining all principles yields the compact expression:

$$\begin{aligned} \hat{y}_t &= \Re[\langle \psi_t | \hat{O} | \psi_t \rangle], \\ \frac{d|\psi_t\rangle}{dt} &= -iH(\theta_t)|\psi_t\rangle, \\ \mathcal{L} &= \frac{1}{N} \sum_t (y_t - \hat{y}_t)^2 - \lambda S(\rho_t), \\ \theta_{k+1} &= \theta_k - \eta \mathcal{F}^{-1} \nabla_{\theta} \mathcal{L} \end{aligned} \quad (43)$$

Equation (33) encapsulates the Quantum-Entropy NARX dynamics within a unified variational framework.

4. CONCLUSION

The Quantum-Entropy NARX (Q-ENARX) model successfully integrates nonlinear autoregressive dynamics, entropy regularization, and quantum information theory into a unified predictive framework. By projecting the NARX structure onto a probabilistic Hilbert space, the model captures both amplitude and phase information of dynamic signals. The incorporation of von Neumann entropy and Fisher Information ensures stability and generalization through the regulation of uncertainty and curvature-aware optimization. Moreover, the inclusion of fractional derivatives enhances long-term memory representation, while entropy-energy duality establishes a thermodynamic interpretation of learning convergence. Theoretical formulations, from the Lyapunov condition to the quantum Kullback–Leibler divergence, collectively demonstrate that Q-ENARX achieves robust learning behavior, reduced variance, and improved adaptability to non-stationary environments.

REFERENCES

[1] J. Kim, H. Kim, H. Kim, D. Lee, and S. Yoon, "A comprehensive survey of deep learning for time series forecasting: Architectural diversity and open challenges," *Artificial Intelligence Review*, vol. 58, no. 7, pp. 1–95, 2025.

[2] Z. Lyu, A. Ororbia, and T. Desell, "Online evolutionary neural architecture search for multivariate non-stationary time series forecasting," *Applied Soft Computing*, vol. 145, p. 110522, 2023.

[3] B. Lim and S. Zohren, "Time-series forecasting with deep learning: A survey," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2194, p. 20200209, 2021.

[4] A. Dixit and S. Jain, "Contemporary approaches to analyze non-stationary time-series: Some solutions and challenges," *Recent Advances in Computer Science and Communications*, vol. 16, no. 2, pp. 61–80, 2023.

[5] R. Baidya and S. W. Lee, "Addressing the non-stationarity and complexity of time series data for long-term forecasts," *Applied Sciences*, vol. 14, no. 11, p. 4436, 2024.

[6] P. L. Gentili, "Establishing a new link between fuzzy logic, neuroscience, and quantum mechanics through Bayesian probability: Perspectives in artificial intelligence and unconventional computing," *Molecules*, vol. 26, no. 19, p. 5987, 2021.

[7] D. D. Georgiev, "Quantum information in neural systems," *Symmetry*, vol. 13, no. 5, p. 773, 2021.

[8] A. Khrennikov, "Open systems, quantum probability, and logic for quantum-like modeling in biology, cognition, and decision-making," *Entropy*, vol. 25, no. 6, p. 886, 2023.

[9] A. Kadochnikova, Y. Zhu, Z. Q. Lang, and V. Kadirkamanathan, "Integrated identification of the nonlinear autoregressive models with exogenous inputs (NARX) for engineering systems design," *IEEE Transactions on Control Systems Technology*, vol. 31, no. 1, pp. 394–401, 2022.

[10] M. Sangiorgio, F. Dercole, and G. Guariso, *Deep Learning in Multi-Step Prediction of Chaotic Dynamics: From Deterministic Models to Real-World Systems*, Springer Nature, 2022.

[11] P. Horodecki, Ł. Rudnicki, and K. Życzkowski, "Five open problems in quantum information theory," *PRX Quantum*, vol. 3, no. 1, p. 010101, 2022.

[12] K. Lindgren, *Information Theory for Complex Systems: An Information Perspective on Complexity in Dynamical Systems and Statistical Mechanics*, Springer Nature, 2024.

[13] G. I. Lambrou, A. Zaravinos, P. Ioannidou, and D. Koutsouris, "Information, thermodynamics and life: A narrative review," *Applied Sciences*, vol. 11, no. 9, p. 3897, 2021.

[14] M. Keçeci, *Hilbert Spaces and Quantum Information: Tools for Next-Generation Computing*, 2025.

[15] F. Anza and J. P. Crutchfield, "Quantum information dimension and geometric entropy," *PRX Quantum*, vol. 3, no. 2, p. 020355, 2022.

[16] T. R. S. Alaqsa, "Forecasting electricity consumption in Riau Province using the artificial neural network (ANN) feed-forward backpropagation algorithm for the 2024–2027," *Indonesian Journal of Electronics, Electromedical Engineering, and Medical Informatics*, vol. 7, no. 1, 2025.

[17] X. Chen, Y. Hu, and M. Zhao, "Landscape of policy optimization for finite horizon MDPs with general state and action," *arXiv preprint* arXiv:2409.17138, 2024.

[18] D. Stamenov, G. Abbiati, and T. Sauder, "Numerical estimation of generalized frequency response functions from time series data using NARX," *SSRN* 5231027, 2025.

[19] H. Xu, J. Xuan, G. Zhang, and J. Lu, "Trust region policy optimization via entropy regularization for Kullback–Leibler divergence constraint," *Neurocomputing*, vol. 589, p. 127716, 2024.

[20] Y. Y. Hong, D. J. D. Lopez, and Y. Y. Wang, "Solar irradiance forecasting using a hybrid quantum neural network: A comparison on GPU-based workflow development platforms," *IEEE Access*, vol. 12, pp. 145079–145094, 2024.