Letter

Nonlinear extensions of linear inverse models under memoryless or persistent random forcing

Justin Lien *
Mathematical Institute, Tohoku University, 2 Chome-1-1 Katahira, Aoba Ward, Sendai, Miyagi 980-8577, Japan

Hiroyasu Ando
Advanced Institute for Materials Research, Tohoku University, 2 Chome-1-1 Katahira, Aoba Ward, Sendai, Miyagi 980-8577, Japan
and Mathematical Institute, Tohoku University, 2 Chome-1-1 Katahira, Aoba Ward, Sendai, Miyagi 980-8577, Japan

(Received 8 April 2025; accepted 28 July 2025; published 20 August 2025)

This study extends the linear inverse modeling (LIM) framework to nonlinear settings by presenting White-nLIM and Colored-nLIM, statistics-based empirical methods that construct approximate stochastic systems incorporating quadratic deterministic dynamics with either memoryless Gaussian white noise or persistent Ornstein-Uhlenbeck colored noise. Beyond the evident improvements over linear models, Colored-nLIM offers a robust approach to parameter estimation and statistical modeling under persistent stochastic forcing. Together with White-nLIM, these methods provide a systematic framework to assess the role of noise persistence in inverse modeling. Applications to the Lorenz 63 system and a simplified El Niño-Southern Oscillation model demonstrate their potential to capture chaotic behavior and climate variability.

DOI: 10.1103/ds1j-fx3v

Introduction. The most fundamental linear inverse model (LIM), referred to here as White-LIM, is widely used to study local dynamical behavior near an equilibrium point [1–4]. More specifically, it extracts the local dynamics from observational data by constructing an approximate linear stochastic differential equation (SDE) driven by Gaussian white noise. This method and its equivalent variants have been extensively applied across mathematical engineering and the earth sciences [5–9]. In particular, White-LIM has proven practical for studying climate phenomena such as the El Niño-Southern Oscillation (ENSO), a major mode of climate variability in the tropical Pacific Ocean that significantly influence global weather patterns through warm (El Niño) and cold (La Niña) phases alternating asymmetrically in amplitude and frequency [10,11]. Though extensively applied to study essential features of ocean-atmosphere coupled variability, White-LIM inherently struggles with extended prediction skills, the ENSO asymmetry, and other nonlinear characteristics [12–22].

Recent advancements have extended stochastic forcing in the White-LIM framework to include noise persistence using Ornstein-Uhlenbeck (OU) colored noise or state dependency via correlated additive-multiplicative (CAM) white noise, leading to the correlation/moment-based (referred to here as statistics-based) formulation known as Colored-LIM and CAM-LIM, respectively [7,20]. Methodologically, Colored-LIM reconstructs the dynamical operator using higher-order derivatives of the correlation function, whereas CAM-LIM relies on higher-order moments. The former enhances prediction skills by modeling memory effects inherent in observational data, while the latter captures skewness through state-dependent random forcing [23,24]. However, both approaches remain linear and thus exhibit limited applicability in strongly nonlinear systems [7,24].

To address these limitations, several nonlinear empirical methods, such as maximum likelihood estimation, multilevel regression models, and operator-theoretic methods, have been developed and applied across scientific studies, including climate variability, biochemical reactions, and epidemic modeling [25–36]. However, while able to reproduce certain statistical quantities such as covariance, many of these models rely on the white-noise assumption, under which parameter estimation is more reliable, and the presence of non-white noise may cause biased parameter estimates [32,37,38]. This challenge has motivated the development of data-driven methods for nonlinear systems in the non-white-noise regime, including the ocean-atmosphere coupled system [39–42].

This study extends both White-LIM and Colored-LIM to a nonlinear framework by incorporating quadratic dynamics, proposing statistics-based modeling approaches, called White-nLIM and Colored-nLIM. Inspired by the methodological advances in Refs. [7,20], we prove that, regardless of the noise type, the higher-order local dynamics of the quadratic SDE is characterized by the local behavior of its correlation functions at 0-lag up to fourth order. While White-nLIM is equivalently effective to methods like multiple polynomial regression by design, Colored-nLIM enables robust estimation of system parameters under persistent noise. More importantly, these two models together

*Contact author: lien.justin.t8@dc.tohoku.ac.jp

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

illustrate how noise persistence influences the resulting empirical system [29,43].

As an application to stochastic modeling, we implement both methods to examine the ENSO recharge oscillator (RO) model and show that they can capture the skewness of key variables, a known limitation of the traditional LIMs. We qualitatively compare our results and methodologies with the extended nonlinear recharge oscillator (XRO) model—a recently proposed regression-based approach that incorporates second-order dynamics and persistent noise, and has demonstrated an extended ENSO forecast range [36]. Finally, their performance in parameter estimation is validated through numerical experiments, detailed in the Supplemental Material [44], demonstrating effectiveness across both nonchaotic and chaotic systems.

In this study, boldface denotes mathematical objects in the forward formulation, while regular font represents estimated quantities from the inverse method. Vectors are assumed to be column vectors. An a-tensor $\mathbf{X}$ is an element in $\mathbb{R}^{\overbrace{n\times\cdots\times n}^{a\text{ times}}}$. The tensor multiplication of an a-tensor $\mathbf{X}$ and a b-tensor $\mathbf{Y}$ produces an $(a+b-2m)$-tensor $\mathbf{X}\times_m\mathbf{Y}$, by summing over the last m indices of $\mathbf{X}$ and the first m indices of $\mathbf{Y}$. For example, given a 3-tensor $\mathbf{X}$ and a 4-tensor $\mathbf{Y}$, their product $\mathbf{X}\times_2\mathbf{Y}$ is a 3-tensor given by $[\mathbf{X}\times_2\mathbf{Y}]_{ijk}=\sum_{p,q}\mathbf{X}_{ipq}\mathbf{Y}_{pqjk}$. The transpose $\mathbf{X}^T$ of $\mathbf{X}$ is given by switching the first and second indices while leaving the remaining indices unchanged (i.e., $[\mathbf{X}^T]_{i_1i_2i_3\cdots i_a}=\mathbf{X}_{i_2i_1i_3\cdots i_a}$) and the symmetrization operator Sym is given by $\mathrm{Sym}(\mathbf{X})=\mathbf{X}^T+\mathbf{X}$.

To support the readers, a review of LIM is included in the Appendix A. Throughout this article, stationarity is assumed implicitly, and the SDE can be interpreted in either the Itô or Stratonovich sense due to additive noise.

Mathematical foundations of nonlinear inverse models. The nonlinear inverse models (nLIMs) describe the n-dimensional observations $\{x(t)\}\subset\mathbb{R}^n$ that may or may not have a zero mean with the approximate SDE

$$\frac{d}{dt}x=B^{\mathrm{nLIM}}\times_2(xx^T)+A^{\mathrm{nLIM}}x+C^{\mathrm{nLIM}}+\sqrt{2Q^{\mathrm{nLIM}}}\cdot\mathrm{noise}. \tag{1}$$

In this formulation, $B^{\mathrm{nLIM}}\in\mathbb{R}^{n\times n\times n}$, $A^{\mathrm{nLIM}}\in\mathbb{R}^{n\times n}$, and $C^{\mathrm{nLIM}}\in\mathbb{R}^n$ are the dynamical tensors corresponding to the quadratic, linear, and constant terms, with $B^{\mathrm{nLIM}}_{ijk}=0$ if $j>k$, and $Q^{\mathrm{nLIM}}\in\mathbb{R}^{n\times n}$ is the positive definite stochastic matrix. The stochastic forcing is either modeled by white noise or colored noise. The superscript $\cdot^{\mathrm{nLIM}}$ represents the general reference for estimated quantities from White-nLIM and Colored-nLIM, specified by $\cdot^{\text{W-nLIM}}$ and $\cdot^{\text{C-nLIM}}$, respectively. We start with the forward formulation before addressing the inverse methods.

For White-nLIM, the governing equation of the state vector $\mathbf{x}$ reads

$$\frac{d}{dt}\mathbf{x}=\mathbf{B}\times_2(\mathbf{x}\mathbf{x}^T)+\mathbf{A}\times_1\mathbf{x}+\mathbf{C}+\sqrt{2\mathbf{Q}}\times_1\boldsymbol{\xi}, \tag{2}$$

where some matrix multiplications are expressed in the tensor form to highlight the hierarchical structures emerging in subsequent equations. The stochastic forcing $\boldsymbol{\xi}$ is the normalized Gaussian white noise with zero mean and unit variance. That is, $\langle\boldsymbol{\xi}(\cdot+\tau)\boldsymbol{\xi}(\cdot)^T\rangle=\delta(\tau)\,\mathbf{I}$, where $\langle\cdot\rangle$, δ, and τ denote the expectation, Dirac delta function, and the lag variable, respectively.

Unlike the linear model, the system dynamics $\mathbf{B}$, $\mathbf{A}$, and $\mathbf{C}$ are not only determined by the correlation function $\mathbf{K}(\tau)=\langle\mathbf{x}(\cdot+\tau)\mathbf{x}(\cdot)^T\rangle$ of $\mathbf{x}$, but also the mean value $\mathbf{E}=\langle\mathbf{x}(\cdot)\rangle$, and higher-order correlation functions $\mathbf{M}_{ijk}(\tau)=\langle\mathbf{x}_i(\cdot+\tau)\mathbf{x}_j(\cdot)\mathbf{x}_k(\cdot)\rangle$ and $\mathbf{S}_{ijkw}(\tau)=\langle\mathbf{x}_i(\cdot+\tau)\mathbf{x}_j(\cdot)\mathbf{x}_k(\cdot)\mathbf{x}_w(\cdot)\rangle$. Despite the lack of explicit forms, these correlation functions and their right derivatives satisfy

$$0=\mathbf{B}\times_2\mathbf{K}(0)+\mathbf{A}\times_1\mathbf{E}+\mathbf{C}, \tag{3}$$

$$\mathbf{K}'(0)=\mathbf{B}\times_2\mathbf{M}(0)+\mathbf{A}\times_1\mathbf{K}(0)+\mathbf{C}\times_0\mathbf{E}, \tag{4}$$

$$\mathbf{M}'(0)=\mathbf{B}\times_2\mathbf{S}(0)+\mathbf{A}\times_1\mathbf{M}(0)+\mathbf{C}\times_0\mathbf{K}(0), \tag{5}$$

and the balance equation describing the equilibrium of the deterministic dynamics and noise forcing is given by

$$0=\mathrm{Sym}(\mathbf{B}\times_2\mathbf{M}(0)+\mathbf{A}\times_1\mathbf{K}(0)+\mathbf{C}\times_0\mathbf{E}+\mathbf{Q}). \tag{6}$$

Therefore, by viewing the mean value as a degenerate first-order correlation function, we may say that the dynamical tensors and the stochastic matrix are characterized by the correlation functions up to fourth order.

On the other hand, the governing equation in the Colored-nLIM framework reads

$$\frac{d}{dt}\mathbf{x}=\mathbf{B}\times_2(\mathbf{x}\mathbf{x}^T)+\mathbf{A}\times_1\mathbf{x}+\mathbf{C}+\sqrt{2\mathbf{Q}}\times_1\boldsymbol{\eta}. \tag{7}$$

The stochastic forcing is modeled by the normalized OU colored noise $\boldsymbol{\eta}$ with the noise correlation time $\boldsymbol{\gamma}>0$, realized by the state steady of $(d/dt)\boldsymbol{\eta}=-(1/\boldsymbol{\gamma})\boldsymbol{\eta}+(1/\boldsymbol{\gamma})\boldsymbol{\xi}$ [45]. As its temporal structure satisfies $\langle\boldsymbol{\eta}(\cdot+\tau)\boldsymbol{\eta}(\cdot)^T\rangle=(1/2\boldsymbol{\gamma})e^{(-|\tau|/\boldsymbol{\gamma})}\mathbf{I}$, $\boldsymbol{\gamma}$ can be physically interpreted as the characteristic time for the past dependence of noise.

The noise-state correlation tensors $\mathbf{K}_{\boldsymbol{\eta}\mathbf{x}}=\langle\boldsymbol{\eta}(\cdot)\mathbf{x}(\cdot)^T\rangle$ and $\mathbf{M}_{\boldsymbol{\eta}\mathbf{x}}$ given by $[\mathbf{M}_{\boldsymbol{\eta}\mathbf{x}}]_{ijk}=\langle\boldsymbol{\eta}_i(\cdot)\mathbf{x}_j(\cdot)\mathbf{x}_k(\cdot)\rangle$ are nontrivial. Similar to White-nLIM, the correlation functions $\mathbf{E}$, $\mathbf{K}$, $\mathbf{M}$, and $\mathbf{S}$ of $\mathbf{x}$ do not admit explicit forms, but satisfy

$$0=\mathbf{B}\times_2\mathbf{K}(0)+\mathbf{A}\times_1\mathbf{E}+\mathbf{C}, \tag{8}$$

$$\mathbf{K}'(0)=\mathbf{B}\times_2\mathbf{M}(0)+\mathbf{A}\times_1\mathbf{K}(0)+\sqrt{2\mathbf{Q}}\times_1\mathbf{K}_{\boldsymbol{\eta}\mathbf{x}}+\mathbf{C}\times_0\mathbf{E}, \tag{9}$$

$$\boldsymbol{\gamma}\,\mathbf{K}''(0)=\boldsymbol{\gamma}\,\mathbf{B}\times_2\mathrm{Sym}(\mathbf{M}'(0))+\boldsymbol{\gamma}\,\mathbf{A}\times_1\mathbf{K}'(0)-\sqrt{2\mathbf{Q}}\times_1\mathbf{K}_{\boldsymbol{\eta}\mathbf{x}}, \tag{10}$$

$$\mathbf{M}'(0)=\mathbf{B}\times_2\mathbf{S}(0)+\mathbf{A}\times_1\mathbf{M}(0)+\sqrt{2\mathbf{Q}}\times_1\mathbf{M}_{\boldsymbol{\eta}\mathbf{x}}+\mathbf{C}\times_0\mathbf{K}(0), \tag{11}$$

$$\boldsymbol{\gamma}\,\mathbf{M}''(0)=\boldsymbol{\gamma}\,\mathbf{B}\times_2\mathrm{Sym}(\mathbf{S}'(0))+\boldsymbol{\gamma}\,\mathbf{A}\times_1\mathbf{M}'(0)-\sqrt{2\mathbf{Q}}\times_1\mathbf{M}_{\boldsymbol{\eta}\mathbf{x}}. \tag{12}$$

Moreover, we have

$$0=\mathbf{M}_{\boldsymbol{\eta}\mathbf{x}}\times_2\mathbf{B}^T+\mathbf{K}_{\boldsymbol{\eta}\mathbf{x}}\times_1\left(\mathbf{A}^T-\frac{1}{\boldsymbol{\gamma}}\mathbf{I}\right)+\frac{1}{2\boldsymbol{\gamma}}\sqrt{2\mathbf{Q}}, \tag{13}$$

and the balance equation reads

$$0 = \mathrm{Sym}(\mathbf{B} \times_2 \mathbf{M}(0) + \mathbf{A} \times_1 \mathbf{K}(0) + \mathbf{C} \times_0 \mathbf{E} + \sqrt{2\mathbf{Q}} \times_1 \mathbf{K}_{\boldsymbol{\eta x}}). \quad (14)$$

Consequently, the dynamics **B**, **A**, and **C** as well as the stochastic matrix **Q** are functions of **E**, **K**, **M**, **S**, and $\boldsymbol{\gamma}$.

Given observations $\{x(t)\}$ (and a prescribed noise correlation time $\gamma^{\text{C-nLIM}}$), let E^{obs}, K^{obs}, M^{obs}, and S^{obs} denote the mean and observed correlation functions. White-nLIM and Colored-nLIM estimate the system dynamics B^{nLIM}, A^{nLIM}, and C^{nLIM} using Eqs. (3)–(5) and Eqs. (8)–(12), and compute an auxiliary stochastic matrix Q_0 using Eq. (6) and Eqs. (9), (11), and (13), respectively.

If $\{x(t)\}$ is sampled from either Eq. (2) or Eq. (7), Q_0 indeed provides a fair estimate. However, in practice, several factors can lead to discrepancies, such as limited sampling, deviations from the ideal quadratic system, or numerical errors in computing derivatives. Consequently, contrary to the linear models, the correlations $K(0)$ and $M(0)$ of the system determined by the estimated dynamics and Q_0 may not be consistent with $K^{\text{obs}}(0)$ and $M^{\text{obs}}(0)$. Therefore, we impose constraints based on the observed moments along with the balance equation, and solve an optimization problem to obtain an improved estimate of the stochastic matrix Q^{nLIM}.

Finally, if $\gamma^{\text{C-nLIM}}$ for Colored-nLIM is not prescribed, it is determined using the γ-selection algorithm similarly described by Eq. (A6), as originally introduced for the linear model [7]. The numerical details are provided in the Supplemental Material [44].

Discussion. The equations characterizing the system dynamics exhibit a hierarchical structure. In White-nLIM, Eqs. (4) and (5) show that the first-order derivatives of nth-order correlation functions ($n = 2, 3$), evaluated at the origin, are composed of three distinct terms: the tensor product of $(k - n + 1)$th-order dynamics and kth-order correlation functions, where $k = n - 1$, n, and $n + 1$. Indeed, this structure naturally generalizes to higher-order derivatives of higher-order correlation functions and persists similarly in Colored-nLIM.

From the perspective of forward formulation, even if the system parameters **B**, **A**, **C**, and **Q** remain unchanged, the derivatives of the correlation functions in the Colored-nLIM framework [i.e., Eqs. (9) and (11)] do not converge to those in the White-nLIM framework [i.e., Eqs. (4) and (5)] in the white-noise limit $\boldsymbol{\gamma} \to 0$. This arises from the noncommutativity between differentiation and the limit operations, as in the linear case reported in Ref. [7]. However, when applied to the same input data, the Colored-nLIM results do converge to the White-nLIM results, as in the white-noise limit, Eqs. (10), (12), and (14) converge to Eqs. (4), (5), and (6), respectively. This suggests that in the inverse problem, the system parameters of the approximate system (1) are continuous in γ over the interval $[0, \infty)$, and that γ can be interpreted as the extent to which concavity of K^{obs} and M^{obs} contributes to the system.

The LIM framework can also be viewed as a special case of the nLIM framework. In the forward formulation, if the quadratic and constant terms vanish, the key equations of nLIMs reduce to those of LIMs shown in the Appendix. Furthermore, for white-noise models, when applied to the same observational data, the stochastic matrices depend only on the first-order derivative of K^{obs}, making them independent of the presence of nonlinear and constant terms if the optimization process in White-nLIM is not implemented. In contrast, for colored-noise models with the same prescribed noise correlation time, the stochastic matrices explicitly depend on the estimated dynamics and therefore differ. Finally, when data are sampled sufficiently close to a stable equilibrium point of the underlying system, LIMs are known to approximate the Jacobian matrix [7]. Under this condition, the nLIM framework could yield an estimated linear term that closely aligns with the result of the linear models.

In practice, nonlinear models must be applied with caution. Since the number of parameters grows cubically with system dimension, estimating them using Eqs. (3)–(5) or Eqs. (8)–(12) may be numerically unstable or sensitive to sampling uncertainty. This suggests the need to prespecify nonlinear terms using, for example, sparse identification techniques like SINDy to enhance model robustness [46–48]. However, despite sparsification, the approximate system is not guaranteed to admit a stationary distribution, as the estimated dynamics may lack global stability or the boundedness of trajectories. In studies aimed at global dynamical structures imposing physics constraints during the estimation of system dynamics helps ensure the existence of a stationary distribution [32,49]. In contrast, when the focus is on local dynamical behavior near a stable equilibrium point, global stability can often be relaxed. Previous studies have addressed this issue by suppressing the unstable nonlinear terms or rewinding the trajectory and then restarting the simulation with a new realization of random forcing to avoid finite-time blowup [29,50].

To ensure theoretical rigor and maintain the well posedness of the proposed nLIMs without introducing discontinuities or using ad hoc procedures, we impose an artificial boundary on Eq. (1) to reflect the escaping trajectories whenever necessary. This is implemented by adding a restoring force **V** given by $\mathbf{V}(\|x\|) = -w(\|x - x_0\| - r) \cdot (x - x_0)$ where r and x_0 define the domain of interest, and the function w is given by

$$w(z) = \begin{cases} 0, & z < 0 \\ e^{mz} \cdot e^{-z^{-1}}, & z \geqslant 0 \end{cases} \quad (15)$$

for some $m \gg 1$. The presence of **V** ensures the existence of a stationary distribution of the modified system, and so long as $\langle w(x) \rangle \ll 1$ (i.e., the distribution near the domain boundary is negligible), the influence of the artificial wall is minimal in practice. As a result, all key equations in White-nLIM and Colored-nLIM remain effective for constructing the approximate system in practical applications. The effectiveness of both White-nLIM and Colored-nLIM in parameter estimation, across nonchaotic and chaotic systems, with or without prescribed nonlinear terms, is validated in the Supplemental Material [44].

An application to the ENSO RO model. We revisit the simple ENSO RO model consisting of the area-averaged sea surface temperature (SST) and thermocline, defined as the 20 °C isotherm depth (D20) [51–53]. The former is a primary indicator of ENSO, driven by oceanic and atmospheric processes, while the latter represents the heat exchange between surface and deeper waters [54]. Previous studies have sug-

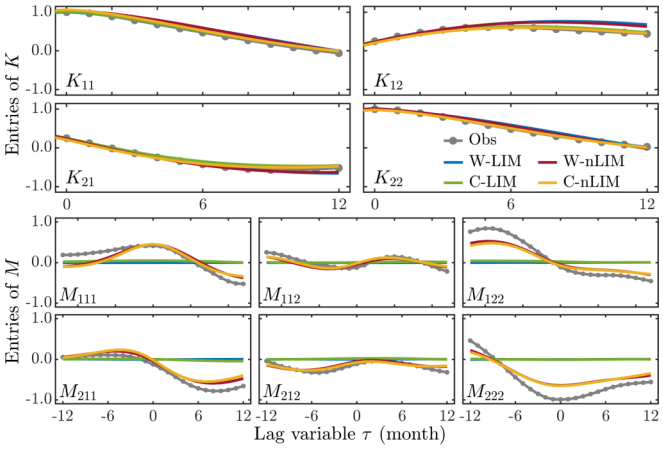


FIG. 1. Entries of the observed correlation function K^{obs} and M^{obs} of the preprocessed SST and D20 anomalies symbolized as 1 and 2, respectively, along with the correlation functions of the inverse models. The correlation functions of LIMs are computed using the analytic formulas (refer to Ref. [7]), while those of nLIMs are determined by a 5000-year simulation.

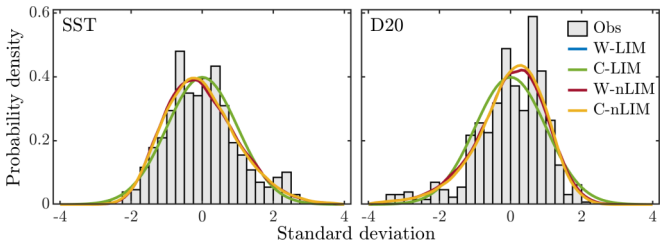


FIG. 2. The observed distributions of preprocessed SST and D20, along with the probability distributions of LIMs and nLIMs. White-LIM and Colored-LIM produce indistinguishable Gaussian distributions whose covariances are the diagonal entries of $K^{\text{obs}}(0)$. The distributions of White-nLIM and Colroed-nLIM are determined by a 5000-year simulation.

gested that their interaction forms the foundation of ENSO evolution [55,56].

The data are collected from the ORAS5 reanalysis from 1979 to 2021 [57]. SST is averaged over 170°W–120°W and 5°S–5°N, while D20 is averaged over 120°E–80°W and 5°S–5°N, following Ref. [7]. The linear trends and climatological means of SST and D20 are subtracted. Since LIMs and nLIMs do not account for seasonality, we take z-scores for each calendar month to eliminate seasonal dependence [7,22,58]. Finally, a 3-month sliding average is applied to smooth the data [59]. The resulting preprocessed dataset, denoted by $\{x(t)\} \subset \mathbb{R}^2$, where SST and D20 are indexed by 1 and 2, respectively, serves as the input (observations) for the linear and nonlinear inverse models. For simplicity, in this application, we refer to each approximate system by the same name as the algorithm used to construct it (e.g., "White-LIM"), with a slight abuse of terminology, when no confusion arises.

The numerical details and the resulting approximate systems are provided in the Supplemental Material [44]. White-nLIM and Colored-nLIM do not admit stationary distributions unless artificial boundaries are imposed. The noise correlation time $\gamma^{\text{C-LIM}} = 1.2$ months in Colored-LIM and $\gamma^{\text{C-nLIM}} = 0.72$ months in Colored-nLIM are determined by the γ-selection algorithms [see Eq. (A6)]. These timescales are of the same order as those in the XRO study [36]. Though the precise physical mechanisms represented by colored noise are unknown, they can be related to atmospheric processes such as the Madden-Julian oscillation [36,60,61]. In addition, regardless of the noise persistence, the entries of the quadratic term B^{nLIM}, related to physical processes such as atmospheric processes through nonlinear SST-wind stress feedback, are not significantly smaller in magnitude than those of A^{nLIM}, indicating that the quadratic dynamics are non-negligible, particularly during major ENSO events [36,62–64].

Figure 1 shows the observed correlation functions K^{obs} and M^{obs} along with the correlation functions of LIMs and nLIMs. As the approximate systems are constructed using K^{obs} by design, the resulting covariance inherently respects the observed one. In contrast, the nonlinear models extend this capability by using M^{obs}, enabling them to capture the variations of M^{obs} in the lag variable—a feature that LIMs cannot represent.

Figure 2 shows the observed distribution and the probability distributions of the approximate systems, and Table I shows their Wasserstein p-distances. The results indicate that the skewed nLIM distributions align more closely with the observed distributions, in contrast to the symmetric Gaussian distributions of LIMs. Notably, these findings remain valid even when the parameter $\gamma^{\text{C-nLIM}}$ is perturbed. A more detailed discussion, including a sensitivity analysis of this parameter, is provided in the Supplemental Material [44]. Given this skewed characteristic, the proposed methods may offer the potential to facilitate the study of asymmetric El Niño and La Niña events, especially in the context of inverse modeling with additive noise.

Concluding remarks. This study extended linear inverse modeling to a nonlinear framework by incorporating quadratic nonlinear dynamics and constant terms, as well as modeling random forcing with either white noise or colored noise, referred to as White-nLIM and Colored-nLIM, respectively. Through idealized numerical experiments, Colored-nLIM has demonstrated robust parameter estimation even under persistent stochastic forcing. Together with White-nLIM, the results demonstrated that the noise correlation time γ affects the approximate system continuously via the curvature of the correlation functions, with this continuity extending to the white-noise case at $\gamma = 0$. Applied to real-world time-series

TABLE I. Wasserstein p-distance[a] between the observed distribution and the probability distributions of LIMs and nLIMs.

	$p = 1$		$p = 2$	
Model	SST	D20	SST	D20
White-LIM	0.0972	0.1868	0.1342	0.2433
Colored-LIM	0.0972	0.1868	0.1342	0.2433
White-nLIM	0.0673	0.0948	0.0980	0.1223
Colored-nLIM	0.0619	0.0727	0.0961	0.1057

[a]Wasserstein p-distance, computed using [65], is unitless.

data for ENSO, both methods accurately reproduced the observed correlation functions and the skewed distributions of SST and D20, an improvement over the conventional linear models with additive noise, which cannot account for such asymmetries.

Although the stationary XRO shares conceptual similarities to Colored-nLIM, the two approaches differ in key aspects. XRO utilizes regression techniques to produce the approximate system [36]. As the colored noise is correlated with the state variables, standard regression methods may lead to biased estimates, requiring more advanced regression approaches [32,37]. In contrast, Colored-nLIM is primarily based on the local behaviors of correlation functions, so a robust algorithm may be needed to compute numerical derivatives. Furthermore, Colored-nLIM includes a constant term in the deterministic dynamics, ensuring consistency with the zero-mean condition often applied to climate anomaly data, an aspect not explicitly incorporated in the original XRO formulation.

The proposed nLIM framework can, in principle, be extended to incorporate polynomial dynamics of any finite order. However, this involves much higher-order correlation functions to estimate the dynamics. In practice, accurately obtaining these correlation functions may require substantially more observational data, which could be challenging. Thus, a trade-off must be carefully considered between the polynomial order (i.e., the number of estimated parameters) and the available sampling.

Finally, we note that this framework has limitations in certain studies, such as seasonal-modulated climate systems [22,36,58,66]. Future work may focus on extending the framework to incorporate periodicity, exploring its use in analyzing the chaotic dynamics of ENSO, or estimating key parameters in nonlinear systems subject to persistent stochastic forcing.

Acknowledgments. The authors would like to thank Yan-Ning Kuo for the valuable discussion that inspired this study. J.L. and H.A. were supported by the Council for Science, Technology and Innovation (CSTI), Cross-ministerial Strategic Innovation Promotion Program (SIP), the third period of SIP "Smart Energy Management System" (Grant No. JPJ012207, funding agency: JST). H.A. also received partial support from Japan Agency for Medical Research and Development, AMED, under Grant No. JP21zf0127005.

Data availability. The data that support the findings of this article are openly available [57], embargo periods may apply.

Appendix A: Review of linear inverse models. LIM describes the observations $\{x(t)\}$ with a zero mean by producing an approximate linear SDE

$$\frac{d}{dt}x = A^{\mathrm{LIM}}x + \sqrt{2Q^{\mathrm{LIM}}}\cdot \text{noise}. \tag{A1}$$

All eigenvalues of the linear dynamical matrix $A^{\mathrm{LIM}} \in \mathbb{R}^{n\times n}$ must have negative real parts, the stochastic matrix $Q^{\mathrm{LIM}} \in \mathbb{R}^{n\times n}$ is positive definite, and the stochastic forcing is modeled by either white or colored noise. Again, we first focus on the forward formulation.

The governing equation of White-LIM is given by

$$\frac{d}{dt}\mathbf{x} = \mathbf{A}\mathbf{x} + \sqrt{2\mathbf{Q}}\boldsymbol{\xi},$$

where $\boldsymbol{\xi}$ is the normalized white noise [1,67]. The correlation function $\mathbf{K}$ of the state vector $\mathbf{x}$ admits an explicit form given by $\mathbf{K}(\tau) = e^{\tau\mathbf{A}}\mathbf{K}(0)$ for $\tau \geqslant 0$, so the dynamics $\mathbf{A}$ can be formally written in terms of the matrix logarithm of $\mathbf{K}(\tau)\mathbf{K}(0)^{-1}$ and the lag variable [3,68]. However, for consistency with other modeling frameworks, the following is adopted:

$$\mathbf{K}'(0) = \mathbf{A}\mathbf{K}(0). \tag{A2}$$

The balance equation is given by

$$0 = \mathrm{Sym}(\,\mathbf{A}\mathbf{K}(0) + \mathbf{Q}\,). \tag{A3}$$

Hence, $\mathbf{A}$ and $\mathbf{Q}$ are characterized by $\mathbf{K}$.

On the other hand, the governing equation in the Colored-LIM framework is given by

$$\frac{d}{dt}\mathbf{x} = \mathbf{A}\mathbf{x} + \sqrt{2\mathbf{Q}}\boldsymbol{\eta},$$

where the stochastic forcing is the OU colored noise $\boldsymbol{\eta}$ with noise correlation time $\boldsymbol{\gamma}$. It has been shown that $\boldsymbol{\gamma}$ and the derivatives up to the third order of the correlation function $\mathbf{K}$ of $\mathbf{x}$ characterize the linear dynamics $\mathbf{A}$ [7]. Although $\mathbf{K}'''$ appears in Ref. [7] in the context of discussing (skew-)symmetry properties, it is not required for parameter estimation in practice. Indeed, the derivatives up to second order are sufficient:

$$\boldsymbol{\gamma}\,\mathbf{K}''(0) = \boldsymbol{\gamma}\,\mathbf{A}\mathbf{K}'(0) - \big(\mathbf{K}'(0) - \mathbf{A}\mathbf{K}(0)\big). \tag{A4}$$

The balance equation is then described by

$$0 = \mathrm{Sym}\big(\mathbf{A}\mathbf{K}(0) + \mathbf{Q}\mathbf{N}^T\big), \tag{A5}$$

where $\mathbf{N} = (\mathbf{I} - \boldsymbol{\gamma}\mathbf{A})^{-1}$. Therefore, $\mathbf{K}$ and $\boldsymbol{\gamma}$ uniquely determine $\mathbf{A}$ and $\mathbf{Q}$.

Let K^{obs} be the observed correlation functions of the given observations $\{x(t)\}$. White-LIM makes the white-noise assumption, estimating the first-order dynamics $A^{\text{W-LIM}}$ and the stochastic matrix $Q^{\text{W-LIM}}$ using Eqs. (A2) and (A3) with K^{obs}. On the contrary, Colored-LIM models the random forcing with an OU colored noise. For a prescribed noise correlation time $\gamma^{\text{C-LIM}}$, it estimates the dynamics $A^{\text{C-LIM}}$ and stochastic matrix $Q^{\text{C-LIM}}$ by solving Eqs. (A4) and (A5) with K^{obs}. Otherwise, $\gamma^{\text{C-LIM}}$ is determined via the γ-selection algorithm

$$\gamma^{\text{C-LIM}} = \arg\min_{\gamma>0}\big\|K(A(\gamma), Q(\gamma), \gamma) - K^{\mathrm{obs}}\big\|, \tag{A6}$$

where $A(\gamma)$ and $Q(\gamma)$ are the solutions of Eqs. (A4) and (A5) with the running varaible γ, K is the corresponding correlation function, and $\|\cdot\|$ is the function norm over the interval $[0, l]$ for some hyperparameter l [7]. Finally, it is important to note that the covariance determined by the approximate system (A1) equals $K^{\mathrm{obs}}(0)$.

Appendix B: Proof of nLIMs. In this Appendix, we give a proof of Colored-nLIM [i.e., Eqs. (8)–(14)]. White-nLIM follows the same workflow and is easier, so we leave it to the readers.

The Fokker-Planck equation of Eq. (7) reads

$$\frac{\partial}{\partial t}P = -\sum_i \frac{\partial}{\partial x_i}F_i(x)P - \sum_{i,j}(\sqrt{2\mathbf{Q}})_{ij}\frac{\partial}{\partial x_i}\eta_j P + \frac{1}{\boldsymbol{\gamma}}\sum_i \frac{\partial}{\partial \eta_i}\eta_i P + \frac{1}{2\boldsymbol{\gamma}^2}\sum_i \frac{\partial^2}{\partial {\eta_i}^2}P$$
$$= \mathbf{L}_{\mathrm{FP}}P,$$

where $\mathbf{L}_{\mathrm{FP}}$ is the Fokker-Planck operator and

$$F_i(x) = \sum_{j,k}\mathbf{B}_{ijk}x_j x_k + \sum_j \mathbf{A}_{ij}x_j + \mathbf{C}_i.$$

Let P_{st} denote a stationary distribution (i.e., $\mathbf{L}_{\mathrm{FP}}P_{st} = 0$) and $\mathbf{L}^*_{\mathrm{FP}}$ denote the adjoint Fokker-Planck operator given by

$$\mathbf{L}^*_{\mathrm{FP}} = \sum_i F_i\frac{\partial}{\partial x_i} + \sum_{i,j}(\sqrt{2\mathbf{Q}})_{ij}\eta_j\frac{\partial}{\partial x_i} - \frac{1}{\boldsymbol{\gamma}}\sum_i \eta_i\frac{\partial}{\partial \eta_i} + \frac{1}{2\boldsymbol{\gamma}^2}\sum_i \frac{\partial^2}{\partial {\eta_i}^2}.$$

Then, following Refs. [69,70], we have

$$\begin{aligned}\langle f(\mathbf{x}(\cdot+\tau),\boldsymbol{\eta}(\cdot+\tau)), g(\mathbf{x}(\cdot),\boldsymbol{\eta}(\cdot))\rangle &= \int_{\mathbb{R}^{2n}} f(x,\eta)e^{\tau\mathbf{L}_{\mathrm{FP}}}P_{st}(x,\eta)g(x,\eta)\,dxd\eta \\ &= \int_{\mathbb{R}^{2n}} g(x,\eta)P_{st}(x,\eta)e^{\tau\mathbf{L}^*_{\mathrm{FP}}}f(x,\eta)\,dxd\eta \\ &= \int_{\mathbb{R}^{2n}} g(x,\eta)P_{st}(x,\eta)\left(\mathbf{I}+\tau\mathbf{L}^*_{\mathrm{FP}}+\frac{\tau^2}{2}\mathbf{L}^*_{\mathrm{FP}}\mathbf{L}^*_{\mathrm{FP}}\right)f(x,\eta)\,dxd\eta + O(\tau^3). \end{aligned} \tag{B1}$$

A straightforward computation shows that

$$\mathbf{L}^*_{\mathrm{FP}}x_p = \sum_{j,k}\mathbf{B}_{pjk}x_j x_k + \sum_j \mathbf{A}_{pj}x_j + \mathbf{C}_p + \sum_j \eta_j(\sqrt{2\mathbf{Q}})_{pj}. \tag{B2}$$

Therefore, by Eqs. (B1) and (B2), we compute the first-order derivative of the (p,q)-entry of $\mathbf{K}$ as follows:

$$\mathbf{K}'_{pq}(0) = \int_{\mathbb{R}^{2n}} x_q P_{st}(x,\eta)\mathbf{L}^*_{\mathrm{FP}}x_p\,dxd\eta = \int_{\mathbb{R}^{2n}} x_q P_{st}(x,\eta)\left(\sum_{j,k}\mathbf{B}_{pjk}x_j x_k + \sum_j \mathbf{A}_{pj}x_j + \mathbf{C}_p + \sum_j \eta_j(\sqrt{2\mathbf{Q}})_{pj}\right)dxd\eta$$
$$= \sum_{j,k}\mathbf{B}_{pjk}\mathbf{M}_{jkq}(0) + \sum_j \mathbf{A}_{pj}\mathbf{K}_{jq}(0) + \mathbf{C}_p\mathbf{E}_q + \sum_j(\sqrt{2\mathbf{Q}})_{pj}(\mathbf{K}_{\boldsymbol{\eta}\mathbf{x}})_{jq}.$$

This is equivalent to Eq. (9) in tensor form.

The computation of $\mathbf{K}''(0)$, $\mathbf{M}'(0)$, and $\mathbf{M}''(0)$ follows a similar procedure. We leave these derivations as an exercise to the readers.

Finally, with the fact that $\langle\boldsymbol{\eta}(\cdot)\rangle = 0$, Eqs. (8), (13), and (14) can be obtained by considering

$$0 = \int_{\mathbb{R}^{2n}} x_j\mathbf{L}_{\mathrm{FP}}P_{st}\,dxd\eta = \int_{\mathbb{R}^{2n}} P_{st}\mathbf{L}^*_{\mathrm{FP}}x_j\,dxd\eta,$$

$$0 = \int_{\mathbb{R}^{2n}} \eta_i x_j\mathbf{L}_{\mathrm{FP}}P_{st}\,dxd\eta = \int_{\mathbb{R}^{2n}} P_{st}\mathbf{L}^*_{\mathrm{FP}}\eta_i x_j\,dxd\eta,$$

and

$$0 = \int_{\mathbb{R}^{2n}} x_i x_j\mathbf{L}_{\mathrm{FP}}P_{st}\,dxd\eta = \int_{\mathbb{R}^{2n}} P_{st}\mathbf{L}^*_{\mathrm{FP}}x_i x_j\,dxd\eta,$$

respectively. They are also left to the readers as an exercise.

[1] C. Penland, Random forcing and forecasting using principal oscillation pattern analysis, Mon. Weather Rev. **117**, 2165 (1989).

[2] C. Penland, A stochastic model of IndoPacific sea surface temperature anomalies, Physica D **98**, 534 (1996).

[3] C. Penland and T. Magorian, Prediction of Niño 3 sea surface temperatures using linear inverse modeling, J. Clim. **6**, 1067 (1993).

[4] C. Penland and L. Matrosova, A balance condition for stochastic numerical models with application to the El Niño-Southern Oscillation, J. Clim. **7**, 1352 (1994).

[5] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, On dynamic mode decomposition: Theory and applications, J. Comput. Dyn. **1**, 391 (2014).

[6] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic Mode Decomposition* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2016).

[7] J. Lien, Y.-N. Kuo, H. Ando, and S. Kido, Colored linear inverse model: A data-driven method for studying dynamical systems with temporally correlated stochasticity, Phys. Rev. Res. **7**, 023042 (2025).

[8] T. W. Muld, G. Efraimsson, and D. S. Henningson, Flow structures around a high-speed train extracted using proper orthogonal decomposition and dynamic mode decomposition, Comput. Fluids **57**, 87 (2012).

[9] C. Pan, D. Yu, and J. Wang, Dynamical mode decomposition of Gurney flap wake flow, Theor. Appl. Mech. Lett. **1**, 012002 (2011).

[10] S.-I. An and F.-F. Jin, Nonlinearity and asymmetry of ENSO, J. Clim. **17**, 2399 (2004).

[11] J. Liang, X.-Q. Yang, and D.-Z. Sun, Factors determining the asymmetry of ENSO, J. Clim. **30**, 6097 (2017).

[12] C. M. Aiken, A. Santoso, S. McGregor, and M. H. England, The 1970's shift in ENSO dynamics: A linear inverse model perspective, Geophys. Res. Lett. **40**, 1612 (2013).

[13] M. A. Alexander, L. Matrosova, C. Penland, J. D. Scott, and P. Chang, Forecasting pacific SSTs: Linear inverse model predictions of the PDO, J. Clim. **21**, 385 (2008).

[14] J. Lou, T. J. O'Kane, and N. J. Holbrook, A linear inverse model of tropical and South Pacific seasonal predictability, J. Clim. **33**, 4537 (2020).

[15] E. Di Lorenzo, G. Liguori, N. Schneider, J. C. Furtado, B. T. Anderson, and M. A. Alexander, ENSO and meridional modes: A null hypothesis for Pacific climate variability, Geophys. Res. Lett. **42**, 9440 (2015).

[16] I. Richter, Y. Kosaka, S. Kido, and H. Tokinaga, The tropical Atlantic as a negative feedback on ENSO, Clim. Dyn. **61**, 309 (2023).

[17] C. Wang, C. Deser, J.-Y. Yu, P. DiNezio, and A. Clement, El Niño and Southern Oscillation (ENSO): A review, in *Coral Reefs of the Eastern Tropical Pacific: Persistence and Loss in a Dynamic Environment*, edited by P. W. Glynn, D. P. Manzello, and I. C. Enochs, Coral Reefs of the World, Vol. 8 (Springer Netherlands, Dordrecht, 2017), pp. 85–106.

[18] T. DelSole, Can quasigeostrophic turbulence be modeled stochastically? J. Atmos. Sci. **53**, 1617 (1996).

[19] T. DelSole, A fundamental limitation of Markov models, J. Atmos. Sci. **57**, 2158 (2000).

[20] C. Martinez-Villalobos, D. J. Vimont, C. Penland, M. Newman, and J. D. Neelin, Calculating state-dependent noise in a linear inverse model framework, J. Atmos. Sci. **75**, 479 (2018).

[21] P. D. Sardeshmukh, G. P. Compo, and C. Penland, Need for caution in interpreting extreme weather statistics, J. Clim. **28**, 9166 (2015).

[22] S.-I. Shin, P. D. Sardeshmukh, M. Newman, C. Penland, and M. A. Alexander, Impact of annual cycle on ENSO variability and predictability, J. Clim. **34**, 171 (2021).

[23] J. Lien, H. Ando, I. Richter, and S. Kido, Noise persistence in a cyclostationary linear inverse model framework and its impact on ENSO evolution and prediction skills, J. Atmos. Sci. (unpublished).

[24] C. Martinez-Villalobos, M. Newman, D. J. Vimont, C. Penland, and J. David Neelin, Observed El Niño-La Niña asymmetry in a linear model, Geophys. Res. Lett. **46**, 9909 (2019).

[25] C. Chen, M. A. Cane, N. Henderson, D. E. Lee, D. Chapman, D. Kondrashov, and M. D. Chekroun, Diversity, nonlinearity, seasonality, and memory effect in ENSO simulation and prediction using empirical model reduction, J. Clim. **29**, 1809 (2016).

[26] S. De, N. K. Agarwal, J. S. Chowdary, and D. Patekar, Scale interactions during ENSO—A new dynamical mechanism on unveiling ENSO-monsoon nonlinearity, Theor. Appl. Climatol. **155**, 8119 (2024).

[27] D. Kondrashov, S. Kravtsov, A. W. Robertson, and M. Ghil, A hierarchy of data-based ENSO models, J. Clim. **18**, 4425 (2005).

[28] D. Kondrashov, M. D. Chekroun, and M. Ghil, Data-driven non-Markovian closure models, Physica D **297**, 33 (2015).

[29] S. Kravtsov, D. Kondrashov, and M. Ghil, Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability, J. Clim. **18**, 4404 (2005).

[30] A. Lück and V. Wolf, Generalized method of moments for estimating parameters of stochastic reaction networks, BMC Syst. Biol. **10**, 98 (2016).

[31] Y. Li and Z. Wei, Dynamics and optimal control of a stochastic coronavirus (COVID-19) epidemic model with diffusion, Nonlinear Dyn. **109**, 91 (2022).

[32] A. J. Majda and J. Harlim, Physics constrained nonlinear regression models for time series, Nonlinearity **26**, 201 (2013).

[33] D. Olabode, J. Culp, A. Fisher, A. Tower, D. Hull-Nye, and X. Wang, Deterministic and stochastic models for the epidemic dynamics of COVID-19 in Wuhan, China, Math. Biosci. Eng. **18**, 950 (2021).

[34] G. Schneider, P. F. Craigmile, and R. Herbei, Maximum likelihood estimation for stochastic differential equations using sequential Gaussian-process-based optimization, Technometrics **59**, 178 (2017).

[35] J. Timmer, Parameter estimation in nonlinear stochastic differential equations, Chaos Solitons Fract. **11**, 2571 (2000).

[36] S. Zhao, F.-F. Jin, M. F. Stuecker, P. R. Thompson, J.-S. Kug, M. J. McPhaden, M. A. Cane, A. T. Wittenberg, and W. Cai, Explainable El Niño predictability from climate mode interactions, Nature (London) **630**, 891 (2024).

[37] L. Ljung, *System Identification: Theory for the User* (Prentice Hall PTR, Upper Saddle River, NJ, 1999), 2nd ed.

[38] X. Wang, J. Feng, Q. Liu, Y. Li, and Y. Xu, Neural network-based parameter estimation of stochastic differential equations driven by Lévy noise, Physica A **606**, 128146 (2022).

[39] K. Lee and E. N. Johnson, State estimation using Gaussian process regression for colored noise systems, in *Proceedings of the 2017 IEEE Aerospace Conference, Big Sky, MT, USA* (IEEE, Piscataway, NJ, 2017), pp. 1–8.

[40] X. Liu and J. Lu, Least squares based iterative identification for a class of multirate systems, Automatica **46**, 549 (2010).

[41] J. Vialard, F. F. Jin, M. J. McPhaden, A. Fedorov, W. Cai, S.-I. An, D. Dommenget, X. Fang, M. F. Stuecker, C. Wang *et al.*, The El Niño Southern Oscillation (ENSO) recharge oscillator conceptual model: Achievements and future prospects, Reviews of Geophysics **63**, e2024RG000843 (2025).

[42] X. Wang, J. Feng, Q. Liu, and Y. Xu, Noise-induced alternations and data-driven parameter estimation of a stochastic perceptual model, Eur. Phys. J.: Spec. Top. **234**, 707 (2025).

[43] P. McCullagh and J. A. Nelder, *Generalized Linear Models*, 2nd ed. (Routledge, New York, 1989), p. 532.

[44] See Supplemental Material at http://link.aps.org/supplemental/10.1103/ds1j-fx3v for numerical details of the nLIMs, validation analyses, and additional results for the ENSO RO model.

[45] G. Gradziuk, G. Torregrosa, and C. P. Broedersz, Irreversibility in linear systems with colored noise, Phys. Rev. E **105**, 024118 (2022).

[46] S. L. Brunton, J. L. Proctor, and J. N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. USA **113**, 3932 (2016).

[47] S. K. Goharoodi, K. Dekemele, L. Dupre, M. Loccufier, and G. Crevecoeur, Sparse identification of nonlinear duffing oscillator from measurement data, IFAC-PapersOnLine **51**, 162 (2018).

[48] T. Ryu and S. Baek, Development of data-driven modeling method for nonlinear coupling components, Sci. Rep. **14**, 14841 (2024).

[49] A. J. Majda and Y. Yuan, Fundamental limitations of ad hoc linear and quadratic multi-level regression models for physical systems, Discrete Contin. Dyn. Syst. B **17**, 1333 (2012).

[50] C. Martinez-Villalobos, B. Dewitte, R. D. Garreaud, and L. Loyola, Extreme coastal El Niño events are tightly linked to the development of the Pacific Meridional Modes, npj Clim. Atmos. Sci. **7**, 123 (2024).

[51] F.-F. Jin and S.-I. An, Thermocline and zonal advective feedbacks within the equatorial ocean recharge oscillator model for ENSO, Geophys. Res. Lett. **26**, 2989 (1999).

[52] F.-F. Jin, H.-C. Chen, S. Zhao, M. Hayashi, C. Karamperidou, M. F. Stuecker, R. Xie, and L. Geng, Simple ENSO models, *El Niño Southern Oscillation in a Changing Climate* (Wiley-American Geophysical Union, Hoboken, NJ, 2020), Chap. 6, pp. 119–151.

[53] C. S. Meinen and M. J. McPhaden, Observations of warm water volume changes in the equatorial Pacific and their relationship to El Niño and La Niña, J. Clim. **13**, 3551 (2000).

[54] C. Wen, A. Kumar, Y. Xue, and M. J. McPhaden, Changes in tropical Pacific thermocline depth and their relationship to ENSO after 1999, J. Clim. **27**, 7230 (2014).

[55] G. Burgers, F.-F. Jin, and G. J. van Oldenborgh, The simplest ENSO recharge oscillator, Geophys. Res. Lett. **32**, 2005GL022951 (2005).

[56] F.-F. Jin, L. Lin, A. Timmermann, and J. Zhao, Ensemble-mean dynamics of the ENSO recharge oscillator under state-dependent stochastic forcing, Geophys. Res. Lett. **34**, L03807 (2007).

[57] Copernicus Climate Change Service, ORAS5 global ocean reanalysis monthly data from 1958 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS), 10.24381/cds.67e8eeb7.

[58] S. Kido, I. Richter, T. Tozuka, and P. Chang, Understanding the interplay between ENSO and related tropical SST variability using linear inverse models, Clim. Dyn. **61**, 1029 (2023).

[59] M. Newman, S.-I. Shin, and M. A. Alexander, Natural variation in ENSO flavors, Geophys. Res. Lett. **38**, L14705 (2011).

[60] A. Capotondi, P. D. Sardeshmukh, and L. Ricciardulli, The nature of the stochastic wind forcing of ENSO, J. Clim. **31**, 8081 (2018).

[61] Y. Peng, W. Duan, and J. Xiang, Effect of stochastic MJO forcing on ENSO predictability, Adv. Atmos. Sci. **28**, 1279 (2011).

[62] K.-Y. Choi, G. A. Vecchi, and A. T. Wittenberg, ENSO transition, duration, and amplitude asymmetries: Role of the nonlinear wind stress coupling in a conceptual model, J. Clim. **26**, 9462 (2013).

[63] T. Geng, W. Cai, and L. Wu, Two types of ENSO varying in tandem facilitated by nonlinear atmospheric convection, Geophys. Res. Lett. **47**, e2020GL088784 (2020).

[64] I.-S. Kang and J.-S. Kug, El Niño and La Niña sea surface temperature anomalies: Asymmetry characteristics associated with their wind stress anomalies, J. Geophys. Res.: Atmos. **107**, ACL 1 (2002).

[65] N. Kolbe, Wasserstein distance, https://github.com/nklb/wasserstein-distance.

[66] D. J. Vimont, M. Newman, D. S. Battisti, and S.-I. Shin, The role of seasonality and the ENSO mode in Central and East Pacific ENSO growth and evolution, J. Clim. **35**, 3195 (2022).

[67] C. Penland and P. D. Sardeshmukh, The optimal growth of tropical sea surface temperature anomalies, J. Clim. **8**, 1999 (1995).

[68] C. Penland, The Nyquist issue in linear inverse modeling, Mon. Weather Rev. **147**, 1341 (2019).

[69] P. Jung and H. Risken, Motion in a double-well potential with additive colored Gaussian noise, Z. Phys. B **61**, 367 (1985).

[70] H. Risken, *The Fokker-Planck Equation: Equation: Methods of Solution and Applications*, Springer Series in Synergetics, 2nd ed., Vol. 18, (Springer, Berlin, 1989), pp. xiv + 472.