A machine learning framework for correcting under-resolved simulations of turbulent systems using nudged datasets

Benedikt Barthel and Themistoklis Sapsis

Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Abstract

Due to the rapidly changing climate, the frequency and severity of extreme weather, such as storms and heatwaves is expected to increase drastically over the coming decades. Accurately quantifying the risk of such events with high spatial resolution is a critical step in the implementation of strategies to prepare for and mitigate the damages. As fully resolved simulations remain computationally out of reach, policy makers must rely on coarse resolution climate models which either parameterize or completely ignore sub-grid scale dynamics. In this work we propose a machine learning framework to debias under-resolved simulations of complex and chaotic dynamical systems such as atmospheric dynamics. The proposed strategy uses "nudged" simulations of the coarse model to generate training data designed to minimize the effects of chaotic divergence. We illustrate through a prototype QG model that the proposed approach allow us to machine learn a map from the chaotic attractor of under-resolved dynamics to that of the fully resolved system. In this way we are able to recover extreme event statistics using a very small training dataset.

1 Introduction

The accurate quantification of the risks of extreme climate events such as heatwaves, droughts, and storms is a critical step in the implementation of strategies to prepare for and mitigate the damage they cause [1, 2]. This requires simulating a large ensemble of climate realizations as these extreme events are, by definition, seldom observed and arise due to a range of physical mechanisms [3, 4]. Unfortunately, the vast range of scales involved in the turbulent atmospheric system make even a single fully resolved simulation over a meaningful time horizon an intractable task [5]. Therefore, scientists and policy makers must resort to under-resolved (coarse-scale) climate models which omit or model the small scale (sub-grid) physics, and despite decades of work, a comprehensive model of these sub-grid dynamics remains elusive.

Furthermore, the chaotic nature of atmospheric dynamics – especially at the sub-grid scale – is not only a challenge for numerical simulations but also for data driven modeling, as it can lead to instabilities of the machine learned models [6, 7] and inaccurate predictions [8, 9]. However, while accurately capturing small-scale statistics remains a challenge, it is the statistics of the large scales that are of primary interest. These are inevitably influenced by the small scales and therefore it is essential to correct the large scale dynamics when these are obtained from an under-resolved climate model. To this end, a large body of work has been dedicated to correcting statistics (debiasing) of under-resolved data [10, 11, 12], and post-processing under-resolved simulations to reflect observed data [13, 10].

Tackling Climate Change with Machine Learning: workshop at NeurIPS 2023.

Here we present a comprehensive machine learning framework for debiasing under-resolved climate simulations. To build a meaningful correction map for the coarse-scale outputs we utilize *nudged* [14, 15] coarse-scale model outputs by the reference data. The result is a training dataset that minimizes the corrupting effects of chaotic dynamics, leading to healthier learning of a correction operator.

2 Methods

Consider the discretized form of a general dynamical system,

$$\frac{\partial \mathbf{q}}{\partial t} = \mathbf{f}(\mathbf{q}) \tag{1}$$

where $\mathbf{q} \in \mathbb{R}^n$ is the vector of the solution evaluated at discrete grid points. Let M be the value of n such that the solution to the discrete system (1) is fully converged. That is to say that there is no change in the statistics of the solution for n > M. Thus we refer to as a solution obtained with n = M to be *fully-resolved*, and any solution with n < M to be *under-resolved*. We will refer to any fully resolved data as *reference data* (RD) and any under-resolved data as *coarse data* (CR).

The aim of our work is to machine learn a data-driven map which takes as its input the underresolved coarse simulation and outputs a solution (on the same coarse grid) whose statistics match those of a fully-resolved reference solution. The primary challenge therein is chaotic divergence. Here we propose a machine learning framework to overcome this challenge. Specifically, the novel aspect of this work is the observation that a reliable map from any trajectory in the chaotic attractor of the under-resolved system to a trajectory in the attractor of the fully-resolved system can in fact be found if the particular under-resolved trajectory is chosen with care.

2.1 Theory

Consider that in our training data set we have one realization of the fully-resolved data which serves as the target output of our model. If we observe some particular realization of the under-resolved model it will in general differ from that fully-resolved realization. This difference is due to two distinct phenomena: first, the effects of the insufficient resolution (the discretized system does not adequately approximate its continuous origin), and second, the effects of chaotic divergence. Our aim is to find, or at least approximate, the one particular realization of the under-resolved model which minimizes the latter difference and thus minimizes the effects of chaotic divergence on our machine learned map.

For a chaotic system many different realizations are possible from infinitesimally similar initial conditions. As the input training data we would like that *one* particular realization of the underresolved system which optimally tracks the qualitative behaviour of our *one* particular realization of the fully-resolved data in our training set. In other words, we seek the particular realization that is maximally correlated with the fully resolved reference solution. We refer to these two realizations (under- and fully- resolved) as being subject to the same "chaotic fate". It is of course infeasible to conduct an infinite amount of simulations of the coarse-scale model and select the realization.

We propose that this realization may be approximated as the solution to the forced coarse-scale dynamics

$$\frac{\partial \mathbf{q}}{\partial t} = \mathbf{f}(\mathbf{q}) - \mathcal{Q}(\mathbf{q}), \text{ where, } \mathcal{Q}(\mathbf{q}) \equiv -\frac{1}{\tau} \left(\mathbf{q} - \mathbf{q}^{RD} \right)$$
(2)

Here \mathbf{q}^{RD} is the known reference solution and τ is a user defined relaxation time scale. This forcing of the governing equations by a term proportional to the deviation from some known reference data is known as *nudging* [16, 17, 18]. The extra term on the right hand side is called the nudging tendency, and acts to penalize the trajectory of the under-resolved solution from straying too far from the known fully-resolved solution. We define the solution to this forced system as the *nudged coarse-scale* (NC) data set. To avoid confusion we will distinguish between nudged and non-nudged solutions by referring to them as *nudged* and *free-running* respectively. When a model is trained on the map from the nudged data to the reference, $q^{NC} \rightarrow q^{RD}$, the model will learn only the underlying transformation from the chaotic attractor of under-resolved dynamics to that of the fullyresolved system – and not the arbitrary chaotic differences between the two arbitrary trajectories which happened to be used for training. When the model is then evaluated on a new realization of *free-running* data it is able to accurately map that trajectory into a new trajectory in the attractor of the fully-resolved system $q^{CR} \rightarrow q^{RD}$. A formal definition and more detailed discussion of the nudging strategy is presented in appendix A.1.

2.2 Dataset

We apply our method to a two-layer incompressible quasi-geostrophic (QG) flow, specifically, a model proposed by [19]. The flow is parameterized by the beta-plane parameter β and the bottom drag coefficient r. The under-resolved (coarse) and fully-resolved (reference) data sets are obtained by solving the governing equations using a spectral resolution of 24×24 and 128×128 respectively. Throughout the following discussion all results will be presented in the form of the stream function – as this uniquely defines the velocity field, this choice incurs no loss of generality. A more detailed overview of the model, the numerical scheme used, and some illustrative examples of the data are included in appendix A.2.

2.3 Machine learning architecture and training strategy

The neural network model we employ takes in as an input (and outputs) the stream function field of both layers - a data set of dimension $24 \times 24 \times 2$. This vector is then compressed through a fully connected layer of dimension 60 and then passed through a long-short-term-memory (LSTM) layer of the same size before being expanded through a second fully connected layer to restore the data to its original size. The fully connected layers utilize hyperbolic tangent activation and the LSTM layer use a hard-sigmoid activation. The model is trained on a semi-physics informed loss function which consists of the L2 norm of the error between the predicted and true stream functions augmented with a second term which penalizes errors in the conservation of mass (see appendix A.3). Training is conducted over 2000 epochs on sequences of 100 data points spanning 10 time units from one realization of the flow with $\beta = 2.0$ and r = 0.1. All results presented below are evaluated on a separate realization of the flow such that none of the data used to generate the results below was in any way used in training. Furthermore, we reiterate that at test time, the model (trained on nudged coarse data) is evaluated on free-running coarse data so that no reference data is used to generate the following results.

3 Results

First, we apply our models, which are trained on data with $\beta = 2.0$ and r = 0.1, to a new realization of the flow with these same parameters. Figure 1a compares the predicted spectra and probability density functions (solid red) to both the fully-resolved reference (solid black) and coarse-scale data (dashed black) for the stream function in both layers. Recall that the machine learned correction takes as an input the coarse scale data and aims to predict the fully resolved reference. For both layers, the model predictions show good agreement with the fully-resolved reference. In terms of the spectra, the model accurately captures the two peaks around f = 0.15 and f = 0.3, and only deviates significantly at very high frequencies. In terms of the probability density functions, the model slightly underpredicts the positive tails, but captures the general shape well.

We reiterate that the only claim we make upon the trajectories predicted by our model is that they reflect the statistical properties of the fully resolved system. However, we expect our predictions to exhibit the qualitative behaviour of the exact solution. To this end we show in figure 1b the zonal average of the predicted solution which indeed shows good qualitative agreement with the fully-resolved simulation.

Next, we apply our model to a realization of the QG model with parameters which were not included in the training data, namely $\beta = 1.1$ and r = 0.5. For these parameter choices the flow lacks the spectral peaks of the β and r_d used to train the model exhibiting much more uniform frequency content. The lack of a dominant time scale makes this a challenging test case to evaluate the generalizability of our model. The results are summarized in figure 2 where we again plot the spectra, probability density function, (a) and zonal mean stream function (b). In both layers the model accurately predicts the spectrum across much of the frequency domain, but underpredicts the spectral decay, and thus over-predicts the strength of the highest frequencies. In terms of the



Figure 1: Model prediction for $\beta = 2.0$ and r = 0.1. Power spectrum and probability density function of stream functions ψ_1 (left) and ψ_2 (right), ML prediction (red), reference (solid black), coarse (dash black) (a). Zonally averaged stream function $\bar{\psi}_1$, reference (upper panel) and ML prediction (lower panel) (b). Training data: $\beta = 2.0$ and r = 0.1.

probability density function, there is excellent agreement in layer 1, while in layer two the model notably over-predicts the tails. Finally, the temporal evolution of the zonal average predicted stream function shows good qualitative agreement with the reference, producing structures of similar characteristic spatio-temporal scales. We note that the model has no intrinsic knowledge of the governing equations to aid in this generalization.



Figure 2: Same as figure 1 but for $\beta = 1.1$ and r = 0.5. Training data: $\beta = 2.0$ and r = 0.1.

4 Discussion

In this work we proposed a general machine learning framework to debias under-resolved simulations of complex and chaotic dynamical systems. The proposed strategy uses "nudged" simulations to generate training data which minimize the effects of chaotic divergence, thereby enabling models which reliably map trajectories from the chaotic attractor of under-resolved dynamics to that of the fully resolved system. The proposed strategy was illustrated on a chaotic quasi-geostrophic reduced order climate model. Our method accurately reproduced first order statistics and the qualitative behaviour of the temporal dynamics of the flow, and also generalized well to parameter regimes not included in the training data set. Applying the proposed framework to more complex climate models will allow for the prediction of accurate ensemble statistics without relying on expensive high resolution simulations. Furthermore, the proposed framework is defined in terms of a general dynamical system, and is thus agnostic to the nature of the problem under investigation or type of model used. Therefore, we believe the proposed method has the potential to be extended to data-driven modeling in a range of fields outside of fluid dynamics and climate science.

5 Acknowledgments

This research has been supported by the DARPA grant HR00112290029

A Appendix

A.1 Spectrum-matched nudged training data

The nudged solution is defined as the solution to an artificially forced form of the coarse-scale dynamics,

$$\frac{\partial \mathbf{q}}{\partial t} = \mathbf{f}(\mathbf{q}) - \mathcal{Q}(\mathbf{q}) \tag{3}$$

where

$$\mathcal{Q}(\mathbf{q}) \equiv -\frac{1}{\tau} \left(\mathbf{q} - \mathbf{q}^{RD} \right). \tag{4}$$

Here \mathbf{q}^{RD} is the known reference solution and τ is a user defined relaxation time scale. The term (4) is called the nudging tendency, and acts to penalize the trajectory of the under-resolved solution from straying too far from the known fully-resolved solution.

The introduction of the nudging term (4) artificially injects energy into the system resulting in a power spectrum which does not match that of the free-running solution. This in turn implies that the nudged solution will not exist within the turbulent attractor of the free-running system. As we intend our machine learned model to be a map from one attractor to another, this discrepancy must be remedied. This is done through the following correction operation which forces the power spectrum of the nudged solution to match that of the free-running solution so that training and testing data come from the same distributions. Consider the Fourier decomposition of the nudged solution solved on the coarse grid

$$\mathbf{q}^{N} = \sum_{\mathbf{k}=0}^{M} \hat{\mathbf{q}}_{\mathbf{k}}^{N} e^{i\mathbf{k}\cdot\mathbf{x}}$$
(5)

We multiply each Fourier mode by a correction term $a_{\mathbf{k}}$ and define the *nudged coarse-scale* (NC) data set.

$$\mathbf{q}^{NC} = \sum_{\mathbf{k}=0}^{M} a_{\mathbf{k}} \hat{\mathbf{q}}_{\mathbf{k}}^{N} e^{i\mathbf{k}\cdot\mathbf{x}}$$
(6)

Here, the coefficients which represents the energy-ratio between the free-running and nudged solutions are defined as

$$a_{\mathbf{k}} \equiv \sqrt{\frac{\mathcal{E}_{\mathbf{k}}^{CR}}{\mathcal{E}_{\mathbf{k}}^{N}}} \tag{7}$$

where

$$\mathcal{E}_{\mathbf{k}} = \int_0^T |\hat{\mathbf{q}}_{\mathbf{k}}|^2 dt.$$
(8)

The resulting dataset shares the qualitative dynamic behaviour of the solution to the nudged coarse system of equations – and thus the reference data – while reflecting the power spectrum of the free running coarse solution.

A.2 Quasi-Geostrophic Model

The two-layer incompressible quasi-geostrophic (QG) flow proposed by [19] is governed by the dimensionless equation

$$\frac{\partial q_j}{\partial t} + \mathbf{u}_j \cdot \nabla q_j + \left(\beta + k_d^2 U_j\right) \frac{\partial \psi_j}{\partial x} = -\delta_{2,j} r \nabla^2 \psi_j - \nu \nabla^4 q_j \tag{9}$$

where j = 1, 2 corresponds to the upper and lower layer respectively, r the bottom-drag coefficient and β is the beta-plane approximation parameter, and k_d^2 denotes the deformation frequency which for this study we fix at 4 – a value consistent with the radius and rotation of the earth and the characteristic length and velocity scales of its atmosphere [19]. This model is intended to approximate mid to high latitude atmospheric flows subject to an imposed shear current. A Taylor expansion of the Coriolis force reveals that for this assumption to hold we require roughly that $\beta \in [1, 2]$, which corresponds to an approximate latitude range of $\phi_0 \in [29^\circ, 64^\circ]$. The flow is defined in the horizontal domain $(x, y) \in [0, 2\pi]$ and is subject to doubly periodic boundary conditions. The state variable is represented in three forms: velocity: \mathbf{u}_j , potential vorticity (PV): q_j and the stream function: ψ_j . The latter are related via the inversion formula

$$q_j = \nabla^2 \psi_j + \frac{k_d^2}{2} \left(\psi_{3-j} - \psi_j \right)$$
(10)

and the velocity is related to the stream function by $\mathbf{u}_j = U_j + \nabla \times \psi_j \hat{\mathbf{k}}$ where $\hat{\mathbf{k}}$ is the unit vector orthogonal to the (x, y) plane and $U_j = -1^{(j+1)}U$, with U = 0.2 represents the imposed mean shear flow. The corresponding nudged system of equations is given by

$$\frac{\partial q_j}{\partial t} + \mathbf{u}_j \cdot \nabla q_j + \left(\beta + k_d^2 U_j\right) \frac{\partial \psi_j}{\partial x} = -\delta_{2,j} r \nabla^2 \psi_j - \nu \nabla^8 q_j - \frac{1}{\tau} \left(q_j - q_j^{RD}\right)$$
(11)

where we fix $\tau = 16$. Furthermore, we note that while the nudging penalty is applied to the vorticity, it could have equivalently been applied to the stream function or velocity. These possibilities are not explored in this work, however, as these three variables are all directly related we would not expect significant differences in the results.

The equations (9) and (11) are solved using a spectral method, with a spectral resolution of 24×24 and 128×128 for the coarse- and fine-scale data respectively. The time integration is evaluated using a 4^{th} order Runga-Kutta scheme with the same temporal resolution used for both the underand fully-resolved simulations. Throughout the following discussion all results will be presented in the form of the stream function – as this uniquely defines the velocity and thus vorticity, this choice incurs no loss of generality. Additionally, we define the zonally averaged stream function as the integral over the x dimension,

$$\bar{\psi}_j(y,t) = \frac{1}{2\pi} \int_0^{2\pi} \psi_j(x,y,t) dx.$$
 (12)

In figure 3 we show some illustrative examples of both the fully- and under-resolved data. We plot the zonal average, $\bar{\psi}_1$, for the reference data (RD) solved on a 128 × 128 grid, the free-running coarse data (CR) solved on a 24 × 24 grid and the nudged coarse data (NC) also solved on the a 24 × 24 grid. The primary qualitative difference between the coarse and fine grid solutions is in their amplitude. This is particularly clear when comparing the tails of the distributions in 3b. Note that the nudged coarse solution reflects the spatio-temporal behavior of the fully resolved solution but due to the spectral correction procedure exhibits the slightly lower magnitude of the coarse solution.



Figure 3: Example solution of zonal average stream function $\bar{\psi}_1$. From top to bottom: reference, free-running coarse, nudged coarse.

A.3 Loss Function

The model is trained on a semi-physics informed loss function which consists of the L2 norm of the error augmented with a second term which penalizes errors in the conservation of mass.

$$L = \sum_{j=1}^{2} \int_{0}^{2\pi} \int_{0}^{2\pi} |\psi_{j}^{ml} - \psi_{j}^{rd}|^{2} dx dy + \lambda \sum_{j=1}^{2} \int_{0}^{2\pi} \int_{0}^{2\pi} \psi_{j}^{ml} dx dy$$
(13)

Here ψ^{ml} and ψ^{rd} denote the machined learned prediction and the reference stream functions respectively. The mass conservation term is derived by noting that the two stream functions are linearly related to the height disturbances of the two layers and that by conservation of volume the integral of all height disturbances must vanish.

A.4 Additional Results

Here we compare the predicted probability distribution of a selection of the individual Fourier modes to the reference solution. The modes are parameterized by the wavenumber vector $\mathbf{k} = [k_x, k_y]$. Figure 4a shows the in-sample results ($\beta = 2.0, r_d = 0.1$) and figure 4b shows the out-of-sample results ($\beta = 1.1, r_d = 0.5$). In the interest of space we consider the barotropic stream function defined as the average of the two layers. In all cases the model prediction captures the probability distribution well, with some discrepancy in the tails observed for $\mathbf{k} = [2, 0]$ for the in-sample results and for $\mathbf{k} = [0, 1]$ and $\mathbf{k} = [2, 0]$ in the out-of-sample results.



Figure 4: Probability density function of individual Fourier modes, model prediction (red), reference (black). In sample results $\beta = 2.0$ and r = 0.1 (a), out-of-sample results $\beta = 1.1$ and r = 0.5 (b).

References

- [1] Simon Allen, Vicente Barros, Ian (Canada, Diarmid (UK, Omar Cardona, Susan Cutter, Opha Pauline Dube, Kristie Ebi, Christopher (USA, John Handmer, Padma (Australia, Allan Lavell, Katharine (USA, Michael Mastrandrea, Gordon Mcbean, Reinhard Mechler, Tom (UK, Neville Nicholls, Karen (Norway, and Thomas (USA. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. Special Report of Working Groups I and II of the Intergovernmental Panel on Climate Change. November 2012.
- [2] Trevor Houser, Solomon Hsiang, Robert Kopp, Kate Larsen, Michael Delgado, Amir Jina, Michael Mastrandrea, Shashank Mohan, Robert Muir-Wood, D. J. Rasmussen, James Rising, Paul Wilson, Karen Fisher-Vanden, Michael Greenstone, Geoffrey Heal, Michael Oppenheimer, Nicholas Stern, Bob Ward, Michael R. Bloomberg, Henry M. Paulson, and Thomas F. Steyer. *Economic Risks of Climate Change: An American Prospectus*. Columbia University Press, 2015.
- [3] Valerio Lucarini, Davide Faranda, A.C.G.M.M. Freitas, Jorge Freitas, M. Holland, Tobias Kuna, Matthew Nicol, Mike Todd, and Sandro Vaienti. *Extremes and Recurrence in Dynamical Systems*. April 2016.
- [4] Themistoklis P. Sapsis. Statistics of Extreme Events in Fluid Flows and Waves. *Annual Review of Fluid Mechanics*, 53(1):85–111, 2021. _eprint: https://doi.org/10.1146/annurev-fluid-030420-032810.
- [5] Geoffrey K. Vallis. *Atmospheric and Oceanic Fluid Dynamics: Fundamentals and Large-Scale Circulation*. Cambridge University Press, Cambridge, 2 edition, 2017.
- [6] Noah D. Brenowitz, Tom Beucler, Michael Pritchard, and Christopher S. Bretherton. Interpreting and Stabilizing Machine-Learning Parametrizations of Convection. *Journal of the Atmospheric Sciences*, 77(12):4357–4375, December 2020.
- [7] Janni Yuval and Paul A. O'Gorman. Stable machine-learning parameterization of subgrid processes for climate modeling at a range of resolutions. *Nature Communications*, 11(1):3295, July 2020.
- [8] Tom Beucler, Michael Pritchard, Stephan Rasp, Jordan Ott, Pierre Baldi, and Pierre Gentine. Enforcing Analytic Constraints in Neural Networks Emulating Physical Systems. *Physical Review Letters*, 126(9):098302, March 2021.
- [9] Ashesh Chattopadhyay, Mustafa Mustafa, Pedram Hassanzadeh, Eviatar Bach, and Karthik Kashinath. Towards physically consistent data-driven weather forecasting: Integrating data assimilation with equivariance-preserving deep spatial transformers, March 2021.
- [10] James Fulton and Ben Clarke. Towards debiasing climate simulations using unsuperviserd image-to-image translation networks. In *Climate Change AI*. Climate Change AI, December 2021.
- [11] Antoine Blanchard, Nishant Parashar, Boyko Dodov, Christian Lessig, and Themis Sapsis. A Multi-Scale Deep Learning Framework for Projecting Weather Extremes. In *Climate Change AI*. Climate Change AI, December 2022.
- [12] Zhong Yi Wan, Ricardo Baptista, Yi-fan Chen, John Anderson, Anudhyan Boral, Fei Sha, and Leonardo Zepeda-Núñez. Debias Coarsely, Sample Conditionally: Statistical Downscaling through Optimal Transport and Probabilistic Diffusion Models, May 2023. arXiv:2305.15618 [physics].
- [13] Zhong Yi Wan, Boyko Dodov, Christian Lessig, Henk Dijkstra, and Themistoklis P. Sapsis. A data-driven framework for the stochastic reconstruction of small-scale features with application to climate data sets. *Journal of Computational Physics*, 442:110484, October 2021.
- [14] Hans von Storch, Heike Langenberg, and Frauke Feser. A Spectral Nudging Technique for Dynamical Downscaling Purposes. *Monthly Weather Review*, 128(10):3664–3673, October 2000. Publisher: American Meteorological Society Section: Monthly Weather Review.

- [15] Gonzalo Miguez-Macho, Georgiy L. Stenchikov, and Alan Robock. Regional Climate Simulations over North America: Interaction of Local Processes with Improved Large-Scale Flow. *Journal of Climate*, 18(8):1227–1246, April 2005. Publisher: American Meteorological Society Section: Journal of Climate.
- [16] Jian Sun, Kai Zhang, Hui Wan, Po-Lun Ma, Qi Tang, and SHIXUAN ZHANG. Impact of Nudging Strategy on the Climate Representativeness and Hindcast Skill of Constrained EAMv1 Simulations. *Journal of Advances in Modeling Earth Systems*, 11, December 2019.
- [17] Ziyu Huang, Lei Zhong, Yaoming Ma, and Yunfei Fu. Development and evaluation of spectral nudging strategy for the simulation of summer precipitation over the Tibetan Plateau using WRF (v4.0). *Geoscientific Model Development*, 14(5):2827–2841, May 2021. Publisher: Copernicus GmbH.
- [18] Alexis-Tzianni Charalampopoulos, Shixuan Zhang, Bryce Harrop, Lai-yung Ruby Leung, and Themistoklis Sapsis. Statistics of extreme events in coarse-scale climate simulations via machine learning correction operators trained on nudged datasets, April 2023.
- [19] Di Qi and Andrew J. Majda. Predicting extreme events for passive scalar turbulence in twolayer baroclinic flows through reduced-order stochastic models. *Communications in Mathematical Sciences*, 16(1):17–51, 2018.