PHYSICAL REVIEW FLUIDS **4**, 013801 (2019)

Editors' Suggestion

Reduced-order modeling of fully turbulent buoyancy-driven flows using the Green's function method

M. A. Khodkar,[1,*] Pedram Hassanzadeh,[1,2,†] Saleh Nabi,[3] and Piyush Grover[3]

[1]*Department of Mechanical Engineering, Rice University, Houston, Texas 77005, USA*

[2]*Department of Earth, Environmental, and Planetary Sciences, Rice University, Houston, Texas 77005, USA*

[3]*Mitsubishi Electric Research Laboratories, Cambridge, Massachusetts 02139, USA*

(Received 8 May 2018; published 11 January 2019)

A one-dimensional (1D) reduced-order model (ROM) is developed for a 3D Rayleigh-Bénard convection system in the turbulent regime with Rayleigh number $\mathrm{Ra} = 10^6$. The state vector of the 1D ROM is horizontally averaged temperature. Using the Green's function (GRF) method, which involves applying many localized weak forcings to the system one at a time and calculating the responses using long-time averaged direct numerical simulations (DNSs), the system's linear response function (LRF) is computed. Another matrix, called the eddy flux matrix (EFM), which relates changes in the divergence of vertical eddy heat fluxes to changes in the state vector, is also calculated. Using various tests, it is shown that the LRF and EFM can accurately predict the time-mean responses of temperature and eddy heat flux to external forcings and that the LRF can predict well the forcing needed to change the mean flow in a specified way (inverse problem). The non-normality of the LRF is discussed and its eigenvectors or singular vectors are compared with the leading proper orthogonal decomposition modes of the DNS data. Furthermore, it is shown that if the LRF and EFM are simply scaled by the square root of the Rayleigh number, they perform equally well for flows at other Ra, at least in the investigated range of $5 \times 10^5 \leqslant \mathrm{Ra} \leqslant 1.25 \times 10^6$. The GRF method can be applied to develop 1D or 3D ROMs for any turbulent flow, and the calculated LRF and EFM can help with better analyzing and controlling the nonlinear system.

DOI: 10.1103/PhysRevFluids.4.013801

I. INTRODUCTION

Buoyancy-driven turbulence plays a key role in various geophysical and environmental flows such as atmospheric and oceanic circulations as well as engineering systems such as wind farms and heating, ventilation, and air conditioning technologies. As a result, understanding, predicting, controlling, and optimizing buoyancy-driven turbulence has been of significant interest to the fluid dynamics and climate science communities. Given that direct numerical simulation (DNS) or large-eddy simulation of the full-dimensional Navier-Stokes equations can become computationally prohibitive for fully turbulent flows, which is the relevant regime in most of the aforementioned problems, considerable attention has been drawn recently to developing reduced-order models (ROMs) for these systems [1–9].

Reduced-order models are low-dimensional models with low computational complexity that retain the necessary dynamics of the turbulent flow and can be as simple as a system of nonlinear

*mkhodkar@rice.edu

†pedram@rice.edu

ordinary differential equations (ODEs) or even simpler linear ODEs, e.g.,

$$\dot{\boldsymbol{x}}(t) = L\boldsymbol{x}(t) + \boldsymbol{f}(t), \tag{1}$$

where $\boldsymbol{x}$ is the state vector, L is the system's evolution operator or linear response function (LRF), and $\boldsymbol{f}(t)$ represents external forcings (actuations) and/or stochastic parametrization of some unresolved physical processes [10–13]. This ROM [Eq. (1)] can be used, for example, to determine the time-mean response of the system to a forcing as $\langle \boldsymbol{x} \rangle = -L^{-1}\langle \boldsymbol{f} \rangle$, where $\langle\,\rangle$ denotes the long-time average, or to find the forcing required to produce a particular response as $\boldsymbol{f} = -L\langle \boldsymbol{x} \rangle$ (inverse problem), which can be used for flow control. Furthermore, the spectral properties of L provide information on the dynamics of the system [the limitations and underlying assumptions of Eq. (1) are discussed in Sec. III].

In the fluid dynamics community, the most common model reduction approach is to identify energetically dominant modes, obtained as top eigenvectors from some variant of proper orthogonal decomposition (POD) on the time series, and project the governing equations onto the subspace spanned by these modes [10,14,15]. The POD-based methods have been used to study various problems such as wall-bounded shear flows [16–18], cavity-driven flows [19,20], and flows past a cylinder [21–23], to name a few. Several studies have employed POD to develop ROMs for buoyancy-driven flows such as the Rayleigh-Bénard (RB) convection system [4,5,24–27], convection in laterally heated cavities [1,2,28,29], gravity currents [6], and turbulence in wind farms [30–32]. However, because the POD leads to a purely energy-based selection of leading modes, the modes may lack any true dynamical relevance. Furthermore, the truncated (low-energy) modes may still play a crucial role in the dynamics, especially for non-normal systems, where the transient growth can be large [16,33]. For instance, for examples of buoyancy-driven turbulence, Bailon-Cuba and Schumacher [4] and Benosman *et al.* [29] showed that, owing to the nonlinear interactions between the retained and excluded POD modes, eddy momentum and heat fluxes are not accurately captured unless some semiempirical mode-dependent closure models for the viscosity and diffusivity coefficients are employed.

As an alternative to POD-based methods, calculating L in Eq. (1) via the modes of the Koopman operator [34,35] or their data-driven approximations obtained from dynamic mode decomposition (DMD) [20,23,36–39] has received significant attention and has been applied to a variety of fluid flows (see, e.g., [33,40] and references therein). These techniques have also been applied to a number of buoyancy-driven turbulent flows. For instance, Kramer *et al.* [9] utilized DMD with sparse sensing to study convection in a laterally heated cavity, Annoni *et al.* [7] and Annoni and Seiler [41] employed this technique to develop ROMs for two-turbine wind farms in the planetary boundary layer, and Giannakis *et al.* [42] conducted Koopman eigenfunction analysis of the three-dimensional (3D) flow in a closed cubic turbulent convection cell. While the Koopman or DMD-based methods have produced promising results in these studies, particularly not far from the onset of linear instability, application of these methods to fully turbulent flows, including buoyancy-driven flows, remains a challenge and subject of extensive research.

Another recently developed framework, known as the resolvent approach, aims to find the perturbations around the turbulent mean flow by knowing the mean profile *a priori* and treating the Reynolds stress term in the Navier-Stokes equations as exogenous forcings [43–46]. This unknown forcing is assumed to be connected to the velocity field response via a linear operator called the resolvent. This method, which does not invoke any assumptions with regard to the amplitude of the perturbations, accounts for the nonlinear interaction between different modes through these forcings.

In the climate community, the most common methods for calculating L in Eq. (1) are the fluctuation-dissipation theorem (FDT) [3,47,48] and linear inverse modeling (LIM) [49,50]; the latter is closely connected to DMD [38,51]. Both LIM and the FDT are data driven and obtained from the Fokker-Planck equation under certain conditions [3,49,51]. While both methods work well when applied to very simple models such as the Lorenz 96 equations, acquiring accurate L for more

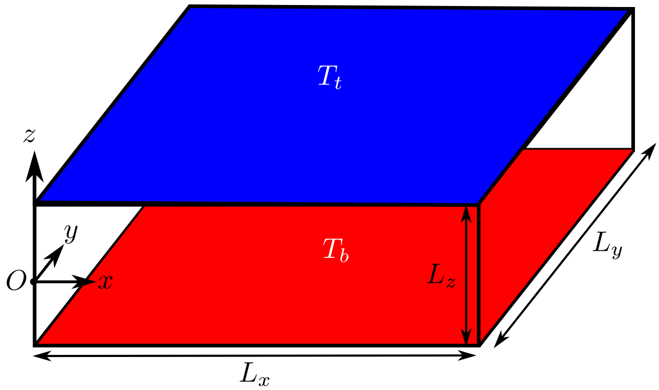

FIG. 1. Schematic of the 3D Rayleigh-Bénard convection system. The temperature at the top and bottom walls is, respectively, maintained at constant values of T_t and T_b, where $\Delta T = T_b - T_t > 0$. The velocity boundary condition at the walls is no slip, $\boldsymbol{u} = 0$. The horizontal directions (x and y) are periodic. In this study, $\mathrm{Pr} = 0.707$ and $5 \times 10^5 \leqslant \mathrm{Ra} \leqslant 1.25 \times 10^6$.

complex systems such as the quasigeostrophic equations or global circulation models (GCMs) has been found challenging [8,52–56].

In a different approach, Kuang [57] introduced the Green's function (GRF) method, which uses simulations with many weak localized forcings to construct L (details are presented in Sec. IV). He showed that the LRF of a cloud-resolving convection model can be accurately calculated using the GRF method. Hassanzadeh and Kuang [58] extended the GRF method to an idealized GCM and found that the calculated LRF was fairly accurate for the fully turbulent large-scale atmospheric circulation. They further showed that an eddy flux matrix (EFM) E that relates changes in the divergence of turbulent eddy momentum and heat fluxes $\boldsymbol{q}$ to a change in the state $\boldsymbol{x}$ via $\boldsymbol{q} = E\boldsymbol{x}$ can be accurately computed as a by-product of calculating L using the GRF method. In a second study, Hassanzadeh and Kuang [8] used this accurate L to identify the source of inaccuracy in the LRF obtained using the FDT as a combination of the GCM operator's non-normality and truncation of the time series to a limited number of POD modes. These accurate LRF and EFM have been also applied to study several aspects of atmospheric circulation in the tropics [59,60] and extratropics [61–63].

Given the success of the GRF method in calculating accurate L and E for fully turbulent atmospheric flows and improving the understanding of the data-driven methods (as mentioned above), it is worthwhile to introduce and examine the GRF method in the context of a canonical fluid dynamics problem that is of broader interest. This is the main purpose of the current study. We also extend the work of Kuang [57] and Hassanzadeh and Kuang [58] by showing that, at least for the problem studied here, the LRF and EFM, calculated at a given parameter, can be simply scaled and applied to a wider parameter regime.

We have applied the GRF method to a 3D RB convection system (Fig. 1) at the Rayleigh number $\mathrm{Ra} = 10^6$, where the flow is far from the onset of linear instability and fully turbulent. The RB convection system is a fitting prototype for buoyancy-driven flows and has been widely used to understand the turbulence physics and to develop techniques for analyzing turbulent systems [64–69]. Focusing on a 1D ROM for the 3D turbulent flow, we have calculated L and E for horizontally averaged temperature and divergence of vertical eddy heat flux at $\mathrm{Ra} = 10^6$. Using several tests, we demonstrate that the calculated L and E can predict the response of the system to external forcings accurately. Furthermore, L can calculate the forcing needed to achieve a specified mean flow. While L and E are obtained for $\mathrm{Ra} = 10^6$, we show that with a scaling factor that is simply proportional to $\sqrt{\mathrm{Ra}}$, these L and E work accurately at least for $5 \times 10^5 \leqslant \mathrm{Ra} \leqslant 1.25 \times 10^6$ as well.

The structure of this paper is as follows. The mathematical formulation of the RB system and the pseudospectral solver used to conduct DNS are described in Sec. II. The 1D ROM is derived in Sec. III. The GRF method is presented in Sec. IV in detail. The accuracy of L and E for Ra $= 10^6$ and for $5 \times 10^5 \leqslant \mathrm{Ra} \leqslant 1.25 \times 10^6$ is discussed in Secs. V and VI, respectively. The spectral properties of the 1D ROM are investigated in Sec. VII. Section VIII concludes the paper with a brief summary of the present investigation and an outlook for future work.

II. BOUSSINESQ EQUATIONS AND NUMERICAL SOLVER

We model the turbulent RB convection system using the 3D Boussinesq equations. We nondimensionalize length with the domain height L_z, temperature with $\Delta T = T_b - T_t$, and time with diffusive timescale $\tau_{\mathrm{diff}} = L_z^2/\kappa$, where κ is the thermal diffusivity, to arrive at the dimensionless equations

$$\nabla^* \cdot \boldsymbol{u}^* = 0, \tag{2}$$

$$\frac{\partial \boldsymbol{u}^*}{\partial t^*} + (\boldsymbol{u}^* \cdot \nabla^*)\boldsymbol{u}^* = -\nabla^* p^* + \mathrm{Pr}\nabla^{*2}\boldsymbol{u}^* + \mathrm{Ra}\,\mathrm{Pr}(T^* - T^*_{\mathrm{cond}})\hat{\mathbf{e}}_z, \tag{3}$$

$$\frac{\partial T^*}{\partial t^*} + (\boldsymbol{u}^* \cdot \nabla^*)T^* = \nabla^{*2}T^*. \tag{4}$$

Here $\boldsymbol{u}^* = (u^*, v^*, w^*)$ represents the 3D velocity field, T^* shows the temperature, and $T^*_{\mathrm{cond}} = 1/2 - z^*$ is the conduction temperature profile. The asterisk superscript denotes dimensionless variables and operators hereafter. It should be noted that while τ_{diff} is used here following convention, we employ the dynamically more relevant advective timescale $\tau_{\mathrm{adv}} = \sqrt{L_z/g\alpha\Delta T}$ to nondimensionalize time and vertical velocity when presenting the results ($\tau_{\mathrm{diff}}/\tau_{\mathrm{adv}} = \sqrt{\mathrm{Ra}\,\mathrm{Pr}} \approx 840$).

The Rayleigh and Prandtl numbers are defined as

$$\mathrm{Ra} = \frac{g\alpha\Delta T L_z^3}{\nu\kappa}, \tag{5}$$

$$\mathrm{Pr} = \frac{\nu}{\kappa}, \tag{6}$$

where g represents gravitational acceleration and α and ν indicate the thermal expansion coefficient and the kinematic viscosity of the fluid, respectively. The boundary conditions are periodic in the horizontal (x and y) directions and fixed temperature and no slip at the top and bottom walls, i.e.,

$$\boldsymbol{u}^*(x^*, y^*, z^* = \pm 1/2, t^*) = 0. \tag{7}$$

In this study we use a fixed Pr $= 0.707$ (air) and develop the LRF and EFM for Ra $= 10^6$, which is $\sim$585 times larger than the critical Rayleigh number for linear instability in this RB setup [70]. The flow is fully turbulent at this Ra (discussed below). A number of additional tests at a range $5 \times 10^5 \leqslant \mathrm{Ra} \leqslant 1.25 \times 10^6$ are also conducted and discussed in Sec. VI.

Direct numerical simulation of Eqs. (2)–(4) is carried out using a pseudospectral Fourier-Chebyshev solver that is based on the code described in [71]. Briefly, the solver uses the second-order Adams-Bashforth and Crank-Nicolson schemes for the time integration of the nonlinear and viscous terms, respectively. The no-slip and fixed-temperature boundary conditions are enforced following Marcus [72]. Variants of this solver have been used in the past to study geophysical and astrophysical turbulence [73–77]. The computational domain is $L_x^* \times L_y^* \times L_z^* = \pi \times \pi \times 1$ and the numerical resolution is $128 \times 128 \times 129$.

For the DNS at Ra $= 10^6$, Fig. 2 exhibits the time series for the anomalous temperature $(\overline{T} - \langle \overline{T} \rangle)$ at $z_0^* = -0.18$ and the principal component of the leading POD (PC1) obtained via the singular-value decomposition of the anomalous temperature. An overbar denotes the spatially averaged variables over the entire x-y plane. These time series illustrate the chaotic nature of the flow, as they show fast oscillations around the mean and peaks that are a few times larger or smaller than the standard deviation. Figure 3(a) demonstrates the power spectra of these two time series,

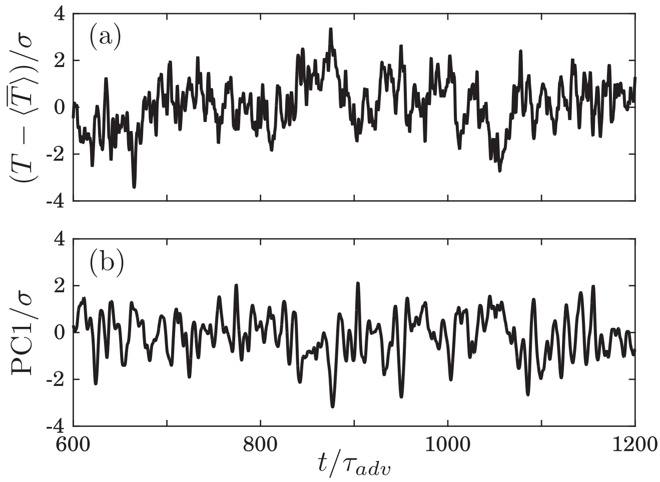

FIG. 2. Time series from DNS at Ra = 10^6 for (a) anomalous temperature $\overline{T} - \langle\overline{T}\rangle$ at $z_0^* = -0.18$ and (b) the principal component of the leading POD (PC1). Both time series are normalized with their standard deviation σ so that they have the same variability. Time is scaled by the advective timescale $\tau_{\rm adv}$.

showing that their spectra are monotonically decaying (red spectrum) and do not show any periodic or quasiperiodic behavior, which indicates that the flow is in the fully turbulent regime. To further demonstrate this point, the singular values s_i of different POD modes and the fraction of variance accumulated up to each POD mode are shown in Figs. 4(a) and 4(b), respectively. As can be seen, no sudden drop in the values of singular numbers occurs, indicating the high dimensionality of the system, even when the flow is horizontally averaged.

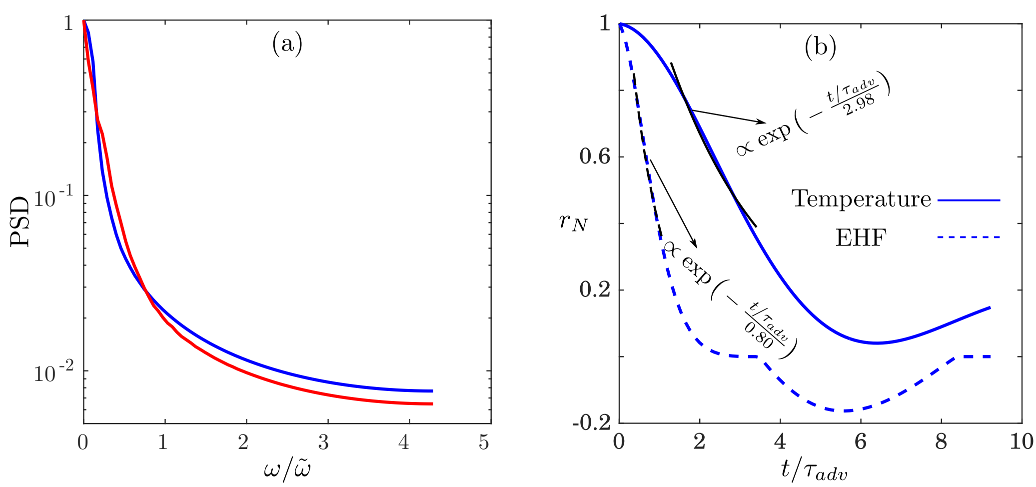

FIG. 3. (a) Power spectral densities (PSDs) of the time series shown in Fig. 2. The PSDs are calculated by dividing all the data consisting of 110 000 samples ($\sim$12 820$\tau_{\rm adv}$) into 500 windows with the same length ($\sim$26$\tau_{\rm adv}$) and carrying out a fast Fourier transform for each window. The results of all windows are then averaged to obtain the plotted PSDs. Frequency ω is normalized by the frequency of advective timescale $\tilde{\omega} = 2\pi/\tau_{\rm adv}$. (b) Autocorrelation r_N of anomalous temperatures (solid line) and EHF (dashed line), shown as a function of time scaled by $\tau_{\rm adv}$. The black lines show exponential fits.

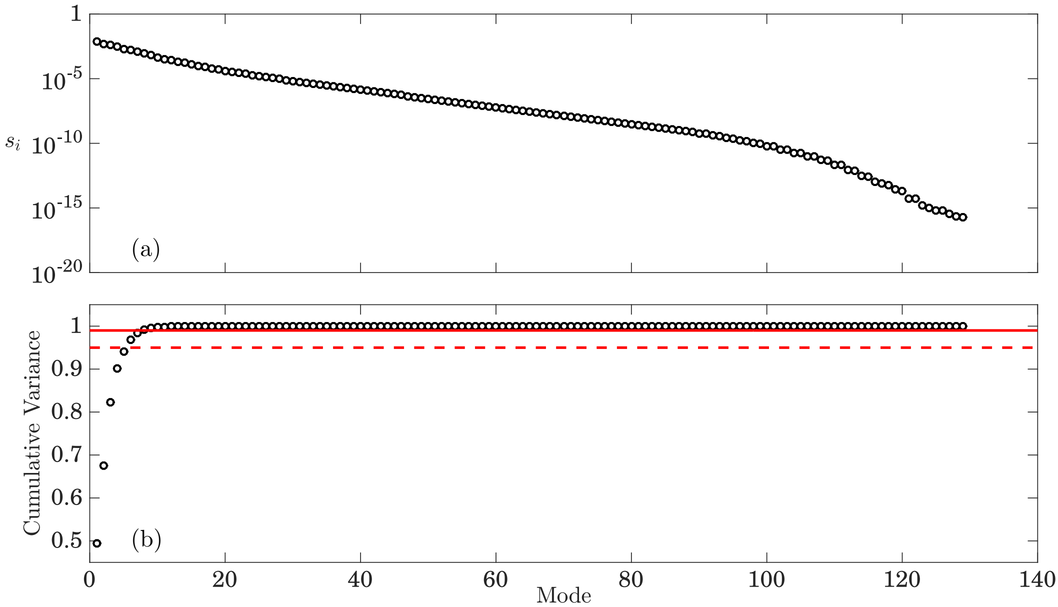

FIG. 4. (a) Singular values s_i of different POD modes obtained by conducting singular-value decomposition on the longest available data set for horizontally averaged anomalous temperature. (b) Cumulative variance explained by the POD modes. The horizontal dashed and solid red lines mark 0.95 and 0.99, respectively.

III. THE 1D ROM FOR 3D RAYLEIGH-BÉNARD TURBULENCE

In the following, we proceed to derive the mathematical formulation of a 1D ROM in the form of Eq. (1) for the 3D RB convection system by first averaging all flow properties and equations of motion in the horizontal (x and y) directions. The horizontally averaged nonlinear Boussinesq equations can be written as

$$\dot{\overline{\boldsymbol{X}}} = F(\overline{\boldsymbol{X}}), \tag{8}$$

where F is a nonlinear functional of the state vector $\overline{\boldsymbol{X}}$, which is a set of horizontally averaged variables describing the system. Suppose the state vector evolves from $\langle\overline{\boldsymbol{X}}\rangle$ at time t to $\langle\overline{\boldsymbol{X}}\rangle + \overline{\boldsymbol{x}}$ at time $t + \delta t$, in response to an external forcing such as $\overline{\boldsymbol{f}}$. Then Eq. (8) yields

$$\dot{\overline{\boldsymbol{x}}} = F(\langle\overline{\boldsymbol{X}}\rangle + \overline{\boldsymbol{x}}) + \overline{\boldsymbol{f}}. \tag{9}$$

If $\overline{\boldsymbol{x}}$ is small, a Taylor expansion of Eq. (9) gives

$$\dot{\overline{\boldsymbol{x}}} = \left.\frac{dF}{d\overline{\boldsymbol{X}}}\right|_{\langle\overline{\boldsymbol{X}}\rangle} \overline{\boldsymbol{x}} + \overline{\boldsymbol{f}} = L\overline{\boldsymbol{x}} + \overline{\boldsymbol{f}}, \tag{10}$$

where the higher-order terms (in $\overline{\boldsymbol{x}}$) are neglected [note that $F(\langle\overline{\boldsymbol{X}}\rangle) = 0$]. Equation (10) shows that the LRF L is the Jacobian of the nonlinear operator F evaluated at mean state $\langle\overline{\boldsymbol{X}}\rangle$. To derive this ROM, we do not ignore the eddy feedback, as we would if we had ignored the nonlinear terms in Eq. (8), but in the same fashion as Ring and Plumb [53] and Hassanzadeh and Kuang [58] we assume that a function relating the eddy fluxes and state vector has been linearized and included in L (see below for further discussion). Because we do not know this function, we cannot calculate Eq. (10) directly from Eqs. (2)–(4).

It is also instructive to formulate the 1D ROM more explicitly from Eq. (4), which combined with Eq. (2) can be rewritten, in dimensional form, as

$$\frac{\partial T}{\partial t} + \nabla_\perp \cdot (\boldsymbol{u}_\perp T) + \frac{\partial (wT)}{\partial z} = \kappa \nabla_\perp^2 T + \kappa \frac{\partial^2 T}{\partial z^2}, \tag{11}$$

where $\boldsymbol{u}_\perp = u\hat{\mathbf{e}}_x + v\hat{\mathbf{e}}_y$ and $\nabla_\perp$ and $\nabla_\perp^2$ act only on the x and y directions. Averaging over the x and y directions and given the periodic boundaries, we find

$$\frac{\partial \overline{T}}{\partial t} + \frac{\partial \overline{Tw}}{\partial z} = \kappa \frac{\partial^2 \overline{T}}{\partial z^2}. \tag{12}$$

We can decompose T and w into horizontally averaged and around-the-mean perturbation components as $T(x,y,z,t) = \overline{T}(z,t) + T'(x,y,z,t)$ and $w(x,y,z,t) = \overline{w}(z,t) + w'(x,y,z,t)$. Note that $\overline{w} = 0$ from continuity. Equation (12) can thus be rewritten as

$$\frac{\partial \overline{T}}{\partial t} + \frac{\partial [\overline{(\overline{T} + T')w'}]}{\partial z} = \kappa \frac{\partial^2 \overline{T}}{\partial z^2}, \tag{13}$$

which further simplifies to

$$\frac{\partial \overline{T}}{\partial t} + \frac{\partial (\overline{T'w'})}{\partial z} = \kappa \frac{\partial^2 \overline{T}}{\partial z^2}. \tag{14}$$

The long-time averaging of Eq. (14) leads to

$$\left\langle \frac{\partial (\overline{T'w'})}{\partial z} \right\rangle = \kappa \left\langle \frac{\partial^2 \overline{T}}{\partial z^2} \right\rangle. \tag{15}$$

Suppose the system evolves from $\langle \overline{T} \rangle(z)$ to $\langle \overline{T} \rangle(z) + \overline{\theta}(z,t)$ in response to the external forcing $\overline{f}(z,t)$. Equations (14) and (15) then show that the state-vector response $\overline{\theta}$, which represents the horizontal average of temperature departure from that of the unforced time-mean flow, is governed by

$$\frac{\partial \overline{\theta}}{\partial t} + E\overline{\theta} = \kappa \frac{\partial^2 \overline{\theta}}{\partial z^2} + \overline{f}. \tag{16}$$

The $E\overline{\theta}$ term represents the change in the divergence of vertical heat flux [second term on the left-hand side of Eq. (14)] caused by a change in the state $\overline{\theta}$. We emphasize that we do not know the EFM E and we are not going to make any assumptions about its properties, but we highlight that the representation of the eddy heat flux change via $E\overline{\theta}$ involves two key assumptions.

(i) The change in the divergence of vertical eddy heat flux, which we denote by $\partial(\overline{\theta' w'})/\partial z$ hereafter, can be fully described by $\overline{\theta}$. This is partly justified if eddies equilibrate rapidly with the new state $\langle \overline{T} \rangle + \overline{\theta}$, which can be evaluated by comparing the autocorrelation timescales of $\overline{\theta}$ and $\partial(\overline{\theta' w'})/\partial z$ [53,58]. Figure 3(b) shows the autocorrelation r_N of the times series obtained from projecting $\overline{\theta}$ and $\partial(\overline{\theta' w'})/\partial z$ onto the leading POD of $\overline{\theta}$ following Ma *et al.* [62]. The results show that the e-folding decorrelation timescale of eddies is $\sim$3.7 times smaller than that of $\overline{\theta}$, suggesting that the eddies decorrelate quickly and equilibrate with the new state, e.g., after $3\tau_{\rm adv}$, the ratio of r_N of eddies, and $\overline{\theta}$ is 0.064.

(ii) The $\partial(\overline{\theta' w'})/\partial z$ changes linearly with $\overline{\theta}$. This is a reasonable assumption if $\overline{\theta}$ has a small amplitude and is consistent with the assumption under which Eq. (10) was derived.

In summary, Eq. (16) shows that state vector $\overline{\boldsymbol{x}} = \overline{\boldsymbol{\theta}}$ describes the response of the system and that the 1D ROM is

$$\dot{\overline{\boldsymbol{\theta}}} = L\overline{\boldsymbol{\theta}} + \overline{\boldsymbol{f}}, \tag{17}$$

where $L = \kappa D^2 - E$. The operator D^2 is the second derivative with respect to z. We show in the next section that the matrix L (and matrix E) in Eq. (17) can be accurately calculated for a fully turbulent flow using the GRF method without any need for explicit knowledge or approximation of E.

IV. GREEN'S FUNCTION METHOD

In order to calculate L and E at $\mathrm{Ra} = 10^6$, we follow the procedure described in [58]. First, we define a set of Gaussian basis functions of the form

$$B_n(z) = \exp\left[-\frac{(z - z_n)^2}{z_w^2}\right], \tag{18}$$

where $z_w = L_z/20$, $z_n = \{-1, -0.95, -0.9, -0.85, -0.8, -0.7, \dots, 0, \dots, 0.7, 0.8, 0.85, 0.9, 0.95, 1\} \times L_z/2$, and $n = 1, 2, \dots, 25$. Simpler choices for z_n such as $z_n = \{-1, -0.9, -0.8, \dots, 0.8, 0.9, 1\} \times L_z/2$ were initially tried, but it was realized that in order to develop a reasonably accurate ROM, basis functions should be denser near the walls to better resolve the sharp gradients in the boundary layers. We calculate L and E in the space of these 25 basis functions rather than for the entire grid space (129 points) to reduce the computational cost.

Second, forcings of the form $f_n(z) = [a_n \Delta T/\tau_{\mathrm{diff}}] B_n(z)$ are added to the right-hand side of Eq. (4) one at a time, and a long DNS is then conducted at $\mathrm{Ra} = 10^6$. The term $\Delta T = T_b - T_t$ is the temperature difference between the bottom and top walls (Fig. 1). The a_n varies with z_n and is stronger near the walls. Its value is chosen, after some trial and error, such that it is not too large to violate the linearity assumption in Eqs. (10) and (17) or too small so that the signal (i.e., $\langle\overline{\theta}\rangle$) to noise (i.e., standard deviation of $\overline{\theta}$) ratio becomes large. To obtain large signal to noise ratio within the linear regime, we have conducted long DNS that are on average nearly 3200 times longer than τ_{adv} after the system reaches quasiequilibrium. The signal to noise ratio and the degree of nonlinearity are quantified using the criteria defined by Hassanzadeh and Kuang [58]. Based on these criteria, a_n is chosen to be 20 for all cases except the first three near-the-wall basis functions for which $a_n = 40$.

Hereinafter, we refer to each forced DNS as a trial. To increase the accuracy of the calculated ROM [57,58], for each f_n, one trial with positive and one trial with negative forcing is conducted and the time-mean response $\langle\overline{\theta}\rangle$ is calculated. Half of the difference between $\langle\overline{\theta}\rangle$ for the positive and negative forcings is used as the net response to f_n (denoted by $\langle\overline{\boldsymbol{\theta}}\rangle_n$). Given the symmetries of Eqs. (2)–(4), we have only conducted the trials for the lower half of the system $-1 \leqslant z_n \leqslant 0$ ($n = 1, 2, \dots, 13$) and just used $\langle\overline{\boldsymbol{\theta}}\rangle_n(z) = \langle\overline{\boldsymbol{\theta}}\rangle_{(26-n)}(-z)$ for $n = 14, 15, \dots, 25$. Therefore, a total of 26 DNSs are needed.

Each $\langle\overline{\theta}\rangle_n$ is projected via least-squares linear regression onto the basis function space. The resulting projection coefficients are $\langle\overline{\boldsymbol{\theta}}\rangle_n$ ($n = 1, 2, \dots, 25$), each of which is a column vector with the length 25. In addition, $\overline{\boldsymbol{f}}_n$ is also a column vector with the same length, whose elements are all zero, except for its nth element, which is equal to the amplitude of the forcing f_n. We can thus construct the following matrices for the time-mean responses and forcings in the reduced dimension of 25:

$$R = [\langle\overline{\boldsymbol{\theta}}\rangle_1 \;\; \langle\overline{\boldsymbol{\theta}}\rangle_2 \;\; \cdots \;\; \langle\overline{\boldsymbol{\theta}}\rangle_{25}], \tag{19}$$

$$F = [\overline{\boldsymbol{f}}_1 \;\; \overline{\boldsymbol{f}}_2 \;\; \cdots \;\; \overline{\boldsymbol{f}}_{25}]. \tag{20}$$

The LRF L of the system is then calculated from the long-time averaged Eq. (17) as

$$L = -FR^{-1}. \tag{21}$$

The EFM E is evaluated from the same simulations using a similar procedure. Here $\langle T'w'\rangle$ is calculated for each trial and the net response to each f_n (denoted by $\langle\theta'w'\rangle_n$) is obtained from the positive- and negative-forcing trials. The vertical derivative of the eddy flux responses is then

TABLE I. Details of the test cases to be discussed in the present and following sections. Here f is the external forcing for each test case and t_{DNS} indicates the length of the DNS data set. For most cases, DNSs with $\pm f$ (as well as control) are conducted to increase the accuracy and ensure the linearity of the response, but to reduce the computational cost, for a few cases, indicated by an asterisk superscript, only positive forcing and control DNSs are conducted. The error for temperature responses is defined as $\|\langle\overline{\boldsymbol{\theta}}\rangle_{\mathrm{GRF}} - \langle\overline{\boldsymbol{\theta}}\rangle_{\mathrm{DNS}}\|_2/\|\langle\overline{\boldsymbol{\theta}}\rangle_{\mathrm{DNS}}\|_2$. A similar formulation is used to find the error in the divergence of the vertical eddy heat flux (EHF) response.

Case	Figures	Ra	$f/(\Delta T/\tau_{\mathrm{diff}})$	θ error (%)	EHF error (%)	$t_{\mathrm{DNS}}/\tau_{\mathrm{adv}}$
C1	5(a) and 5(b)	10^6	$10\exp\big[-\frac{(z^*-0.2)^2}{0.1^2}\big]$	2.28	16.35	2307
C2	not shown	10^6	$20\cos(2\pi z^*)$	5.35	6.08	2496
C3	5(c) and 5(d)	10^6	$10\sin(2\pi z^*)$	9.32	4.84	2460
C4	5(e) and 5(f)	10^6	$10\cos(8\pi z^*)$	9.45	16.82	3133
C5	6(a) and 6(b)	10^6	as in Fig. 7(a)	6.07	8.28	2554
C6	6(c) and 6(d)	10^6	as in Fig. 7(b)	19.41	11.05	2772
C7*	not shown	5×10^5	$10\exp\big[-\frac{(z^*-0.2)^2}{0.1^2}\big]$	3.49	13.44	2135
C8*	8(a) and 8(b)	5×10^5	$10\cos(2\pi z^*)$	4.36	7.02	2150
C9*	8(c) and 8(d)	7.5×10^5	$20\cos(2\pi z^*)$	4.87	8.92	2543
C10	8(e) and 8(f)	1.25×10^6	$20\exp\big[-\frac{(z^*-0.2)^2}{0.15^2}\big]$	3.91	17.22	2719

calculated and projected onto the basis functions to obtain

$$Q = \left[\left\langle\frac{\partial(\overline{\boldsymbol{\theta}'\boldsymbol{w}'})}{\partial z}\right\rangle_1 \quad \left\langle\frac{\partial(\overline{\boldsymbol{\theta}'\boldsymbol{w}'})}{\partial z}\right\rangle_2 \quad \cdots \quad \left\langle\frac{\partial(\overline{\boldsymbol{\theta}'\boldsymbol{w}'})}{\partial z}\right\rangle_{25}\right]. \tag{22}$$

Then E is computed as

$$E = QR^{-1}. \tag{23}$$

The accuracy and predictive capabilities of L and E presented here are examined in detail for several test cases in Sec. V.

V. COMPARISON OF THE GRF-BASED ROM AND DNS AT Ra = 10^6

In the following section, we assess the accuracy of L and E obtained using the GRF method by examining their capabilities to predict the time-mean response of temperature and vertical eddy heat flux to external forcings. Furthermore, we study the performance of L in calculating the forcing required for the control of time-mean flow. To find the "true" responses or to evaluate the accuracy of the calculated forcing, long, forced DNSs are conducted. The details of all these test cases (denoted by C) are presented in Table I. Some of the forcings used in these test cases are localized, e.g., the Gaussian forcing of C1, but most of them are in the form of cosine or sinusoidal functions, which excite the flow along the z direction and can lead to complex responses. For example, cosine forcings are strong at the boundaries and lead to large eddy heat flux responses at the boundary layers, and forcings with high wave numbers create multiple contiguous stabilized and destabilized regions in the domain.

Figure 5 compares the ROM and DNS results for the time-mean response of horizontally averaged temperature $\langle\overline{\theta}\rangle$, scaled by ΔT, and eddy heat fluxes $\langle\partial(\overline{\theta' w'})/\partial z\rangle$ scaled by $\Delta T/\tau_{\mathrm{adv}}$, for three different cases (C1, C3, and C4; C2 is not shown for brevity). The red shadings in this figure and other figures demonstrate the uncertainty in the time-mean responses calculated from DNS. To find this uncertainty, each DNS time series is divided into eight segments with equal length and the standard deviation σ of the time mean of these segments are calculated. The solid blue lines show the mean of these eight segments, while the shading shows $\pm\sigma$. As shown in Fig. 5, despite the notable complexity of some of the responses such as sharp gradients in the boundary layers and

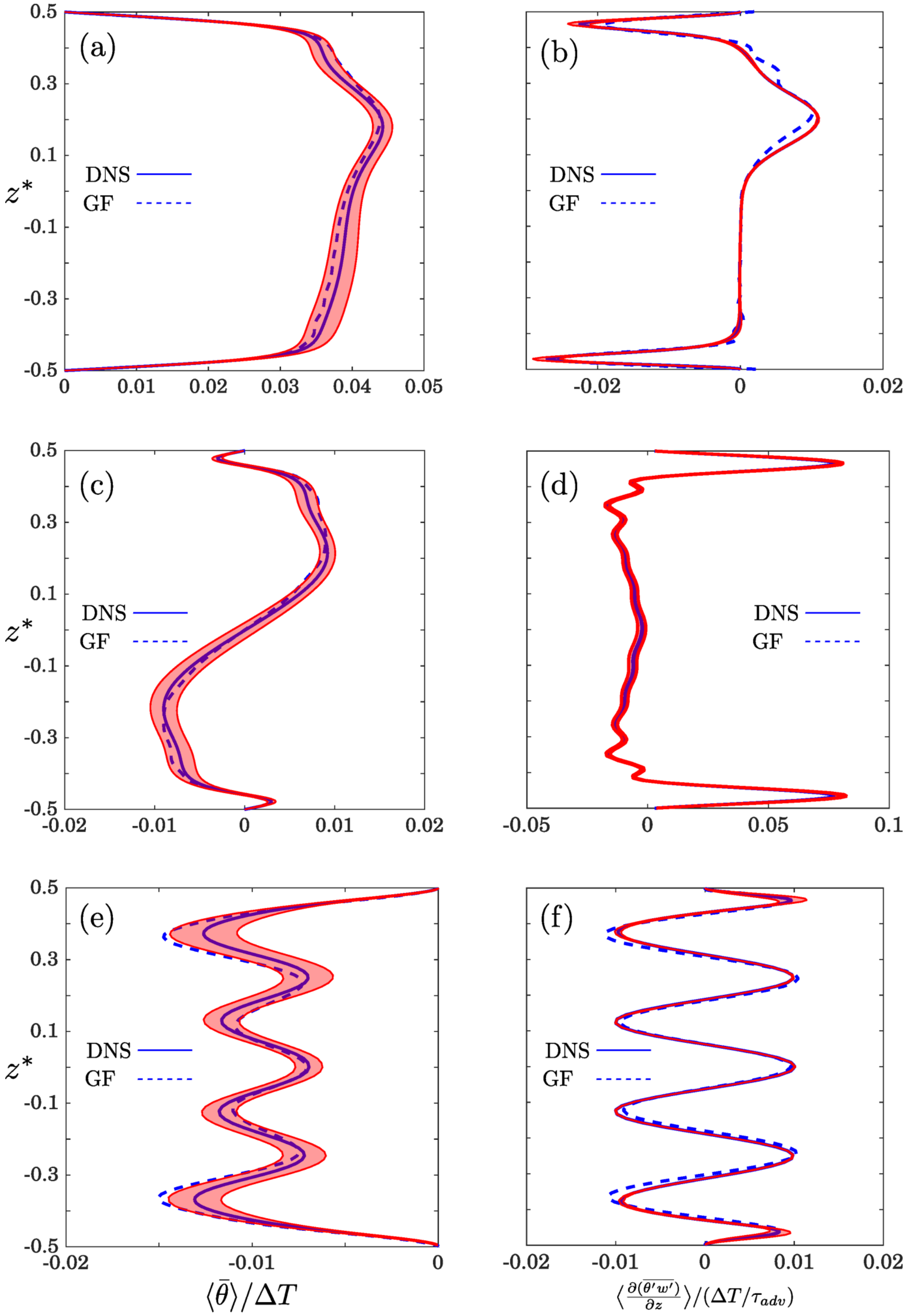

FIG. 5. Time-mean responses of temperature (left column) and eddy heat flux (right column) to forcings (a) and (b) f_1, (c) and (d) f_3, and (e) and (f) f_4 (see Table I). Solid lines show the true response obtained from a long forced DNS and dashed lines show the predictions from L and E obtained using the GRF method. The red shading shows the uncertainty in the time-mean responses calculated from the DNS data (see the text for more details). For all cases, $\mathrm{Ra} = 10^6$.

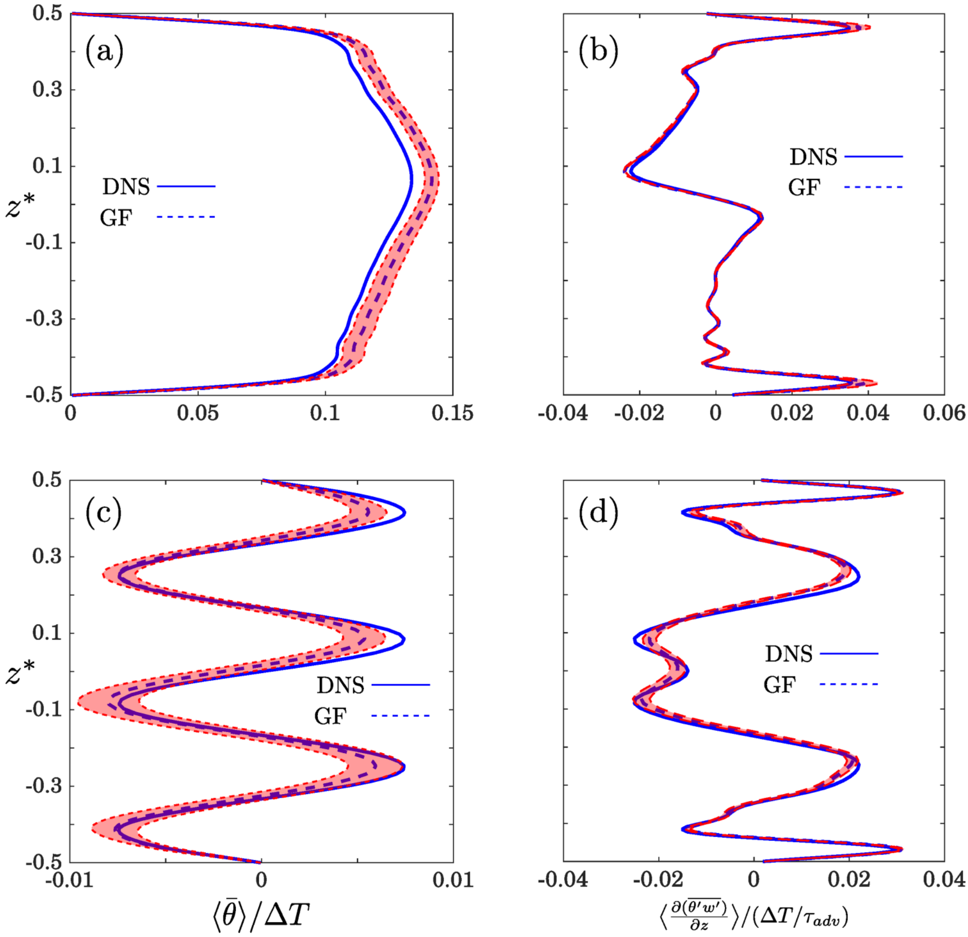

FIG. 6. Cases (a) and (b) C5 and (c) and (d) C6. (a) and (c) Solid lines show the target change in the mean flow θ_{target}, while the dashed lines demonstrate the time-mean responses to forcings f_5 [shown in Fig. 7(a)] and f_6 [shown in Fig. 7(b)] obtained from a long forced DNS. (c) and (d) Changes in vertical eddy heat flux obtained from a long forced DNS (solid lines) or calculated by E (dashed lines). As before, the red shading shows the uncertainty in the time-mean responses calculated from the DNS data. For all cases, $\text{Ra} = 10^6$. More details are in Table I.

multiple extrema, the pattern and amplitude of the temperature and eddy heat flux responses are well predicted by L and E in all these test cases.

Knowing L also enables us to find the required forcing to produce a desired change in the time-mean flow. This is of particular interest when flow control is intended. To test the skill of L for such inverse problems, we have chosen two target profiles θ_{target} shown with solid lines in Fig. 6(a) (C5) and Fig. 6(c) (C6). The forcings needed to change the mean flow by θ_{target} for these two cases are calculated as $f = -L\theta_{\text{target}}$ and shown in Fig. 7. The forcing profiles are not trivial, particularly near the walls, even for the simpler θ_{target} of C5. To evaluate the accuracy of these predicted forcings, forced DNSs with f_5 and f_6 are conducted and the mean-flow changes are shown in Figs. 6(a) and 6(c) (dashed lines), which match the target well, although the amplitude is larger for C5. The accuracy of E can further be examined using these test cases as shown in Figs. 6(b) and 6(d). As before, we find that E can well capture changes in the vertical eddy heat flux even for complex $\langle\partial(\overline{\theta' w'})/\partial z\rangle$ profiles.

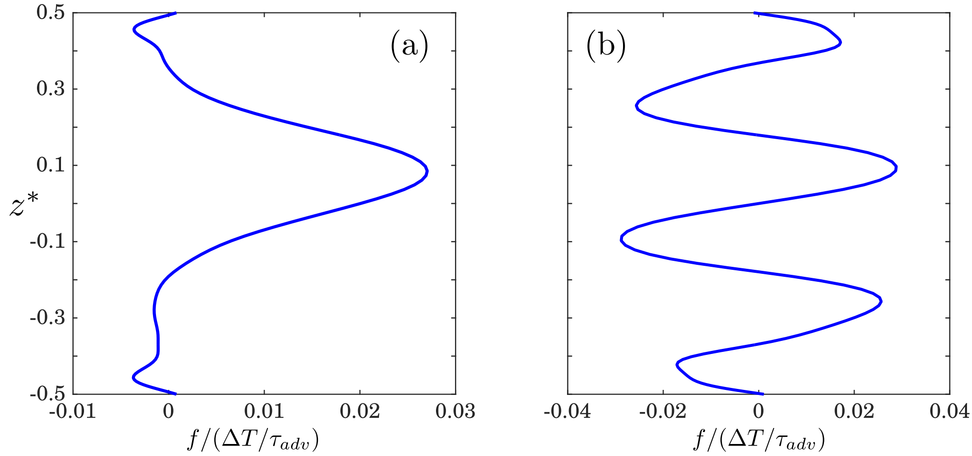

FIG. 7. Forcing that is predicted as $f = -L\theta_{\text{target}}$ to lead to the target time-mean responses θ_{target}: (a) f_5 [θ_{target} is shown in Fig. 6(a)] and (b) f_6 [θ_{target} is shown in Fig. 6(c)].

VI. EXTENDING THE 1D ROM TO OTHER VALUES OF Ra

In the preceding section we showed that L and E that are calculated using the GRF method at $\text{Ra} = 10^6$ work well in predicting the response or forcing at this value of Ra. As will be discussed in Sec. VIII, the main drawback of the GRF method is that it is computationally expensive, therefore it is worthwhile to explore how the L and E calculated for one value of Ra can be used for other Ra numbers.

We have conducted several more forced DNSs within the range of $5 \times 10^6 \leqslant \text{Ra} \leqslant 1.25 \times 10^6$. Details of some of these simulations are presented in Table I (C7–C10). The solid lines in Fig. 8 show the time-mean responses in temperature and vertical eddy heat flux while the dotted lines show the predictions when the LRF and EFM of $\text{Ra} = 10^6$ are used. For $\langle\overline{\theta}\rangle$, while the general shapes of the profiles are well captured by $L(10^6)$, the amplitudes are under- or overestimated, depending on Ra. These results suggest that the eigenvectors of L have remained fairly unchanged for this range of Ra and that only its eigenvalues have varied. For the eddy heat flux, we find that if the response is calculated as $E(10^6)L^{-1}(10^6)f$, then the prediction is surprisingly accurate [Figs. 8(b), 8(d), and 8(f)], indicating that EL^{-1} remains approximately constant for the aforementioned range of Ra. We highlight that not only is this hypothesis based on these four observations, but more simulations in the range of $5 \times 10^5 \leqslant \text{Ra} \leqslant 1.25 \times 10^6$ confirmed this hypothesis as well.

Based on the these observations, we postulate that $L(10^6)$ and $E(10^6)$ can be simply scaled to find the LRF and EFM at a new Ra,

$$L(\text{Ra}) = c_f L(10^6), \quad E(\text{Ra}) = c_e E(10^6). \tag{24}$$

Furthermore, the fact that EL^{-1} remains nearly constant suggests that the scaling factors are the same: $c_f = c_e$.

To validate Eq. (24), scaling factors are found as

$$c_f(\text{Ra}) = \frac{\|\langle\overline{\boldsymbol{\theta}}\rangle_{\text{GRF}}\|_\infty}{\|\langle\overline{\boldsymbol{\theta}}\rangle_{\text{DNS}}\|_\infty}, \tag{25}$$

$$c_e(\text{Ra}) = \frac{\|\langle\partial(\overline{\boldsymbol{\theta}'\boldsymbol{w}'})/\partial z\rangle_{\text{GRF}}\|_\infty}{\|\langle\partial(\overline{\boldsymbol{\theta}'\boldsymbol{w}'})/\partial z\rangle_{\text{DNS}}\|_\infty}, \tag{26}$$

where the subscript GRF in the numerators indicates that $L(10^6)$ and $E(10^6)$ are employed and the subscript DNS in the denominators shows results from long forced DNSs at Ra are used.

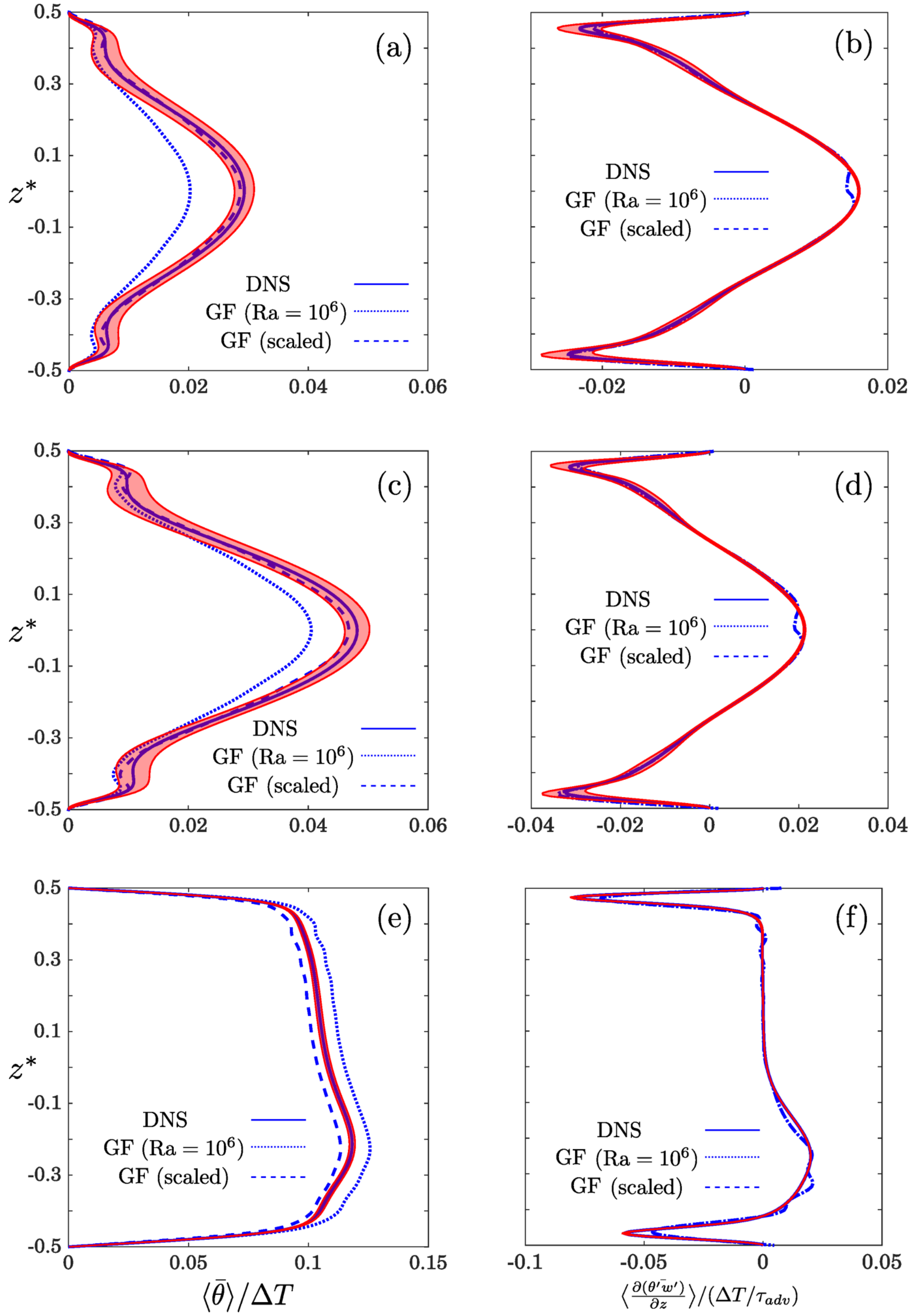

FIG. 8. Time-mean responses of temperature (left column) and eddy heat flux (right column) to forcings (a) and (b) f_8 (Ra = 5×10^5), (c) and (d) f_9 (Ra = 7.5×10^5), and (e) and (f) f_{10} (Ra = 1.25×10^6) (see Table I). Solid lines show the true response obtained from a long forced DNS, dotted lines show the predictions from $L(10^6)$ and $E(10^6)$, and dashed lines show the predictions obtained by the LRF and EFM scaled according to Eq. (27).

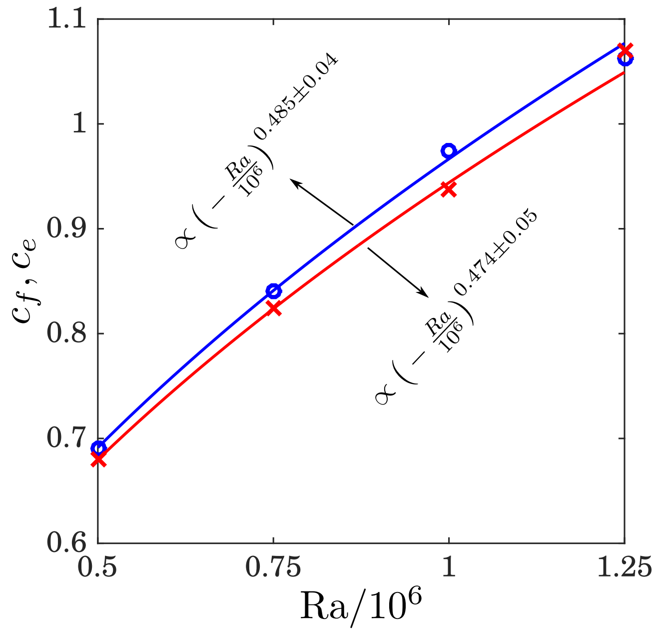

FIG. 9. Scaling factors c_f and c_e demonstrated by blue circles and red crosses, respectively. The solid lines are the power fits to the corresponding discrete data whose functions are shown in the figure.

Figure 9 shows the scaling factors calculated for C2 and C8–C10 along with the power fit to each of them. We find that $c_f \sim c_e$ and that both are approximately 0.5, suggesting the scaling with $\sqrt{\mathrm{Ra}}$. Therefore, we can reasonably approximate the scaled L and E as

$$L(\mathrm{Ra}_2) = \sqrt{\frac{\mathrm{Ra}_2}{\mathrm{Ra}_1}}L(\mathrm{Ra}_1), \quad E(\mathrm{Ra}_2) = \sqrt{\frac{\mathrm{Ra}_2}{\mathrm{Ra}_1}}E(\mathrm{Ra}_1). \tag{27}$$

Dashed lines in Fig. 8 demonstrate the performance of L and E calculated using Eqs. (27), for three different test cases with $\mathrm{Ra}_1 = 10^6$. As shown in this figure, predicted responses agree closely with those of the DNS results, which substantiates the validity of the scaling argument presented earlier for a fairly broad range of Rayleigh numbers. We also highlight that the accuracy of the scaled LRFs and EFMs for a given Ra is comparable to the accuracy of the $L(10^6)$ and $E(10^6)$. Whether this scaling holds for a larger range of Ra is computationally expensive to test and is left for future work.

VII. SPECTRAL PROPERTIES OF THE 1D ROM

As shown in previous sections, the L obtained using the GRF method can predict the time-mean response of the 3D RB system to external forcings, or the forcing needed for a given change in the time-mean flow, with high accuracy. In the present section we study some of the spectral properties of L. Figures 10 and 11 show the four slowest-decaying eigenvectors of L and its eigenvalues, respectively.

The slowest-decaying mode is real, mostly in the interior (outside the boundary layers), and decays with a timescale of $\sim 17\tau_{\mathrm{adv}}$. This eigenvector coincides with L's neutral vector, which is the right singular vector with smallest singular number and the system's most excitable dynamical mode because it is the largest time-mean response to external forcings [78,79]. The leading POD of a turbulent flow (POD1) is expected to be identical to its neutral vector if the forcing from turbulent eddies is spatially uncorrelated and has uniform variance everywhere [79]. Figure 10(a)

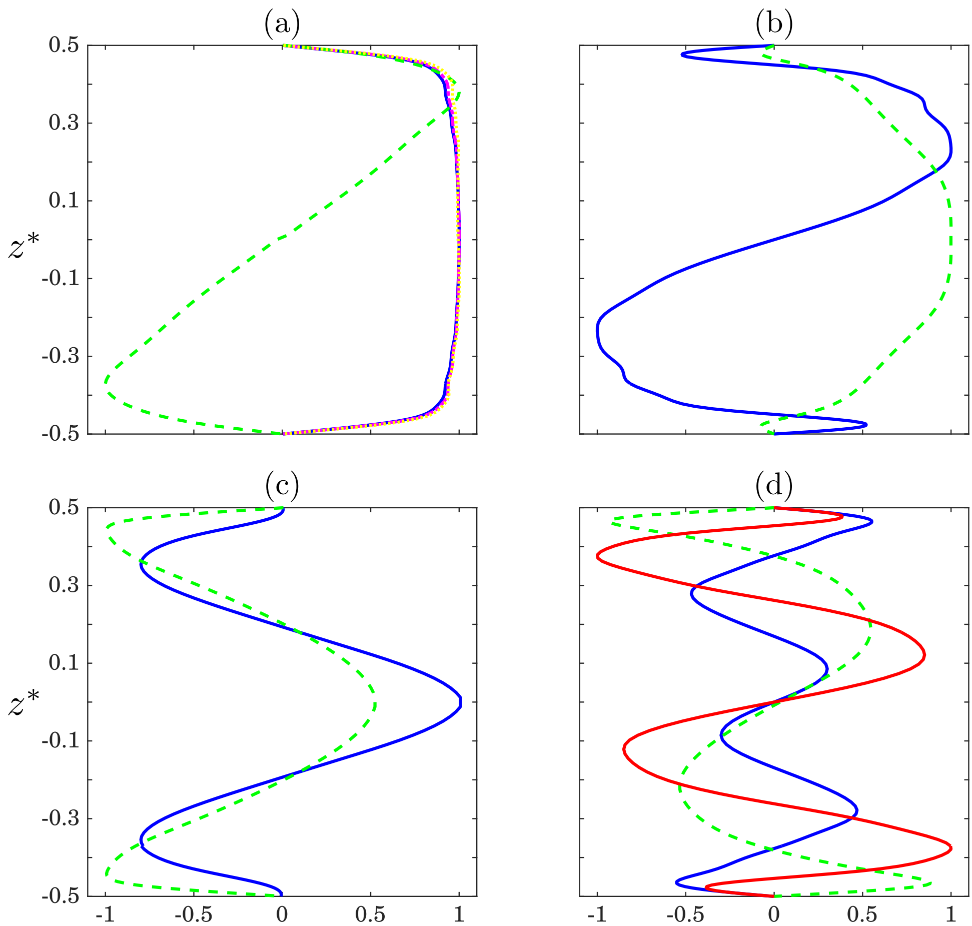

FIG. 10. The first four leading (i.e., slowest-decaying) modes and neutral vector of L and the leading POD modes. Solid blue (red) lines show the real (imaginary) part of L's eigenvectors. The green dashed lines show the leading POD modes (time mean removed) obtained from the unforced DNS. The first four leading POD modes of unforced DNS explain around 48%, 20%, 15%, and 7.5%. In (a), the magenta dash-dotted and yellow dotted lined display, respectively, the neutral vector of L and POD1 of the data obtained from the stochastic ODE (28); both lines coincide with the eigenvector. The eigenvalues of the shown eigenmodes are (a) $-0.058(1/\tau_{\mathrm{adv}})$, (b) $-1.303(1/\tau_{\mathrm{adv}})$, (c) $-1.506(1/\tau_{\mathrm{adv}})$, and (d) $(-2.726+0.504i)(1/\tau_{\mathrm{adv}})$ (see Fig. 11). All these modes are projected onto the grid space and normalized to have the magnitude of one.

shows that the POD1 of the (unforced) DNS and L's neutral vector are different, which is not surprising given that the presence of the boundary layers and turbulent plumes makes the flow anisotropic and spatially correlated. Just to demonstrate this point, for the L calculated using the GRF method and Gaussian white noise $\boldsymbol{\xi}(t)$, we have integrated

$$\dot{\boldsymbol{x}}(t) = L\boldsymbol{x}(t) + \boldsymbol{\xi}(t), \tag{28}$$

using the Euler-Maruyama method. The leading POD of this data set is shown in Fig. 10(a), which, unlike the POD1 of DNS, agrees with the neutral vector of L.

The second slowest-decaying mode [Fig. 10(b)] is real as well but spans both the interior and boundary layers and decays faster than $1/\tau_{\mathrm{adv}}$. The third slowest-decaying mode [Fig. 10(c)] is real and mostly varies in the interior. The fourth slowest-decaying mode [Fig. 10(d)] is complex with

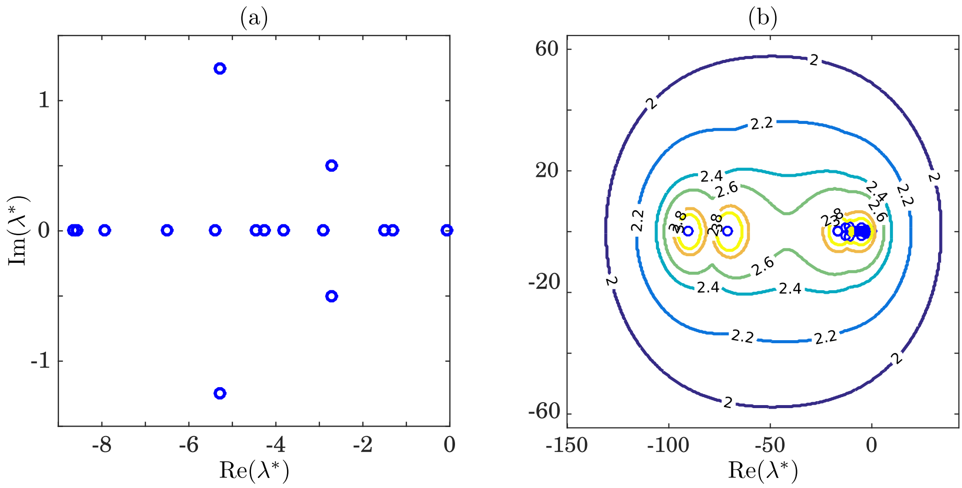

FIG. 11. (a) Eigenvalues of L nondimensionalized using the advective timescale τ_{adv}. (b) The ϵ pseudospectrum of L [Eq. (29)] calculated following Trefethen and Embree [80], which shows the non-normality of L. The numbers on the isolines are $-\log(\epsilon)$. In (a), only the first 15 eigenvalues (out of 25) are shown for better illustration and to keep the focus on the eigenvalues corresponding to the slowest-decaying modes. This has caused the large difference in the range of x axis of the two panels.

both real and imaginary parts of the eigenvector changing across the interior and boundary layers. This mode decays with the timescale of $\sim 0.37\tau_{\text{adv}}$ and oscillates with the frequency of around $2\tilde{\omega}$.

Figure 11(a) shows the eigenvalues of L, which all have negative real parts (i.e., decaying). Except for the slowest-decaying mode, all other eigenmodes decay with timescales faster than τ_{adv}; all eigenmodes of L decay faster than the diffusive timescale τ_{diff} ($\sim 840\tau_{\text{adv}}$). Figure 11(b) depicts the ϵ pseudospectrum of L [$\Lambda_\epsilon(L)$] given by [80]

$$\Lambda_\epsilon(L) = \{z \in C : \|(zI - L)^{-1}\|_2 \geqslant \epsilon^{-1}\}. \tag{29}$$

Here ϵ is the measure of proximity of a point in the complex plane C to the spectrum of L. The calculated pseudospectrum shows that L is non-normal and supports transient growth [81,82]. The non-normality of L also suggests that estimating L accurately using data-driven techniques such as the FDT can face similar challenges reported in [66] if POD modes are used as basis functions. In fact, recently Khodkar and Hassanzadeh [51] have shown that for this system, the POD-based FDT does not provide an accurate L. Guided by the results of Fig. 11(b), they proposed using DMD modes as the basis functions instead and showed that the DMD-enhanced FDT provides an accurate L, as accurate as the GRF-based LRF, for this system.

VIII. CONCLUSION

We have developed a 1D linear ROM in the form of Eq. (17) for a 3D Rayleigh-Bénard convection system, which is a fitting prototype for buoyancy-driven turbulence in various natural and engineering flows. Using the Green's function method, we have calculated the LRF L and EFM E at $\text{Ra} = 10^6$. The EFM E is basically a matrix that parametrizes changes in the divergence of vertical eddy heat flux based on changes in the temperature profile. In Sec. V, using several tests at $\text{Ra} = 10^6$, we have shown that L and E can accurately predict the time-mean responses of temperature and eddy heat flux to external forcings and that L can well predict the forcing needed to change the mean flow in a specified way (inverse problem). Furthermore, we have shown in Sec. VI that once these L and E are simply scaled by $\sqrt{\text{Ra}/10^6}$, they work equally well for flows at other Ra, at least in the investigated range of $5 \times 10^5 \leqslant \text{Ra} \leqslant 1.25 \times 10^6$.

The GRF method can be readily extended to use forcings that vary in the horizontal directions (e.g., applied at different Fourier modes one at a time) and are time dependent (e.g., applied at different frequencies one at a time). Such 3D ROMs, while computationally more expensive to calculate, can provide further insight into the spatiotemporal characteristics of buoyancy-driven turbulence.

The GRF method shows a promising performance for high-Ra turbulence, however there are two issues that should be highlighted. First, a key assumption in developing the 1D ROM is linearity of the response. While it has been shown that, at least for the large-scale atmospheric turbulence, L and E work well for responses or forcings that are large enough to be useful for various practical purposes [58,61–63], the limitations of the linearity assumption for the RB system and other problems should be explored in future studies. Second, the GRF method is computationally demanding because of the need for many forced full-dimensional simulations (although these simulations are needed only once, e.g., for the purpose of online flow control or optimization, the calculations can be done offline and the calculated LRF can then be used online with negligible computational cost). While the simple scaling found here suggests that the LRF and EFM do not have to be calculated for every Rayleigh number (at least for a range of Ra), the numerical cost can limit its use as a generally applicable method (particularly to build 3D ROMs). Still, calculating the accurate 1D and 3D ROMs using the GRF method for some turbulent systems has the following major advantages.

(i) Knowing the accurate L can guide developing better data-driven techniques, as done, for example, by Hassanzadeh and Kuang [8]. In particular, comparing the flow's Koopman-DMD modes with the eigenvectors or singular vectors of the L calculated here might be informative. In another direction, while we have not attempted to optimize the basis functions used in the GRF method in this work, the Koopman-DMD modes might provide some insight into better or optimal basis functions for the GRF method, which can reduce the computational cost and improve the accuracy.

(ii) Analyzing the spectral properties of E can help with better understanding the physics of eddy fluxes and improving the turbulence closure schemes, which connects with the ongoing efforts in developing better deterministic and stochastic parametrizations for geophysical turbulence [4,29,83,84].

ACKNOWLEDGMENTS

We thank Thanos Antoulas and Matthias Heinkenschloss for fruitful discussions, Arthi Appathurai for help with conducting some of the simulations, and two anonymous reviewers for insightful comments. This work was supported by funding provided to P.H. by NASA Grant No. 80NSSC17K0266, NSF Grant No. AGS-1552385, a Faculty Initiative Fund award from the Rice University Creative Ventures, and the Mitsubishi Electric Research Laboratories. This work used the Extreme Science and Engineering Discovery Environment Stampede2 through allocation ATM170020, the Yellowstone high-performance computing system provided by NCAR's Computational and Information Systems Laboratory through allocation NCAR0462, and the DAVinCI cluster of the Rice University Center for Research Computing.

---

[1] A. Liakopoulos, P. A. Blythe, and H. Gunes, A reduced dynamical model of convective flows in tall laterally heated cavities, Proc. R. Soc. A **453**, 663 (1997).

[2] B. Podvin and P. Le Quéré, Low-order models for the flow in a differentially heated cavity, Phys. Fluids **13**, 3204 (2001).

[3] A. Majda, R. V. Abramov, and M. J. Grote, *Information Theory and Stochastics for Multiscale Nonlinear Systems* (American Mathematical Society, Providence, 2005), Vol. 25.

[4] J. Bailon-Cuba and J. Schumacher, Low-dimensional model of turbulent Rayleigh-Bénard convection in a Cartesian cell with square domain, Phys. Fluids **23**, 077101 (2011).

[5] B. Podvin and A. Sergent, Proper orthogonal decomposition investigation of turbulent Rayleigh-Bénard convection in a rectangular cavity, Phys. Fluids **24**, 105106 (2012).
[6] O. San and J. Borggaard, Principal interval decomposition framework for the POD reduced-order modeling of convective Boussinesq flows, Int. J. Numer. Meth. Fluids **78**, 37 (2015).
[7] J. Annoni, P. Gebraad, and P. Seiler, *Proceedings of the 2016 American Control Conference, Boston, 2016* (IEEE, Piscataway, 2016), pp. 506–512.
[8] P. Hassanzadeh and Z. Kuang, The linear response function of an idealized atmosphere. Part II: Implications for the practical use of the fluctuation-dissipation theorem and the role of operator's non-normality, J. Atoms. Sci. **73**, 3441 (2016).
[9] B. Kramer, P. Grover, P. Boufounos, S. Nabi, and M. Benosman, Sparse sensing and DMD-based identification of flow regimes and bifurcations in complex flows, SIAM J. Appl. Dyn. Syst. **16**, 1164 (2017).
[10] P. Holmes, J. L. Lumley, G. Berkooz, and C. W. Rowley, *Turbulence, Coherent Structures, Dynamical Systems and Symmetry* (Cambridge University Press, Cambridge, 2012).
[11] B. R. Noack, M. Morzynski, and G. Tadmor, *Reduced-Order Modeling for Flow Control* (Springer Science + Business Media, New York, 2011), Vol. 528.
[12] T. N. Palmer, A nonlinear dynamical perspective on climate prediction, J. Clim. **12**, 575 (1999).
[13] C. Penland, A stochastic approach to nonlinear dynamics: A review, Bull. Am. Meteorol. Soc. **84**, 925 (2003).
[14] G. Berkooz, P. Holmes, and J. L. Lumley, The proper orthogonal decomposition in the analysis of turbulent flows, Annu. Rev. Fluid Mech. **25**, 539 (1993).
[15] C. W. Rowley, *Modeling and Computations in Dynamical Systems: In Commemoration of the 100th Anniversary of the Birth of John von Neumann* (World Scientific, Singapore, 2006), pp. 301–317.
[16] N. Aubry, P. Holmes, J. L. Lumley, and E. Stone, The dynamics of coherent structures in the wall region of a turbulent boundary layer, J. Fluid Mech. **192**, 115 (1988).
[17] G. Berkooz, P. Holmes, and J. L. Lumley, On the relation between low-dimensional models and the dynamics of coherent structures in the turbulent wall layer, Theor. Comput. Fluid Dyn. **4**, 255 (1993).
[18] J. Moehlis, T. R. Smith, P. Holmes, and H. Faisst, Models for turbulent plane Couette flow using the proper orthogonal decomposition, Phys. Fluids **14**, 2493 (2002).
[19] W. Cazemier, R. W. C. P. Verstappen, and A. E. P. Veldman, Proper orthogonal decomposition and low-dimensional models for driven cavity flows, Phys. Fluids **10**, 1685 (1998).
[20] H. Arbabi and I. Mezić, Ergodic theory, dynamic mode decomposition, and computation of spectral properties of the Koopman operator, SIAM J. Appl. Dyn. Syst. **16**, 2096 (2017).
[21] X. Ma and G. E. Karniadakis, A low-dimensional model for simulating three-dimensional cylinder flow, J. Fluid Mech. **458**, 181 (2002).
[22] B. R. Noack, K. Afanasiev, G. E. Morzyński, G. Tadmor, and F. A. Thiele, A hierarchy of low-dimensional models for the transient and post-transient cylinder wake, J. Fluid Mech. **497**, 335 (2006).
[23] C. W. Rowley, I. Mezić, S. Bagheri, P. Schlatter, and D. S. Henningson, Spectral analysis of nonlinear flows, J. Fluid Mech. **641**, 115 (2009).
[24] L. Sirovich and H. Park, Turbulent thermal convection in a finite domain: Part I. Theory, Phys. Fluids **2**, 1649 (1990).
[25] H. Park and L. Sirovich, Turbulent thermal convection in a finite domain: Part II. Numerical results, Phys. Fluids **2**, 1659 (1990).
[26] A. E. Deane and L. Sirovich, A computational study of Rayleigh-Bénard convection. Part 1. Rayleigh-number scaling, J. Fluid Mech. **222**, 231 (1991).
[27] L. Sirovich and A. E. Deane, A computational study of Rayleigh-Bénard convection. Part 2. Dimension considerations, J. Fluid Mech. **222**, 251 (1991).
[28] H. Gunes, A. Luiakopoulos, and R. A. Sahan, Low-dimensional oscillatory description of thermal convection: The small Prandtl number limit, Theor. Comput. Fluid Dyn. **9**, 1 (1997).
[29] M. Benosman, J. Borggaard, and B. Kramer, Learning-based robust stabilization for reduced-order models of 2D and 3D Boussinesq equations, Appl. Math. Model. **49**, 162 (2017).

[30] S. J. Andersen, J. N. Sørensen, and R. Mikkelsen, Simulation of the inherent turbulence and wake interaction inside an infinitely long row of wind turbines, J. Turbul. **14**, 1 (2013).
[31] J. Annoni, P. Gebraad, and P. Seiler, *Proceedings of the 2015 American Control Conference, Chicago, 2015* (IEEE, Piscataway, 2015), pp. 1721–1727.
[32] N. Hamilton, M. Tutkun, and R. B. Cal, Low-order representations of the canonical wind turbine array boundary layer via double proper orthogonal decomposition, Phys. Fluids **28**, 025103 (2016).
[33] C. W. Rowley and S. T. M. Dawson, Model reduction for flow analysis and control, Annu. Rev. Fluid Mech. **49**, 387 (2017).
[34] B. O. Koopman, Hamiltonian systems and transformation in Hilbert space, Proc. Natl. Acad. Sci. USA **17**, 315 (1931).
[35] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, Nonlinear Dyn. **41**, 309 (2005).
[36] P. J. Schmid and J. L. Sesterhenn, Proceedings of the 61st Annual Meeting of the APS Division of Fluid Dynamics, San Antonio, 2008 [Bull. Am. Phys. Soc. **53**, 208 (2008)].
[37] P. J. Schmid, Dynamic mode decomposition of numerical and experimental data, J. Fluid Mech. **656**, 5 (2010).
[38] 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).
[39] M. O Williams, I. G. Kevrekidis, and C. W. Rowley, A data-driven approximation of the Koopman operator: Extending dynamic mode decomposition, J. Nonlinear Sci. **25**, 1307 (2015).
[40] I. Mezić, Analysis of fluid flows via spectral properties of the Koopman operator, Annu. Rev. Fluid Mech. **45**, 357 (2013).
[41] J. Annoni and P. Seiler, A method to construct reduced-order parameter-varying models, Int. J. Robust. Nonlinear Control **27**, 582 (2017).
[42] D. Giannakis, A. Kolchinskaya, D. Krasnov, and J. Schumacher, Koopman analysis of the long-term evolution in a turbulent convection cell, J. Fluid Mech. **847**, 735 (2018).
[43] B. J. McKeon and A. S. Sharma, A critical layer framework for turbulent pipe flow, J. Fluid Mech. **658**, 336 (2010).
[44] A. S. Sharma and B. J. McKeon, On coherent structure in wall turbulence, J. Fluid Mech. **728**, 196 (2013).
[45] R. Moarref, M. R. Jovanović, J. A. Tropp, B. J. McKeon, and A. S. Sharma, A low-order decomposition of turbulent channel flow via resolvent analysis and convex optimization, Phys. Fluids **26**, 051701 (2014).
[46] F. Gómez, H. M. Blackburn, M. Rudman, B. J. McKeon, and A. S. Sharma, A reduced-order model of three-dimensional unsteady flow in a cavity based on the resolvent operator, J. Fluid Mech. **798**, R2 (2016).
[47] R. Kubo, The fluctuation-dissipation theorem, Rep. Prog. Phys. **29**, 255 (1966).
[48] C. E. Leith, Climate response and fluctuation dissipation, J. Atoms. Sci. **32**, 2022 (1975).
[49] C. Penland, Random forcing and forecasting using principal oscillation pattern analysis, Mon. Weather Rev. **117**, 2165 (1989).
[50] C. Penland, *Nonlinear Dynamics in Geosciences* (Springer, Berlin, 2007), pp. 485–515.
[51] M. A. Khodkar and P. Hassanzadeh, Data-driven reduced modeling of turbulent Rayleigh-Bénard convection using DMD-enhanced fluctuation-dissipation theorem, J. Fluid Mech. **852**, R3 (2018).
[52] A. Gritsun and G. Branstator, Climate response using a three-dimensional operator based on the fluctuation-dissipation theorem, J. Atmos. Sci. **64**, 2558 (2007).
[53] M. J. Ring and R. A. Plumb, The response of a simplified GCM to axisymmetric forcings: Applicability of fluctuation-dissipation theorem, J. Atoms. Sci. **65**, 3880 (2008).
[54] F. C. Cooper and P. H. Haynes, Climate sensitivity via a nonparametric fluctuation-dissipation theorem, J. Atmos. Sci. **68**, 937 (2011).
[55] F. C. Cooper, J. G. Esler, and P. H. Haynes, Estimation of the local response to a forcing in a high dimensional system using the fluctuation-dissipation theorem, Nonlinear Process. Geophys. **20**, 239 (2013).

[56] N. J. Lutsko, I. M. Held, and P. Zurita-Gotor, Applying the fluctuation-dissipation theorem to a two-layer model of quasigeostrophic turbulence, J. Atmos. Sci. **72**, 3161 (2015).
[57] Z. Kuang, Linear response functions of a cumulus ensemble to temperature and moisture perturbations and implications for the dynamics of convectively coupled waves, J. Atoms. Sci. **67**, 941 (2010).
[58] P. Hassanzadeh and Z. Kuang, The linear response function of an idealized atmosphere. Part I: Construction using Green's functions and applications, J. Atoms. Sci. **73**, 3423 (2016).
[59] Z. Kuang, Weakly forced mock Walker cells, J. Atmos. Sci. **69**, 2759 (2012).
[60] M. J. Herman and Z. Kuang, Linear response functions of two convective parametrization schemes, J. Adv. Model. Earth Syst. **5**, 510 (2013).
[61] P. Hassanzadeh and Z. Kuang, Blocking variability: Arctic amplification versus Arctic oscillation, Geophys. Res. Lett. **42**, 8586 (2015).
[62] D. Ma, P. Hassanzadeh, and Z. Kuang, Quantifying the eddy-jet feedback strength of the annular mode in an idealized GCM and reanalysis data, J. Atmos. Sci. **74**, 393 (2017).
[63] P. Hassanzadeh and Z. Kuang, Quantifying the annular mode dynamics in an idealized atmosphere, arXiv:1809.01054.
[64] G. Ahlers, S. Grossmann, and D. Lohse, Heat transfer and large scale dynamics in turbulent Rayleigh-Bénard convection, Rev. Mod. Phys. **81**, 503 (2009).
[65] C. R. Doering and P. Constantin, Variational bounds on energy dissipation in incompressible flows. III. Convection, Phys. Rev. E **53**, 5957 (1996).
[66] P. Hassanzadeh, G. P. Chini, and C. R. Doering, Wall to wall optimal transport, J. Fluid Mech. **751**, 627 (2014).
[67] A. Farhat, E. Lunasin, and E. S. Titi, Continuous data assimilation for a 2D Bénard convection system through horizontal velocity measurements alone, J. Nonlinear Sci. **27**, 1065 (2017).
[68] B. Wen, L. T. Corson, and G. P. Chini, Structure and stability of steady porous medium convection at large Rayleigh number, J. Fluid Mech. **772**, 197 (2015).
[69] B. Wen and G. P. Chini, Reduced modeling of porous media convection in a minimal flow unit at large Rayleigh number, J. Comput. Phys. **371**, 551 (2018).
[70] P. G. Drazin and W. H. Reid, *Hydrodynamic Stability* (Cambridge University Press, Cambridge, 2004).
[71] J. A. Barranco and P. S. Marcus, A 3D spectral anelastic hydrodynamic code for shearing, stratified flows, J. Comput. Phys. **219**, 21 (2006).
[72] P. S. Marcus, Simulation of Taylor-Couette flow. Part 1. Numerical methods and comparison with experiment, J. Fluid Mech. **146**, 45 (1984).
[73] J. A. Barranco and P. S. Marcus, Three-dimensional vortices in stratified protoplanetary disks, Astrophys. J. **623**, 1157 (2005).
[74] P. Hassanzadeh, P. S. Marcus, and P. Le Gal, The universal aspect ratio of vortices in rotating stratified flows: Theory and simulation, J. Fluid Mech. **706**, 46 (2012).
[75] P. Hassanzadeh, Baroclinic vortices in rotating stratified shearing flows: Cyclones, anticyclones, and zombie vortices, Ph.D. thesis, University of California at Berkeley, 2013.
[76] P. S. Marcus, S. Pei, C.-H. Jiang, and P. Hassanzadeh, Three-Dimensional Vortices Generated by Self-Replication in Stably Stratified Rotating Shear Flows, Phys. Rev. Lett. **111**, 084501 (2013).
[77] P. S. Marcus, S. Pei, C.-H. Jiang, J. A. Barranco, P. Hassanzadeh, and D. Lecoanet, Zombie vortex instability. I. A purely hydrodynamic instability to resurrect the dead zones of protoplanetary disks, Astrophys. J. **808**, 87 (2015).
[78] J. Marshall and F. Molteni, Toward a dynamical understanding of planetary-scale flow regimes, J. Atoms. Sci. **50**, 1792 (1993).
[79] J. C. Goodman and J. Marshall, Using neutral singular vectors to study low-frequency atmospheric variability, J. Atoms. Sci. **59**, 3206 (2002).
[80] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators* (Princeton University Press, Princeton, 2005).
[81] B. F. Farrell and P. J. Ioannou, Generalized stability theory. Part I: Autonomous operators, J. Atmos. Sci. **53**, 2025 (1996).

[82] L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, Hydrodynamic stability without eigenvalues, Science **261**, 578 (1993).
[83] F. C. Cooper and L. Zanna, Optimisation of an idealised ocean model, stochastic parametrization of sub-grid eddies, Ocean Model. **88**, 38 (2015).
[84] Z. Tan, C. M. Kaul, K. G. Pressel, Y. Cohen, T. Schneider, and J. Teixeira, An extended eddy-diffusivity mass-flux scheme for unified representation of subgrid-scale turbulence and convection, J. Adv. Model. Earth Syst. **10**, 770 (2018).