Stochastic data-driven POD-based modeling for high-fidelity coarsening of two-dimensional Rayleigh-Bénard turbulence

S.R. Ephrati, P. Cifani, and B.J. Geurts

1 Introduction

The aim of LES is that of performing accurate numerical simulations at a moderate computational cost. The challenge is to devise a significant coarsening of the model while retaining an adequate level of reliability. In recent years, due to ever-growing computational resources, it has become possible to perform high-resolution direct numerical simulations (DNS) from which data are used to build coarse-grid fluid models. The latter allows for high-fidelity simulation of the dynamics of large-scale flow structures over long time horizons, otherwise unreachable by DNS. In particular, data is used to find LES closure models [3] or to alter governing equations [2]. In addition, observational data can be used as an alternative or to complement a DNS dataset. Data assimilation, in fact, has been used for several decades in the context of geophysical fluid dynamics [1, 7] and is of particular interest to this work.
Here we propose a data-driven approach to augment the coarse-grained numerical predictions of two-dimensional Rayleigh-Bénard convection. Proper orthogonal decomposition (POD) is used to decompose fine-grid reference data into spatial fields and temporal coefficients. These coefficients are modelled as stochastic processes whose variance equals the POD spectrum of the flow. We found that a comparably small number of POD modes is sufficient to attain an accurate estimation of the Nusselt number even at high Rayleigh number, thereby enabling studies in the range of 'ultimate turbulence'. This strong coarsening makes the model efficient and accurate over long time.

S.R. Ephrati · P. Cifani · B.J. Geurts
University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands, e-mail: {s.r.ephrati,p.cifani,b.j.geurts}@utwente.nl
Gran Sasso Science Institute, Viale F. Crispi, 7 67100 L'Aquila, Italy
Eindhoven University of Technology, P.O. Box 213, 5600 MB Eindhoven, The Netherlands

2 Data generation and processing

Rayleigh-Bénard convection is described by the Navier-Stokes equations coupled to temperature transport under the Boussinesq approximation. The governing equations, in non-dimensional form, are the following:

$$\frac{\partial u}{\partial t} + u \cdot \nabla u = \sqrt{\frac{Pr}{Ra}} \nabla^2 u - \nabla p + T e_z, \tag{1}$$

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

$$\frac{\partial T}{\partial t} + u \cdot \nabla T = \frac{1}{\sqrt{PrRa}} \nabla^2 T. \tag{3}$$

Here, we restrict to two spatial dimensions. We adopt u, p and T to denote velocity, pressure and temperature, respectively. Buoyancy effects due to temperature gradients are included by means of the term Te_z, where e_z is the vertical unit vector. The dimensionless parameters that determine the flow are the Rayleigh number $Ra = g\beta\Delta L_y^3/(\nu\kappa)$ and the Prandtl number $Pr = \nu/\kappa$, set to 10^{10} and 1, respectively. Here g is the gravitational acceleration, β the thermal expansion coefficient, Δ the temperature difference between the boundaries, ν the kinematic viscosity and κ the thermal expansion coefficient. The computational domain is a rectangle of size $L_x = 2, L_y = 1$. Periodic boundary conditions are imposed in the horizontal direction while the top and bottom boundary are no-slip walls. The temperature is prescribed at 1 at the top wall and 0 at the bottom wall.

The time integration is carried out by employing an RK3 method. An energy-conserving finite difference method is employed for the spatial discretization [10] and parallelized as in [5]. A uniform grid spacing is adopted along the periodic direction while grid refinement close to the wall is realized by means of a hyperbolic tangent profile to properly resolve the boundary layer.

The DNS simulation is performed over a 4096×2048 grid. This resolution has been shown to be adequate for the selected Ra number, as indicated in [11]. Coarse-grid runs are carried out on a 64×32 grid. The rather strong coarsening allows to identify clear differences between the proposed model and the solution of system (1-3) solved on the fine grid without model, as shown in Fig. 1.

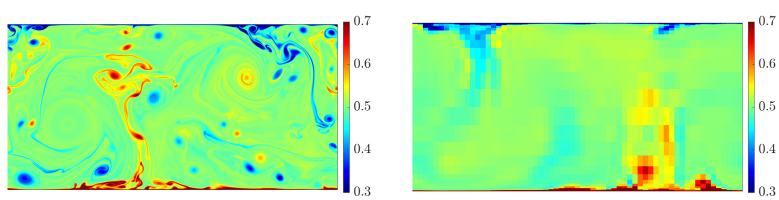

Fig. 1 Snapshots of the temperature field of the DNS (left) and the numerical simulation on the coarse grid using 100 POD modes (right).

Proper orthogonal decomposition (POD) is applied to a sequence of equispaced and uncorrelated DNS snapshots. The spatial and temporal coefficients, indicated by $\sigma_i(x)$ and $a_i(t)$, respectively, follow from the minimization problem

$$\min \int_\Omega \int_T \left(f(x,t) - \sum_{i=1}^{N} a_i(t)\sigma_i(x) \right)^2 \mathrm{d}t\mathrm{d}\Omega, \tag{4}$$

where $f(x,t)$ is either the velocity or the temperature field. The fraction of energy related to each of the modes is given by the variance of the corresponding temporal coefficient. The solution of (4) is given by the eigenvectors of the spatial-spatial correlation tensor:

$$\mu_i a_i(t) = \int_T \langle f(x,t) f(x,t') \rangle_\Omega a_i(t')\mathrm{d}t', \tag{5}$$

where $\langle \cdot \rangle_\Omega$ denotes integration over the spatial domain and μ_i is the standard deviation of the time series $a_i(t)$. The spatial profiles $\sigma_i(x)$ then follow from orthogonality:

$$\sigma_i(x) = \frac{\langle f(x,t) a_i(t) \rangle_T}{\langle a_i(t)^2 \rangle_T}, \tag{6}$$

where $\langle \cdot \rangle_T$ indicates integration over time. For an in-depth description of this procedure the reader is referred to [4].

3 Stochastic modelling

3.1 Model description

The momentum and temperature equations (1, 3) can be written as a system of ODEs for the temporal coefficients by projection onto the POD basis. For each POD mode one has

$$\frac{\mathrm{d}a_{i,f}}{\mathrm{d}t} = L_f(\{a_{j,f}(t)\}), \quad i = 0, ..., N, \tag{7}$$

where L_f is formally the spectral representation of $\mathbf{u} \cdot \nabla$, ∇, ∇^2 and the source term appearing in (1-3). This system of ODEs involves all coefficients. In actual models we truncate this formulation to include only a certain number N of modes.
Representing the fine-scale dynamics in the coarsened model using the POD time series has been shown to be an efficient approach to improve coarse-grid simulations [6]. In the same spirit, here we regard the evolution of the coefficients as a stochastic process. First, Eq. (7) is integrated in physical space to find an intermediate value $\tilde{a}_{i,f}$. Subsequently, after projection, the temporal coefficients at the new time level $n+1$ are updated by

$$a_{i,f}^{n+1} = \widetilde{a}_{i,f}^{n+1}\left(1 - \frac{\Delta t}{\tau_i}\right) + \mu_i \sqrt{1 - \left(1 - \frac{\Delta t}{\tau_i}\right)^2}\, r_i^n \tag{8}$$

where μ_i and τ_i are the standard deviation and time correlation of the temporal coefficients computed from the DNS dataset. Moreover, r^n is a random variable drawn from a standard normal distribution. The formulation (8) is a first-order autoregressive model with drift coefficient $1 - \Delta t/\tau_i$ and variance μ_i^2. Once $a_{i,f}^{n+1}$ is computed from (8), the full spatial field $f(x,t)$ is reconstructed in physical space and the time step is completed. The projection and reconstruction of the field from/to physical space to/from spectral space is a computationally intensive operation. However, as we will show in the next section, only a few POD modes are necessary to significantly improve the predictions of the flow statistics compared to adopting no model at all. Thus, the overall cost of the model is comparable with that of solving the coarse-grained PDE without the model.

3.2 Nusselt number predictions

First we verify that the simulated coarse-grid POD spectrum is well represented when using (8). Preliminary tests have shown that the temperature spectrum simulated without model matches the DNS spectrum well, in particular at large scales. However, the velocity spectrum showed considerable differences at all scales. Given these findings, we apply the model to the momentum equation only. In Fig. 2 and 3 the POD spectrum of both the velocity and the temperature is shown when using the primary 20 and 100 modes in the forcing. The velocity spectra match the reference spectrum up to the forced mode. Furthermore, an overall improvement is observed in the temperature spectra compared to the no-model case.

Next we predict the Nusselt number and compare it with its reference value obtained from DNS. We adopt the following definition

$$Nu = 1 + \sqrt{PrRa}\langle u_y T\rangle_\Omega \tag{9}$$

where u_y denotes the vertical velocity and $\langle\ ,\rangle_\Omega$ denotes the volume average. This definition is particularly suitable for the computation of the Nusselt number since it does not involve gradients, which would introduce an additional source of error when discretized on a coarse grid. Fig. 4 shows the time averages of the Nusselt number as measured on the reference fine grid, on the coarse grid without model and on the coarse grid with the stochastic model. Increasing the number of POD modes clearly results in a more accurate estimation of Nu. Even at the lowest number of POD modes simulated, a significant improvement is clearly visible. This suggests that a correction on the large-scale spectrum of velocity is beneficial for the correct prediction of mean heat transfer.

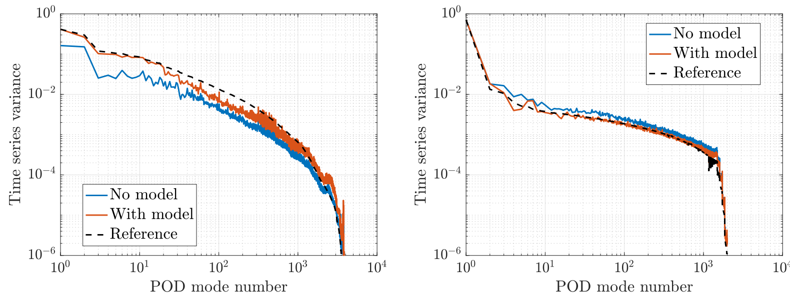

Fig. 2 POD Spectra of the velocity (left) and temperature (right). The spectra of the reference solution (dashed line), the no-model (blue line) and the stochastic model using 20 POD modes (red line) are shown.

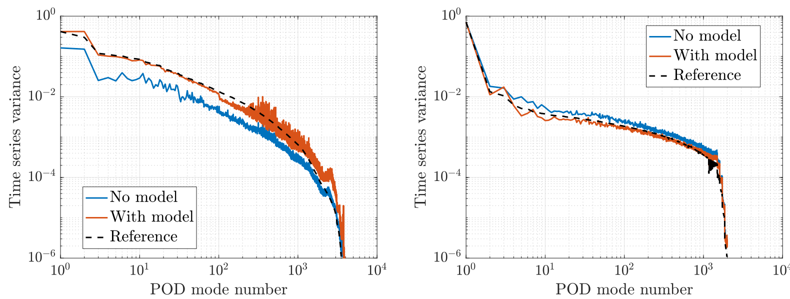

Fig. 3 POD Spectra of the velocity (left) and temperature (right). The spectra of the reference solution (dashed line), the no-model (blue line) and the stochastic model using 100 POD modes (red line) are shown.

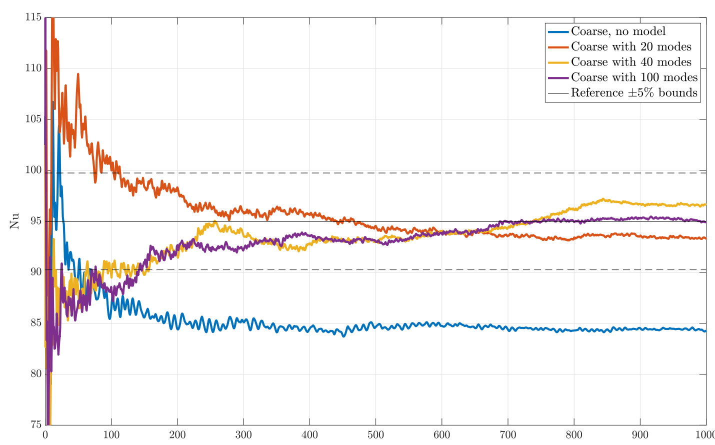

Fig. 4 Rolling time averages of the Nusselt number obtained from the no-model and the stochastic model using 20, 40 and 100 POD modes. The reference DNS result is represented by the solid line together with a 5% uncertainty interval.

4 Conclusions

In this work, a stochastic data-driven model for coarse-grained numerical simulations for two-dimensional Rayleigh-Bénard turbulence has been proposed. The model is based on the proper orthogonal decomposition (POD) of high-resolution numerical data. In particular, parameters are extracted from the temporal coefficients of the POD basis and are used to augment the coarse-grained numerical predictions via stochastic forcing. The resulting model requires only two parameters per POD mode, making it straightforward to interpret and to implement. In fact, these two parameters, i.e., (τ_i, μ_i) for the time correlation and the standard deviation, can be computed per POD mode from the DNS data and are not free parameters. It was found that the Nusselt number estimates on the coarse grid with stochastic forcing improved significantly compared to those obtained without a model. The method presented here is general and may be extended to different flows.

Acknowledgements This work was performed in the context of the SPRESTO project, funded by the Dutch Science Foundation (NWO) in their TOP1 program.

References

1. Azouani, A., Olson, E., Titi, E.S.: Continuous data assimilation using general interpolant observables. Journal of Nonlinear Science, 24(2), 277-304 (2014).
2. Bai, Z.,: Data-driven methods in fluid dynamics: Sparse classification from experimental data. In: Pollard, A., Castillo, L., Danaila, L., Glauser, M. (eds) Whither turbulence and big data in the 21st century?, pp. 323-342. Springer, Cham. (2017)
3. Beck, A., Kurz, M.: A perspective on machine learning methods in turbulence modeling. GAMM-Mitteilungen, 44(1), e202100002 (2021).
4. Cazemier, W., Verstappen, R.W.C.P., Veldman, A.E.P.: Proper orthogonal decomposition and low-dimensional models for driven cavity flows. Physics of fluids, 10(7), 1685-1699 (1998).
5. Cifani, P., Kuerten, J.G.M. and Geurts, B.J., 2018. Highly scalable DNS solver for turbulent bubble-laden channel flow. Computers & Fluids, 172, pp.67-83.
6. Ephrati, S.R., Luesink, E., Wimmer, G., Cifani, P., Geurts, B.J.: Computational modeling for high-fidelity coarsening of shallow water equations based on subgrid data. SIAM Multiscale Modeling and Simulation, accepted for publication.
7. Gesho, M., Olson, E., Titi, E.S.: A computational study of a data assimilation algorithm for the two-dimensional Navier-Stokes equations. Communications in computational physics, 19(4), 1094-1110 (2016).
8. Johnston, H., Doering, C.R.: Comparison of turbulent convection between conditions of constant temperature and constant flux. Physical review letters, 102(6), 064501 (2009).
9. Stevens, R.J., Verzicco, R., Lohse, D.: Radial boundary layer structure and Nusselt number in Rayleigh-Bénard convection. Journal of fluid mechanics, 643, 495-507 (2010).
10. Vreman, A.W., 2014. The projection method for the incompressible Navier–Stokes equations: the pressure near a no-slip wall. Journal of Computational Physics, 263, pp.353-374.
11. Zhu, X., Mathai, V., Stevens, R. J., Verzicco, R., & Lohse, D. (2018). Transition to the ultimate regime in two-dimensional Rayleigh-Bénard convection. Physical review letters, 120(14), 144502.