UCLA

UCLA Previously Published Works

Title

Stochastic rectification of fast oscillations on slow manifold closures

Permalink

Journal

ISSN

Authors

Chekroun, Mickaël D
Liu, Honghu
McWilliams, James C

Publication Date

DOI

Peer reviewed

Stochastic rectification of fast oscillations on slow manifold closures

Mickaël D. Chekroun[a,b] [ID], Honghu Liu[c] [ID], and James C. McWilliams[a,1]

[a]Department of Atmospheric & Oceanic Sciences, University of California, Los Angeles, CA 90095-1565; [b]Department of Earth and Planetary Sciences, Weizmann Institute, Rehovot 76100, Israel; and [c]Department of Mathematics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

The problems of identifying the slow component (e.g., for weather forecast initialization) and of characterizing slow–fast interactions are central to geophysical fluid dynamics. In this study, the related rectification problem of slow manifold closures is addressed when breakdown of slow-to-fast scales deterministic parameterizations occurs due to explosive emergence of fast oscillations on the slow, geostrophic motion. For such regimes, it is shown on the Lorenz 80 model that if 1) the underlying manifold provides a good approximation of the optimal nonlinear parameterization that averages out the fast variables and 2) the residual dynamics off this manifold is mainly orthogonal to it, then no memory terms are required in the Mori–Zwanzig full closure. Instead, the noise term is key to resolve, and is shown to be, in this case, well modeled by a state-independent noise, obtained by means of networks of stochastic nonlinear oscillators. This stochastic parameterization allows, in turn, for rectifying the momentum-balanced slow manifold, and for accurate recovery of the multiscale dynamics. The approach is promising to be further applied to the closure of other more complex slow–fast systems, in strongly coupled regimes.

multiscale chaos | stochastic parameterization | multiscale closure | slow manifold | fast oscillations

For complex systems, most models only resolve spatiotemporal scales, processes, and variables to a certain level of accuracy because of the high computational costs associated with high-fidelity simulations. Such truncations of scales, processes, or fast variables often limit the reliability and usefulness of simulations, especially for scientific, engineering, and societal applications where longer-term model predictions are needed to guide decisions. In many applications, the neglected and unresolved variables along with their interactions with the resolved ones are often key to include, and parameterizations of these ingredients are sought to provide "closure models" aimed at faithfully emulating solutions of the fully resolved problem while solving only a subset of relevant variables (1).

The new advances in data-driven stochastic modeling techniques over the last two decades open up prodigious perspectives for finding, from observational or high-fidelity simulations data, a few dynamical equations able to simulate, with efficiency, datasets with complex spatiotemporal variability such as that arising in various geophysical applications (see, e.g., refs. 2–7). The diversity of these applications is explained by a common thread which is necessary for the derivation of a good data-driven closure model. In the context of dynamical systems with both slow and fast variations, its derivation relies on the efficient learning of fast, stochastic variables along with their interaction laws with the slow variables, to emulate the missing dynamical information which is central in the organization of, for example, the observed larger-scale patterns. This operation is accomplished within a theoretical framework that guides the search for these hidden variables and interactions elements, namely, the Mori–Zwanzig (MZ) formalism (8, 9) from statistical physics, extended to a modern and general framework by refs. 10–12. This framework predicts that the theoretical optimal closure model in its general formulation involves nonlinear terms that provide the average

motion of the observed variables while the (usually fast) fluctuations are parameterized by (possibly nonlinear) memory terms plus noise terms.

The purpose of this article is to identify 1) situations for which no memory but noise terms are crucial for accurate closure and 2) how to model the noise term by a dynamics-informed approach. To serve and illustrate our purpose, we restrict our attention to the Lorenz 80 (L80) model from atmospheric science (13–17), in highly nonlinear regimes in which the rotational Rossby variables are populated by fast gravity waves which are not only no longer a functional of the former but correspond to an "explosive" breakdown of slow-to-fast scales deterministic parameterizations (16, 17), in contrast with other breakdowns associated with, for example, exponentially small errors (15, 18). We believe the approach presented here is not limited to the L80 model and would actually apply to any slow–fast system for which 1) a manifold filtering out the fast oscillations in some optimal sense is known (typically least-square sense) and 2) the residual dynamics off this manifold is prominently orthogonal to it, namely, mainly uncorrelated to the slow dynamics.

The Slow Optimal Parameterizing Manifold and Its Relation to the Conditional Expectation

Recently, in the context of the slow manifold problem for the atmosphere, the concept of a slow manifold has been revisited as that of a slow optimal parameterizing manifold (OPM) (16).

Significance

Slow–fast systems arise in many scientific applications, in particular in atmospheric and oceanic flows with fast inertia–gravity waves and slow geostrophic motions. When the slow and fast variables are strongly coupled—symptomatic of breakdown of slow-to-fast scales deterministic parameterizations—it remains a challenge to derive reduced systems able to capture the dynamics. Here, generic ingredients for successful reduction of such systems are identified and illustrated for the paradigmatic atmospheric Lorenz 80 model. The approach relies on a filtering operated through a nonlinear parameterization that separates the full dynamics into its slow motion and fast residual dynamics. The latter is mainly orthogonal to the former and is modeled via networks of stochastic nonlinear oscillators, independent of the slow dynamics.

[1]To whom correspondence may be addressed. Email: jcm@atmos.ucla.edu.

Dynamically, the slow OPM provides the manifold on which lies the average motion of the neglected fast variables as a function of the resolved scales (ref. 19, theorem 4). It does not assume nor require any exact relationships between the slow and fast scales. In that sense, the slow OPM relates, for slow–fast systems, to the conditional expectation in optimal prediction theories (10) by providing the optimal closure, when the latter is sought by conditioning only on the current time value of the resolved variables (ref. 19, theorem 5).

To make more precise this statement, consider an evolution equation, inspired from fluid motion, of the form

$$\frac{\mathrm{d}u}{\mathrm{d}t} = Au + B(u, u) + F, \qquad u \in E, \qquad \textbf{[1]}$$

in which E denotes a vector space, either infinite or finite dimensional, assumed to be the sum of a reduced state space, $E_{\mathfrak{c}}$, spanning the "slow" variables, and of its complement, $E_{\mathfrak{s}}$, spanning the "fast" variables. The forcing term F is considered here to be deterministic, while A denotes a linear operator not necessarily self-adjoint (but diagonalizable over the set of complex numbers $\mathbb{C}$) and B denotes a quadratic operator which may account for a loss of regularity (such as for nonlinear advection) in the infinite-dimensional setting (20).

Assuming that Eq. 1 possesses an invariant probability measure μ that satisfies the right ergodicity assumption (19), the slow OPM is then obtained as the graph of the mapping $\Psi^* : E_{\mathfrak{c}} \to E_{\mathfrak{s}}$ which solves the following minimization problem:

$$\min_{\Psi \in \mathcal{E}} \int_{(X, Y) \in E_{\mathfrak{c}} \times E_{\mathfrak{s}}} \| Y - \Psi(X) \|^2 \, \mathrm{d}\mu, \qquad \textbf{[2]}$$

where $\mathcal{E}$ denotes the space of functions from $E_{\mathfrak{c}}$ to $E_{\mathfrak{s}}$, which are square-integrable with respect to the "projection" (push forward) $\mathfrak{m}$ of the probability measure μ onto $E_{\mathfrak{c}}$.

Under these conditions, the minimization problem Eq. 2 possesses a unique solution in $\mathcal{E}$: the slow OPM given by

$$\Psi^*(X) = \int_{E_{\mathfrak{s}}} Y \, \mathrm{d}\mu_X(Y), \qquad X \in E_{\mathfrak{c}}, \qquad \textbf{[3]}$$

where μ_X denotes the disintegrated probability measure of μ conditioned on X, the latter denoting, roughly speaking, the probability distribution (in $E_{\mathfrak{s}}$) of the variables Y that project onto X; see ref. 19 for more details.

Denoting by $\mathcal{F}$ the vector field associated with Eq. 1, that is, $\mathcal{F}(u) = Au + B(u, u) + F$, by $\Pi_{\mathfrak{c}}$ (resp. $\Pi_{\mathfrak{s}}$), the projector onto $E_{\mathfrak{c}}$ (resp. $E_{\mathfrak{s}}$), and assuming that B is symmetric for the sake of simplicity, it is easy to show that the conditional expectation, $\mathbb{E}_{\mu}[\mathcal{F}|\Pi_{\mathfrak{c}}]$, satisfies

$$\begin{aligned} \mathbb{E}_{\mu}[\mathcal{F}|\Pi_{\mathfrak{c}}](X) = A_{\mathfrak{c}} X + \Pi_{\mathfrak{c}} B(X, X) \\ + 2\Pi_{\mathfrak{c}} B(X, \Psi^*(X)) + F_{\mathfrak{c}} + \zeta(X), \end{aligned} \qquad \textbf{[4]}$$

with $\zeta(X) = \int_{Y \in E_{\mathfrak{s}}} \Pi_{\mathfrak{c}} B(Y, Y) \, \mathrm{d}\mu_X(Y)$. This latter term provides the average contribution of the high–high interactions to the evolution of the low modes. When $\zeta \approx 0$ [also known as the centering condition (1)], we observe thus, from Eq. 4, that determining the conditional expectation of $\mathcal{F}$ in the MZ expansion boils down to determining the OPM Ψ^*.

As shown in ref. 19, the interest of this remark is that it allows for envisioning the determination of the conditional expectation, via approximations to the (slow) OPM. The latter have been shown to be efficiently sought as homotopic deformations of parameterizations known to be relevant for certain regimes (such as, e.g., invariant/slow manifolds). We then seek for optimal homotopic deformations to reach dynamical regimes beyond the domain of validity of the original parameterization, by optimizing Eq. 2 in the corresponding class of parameterizations. Once a

relevant family $\mathcal{E}' = \{\Psi_\tau\}$ of such parameterizations is identified with τ, the parameter that controls the continuous deformations (see ref. 19, sections 4.3 and 4.4), the ergodicity assumption and mixing properties of the system allows us often to bypass the demanding (and often out-of-reach) knowledge of μ, required for the minimization of Eq. 2, by minimizing instead the parameterization defect

$$\mathcal{Q}_T(\Psi_\tau) = \overline{\| u_{\mathfrak{s}}(t) - \Psi_\tau(u_{\mathfrak{c}}(t)) \|^2}, \qquad \textbf{[5]}$$

estimated from high-fidelity trajectories obtained from integration of Eq. 1. Here, $\overline{(\cdot)}$ denotes the time mean over a training interval of length T, while $u_{\mathfrak{s}}(t)$ and $u_{\mathfrak{c}}(t)$ denote the projections of $u(t)$ onto $E_{\mathfrak{s}}$ and the reduced state space $E_{\mathfrak{c}}$, respectively.

Obviously, shifting the full minimization problem Eq. 2 to the data-driven minimization of $\mathcal{Q}_T(\Psi_\tau)$ is subject to a good choice of T (the length of the training interval), and a good choice of the parametric family of parameterizations $\{\Psi_\tau\}$. When these two ingredients are well chosen (see ref. 19 for a discussion), the optimal deformation parameter τ^* is found by minimizing $\mathcal{Q}_T(\Psi_\tau)$ in τ. The corresponding optimal parameterization Ψ_{τ^*} (in $\mathcal{E}'$) approximates then the OPM given by Eq. 3. We refer to ref. 19 for OPM examples as obtained by this approach. As a common denominator, it has been shown in ref. 19 that, by exploiting homotopic deformations of invariant or slow manifold approximations, analytic formulas for the approximation of OPMs are accessible for a broad class of nonlinear systems, and in a variety of dynamical regimes.

The Memory and Noise Correction Terms

Often, the conditional expectation alone (and thus OPM) is insufficient to provide a satisfactory closure. In this case, the neglected variables exert fluctuating driving forces on the explicit variables which are distinct from any exact relationships between the slow and fast scales or nonlinear functional dependence that would be conditioned only on the current state of the slow variables.

One needs then to rely on guidance as provided by the MZ theory, namely on the MZ closure of Eq. 1 in $E_{\mathfrak{c}}$ which is written formally as

$$\dot{X} = \mathbb{E}_{\mu}[\mathcal{F}|\Pi_{\mathfrak{c}}](X) + \int_0^t \Gamma(t - s; \mathcal{F}) [X(s)] \, \mathrm{d}s + \eta(t). \qquad \textbf{[6]}$$

In Eq. 6, the term $\eta(t)$ is void of slow oscillations (transversal dynamics to the OPM), and the kernel $\Gamma(t - s; \mathcal{F})$ is a time-lagged damping operator that depends possibly nonlinearly on $\mathcal{F}$, while $[X(s)]$ is used to indicate the dependence of this operator on the past history of the slow dynamics. Typically, $\eta(t)$ is sought as a random force that is uncorrelated with $X(t)$. These memory and noise terms, when needed to derive an efficient closure, are often quite challenging to calculate.

Already for weakly coupled nonlinear systems, that is, systems in which coupled terms are order(s) of magnitude smaller than the uncoupled terms, these memory and noise terms take a hierarchical form in which repeated temporal convolutions are involved. These repeated convolutions, albeit possibly deep in depth, can be approximated by means of multilayer stochastic models (MSMs) that add hidden layers of stochastic variables aimed at learning the corresponding (exponentially decaying) memory kernels (21, 22). However, for strongly coupled systems for which the coupled and uncoupled terms are of the same order of magnitude, and the disparate-scale interactions cover a wide range of scales, the constitutive elements of MSMs become challenging to determine, and to find a generic approach for determining the constitutive terms in the MZ closure Eq. 6 remains still widely open.

In what follows, we propose a general approach to determine the noise term in Eq. **6** for closing systems in which the conditional expectation is accurately determined via a slow OPM which captures most of the slow motion, while the residual dynamics is mainly fast and "orthogonal" to this OPM. For such situations, we argue that the memory terms are not required in the MZ decomposition Eq. **6**, while the design of the noise term $\eta(t)$ plays a key role to emulate the full dynamics with high fidelity. The following L80 model (13) serves as an emblematic example to illustrate these features.

The L80 Model and the Balance Equation Closure

Atmospheric and oceanic flows constrained by Earth's rotation satisfy an approximately geostrophic momentum balance on larger scales, associated with slow evolution on time scales of days, but they also exhibit fast inertia–gravity wave oscillations. The problems of identifying the slow component (e.g., for weather forecast initialization) and of characterizing slow–fast interactions are central to geophysical fluid dynamics, and the former was first coined as a slow manifold problem by Leith (23). The L63 model (24) (famous for its chaotic strange attractor) is a paradigm for the geostrophic component, while the L80 model (13) is its paradigmatic successor both for the generalization of slow balance and for slow–fast coupling.

The idea of Leith was to filter out, on an analytical basis, the fast gravity waves for the initialization of the primitive equations of the atmosphere. The motivation was that small errors in a "proper balance" between the fast time scale motion associated with gravity waves and slower motions such as those associated with the Rossby waves lead typically to an abnormal evolution of gravity waves, which, in turn, can cause appreciable deviations in weather forecasts. This filtering approach has a long history in forecast initialization (e.g., refs. 25 and 26).

Leith's idea was appealing for dealing with this filtering problem, but uncertainty in the definition of a slow manifold for finite Rossby number has led to a proliferation of different schemes, on one hand, and to the question of whether a slow invariant manifold even exists at finite Rossby number, on the other (e.g., refs. 15 and 27–30). In this atmospheric context, it was, in fact, shown that the generation of exponentially small inertia–gravity oscillations ($\sim \exp(-c/R_o)$ for c a constant) takes place for finite times, synonymous with the breakdown of quasigeostrophic balance (31) and of other slow-to-fast parameterizations that would be deterministic.

Actually, even more severe breakdowns of slow-to-fast deterministic parameterizations than through exponential smallness may occur. This is what was observed in ref. 16 for the L80 model examined in dynamical regimes only explored by a few by then (32), for which there is ultimately a finite R_o critical transition that explosively exhibits fast oscillations beyond the parameter range Lorenz explored (17); see Fig. 1. As discussed in refs. 16 and 17 and recalled below, such regimes constitute a harsh impediment to close the L80 model in its "slow" variables. Whereas these difficult closure problems were identified in ref. 16, they remained open. The present study provides a solution to such problems, while identifying the generic elements for closing other slow–fast systems in such regimes.

To better appreciate this solution, recall that the L80 model, obtained as a nine-dimensional (9D) truncation of the shallow water equations onto three Fourier modes with low wavenumbers (13), is written as

$$a_i \frac{\mathrm{d}x_i}{\mathrm{d}\tau} = a_i b_i x_j x_k - c(a_i - a_k)x_j y_k + c(a_i - a_j)y_j x_k$$
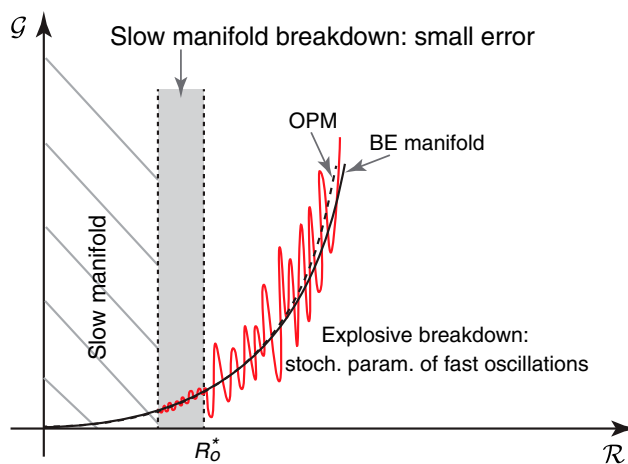$$- 2c^2 y_j y_k - \nu_0 a_i^2 x_i + a_i(y_i - z_i),$$

Fig. 1. Schematic of transitions in the L80 model. For $R > R_o^*$, an explosive breakdown of slow-to-fast deterministic parameterizations is observed (16). After this transition, a stochastic parameterization of the fast oscillations is required. It is successfully achieved here thanks to the filtering operated through the BE manifold that separates (nearly optimally) the full dynamics into its slow motion and a fast residual dynamics.

$$a_i \frac{\mathrm{d}y_i}{\mathrm{d}\tau} = -a_k b_k x_j y_k - a_j b_j y_j x_k + c(a_k - a_j)y_j y_k$$
$$- a_i x_i - \nu_0 a_i^2 y_i, \qquad [7]$$
$$\frac{\mathrm{d}z_i}{\mathrm{d}\tau} = -b_k x_j(z_k - h_k) - b_j(z_j - h_j)x_k + cy_j(z_k - h_k)$$
$$- c(z_j - h_j)y_k + g_0 a_i x_i - \kappa_0 a_i z_i + F_i.$$

The above equations are written for each cyclic permutation of the set of indices $\{(1, 2, 3)\}$, namely, for

$$(i, j, k) \in \{(1, 2, 3), (2, 3, 1), (3, 1, 2)\}. \qquad [8]$$

The parameters are those used in Lorenz's original paper (13) except for F_1 (see refs. 14 and 16 and *Materials and Methods*). In Eq. **7**, and for small Rossby numbers, the slow variables are made of the components of $\mathbf{y}(t)$, whereas the fast ones are made of those of $\mathbf{x}(t)$ and $\mathbf{z}(t)$.

As shown in ref. 16, whereas fast oscillations are absent from the solutions to Eq. **7** for small and moderate R_o, they develop brutally, in the course of time, once a critical Rossby number R_o^* is crossed, in contradistinction with fast oscillations emerging according to an exponential smallness scenario such as mentioned above. This transition corresponds to the emergence of fast gravity waves that can contain a significant fraction of the energy (up to $\sim$40%) as time evolves and that may either populate transient behaviors of various lengths or persist in an intermittent way. Such a substantial transfer of energy between the "slow" and "fast" variables is acting against the existence of deterministic parameterizations, unlike what is observed for slow–fast systems often placed in regimes for which the fast variables represent only a small fraction of the energy (19, 33). Past this transition, the consequences of the closure problem for the slow rotational variables are thus drastic. It was shown indeed, in ref. 16, that parameterizations conditioned on the current state of the slow variables are no longer sufficient to close the system: The fast oscillations need to be parameterized differently for $R > R_o^*$.

Nevertheless, among the possible deterministic, nonlinear parameterizations, the balance equations (BE) manifold, based on a minimalistic simplification of the horizontal momentum curl and divergence equations, plus hydrostatic balance (14, 34), is conspicuously more accurate than the others in providing an
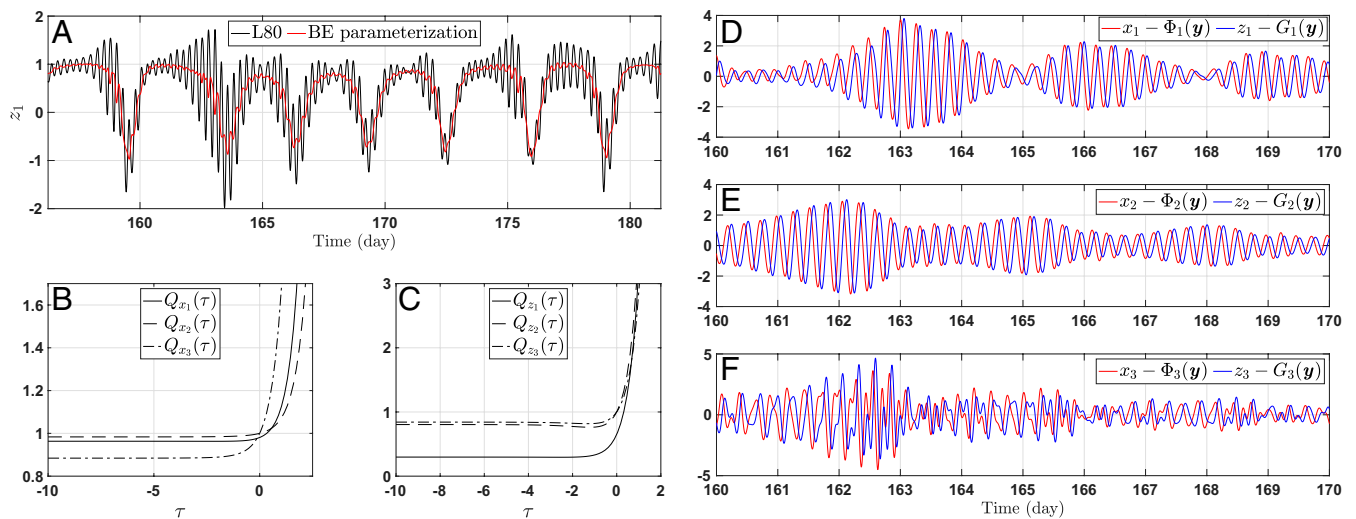
Chekroun et al.
Stochastic rectification of fast oscillations on slow manifold closures

PNAS | 3 of 9
https://doi.org/10.1073/pnas.2113650118

Fig. 2. (*A*) Illustration of the BE parameterization (red curve) averaging out the fast oscillations for regimes of spontaneous generation of "explosive" fast oscillations on the **x** and **z** variables, here shown for the z_1 variable (black curve). (*B* and *C*) Normalized parameterization defects $Q_{x_j}(\tau) = \overline{|x_j(t) - \Psi_j(\tau, (\mathbf{y}(t))|^2}/\overline{|x_j|^2}$ for continuous deformations of the BE parameterization. The parameter τ is here a dummy parameter that controls the deformations according to $\Psi_j(\tau, y) = m + (1 - \exp(\tau))\Phi_j(\mathbf{y})$ (as inspired by ref. 19, section 4.4), where Φ_j denotes the jth component of the BE parameterization for the **x** variable, while $m = \overline{p} - (1 - \exp(\tau))\overline{\Phi_j(\mathbf{y})}$, with p being the x_j variable which is parameterized. The same applies when x_j is replaced by z_j and Φ_j is replaced by G_j. We observe that the minima of the $Q_{x_j}(\tau)$ and $Q_{z_j}(\tau)$ are close to the values for $\tau \to -\infty$ corresponding to the BE manifold. The BE manifold is thus nearly optimal in this class of deformations; see ref. 16 for more elements about "BE ≈ OPM." (*D–F*) Times series composing the residual dynamics $\mathbf{r}(t)$ defined in Eq. **9**.

approximation of the slow OPM beyond the critical Rossby number, that is, for $R > R_o^*$, when the fast gravity waves contain a large fraction of the energy. This was shown in ref. 16 via detailed numerical computations (see also ref. 19, section 3.4) and rigorous error estimates (ref. 16, proposition 3.1), from which we report here a homotopic deformation analysis (Fig. 2 *B* and *C*) to further support that the BE manifold is indeed close to the OPM that averages out the fast oscillations for regimes of spontaneous generation of "explosive" fast oscillations on the **x** and **z** variables (Fig. 2*A*). Details about the BE closure based on the BE manifold are provided in *Materials and Methods*.

By forming the corresponding residual dynamics $\mathbf{r}(t)$, whose components are given by

$$\mathbf{r}_j = (x_j - \Phi_j(\mathbf{y}), z_j - G_j(\mathbf{y}))^T, \quad j = 1, 2, 3, \qquad [9]$$
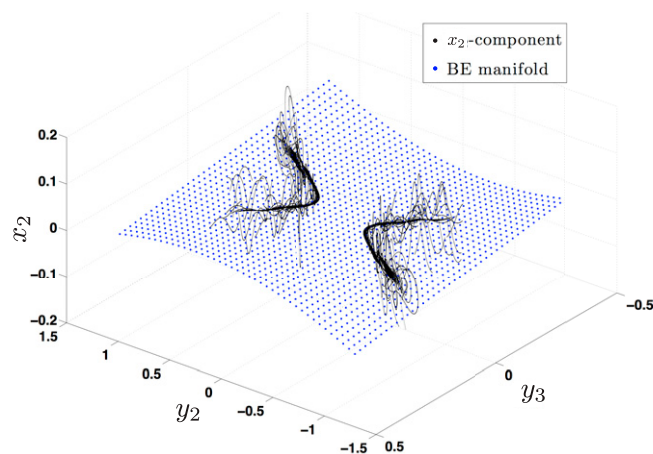

Fig. 3. Representation of the BE manifold and the transversal residual dynamics $\mathbf{r}(t)$. A substantial part of the L80 dynamics (black curve) lies very close to the BE manifold shown by blue dots, while the residual dynamics $\mathbf{r}(t)$ (defined in Eq. **9**) is mainly transversal to this manifold. It is this residual dynamics that is sought as the noise term $\eta(t)$ in Eq. **6**. We refer to ref. 19, section 3.4 for further details about the representation shown here.

we observe that the deviation from the BE manifold (also measured by the parameterization defect) is mainly transversal, even orthogonal to this BE manifold; see Fig. 3 and Table 1. The corresponding time series constituting the $\mathbf{r}_j(t)$ are shown in Fig. 2 *D–F*, after normalization by their respective SDs. It is this residual dynamics that is sought as the noise term $\eta(t)$ in Eq. **6**.

We emphasize that it is the prominent orthogonality of the residual dynamics $\mathbf{r}(t)$ that makes its identification with $\eta(t)$ in Eq. **6** possible. Otherwise, if $\mathbf{r}(t)$ would display prominent correlations with $\mathbf{y}(t)$, the memory terms in the MZ decomposition Eq. **6** should be part of the constitutive ingredients of an emulator of $\mathbf{r}(t)$, or, alternatively, the latter should be sought as a noise depending on the state $\mathbf{y}(t)$. Here, we thus only need to look for a stochastic generator of $\eta(t)$, which is independent of the slow rotational variables **y**. This is accomplished via analysis of the Ruelle–Pollicot (RP) resonances to infer the corresponding stochastic differential equations (SDEs).

Inferring Noise: RP Resonances Analysis

The theory of RP resonances is encountered in many branches of physics (scattering resonances, statistical mechanics) and mathematics (zeta functions, dynamical systems) (e.g., refs. 35–40).

These resonances, obtained as eigenvalues of the Perron–Frobenius operator [or its adjoint, the Koopman operator (41)], characterize fundamental properties of dynamical systems, such as mixing properties and decay of correlations (37, 39, 41), coherent structures (41, 42), or sensitivity to perturbations (43, 44). They also inform on the generator of the dynamics, as

Table 1. Orthogonality between the BE manifold and the residual r(t)

	Mean	SD
$\theta_{\mathbf{x}}$	89.89°	19.05°
$\theta_{\mathbf{z}}$	97.57°	20.71°

Here $\theta_{\mathbf{x}}(t)$ (resp. $\theta_{\mathbf{z}}(t)$) denotes the angle between the vectors $\Phi(\mathbf{y}(t))$ and $\mathbf{x}(t) - \Phi(\mathbf{y}(t))$ (resp. $G(\mathbf{y}(t))$ and $\mathbf{z}(t) - G(\mathbf{y}(t))$), with Φ (resp. G) defined in Eq. **15** (resp. Eq. **14**) in *Materials and Methods*.

they naturally relate to the Liouville operator in the case of deterministic (autonomous) flows (43, 45).

The estimation of these resonances is accessible from time series, as eigenspectrum of the underlying transition probability (or Markov) matrix, and several schemes exploiting the Ulam's method (46) have been designed to do so (e.g., refs. 45, 47, and 48).

When the observed time series are sought to be obtained from an SDE,

$$\mathrm{d}X = F(X)\,\mathrm{d}t + G(X)\,\mathrm{d}W_t, \quad X \in \mathbb{R}^d, \qquad [10]$$

with W_t, a Wiener process, the RP resonances inform on the drift F and the diffusion matrix $G(X)$, for $d = 1, 2$ (49).

When the observations of X lie in a reduced state space V, while X evolves in a higher-dimensional state space, the estimated resonances from time series become reduced RP resonances (50). The latter may still approximate the RP resonances, but they more generally inform on a coarse-grained version of the SDE generator in V and actually characterize the generator of the optimal reduced system in V obtained by averaging out the contribution of the unobserved variables (ref. 50, theorem 3). These reduced resonances can be also estimated using spectral data from time series following the Ulam's method; see *Materials and Methods*.

Thus, reduced RP resonances provides an efficient tool to address a "dynamics-informed" inference of stochastic processes aimed at modeling a set of observed time series. These features of (reduced) RP resonances have been recently illustrated in ref. 51 for characterizing, from time series, nonlinear stochastic oscillations, such as those produced from stochastic Stuart–Landau oscillators (SLOs) (see also ref. 52). In its simplest form, a prototypal SLO is given by

$$\dot{z} = (\alpha + i\omega)z - \beta z|z|^2 + \sigma \dot{W}, \qquad [11]$$

where $\dot{W}$ denotes a complex white noise, while σ, α, and ω denote real scalars, and β denotes a complex coefficient, with a positive real part. Typically, $\Re z(t)$ and $\Im z(t)$ are time series which are in phase quadrature, modulated in amplitude, and narrowband in the Fourier domain. However, other stochastic processes may share similar temporal (and

Fourier) features such as, for instance, rotational but damped (linear) Ornstein–Uhlenbeck (OU) processes, and thus a better discriminatory criterion is required to identify an SLO from time series.

This is what RP resonances allow for: An SLO has a clear signature in terms of the geometric organization of the (reduced) RP resonances in the complex left half-plane. It has been shown, in ref. 51, proposition 5, that RP resonances for SLOs are organized in the left half-plane, into an array of shifted parabolas, at least for most of these eigenvalues, whose imaginary parts are given as multiples of the fundamental frequency contained in the signal (see also refs. 52 and 7, appendix C). The first parabola passes through the eigenvalue zero, which is the only eigenvalue with zero real part. The RP resonances of any OU process are instead organized into a triangular array of resonances (53).

It turns out that such SLOs are actually key to provide a stochastic model of the time series $\mathbf{r}(t)$, obtained as the residual between the BE manifold and a high-fidelity solution to the L80 model. Indeed, for each 2D residual variable $\mathbf{r}_j(t)$ from Eq. 9, the estimated reduced RP resonances in their corresponding 2D reduced state space V show unambiguous signatures for modeling $\mathbf{r}_1$ and $\mathbf{r}_2$, as these resonances are organized according to parabolic structures, symptomatic of SLOs associated here with the same frequency ω_1; see Fig. 4. The latter is obtained as the fundamental frequency corresponding to the smallest nonzero imaginary part of the (discrete) parabola of resonances passing the eigenvalue zero (see ref. 51, equation 4.7). In the case of $\mathbf{r}_3$, we observed, however, some slight deviations from a parabolic arrangement of resonances which, as noted in other studies (ref. 7, appendix D), serve as an indication of the need for coupling terms to the SLO formulation.

The BE–SLO Closure of the L80 Model

Many types of networks of SLOs have been proposed in the literature to address coupling (e.g., refs. 54 and 55). We adopt here the coupling framework of ref. 56 that assumes linear coupling between the SLOs, and coupling through the driving white noise (via a covariance matrix), each aimed at respecting phase coherence among the different time series constituting $\mathbf{r}(t)$. Furthermore, as the pattern of the time series of $\mathbf{r}_3(t)$
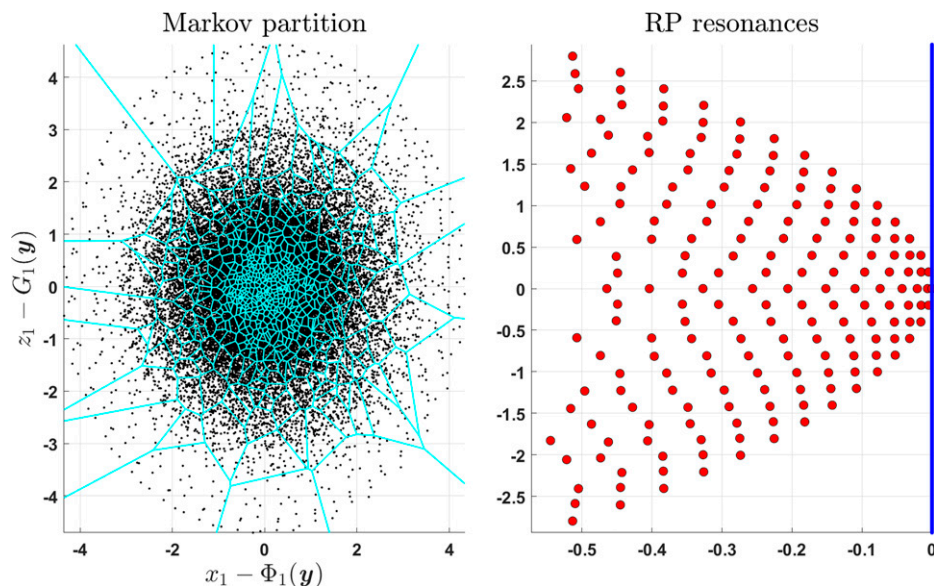
Fig. 4. (*Left*) The 2D time series $r_1(t) = (x_1(t) - \Phi_1(\mathbf{y}(t)), z_1(t) - G_1(\mathbf{y}(t)))^{\mathrm{T}}$ (black points), along with its Markov partition according to Voronoi cells (cyan lines). (*Right*) Reduced RP resonances estimated from this Voronoi tessellation; see *Materials and Methods*.

Chekroun et al.
Stochastic rectification of fast oscillations on slow manifold closures

PNAS | 5 of 9
https://doi.org/10.1073/pnas.2113650118

suggests a combination of tones (Fig. 2*F*), we propose to model the residual dynamics $\mathbf{r}(t)$ by the following system of SLOs:

$$\dot{\eta}_1 = (\alpha_1 + i\omega_1)\eta_1 - \beta_1\eta_1|\eta_1|^2 + \sum_{j\neq 1} c_{1j}\eta_j + \sum_{j=1}^{3} Q_{1j}\dot{W}_t^j$$

$$\dot{\eta}_2 = (\alpha_2 + i\omega_1)\eta_2 - \beta_2\eta_2|\eta_2|^2 + \sum_{j\neq 2} c_{2j}\eta_j + \sum_{j=1}^{3} Q_{2j}\dot{W}_t^j$$

$$\dot{\eta}_3 = (\alpha_3 + i\omega_2)\eta_3 - \beta_3\eta_3|\eta_3|^2 + \sum_{j\neq 3} c_{3j}\eta_j + \sum_{j=1}^{3} Q_{3j}\dot{W}_t^j.$$

[12]

Here, the first two equations are associated with one frequency ω_1, whereas ω_2 is passed into the third equation to account for the combination of tones. Each $\eta_j(t)$ is a complex-valued time series such that $\Re\eta_j(t)$ (resp. $\Im\eta_j(t)$) is aimed at emulating $x_j(t) - \Phi_j(\mathbf{y}(t))$ (resp. $z_j(t) - G_j(\mathbf{y}(t))$), after normalization by the SD. The value of ω_1 is estimated by looking at the dominant peak of the power spectrum of $\mathbf{r}_1(t)$, while ω_2 is taken to be $\omega_2 = 2\omega_1$ which corresponds to the second dominant peak in $\mathbf{r}_3(t)$. The rest of the coefficients in Eq. **12** are then simply estimated via regression of the right-hand side against $d\mathbf{r}/dt$, over the training period; see *Materials and Methods*. Once these coefficients are estimated, Eq. **12** is used to simulate (out of sample) the high-frequency residual dynamics by $\eta(t)$ solving Eq. **12**.

Taking into account the structure of the $\mathbf{y}$ equation in Eq. **7**, the resulting BE–SLO closure is then given by

$$\frac{dY_i}{dt} = -a_i^{-1}a_k b_k \big[\Phi_j(\mathbf{Y}) + \Re\eta_j(t)\big] Y_k$$
$$- a_i^{-1}a_j b_j Y_j \big[\Phi_k(\mathbf{Y}) + \Re\eta_k(t)\big]$$
$$+ ca_i^{-1}(a_k - a_j) Y_j Y_k - \big[\Phi_i(\mathbf{Y}) + \Re\eta_i(t)\big] - \nu_0 a_i Y_i,$$
with $\mathbf{Y} = (Y_1, Y_2, Y_3)^\mathsf{T}$,

[13]

for any (i, j, k) as in Eq. **8**, in which the η_ℓ are rescaled by the corresponding SDs. Note that, in Eq. **13**, the nonlinear stochastic parameterization, $\Phi_\ell(\mathbf{Y}) + \Re\eta_\ell(t)$, is used to emulate x_ℓ in the $\mathbf{y}$ equation. This BE–SLO parameterization provides a stable and accurate closure of the $\mathbf{y}$ variable. Not only the global geometry of the (projected) attractor in the $\mathbf{y}$ variable is very well reproduced (Fig. 5) but also the statistics accounting for the time variability such as the power spectral density (PSD), capturing, in particular, very well the dominant peaks associated with the Rossby and gravity waves (Fig. 6). In contradistinction, the BE attractor from the BE closure Eq. **16** alone (without SLO noise rectification) fails to even capture the coarse skeleton of the L80 attractor (blue curve in Fig. 5), unlike for other regimes (cf. ref. 16, figure 9). It shows how crucial are the correction features brought by the η-noise term as simulated by Eq. **12**, in order to rectify the BE closure. This rectification is even more
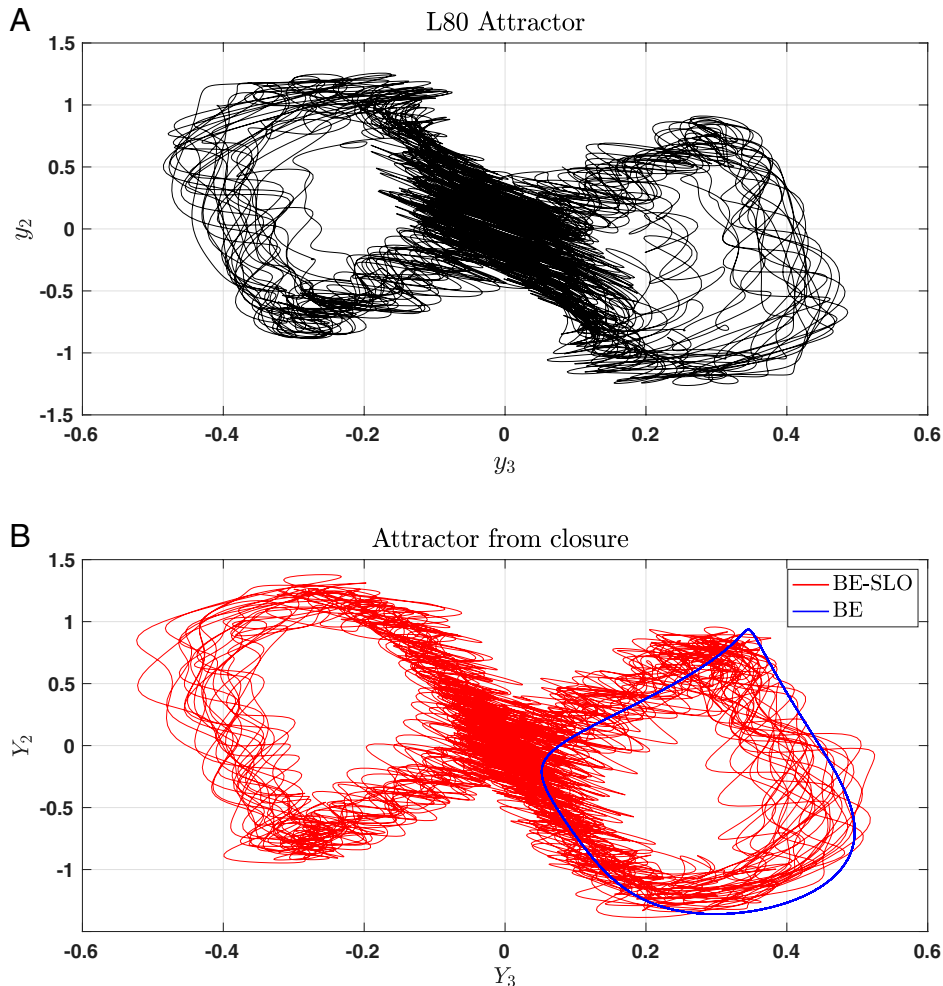

Fig. 5. Attractors comparison: *A* shows the attractor projected onto the slow variables (y_2, y_3) for the L80 model (black curve), while *B* shows the attractor from the BE–SLO closure Eq. **13** (red curve) and the BE closure Eq. **16** (blue curve), projected onto the corresponding slow variables. The BE–SLO closure reproduces the L80 chaotic dynamics, whereas the BE dynamics is only periodic.
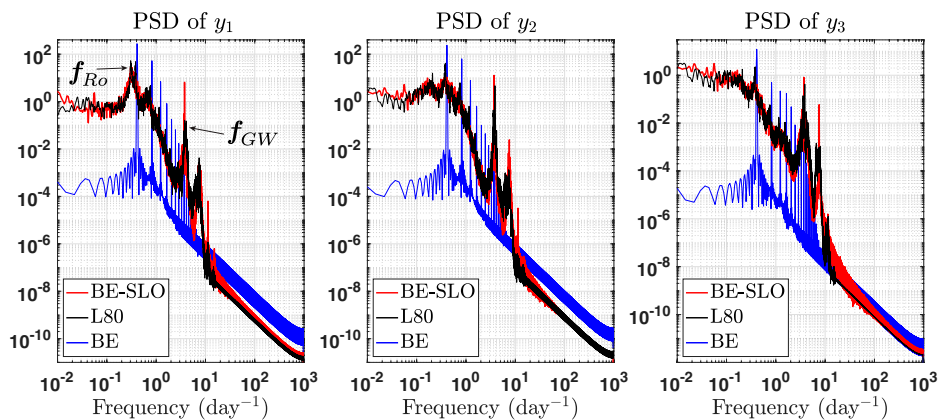
Fig. 6. PSD comparison for the **y** variable. The dominant frequency of the Rossby wave is $f_{Ro} = 0.31\,\mathrm{d}^{-1}$, and that of the gravity waves is $f_{GW} = 3.76\,\mathrm{d}^{-1}$, corresponding, respectively, to a period of $T_{Ro} = 3.2\,\mathrm{d}$ and $T_{GW} = 6.3\,\mathrm{h}$. As a comparison, the BE closure (blue curves) only captures f_{Ro} (and its subharmonics), while the BE–SLO closure recovers the full multiplicity of time scales.

striking when looking at the time variability details of the BE–SLO solutions which reproduce to a high fidelity the multiplicity of time scales of the **y** solution obtained from the full L80 model; see Fig. 7.

Discussion

Thus, the L80 model has been placed in a hard regime for closure for which we have presented an accurate solution based on a stochastic SLO rectification of the BE parameterization, and where the latter alone (or any other nonlinear parameterization) is insufficient for closure. The reasons underpinning the success of the resulting stochastic BE–SLO parameterization lies in two main elements: 1) the ability of the BE in capturing the slow motion of the coupled dynamics; 2) the prominence of the residual dynamics to be orthogonal to the BE manifold (Table 1), displacing the learning efforts for the MZ decomposition to the noise term only, and, of lesser importance, the centering condition $\zeta = 0$ in Eq. **4** naturally satisfied for the **y** equation of the L80 model.

It was shown in ref. 16 and recalled here that the BE manifold approximately equals the OPM, and thus allows for approximating the conditional expectation via Eq. **4**. Indeed, since $\zeta = 0$ for the L80 model, Eq. **4** is satisfied (up to a negligible error) when the OPM, Ψ^*, therein is replaced by BE. Nevertheless,
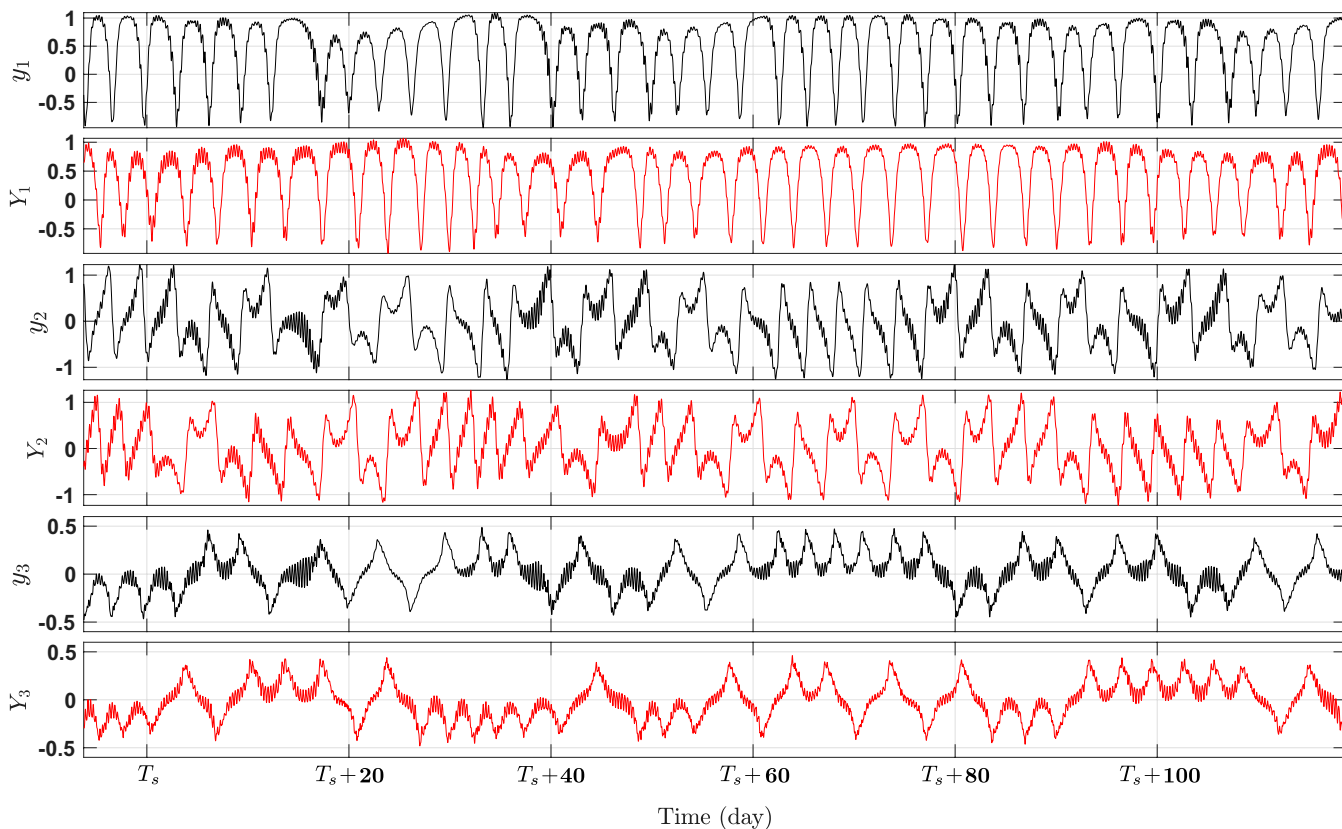

Fig. 7. Rotational **y**(t) variable from the L80 model (black curves) Eq. **7** and $Y(t)$ variable from the BE–SLO closure (red curves) Eq. **13**. Here $T_s = 420$ for the L80 model, and $T_s = 500$ for BE–SLO closure Eq. **13**, to show episodes with comparable patterns. The BE–SLO closure is able to reproduce accurately the multiscale chaotic dynamics of the L80 model; see also Fig. 6.

Chekroun et al.
Stochastic rectification of fast oscillations on slow manifold closures

PNAS | 7 of 9
https://doi.org/10.1073/pnas.2113650118

we emphasize that, when the conditional expectation is poorly approximated by, say, another nonlinear parameterization $\Psi \neq$ BE, the residual dynamics is no longer orthogonal to the manifold formed by this parameterization, and is thus contaminated by slow frequencies. In such a situation, the residual contains too many frequencies, which makes difficult their resolution by (blind) data-driven modeling techniques (57) or other methods based on Ruelle response theory (58), exploiting, essentially, either polynomial libraries of functions or specific interactions laws between the slow and fast variables to learn the ingredients of the MZ decomposition. Another impediment for these techniques to operate here lies in the strongly coupling nature of the slow–fast variables in the dynamical regime considered here. As pointed out in ref. 21, such strong couplings require, indeed, going beyond these approaches.

It is thus this remarkable ability of the BE manifold to capture the slow motion which allows for filtering out the residual dynamics from slow frequencies which is the basis of the success of its efficient parameterization by the proposed networks of stochastic oscillators. This fundamental understanding goes beyond the L80 model analyzed here as soon as, for other slow–fast systems, the BE manifold is replaced by a nonlinear parameterization that allows for approximating the dynamics' slow motion, while operating a separation of variables, the residual dynamics being essentially void of slow oscillations. Lately, much effort has been devoted to the learning of memory and noise terms in MZ decompositions (59–61). Our results teach us that, for slow–fast systems, the good capture of the slow motion is key to achieve before embarking on sophisticated learning of the memory and noise terms, and that, actually, the learning of the latter terms can be greatly simplified once a good approximation of the conditional expectation is known.

It is also worth noting that our closure approach relates to strategies other than MZ closures. Thinking of the quadratic terms in the L80 model as proceeding from nonlinear advective terms, we may interpret the corresponding nonlinear terms in the stochastic BE–SLO closure Eq. **13** as stochastic advective terms compared to the BE closure Eq. **16**. Other recent approaches have shown the relevance of such stochastic advective terms to derive stochastic formulations of fluid flows as well as for emulating, suitably, the coarse-grained dynamics (62–64).

Finally, as a low-dimensional model, we think that the L80 model for regimes such as those analyzed in this study can serve to test other closure methods and ideas for strongly coupled slow–fast systems and to gain usefulness as the more known Lorenz 96 model did over the last two decades (see, e.g., refs. 65–67).

The latter has been, indeed, pointed out recently in ref. 68 as too simple for closure, emphasizing the need for more challenging low-dimensional models to test and analyze emerging machine learning methods for closure. More generally, we believe that regimes exhibiting a mixture of fast oscillations superimposed on slower time scales, such as those displayed on the **y** solutions shown in Fig. 7 for the L80 model, will seemingly provide a challenging ground for closure in more-sophisticated fluid problems. Such regimes are known to arise in (multilayer) shallow water models (see, e.g., ref. 69, figure 5). It is theoretically expected that, for rotating stratified flows, the slow manifold, in particular as represented by the BE, will fail to provide (close to) exact parameterizations at some finite value of R_o (70, 71), even if the particular form of the breakdown manifested in the L80 system might not be generic.

Materials and Methods

The L80 Data. The parameters are chosen such that $a_1 = a_2 = 1$, $a_3 = 3$, $\nu_0 = \kappa_0 = 1/48$, $g_0 = 8$, $b_i = (a_i - a_j - a_k)/2$, and $c = \sqrt{b_1 b_2 + b_2 b_3 + b_3 b_1}$. Finally, $h_1 = -1$, and $h_2 = h_3 = F_2 = F_3 = 0$. These parameter values correspond to those used in ref. 13. Our analysis takes place for F_1 corresponding to a parameter regime for which $R_o > R_o^*$, on a rescaled version of Eq. **7**

(see ref. 16, equation 2.5). More precisely, the regime of ref. 16, figure 11 is considered here, which corresponds to $F_1 = 0.3027$ by using the rescaling given by ref. 16, equation 2.4 mapping the L80 model onto this rescaled version.

This system is numerically integrated using a standard fourth-order Runge–Kutta method with time stepping given by $\delta t = 4.2 \times 10^{-3}$ (corresponding to 45 s). Throughout the numerical experiments (including for the BE–SLO closure), we have taken the initial data to be very close to the Hadley fixed point (see ref. 16, section 2.5). The training period consists of 5×10^5 data points corresponding to 260 d (after removal of 2×10^5 transient data points) to learn the coefficients of the SLO system Eq. **12** via regression after normalization by the SD of the r_j. The validation period consists of 5×10^6, out-of-sample, simulated data points from the BE–SLO closure model Eq. **13**, corresponding to 10×260 d. The coefficients c_{ij} in Eq. **12** are complex, and the real part of each β_j is enforced to be larger than a small positive threshold chosen here to be 10^{-3} to ensure stability of the BE–SLO closure model Eq. **13**, while the Q_{ij} forms a positive definite matrix and is obtained as the Cholesky decomposition of the covariance matrix of the noise. The BE–SLO closure Eq. **13** is integrated using an Euler scheme. The L80 data and BE parameterization (see next subsection) are provided in *SI Appendix*.

The BE Manifold and BE Closure. Mathematically, the BE manifold aims at reducing the 9D system of ordinary differential equations (ODEs) Eq. **7** to a 3D system of ODE, by means of nonlinear parameterization of the variables $\mathbf{x} = (x_1, x_2, x_3)^T$ and $\mathbf{z} = (z_1, z_2, z_3)^T$, in terms of the variable $\mathbf{y} = (y_1, y_2, y_3)^T$ (see ref. 14). By analyzing the order of magnitudes of the different terms in the x_i equations [and after rescaling (16)], we arrive at the following parameterization of the **z** variable in terms of the rotational **y** variable,

$$z_i = G_i(\mathbf{y}) = y_i - \frac{2c^2}{a_i} y_j y_k. \qquad [14]$$

Further algebraic manipulations show that, under an invertibility condition of a matrix $M(\mathbf{y}, G(\mathbf{y}))$ conditioned on the **y** variable, one obtains (implicitly) **x** as a function Φ of **y** given by

$$\Phi(\mathbf{y}) = [M(\mathbf{y}, G(\mathbf{y}))]^{-1} \begin{pmatrix} d_{1,2,3}(\mathbf{y}, G(\mathbf{y})) \\ d_{2,3,1}(\mathbf{y}, G(\mathbf{y})) \\ d_{3,1,2}(\mathbf{y}, G(\mathbf{y})) \end{pmatrix}, \qquad [15]$$

where the $d_{i,j,k}$ are given explicitly (see refs. 14 and 16). The function $\Phi(\mathbf{y}) = (\Phi_1(\mathbf{y}), \Phi_2(\mathbf{y}), \Phi_3(\mathbf{y}))^T$ corresponds to the BE manifold; it aims to provide a nonlinear parameterization between **x** and **y** when the latter exists.

The BE closure is then

$$\begin{aligned} \frac{dy_i}{d\tau} = &-a_i^{-1} a_k b_k \Phi_j(\mathbf{y}) y_k - a_i^{-1} a_j b_j y_j \Phi_k(\mathbf{y}) \\ &+ c a_i^{-1} (a_k - a_j) y_j y_k - \Phi_i(\mathbf{y}) - \nu_0 a_i y_i, \end{aligned} \qquad [16]$$

for which (i, j, k) once more is as in Eq. **8**.

Estimation of Reduced RP Resonances. We recall below the estimation procedure of reduced RP resonances (see ref. 50, section 3.3). Here, the reduced state space V consists of the plane in which evolves the 2D component $r_j(t)$ of residual $\mathbf{r}(t)$ given by Eq. **9**. This projected trajectory lies within a domain $\mathcal{D}$ that is discretized as the union of M disjoint boxes or cells B_j, forming thus a partition. Unlike previous studies such as refs. 43 and 50 using a uniform meshing, we adopt a Voronoi tessellation to generate this partition. The interest in doing so is that the Voronoi cells become naturally more numerous over, for example, small regions with a high-density population of points, resulting, in turn, in better estimation of transitions matrices.

We assume that our observations are made at discrete time instants $t = t_n$, given as a multiple of a transition time τ, that is, $t_n = n\tau$ with $1 \leq n \leq N$, with N assumed to be large enough. We denote, by Y_n, any of the $r_j(t_n)$.

In practice, the counting of the transitions between the cells is highly dependent on the transition time τ, and the latter has to be chosen as guided by, for example, the physics of the problem to obtain interpretable results (50, 51). In the case of the L80 model, we used a τ-value corresponding to 20 min. The entries of the $M \times M$ transition matrix Γ_τ are then estimated according to

$$(\Gamma_\tau)_{ij} = \frac{\#\left\{(Y_n \in B_j) \wedge (Y_{n+1} \in B_i)\right\}}{\#\left\{Y_n \in B_j\right\}}, \qquad [17]$$

where the B_j form the Voronoi partition (composed of M disjoint Voronoi cells) of the domain $\mathcal{D}$ in V (see, e.g., refs. 43 and 47). In Eq. **17**, the notation $\#\{(Y_n \in B_k)\}$ gives the number of observations Y_n visiting the cell B_k, and the logical symbol "$\wedge$" means "and."

The reduced RP resonances are then obtained as the eigenvalues $\lambda_k(\tau)$ obtained from the eigenvalues $\zeta_k(\tau)$ of the Markov matrix Γ_τ, according to

$$\lambda_k(\tau) = \frac{\log\left(|\zeta_k(\tau)|\right)}{\tau} + i\frac{\arg\left(\zeta_k(\tau)\right)}{\tau}, \qquad 1 \leq k \leq M, \qquad [18]$$

where $\arg(z)$ (resp. $\log(z)$) denotes the principal value of the argument (that we adopt to lie in $[-\pi, \pi)$ in this article) (resp. logarithm) of the complex number z. Note that $|\zeta_k(\tau)| \leq 1$ as eigenvalues of the Markov matrix Γ_τ,

and thus the $\lambda_k(\tau)$ given by Eq. 18 lie naturally within the left half complex plane.

Data Availability. All study data are included in the article and/or *SI Appendix*.

1. G. Pavliotis, A. Stuart, *Multiscale Methods: Averaging and Homogenization* (Texts in Applied Mathematics, Springer Science & Business Media, 2008), vol. 53.
2. A. G. Barnston, M. K. Tippett, M. L. L'Heureux, S. Li, D. G. DeWitt, Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Am. Meteorol. Soc.* **93**, 631–651 (2012).
3. A. J. Majda, J. Harlim, Physics constrained nonlinear regression models for time series. *Nonlinearity* **26**, 201 (2012).
4. C. Franzke, T. O'Kane, J. Berner, P. Williams, V. Lucarini, Stochastic climate theory and modeling. *Wiley Interdiscip. Rev. Clim. Change* **6**, 63–78 (2015).
5. N. Boers *et al.*, Inverse stochastic-dynamic models for high-resolution Greenland ice-core records. *Earth Syst. Dyn.* **8**, 1171–1190 (2017).
6. D. Kondrashov, M. D. Chekroun, X. Yuan, M. Ghil, "Data-adaptive harmonic decomposition and stochastic modeling of Arctic sea ice" in *Advances in Nonlinear Geosciences*, A. Tsonis, Ed. (Springer, 2018), pp. 179–205.
7. D. Kondrashov, M. D. Chekroun, P. Berloff, Multiscale Stuart-Landau emulators: Application to wind-driven ocean gyres. *Fluids* **3**, 21 (2018).
8. R. Zwanzig, Memory effects in irreversible thermodynamics. *Phys. Rev.* **124**, 983–992 (1961).
9. H. Mori, Transport, collective motion, and Brownian motion. *Prog. Theor. Phys.* **33**, 423–455 (1965).
10. A. J. Chorin, O. H. Hald, R. Kupferman, Optimal prediction with memory. *Physica D* **166**, 239–257 (2002).
11. A. J. Chorin, O. H. Hald, *Stochastic Tools in Mathematics and Science* (Surveys and Tutorials in the Applied Mathematical Sciences, Springer, New York, NY, 2006), no. 147.
12. D. Givon, R. Kupferman, A. Stuart, Extracting macroscopic dynamics: Model problems and algorithms. *Nonlinearity* **17**, R55–R127 (2004).
13. E. N. Lorenz, Attractor sets and quasi-geostrophic equilibrium. *J. Atmos. Sci.* **37**, 1685–1699 (1980).
14. P. R. Gent, J. C. McWilliams, Intermediate model solutions to the Lorenz equations: Strange attractors and other phenomena. *J. Atmos. Sci.* **39**, 3–13 (1982).
15. J. Vanneste, Balance and spontaneous wave generation in geophysical flows. *Annu. Rev. Fluid Mech.* **45**, 147–172 (2013).
16. M. D. Chekroun, H. Liu, J. C. McWilliams, The emergence of fast oscillations in a reduced primitive equation model and its implications for closure theories. *Comput. Fluids* **151**, 3–22 (2017).
17. J. C. McWilliams, A perspective on the legacy of Edward Lorenz. *Earth Space Sci.* **6**, 336–350 (2019).
18. R. Temam, D. Wirosoetisno, Slow manifolds and invariant sets of the primitive equations. *J. Atmos. Sci.* **68**, 675–682 (2011).
19. M. D. Chekroun, H. Liu, J. C. McWilliams, Variational approach to closure of nonlinear dynamical systems: Autonomous case. *J. Stat. Phys.* **179**, 1073–1160 (2020).
20. R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics* (Applied Mathematical Sciences, Springer Science & Business Media, 2012), vol. 68.
21. M. Santos Gutiérrez, V. Lucarini, M. D. Chekroun, M. Ghil, Reduced-order models for coupled dynamical systems: Data-driven methods and the Koopman operator. *Chaos* **31**, 053116 (2021).
22. D. Kondrashov, M. D. Chekroun, M. Ghil, Data-driven non-Markovian closure models. *Physica D* **297**, 33–55 (2015).
23. C. E. Leith, Nonlinear normal mode initialization and quasi-geostrophic theory. *J. Atmos. Sci.* **37**, 958–968 (1980).
24. E. N. Lorenz, Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130–141 (1963).
25. F. Baer, J. J. Tribbia, On complete filtering of gravity modes through nonlinear initialization. *Mon. Weather Rev.* **105**, 1536–1539 (1977).
26. B. Machenhauer, On the dynamics of gravity oscillations in a shallow water model, with applications to normal mode initialization. *Beitr. Phys. Atmos.* **50**, 253–271 (1977).
27. E. N. Lorenz, On the existence of a slow manifold. *J. Atmos. Sci.* **43**, 1547–1558 (1986).
28. E. N. Lorenz, V. Krishnamurthy, On the nonexistence of a slow manifold. *J. Atmos. Sci.* **44**, 2940–2950 (1987).
29. E. N. Lorenz, The slow manifold—What is it? *J. Atmos. Sci.* **49**, 2449–2451 (1992).
30. R. Camassa, On the geometry of an atmospheric slow manifold. *Physica D* **84**, 357–397 (1995).
31. J. Vanneste, I. Yavneh, Exponentially small inertia-gravity waves and the breakdown of quasigeostrophic balance. *J. Atmos. Sci.* **61**, 211–223 (2004).
32. R. Vautard, B. Legras, Invariant manifolds, quasi-geostrophy and initialization. *J. Atmos. Sci.* **43**, 565–584 (1986).
33. C. Kuehn, *Multiple Time Scale Dynamics* (Applied Mathematical Sciences, Springer, 2015), vol. **191**.
34. J. McWilliams, P. Gent, Intermediate models of planetary circulations in the atmosphere and ocean. *J. Atmos. Sci.* **37**, 1657–1678 (1980).
35. D. Ruelle, Locating resonances for axiom a dynamical systems. *J. Stat. Phys.* **44**, 281–292 (1986).
36. M. Pollicott, Meromorphic extensions of generalised zeta functions. *Invent. Math.* **85**, 147–164 (1986).
37. V. Baladi, *Positive Transfer Operators and Decay of Correlations* (Advanced Series in Nonlinear Dynamics, World Scientific, 2000), vol. 16.
38. P. Gaspard, *Chaos, Scattering and Statistical Mechanics* (Cambridge Nonlinear Science Series, Cambridge University Press, 2005), no. 9.
39. A. Lasota, M. C. Mackey, *Chaos, Fractals, and Noise: Stochastic Aspects of Dynamics* (Applied Mathematical Sciences, Springer Science & Business Media, 2013), vol. 97.
40. S. Dyatlov, M. Zworski, *Mathematical Theory of Scattering Resonances* (Graduate Studies in Mathematics, American Mathematical Society, 2019), **vol. 200**.
41. M. Budišić, R. Mohr, I. Mezić, Applied Koopmanism. *Chaos* **22**, 047510 (2012).
42. G. Froyland, K. Padberg, Almost-invariant sets and invariant manifolds—Connecting probabilistic and geometric descriptions of coherent structures in flows. *Physica D* **238**, 1507–1523 (2009).
43. M. D. Chekroun, J. D. Neelin, D. Kondrashov, J. C. McWilliams, M. Ghil, Rough parameter dependence in climate models and the role of Ruelle-Pollicott resonances. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1684–1690 (2014).
44. V. Lucarini, Response operators for Markov processes in a finite state space: Radius of convergence and link to the response theory for axiom A systems. *J. Stat. Phys.* **162**, 312–333 (2016).
45. G. Froyland, O. Junge, P. Koltai, Estimating long-term behavior of flows without trajectory integration: The infinitesimal generator approach. *SIAM J. Numer. Anal.* **51**, 223–247 (2013).
46. S. M. Ulam, *A Collection of Mathematical Problems* (Interscience Tracts in Pure and Applied Mathematics, Interscience, 1960), no. 8.
47. C. Schütte, A. Fischer, W. Huisinga, P. Deuflhard, A direct approach to conformational dynamics based on hybrid Monte Carlo. *J. Stat. Phys.* **151**, 146–168 (1999).
48. S. Klus *et al.*, Data-driven model reduction and transfer operator approximation. *J. Nonlinear Sci.* **28**, 985–1010 (2018).
49. D. Crommelin, E. Vanden-Eijnden, Reconstruction of diffusions using spectral data from timeseries. *Commun. Math. Sci.* **4**, 651–668 (2006).
50. M. D. Chekroun, A. Tantet, H. A. Dijkstra, J. D. Neelin, Ruelle–Pollicott resonances of stochastic systems in reduced state space. Part I: Theory. *J. Stat. Phys.* **179**, 1366–1402 (2020).
51. A. Tantet, M. D. Chekroun, H. A. Dijkstra, J. D. Neelin, Ruelle-Pollicott resonances of stochastic systems in reduced state space. Part II: Stochastic Hopf bifurcation. *J. Stat. Phys.* **179**, 1403–1448 (2020).
52. S. Bagheri, Effects of weak noise on oscillating flows: Linking quality factor, Floquet modes, and Koopman spectrum. *Phys. Fluids* **26**, 094104 (2014).
53. G. Metafune, D. Pallara, E. Priola, Spectrum of Ornstein-Uhlenbeck operators in Lp spaces with respect to invariant measures. *J. Funct. Anal.* **196**, 40–60 (2002).
54. A. Koseska, E. Volkov, J. Kurths, Oscillation quenching mechanisms: Amplitude vs. oscillation death. *Phys. Rep.* **531**, 173–199 (2013).
55. A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge Nonlinear Science Series, Cambridge University Press, 2003), no. 12.
56. M. D. Chekroun, D. Kondrashov, Data-adaptive harmonic spectra and multilayer Stuart-Landau models. *Chaos* **27**, 093110 (2017).
57. S. Kravtsov, D. Kondrashov, M. Ghil, Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability. *J. Clim.* **18**, 4404–4424 (2005).
58. J. Wouters, V. Lucarini, Multi-level dynamical systems: Connecting the Ruelle response theory and the Mori-Zwanzig approach. *J. Stat. Phys.* **151**, 850–860 (2013).
59. X. Fu, L. B. Chang, D. Xiu, Learning reduced systems via deep neural networks with memory. *J. Mach. Learn. for Model. Comput.* **1**, 97–118 (2020).
60. Q. Wang, N. Ripamonti, J. S. Hesthaven, Recurrent neural network closure of parametric POD-Galerkin reduced-order models based on the Mori-Zwanzig formalism. *J. Comput. Phys.* **410**, 109402 (2020).
61. A. Gupta, P. F. J. Lermusiaux, Neural closure models for dynamical systems. *Proc. Royal Soc. A* **477**, 20201004 (2021).
62. D. D. Holm, Variational principles for stochastic fluid dynamics. *Proc. Math. Phys. Eng. Sci.* **471**, 20140963 (2015).
63. C. Cotter, D. Crisan, D. D. Holm, W. Pan, I. Shevchenko, Numerically modeling stochastic Lie transport in fluid dynamics. *Multiscale Model. Simul.* **17**, 192–232 (2019).
64. V. Resseguier, E. Mémin, B. Chapron, Geophysical flows under location uncertainty, Part I. Random transport and general models. *Geophys. Astrophys. Fluid Dyn.* **111**, 149–176 (2017).
65. D. S. Wilks, Effects of stochastic parametrizations in the Lorenz'96 system. *Q. J. R. Meteorol. Soc.* **131**, 389–407 (2005).
66. T. P. Sapsis, A. J. Majda, Statistically accurate low-order models for uncertainty quantification in turbulent dynamical systems. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 13705–13710 (2013).
67. G. Vissio, V. Lucarini, A proof of concept for scale-adaptive parametrizations: The case of the Lorenz'96 model. *Q. J. R. Meteorol. Soc.* **144**, 63–75 (2018).
68. S. Rasp, Coupled online learning as a way to tackle instabilities and biases in neural network parameterizations: General algorithms and Lorenz 96 case study. *Geosci. Model Dev.* **13**, 2185–2196 (2020).
69. E. Simonnet, M. Ghil, K. Ide, R. Temam, S. Wang, Low-frequency variability in shallow-water models of the wind-driven ocean circulation. Part II: Time-dependent solutions. *J. Phys. Oceanogr.* **33**, 729–752 (2003).
70. J. C. McWilliams, I. Yavneh, M. J. P. Cullen, P. R. Gent, The breakdown of large-scale flows in rotating, stratified fluids. *Phys. Fluids* **10**, 3178–3184 (1998).
71. M. J. Molemaker, J. C. McWilliams, I. Yavneh, Baroclinic instability and loss of balance. *J. Phys. Oceanogr.* **35**, 1505–1517 (2005).

EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES

Chekroun et al.
Stochastic rectification of fast oscillations on slow manifold closures

PNAS | 9 of 9
https://doi.org/10.1073/pnas.2113650118