Low-Dimensional Reduced-Order Models for Statistical Response and Uncertainty Quantification in Turbulent Systems

Andrew J Majda
joint work with Di Qi

Morse Professor of Arts and Science
Department of Mathematics
Founder of Center for Atmosphere Ocean Science (CAOS)
Courant Institute, New York University
and
Principal Investigator, Center for Prototype Climate Modeling (CPCM)
NYU – Abu Dhabi (NYUAD)

CAOS Colloquium, October 18, 2017

Challenges for turbulent dynamical systems

Uncertainty quantification (UQ) deals with the probabilistic characterization of all the possible evolutions of a dynamical system given an *initial set of possible states* as well as *the random forcing or parameters*.

- Turbulent dynamical systems are characterized by *a large dimensional phase space* and *high degrees of internal instability*.
- Instabilities through *energy-conserving nonlinear interactions* result in a statistical steady state that is usually *non-Gaussian*.
- Accurate quantification for the statistical variability to *general external perturbations* is important in climate change sciences.

Major Task of this work:

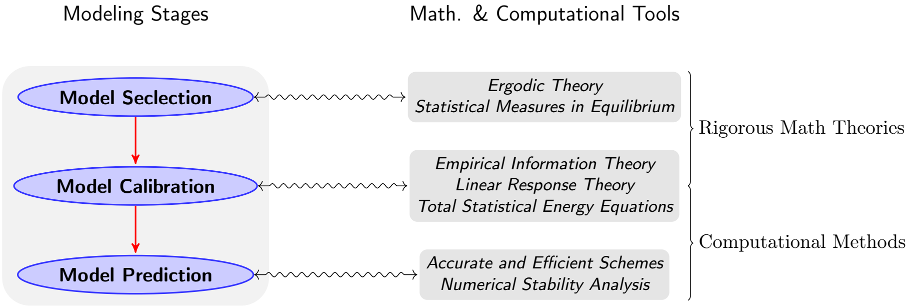
Investigate a concise systematic framework for measuring and optimizing consistency and sensitivity of imperfect dynamical models.

General framework for statistical modeling

The system setup will be a finite-dimensional system of, $\mathbf{u} \in \mathbb{R}^N$, with linear dynamics and an energy preserving quadratic part

$$\frac{d\boldsymbol{u}}{dt} = \mathscr{L}[\boldsymbol{u}] = (L+D)\boldsymbol{u} + B(\boldsymbol{u},\boldsymbol{u}) + \boldsymbol{F}(t) + \sigma_k(t)\dot{W}_k(t;\omega), \tag{1}$$

$$L^* = -L; \quad D \leq 0; \quad \mathbf{u}\cdot B(\boldsymbol{u},\boldsymbol{u}) \equiv 0.$$



Related Works and Papers

- Recent new developments
  - Majda, *Introduction to turbulent dynamical systems in complex systems*, Springer, 2016.
  - Majda and Qi, *New strategies for reduced-order models for predicting the statistical responses and uncertainty quantification in complex turbulent dynamical systems*, SIAM Review, 2017.
- Statistical theories
  - Majda, *Statistical energy conservation principle for inhomogeneous turbulent dynamical systems*. PNAS, 2015.
  - Majda and Gershgorin, *Link between statistical equilibrium fidelity and forecasting skill for complex systems with model error*. PNAS, 2011.
  - Majda and Wang, *Linear response theory for statistical ensembles in complex systems with time-periodic forcing*. CMS, 2010.
- Improving imperfect model skill
  - Majda and Qi, *Improving prediction skill of imperfect turbulent models through statistical response and information theory*, Journal of Nonlinear Science, 2015.
  - Qi and Majda, *Low-dimensional reduced-order models for statistical response and uncertainty quantification: two-layer baroclinic turbulence,* JAS, 2016.
  - Qi and Majda, *Low-dimensional reduced-order models for statistical response and uncertainty quantification: barotropic turbulence with topography,* Physica D, 2016.

Outline

1. A two-layer quasi-geostrophic model for baroclinic turbulence
   - Two-layer baroclinic turbulence in ocean and atmosphere regimes
2. Reduced-order statistical models for general turbulent systems
   - Formulation of the exact statistical moment dynamics
   - A reduced-order statistical model with consistency and sensitivity
   - Model calibration for optimal performance
3. Low-dimensional reduced-order models for the two-layer system
   - Statistical equations for the two-layer model
   - Tuning imperfect model parameters in the training phase
   - Imperfect model prediction in various dynamical regimes

Outline

1. A two-layer quasi-geostrophic model for baroclinic turbulence
   - Two-layer baroclinic turbulence in ocean and atmosphere regimes
2. Reduced-order statistical models for general turbulent systems
   - Formulation of the exact statistical moment dynamics
   - A reduced-order statistical model with consistency and sensitivity
   - Model calibration for optimal performance
3. Low-dimensional reduced-order models for the two-layer system
   - Statistical equations for the two-layer model
   - Tuning imperfect model parameters in the training phase
   - Imperfect model prediction in various dynamical regimes

The two-layer flow with forcing and dissipation

The **two-layer quasi-geostrophic model** with baroclinic instability is one simple but fully nonlinear fluid model capable in capturing the essential physics *in ocean and atmosphere science*.

Two-layer model

$$\frac{\partial q_\psi}{\partial t} + J(\psi, q_\psi) + J(\tau, q_\tau) + \beta \frac{\partial \psi}{\partial x} + U \frac{\partial \Delta \tau}{\partial x} = -\frac{\kappa}{2}\Delta(\psi - \tau) - \nu \Delta^s q_\psi + \mathscr{F}_\psi,$$

$$\frac{\partial q_\tau}{\partial t} + J(\psi, q_\tau) + J(\tau, q_\psi) + \beta \frac{\partial \tau}{\partial x} + U \frac{\partial}{\partial x}\left(\Delta \psi + k_d^2 \psi\right) = -\frac{\kappa}{2}\Delta(\tau - \psi) - \nu \Delta^s q_\tau + \mathscr{F}_\tau.$$

Barotropic and baroclinic modes:

$$q_\psi = \Delta \psi,\ \psi = \frac{1}{2}(\psi_1 + \psi_2),$$

$$q_\tau = \Delta \tau - k_d^2 \tau,\ \tau = \frac{1}{2}(\psi_1 - \psi_2).$$

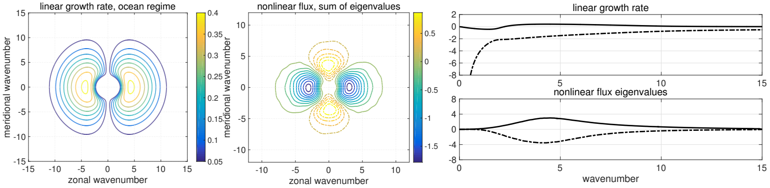
Normalized energy-consistent modes:

$$p_{\psi,\mathbf{k}} = \frac{q_{\psi,\mathbf{k}}}{|\mathbf{k}|} = -|\mathbf{k}|\,\psi_{\mathbf{k}},$$

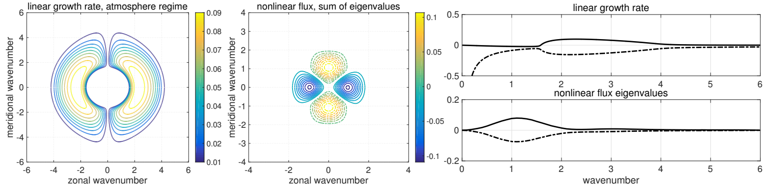
$$p_{\tau,\mathbf{k}} = \frac{q_{\tau,\mathbf{k}}}{\sqrt{|\mathbf{k}|^2 + k_d^2}} = -\sqrt{|\mathbf{k}|^2 + k_d^2}\,\tau_{\mathbf{k}}.$$

Flow in high-latitude homogeneous regimes

regime	N	β	k_d	U	κ	$(k_{\min}, k_{\max})$	$\sigma_{\max}$	$(k_x, k_y)_{\max}$
ocean, high lat.	256	10	10	1	9	$(2.25, 14.61)$	0.411	(4, 0)
atmosphere, high lat.	256	1	4	0.2	0.2	$(1.58, 6.78)$	0.099	(2, 0)

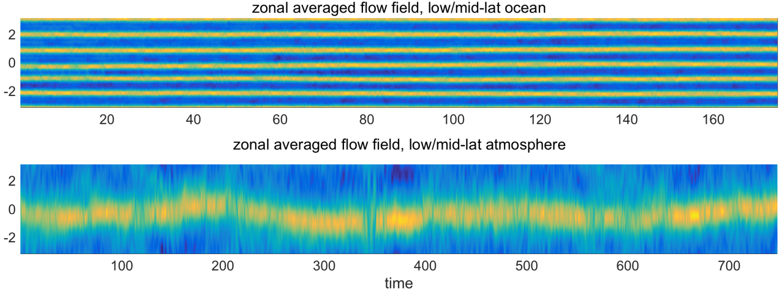


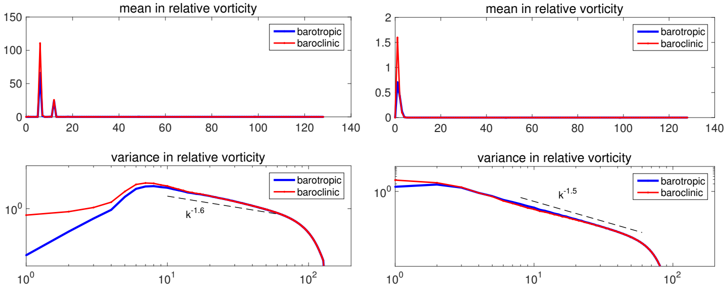
(a) high-latitude ocean regime



(b) high-latitude atmosphere regime

Flow in low/mid-latitude regimes with zonal jets





Outline

General setup of turbulent systems with quadratic nonlinearities

The system setup will be a finite-dimensional system with linear dynamics and an energy preserving quadratic part with $\boldsymbol{u} \in \mathbb{R}^N$

$$\frac{d\boldsymbol{u}}{dt} = \mathscr{L}\left[\boldsymbol{u}(t;\omega);\omega\right] = (L+D)\boldsymbol{u} + B(\boldsymbol{u},\boldsymbol{u}) + \boldsymbol{F}(t) + \sigma_k(t)\dot{W}_k(t;\omega), \tag{2}$$

$$\boldsymbol{u}(t_0;\omega) = \boldsymbol{u}_0(\omega). \tag{3}$$

- L being a **skew-symmetric linear operator** $L^* = -L$, representing the β-effect of Earth's curvature, topography etc.
- D being a **negative definite symmetric operator** $D^* = D$, representing dissipative processes such as surface drag, radiative damping, viscosity etc.
- $B(\boldsymbol{u},\boldsymbol{u})$ being a **quadratic operator** which conserves the energy by itself so that it satisfies $B(\boldsymbol{u},\boldsymbol{u})\cdot\boldsymbol{u} = 0$.
- $\boldsymbol{F}(t) + \sigma_k(t)\dot{W}_k(t;\omega)$ being the **effects of external forcing**, i.e. solar forcing, seasonal cycle, which can be split into a mean component $\boldsymbol{F}(t)$ and a stochastic component with white noise characteristics.

Exact statistical moment equations

Statistical mean and covariance dynamics, $\boldsymbol{u} = \bar{\boldsymbol{u}} + Z_i \boldsymbol{v}_i$, $R_{ij} = \left\langle Z_i Z_j^* \right\rangle$,

$$\frac{d\bar{\boldsymbol{u}}}{dt} = (L+D)\,\bar{\boldsymbol{u}} + B\left(\bar{\boldsymbol{u}},\bar{\boldsymbol{u}}\right) + R_{ij}B\left(\boldsymbol{v}_i,\boldsymbol{v}_j\right) + \boldsymbol{F}(t),$$

$$\frac{dR}{dt} = L_v R + R L_v^* + Q_F + Q_\sigma.$$

- the linear dynamics operator L_v expressing energy transfers between the *mean field and the stochastic modes* (B), as well as *energy dissipation* (D), and *non-normal dynamics* (L)

$$\left\{L_v\right\}_{ij} = \left[(L+D)\,\boldsymbol{v}_j + B\left(\bar{\boldsymbol{u}},\boldsymbol{v}_j\right) + B\left(\boldsymbol{v}_j,\bar{\boldsymbol{u}}\right)\right]\cdot \boldsymbol{v}_i.$$

- the positive definite operator Q_σ expressing energy transfer due to external stochastic forcing

$$\left\{Q_\sigma\right\}_{ij} = \boldsymbol{v}_i^* \sigma_k^* \sigma_k \boldsymbol{v}_j.$$

- the third-order moments expressing the energy flux between different modes due to non-linear terms

$$Q_F = \left\langle Z_m Z_n Z_j \right\rangle B\left(\boldsymbol{v}_m,\boldsymbol{v}_n\right)\cdot \boldsymbol{v}_i + \left\langle Z_m Z_n Z_i \right\rangle B\left(\boldsymbol{v}_m,\boldsymbol{v}_n\right)\cdot \boldsymbol{v}_j.$$

  note that energy is still conserved in this nonlinear interaction part

$$\begin{aligned}\mathrm{Tr}\left[Q_F\right] &= 2\left\langle Z_m Z_n Z_i \right\rangle B\left(\boldsymbol{v}_m,\boldsymbol{v}_n\right)\cdot \boldsymbol{v}_i \\ &= 2\left\langle B\left(Z_m\boldsymbol{v}_m, Z_n\boldsymbol{v}_n\right)\cdot Z_i\boldsymbol{v}_i\right\rangle = 2\left\langle B\left(\boldsymbol{u}',\boldsymbol{u}'\right)\cdot \boldsymbol{u}'\right\rangle = 0.\end{aligned}$$

Reduced-Order Statistical Energy Closure

The true statistical model

$$\frac{d\bar{\boldsymbol{u}}}{dt} = (L+D)\,\bar{\boldsymbol{u}} + B(\bar{\boldsymbol{u}},\bar{\boldsymbol{u}}) + R_{ij}B\left(\boldsymbol{v}_i,\boldsymbol{v}_j\right) + \boldsymbol{F}(t),\ \bar{\mathbf{u}} \in \mathbb{R}^N,$$

$$\frac{dR}{dt} = L_v(\bar{\mathbf{u}})\,R + RL_v^*(\bar{\mathbf{u}}) + Q_F + Q_\sigma, \quad R \in \mathbb{R}^{N\times N}.$$

$$\boxed{Q_{F,ij} = \left\langle Z_m Z_n Z_j\right\rangle B(\boldsymbol{v}_m,\boldsymbol{v}_n)\cdot\boldsymbol{v}_i + \left\langle Z_m Z_n Z_i\right\rangle B(\boldsymbol{v}_m,\boldsymbol{v}_n)\cdot\boldsymbol{v}_j}$$

Reduced-Order Statistical Energy Closure

The reduced-order approximation $\bar{\mathbf{u}}_M \in \mathbb{R}^M$, $M \ll N$

$$\begin{aligned}
\frac{d\bar{\mathbf{u}}_M}{dt} &= (L+D)\,\bar{\mathbf{u}}_M + B\left(\bar{\mathbf{u}}_M, \bar{\mathbf{u}}_M\right) + R_{M,ij}B\left(\mathbf{v}_i, \mathbf{v}_j\right) + \mathbf{F}, \\
\frac{dR_M}{dt} &= L_v R_M + R_M L_v^* + Q_F^M + Q_\sigma,
\end{aligned}$$

A preferred approach for the nonlinear flux Q_F^M combining both the *detailed model energy mechanism* and *control over model sensitivity* is proposed

$$\boxed{Q_F^M = Q_F^{M,-} + Q_F^{M,+} = f_1(E)\left[-\left(N_{M,\mathrm{eq}} + d_M I_N\right) R_M\right] + f_2(E)\left[Q_{F,\mathrm{eq}}^{+} + \Sigma_M\right].}$$

Reduced-Order Statistical Energy Closure

A preferred approach for the nonlinear flux Q_F^M combining both the *detailed model energy mechanism* and *control over model sensitivity* is proposed

$$Q_F^M = Q_F^{M,-} + Q_F^{M,+} = f_1(E)\left[-\left(N_{M,\mathrm{eq}} + d_M I_N\right) R_M\right] + f_2(E)\left[Q_{F,\mathrm{eq}}^{+} + \Sigma_M\right].$$

- *Higher-order corrections from equilibrium statistics*:

$$Q_{F,\mathrm{eq}} = Q_{F,\mathrm{eq}}^{-} + Q_{F,\mathrm{eq}}^{+} = -L_v(\bar{\mathbf{u}}_{\mathrm{eq}}) R_{\mathrm{eq}} - R_{\mathrm{eq}} L_v^{*}(\bar{\mathbf{u}}_{\mathrm{eq}}) - Q_\sigma, \quad N_{M,\mathrm{eq}} = \frac{1}{2} Q_{F,\mathrm{eq}}^{-} R_{\mathrm{eq}}^{-1}.$$

- *Additional damping and noise to model nonlinear flux*:

$$Q_M^{\mathrm{add}} = -d_M R_M + \Sigma_M.$$

- *Statistical energy-consistent scaling to improve model sensitivity*:

$$f_1(E) = \left(\frac{E}{E_{\mathrm{eq}}}\right)^{1/2}, \quad f_2(E) = \left(\frac{E}{E_{\mathrm{eq}}}\right)^{3/2}.$$

Climate fidelity for equilibrium

Equilibrium fidelity refers to the convergence to the same final unperturbed statistical equilibrium R_{eq} in the reduced-order models R_M in each resolved component.

Specifically, it requires that the model nonlinear flux correction term Q_M converges to the truth, $Q_M \to Q_{F,\mathrm{eq}}$, when no external perturbation is added

$$\frac{dR_{M,\mathrm{eq}}}{dt} = 0 = L_v\left(\bar{\mathbf{u}}_{\mathrm{eq}}\right) R_{M,\mathrm{eq}} + R_{M,\mathrm{eq}} L_v^*\left(\bar{\mathbf{u}}_{\mathrm{eq}}\right) + Q_{F,\mathrm{eq}}^M + Q_\sigma \ \to \ R_{M,\mathrm{eq}} = R_{\mathrm{eq}}.$$

- the first component $\left(N_{M,\mathrm{eq}}, Q_{F,\mathrm{eq}}^+\right)$ comes from the true equilibrium statistics.
- climate consistency requires the second component correction makes no contribution in the unperturbed case

$$\Sigma_M = \frac{1}{2} d_M R_{\mathrm{eq}}, \quad f_1\left(E_{\mathrm{eq}}\right) = 1,\ f_2\left(E_{\mathrm{eq}}\right) = 1.$$

Statistical energy conservation principle[1]

Theorem

*(**Statistical Energy Conservation Principle**) Under the structural symmetries on the basis $\mathbf{v}_i$, for any turbulent dynamical systems in the form (1) the total statistical energy, $E = \bar{E} + E' = \frac{1}{2}\bar{\mathbf{u}}\cdot\bar{\mathbf{u}} + \frac{1}{2}\mathrm{tr}R$, satisfies*

$$\frac{dE}{dt} = \bar{\mathbf{u}}\cdot D\bar{\mathbf{u}} + \bar{\mathbf{u}}\cdot\mathbf{F} + \mathrm{tr}\left(DR\right) + \frac{1}{2}\mathrm{tr}Q_\sigma,$$

where R satisfies the exact covariance equation.

Corollary

Under the assumption of the Theorem, assume $D = -dI$, with $d > 0$, then the turbulent dynamical system satisfies the closed statistical energy equation for $E = \frac{1}{2}\bar{\mathbf{u}}\cdot\bar{\mathbf{u}} + \frac{1}{2}\mathrm{tr}R$,

$$\frac{dE}{dt} = -2dE + \bar{\mathbf{u}}\cdot\mathbf{F} + \frac{1}{2}\mathrm{tr}Q_\sigma.$$

[1] Majda, *Statistical energy conservation principle for inhomogeneous turbulent dynamical systems*, PNAS, 2015.

Model calibration blending *statistical response* and *information theory*

Accurate modeling about the model sensitivity to various external perturbations requires the imperfect reduced-order models to correctly reflect the true system's "memory" about its previous states.

- *the linear response operator* can characterize the model sensitivity involving the nonlinear effects in the system regardless of the specific forms of the external perturbations.
- *empirical information theory* can be used as the distance between these two operators to calculate the unbiased and invariant measure for model distributions.

Linear response operator $\mathscr{R}_A(t)$

Linear response theory: The linear response of a system, $\mathbf{u}_t = \mathbf{f}(\mathbf{u})$, $\mathbf{f}^\delta = \mathbf{f} + \delta\mathbf{f}'(t)$, can be predicted by observing appropriate statistics of the system in equilibrium π_{eq}

$$\mathbb{E}^\delta A(\mathbf{u}) = E_{\mathrm{eq}}(A) + \delta E'_A + O\left(\delta^2\right),$$

$$\delta E'_A = \int_0^t \mathscr{R}_A(t-s)\delta f'(s)\,ds.$$

without the need of applying any perturbations.

kicked response: For δ small enough, the linear response operator $\mathscr{R}_A(t)$ can be calculated by solving the unperturbed system with a perturbed initial distribution

$$\pi\mid_{t=0} = \pi_{\mathrm{eq}}(\mathbf{u} - \delta\mathbf{u}) = \pi_{\mathrm{eq}} - \delta\mathbf{u}\cdot\nabla\pi_{\mathrm{eq}} + O\left(\delta^2\right).$$

$$\delta\mathscr{R}_A(t) \equiv \delta\mathbf{u}\cdot\mathscr{R}_A = \int A(\mathbf{u})\,\delta\pi' + O\left(\delta^2\right).$$

Link between equilibrium fidelity and forecasting skill

Given the optimal model for the unperturbed climate π_{eq}, how can we assess the error in the climate change prediction

$$\mathscr{P}\left(\pi,\pi^M\right)=\int\pi\ln\frac{\pi}{\pi^M},$$

based on the unperturbed climate?

Under assumptions with diagonal covariance matrices $R=\mathrm{diag}\left(R_k\right)$ and equilibrium model fidelity $\mathscr{P}\left(\pi_G,\pi_G^M\right)=0$

$$\begin{aligned}\mathscr{P}\left(\pi_\delta,\pi_\delta^M\right) &= \mathscr{S}\left(\pi_{G,\delta}\right)-\mathscr{S}\left(\pi_\delta\right)\\ &+\frac{1}{2}\sum_k\left(\delta\bar{u}_k-\delta\bar{u}_{M,k}\right)R_k^{-1}\left(\delta\bar{u}_k-\delta\bar{u}_{M,k}\right)\\ &+\frac{1}{4}\sum_k R_k^{-2}\left(\delta R_k-\delta R_{M,k}\right)^2+O\left(\delta^3\right).\end{aligned}$$

R_k is the equilibrium variance in k-th component, and $\delta\bar{u}_k$ and δR_k are the linear response operators for the mean and variance in k-th component. (Majda & Gershogorin, PNAS, 2011)

Optimization in the training phase

To summarize, consider a class of imperfect models, $\mathscr{M}$. The *optimal model* $M^* \in \mathscr{M}$ that ensures best information consistent responses is characterized with the smallest additional information in the linear response operator $\mathscr{R}_A$, such that

$$\left\| \mathscr{P}\left(p_\delta, p_\delta^{M^*}\right) \right\|_{L^1([0,T])} = \min_{M \in \mathscr{M}} \left\| \mathscr{P}\left(p_\delta, p_\delta^{M}\right) \right\|_{L^1([0,T])},$$

- p_δ^M can be achieved through a kicked response procedure in the training phase compared with the actual observed data p_δ in nature;
- the information distance between perturbed responses $\mathscr{P}\left(p_\delta, p_\delta^M\right)$ can be calculated through the expansion formula;
- the information distance $\mathscr{P}\left(p_\delta(t), p_\delta^M(t)\right)$ is measured at each time instant, so the entire error is averaged under the L^1-norm inside a time window $[0,T]$.

Outline

1. A two-layer quasi-geostrophic model for baroclinic turbulence
   - Two-layer baroclinic turbulence in ocean and atmosphere regimes
2. Reduced-order statistical models for general turbulent systems
   - Formulation of the exact statistical moment dynamics
   - A reduced-order statistical model with consistency and sensitivity
   - Model calibration for optimal performance
3. Low-dimensional reduced-order models for the two-layer system
   - Statistical equations for the two-layer model
   - Tuning imperfect model parameters in the training phase
   - Imperfect model prediction in various dynamical regimes

Exact statistical moment equations for the two-layer model

The rescaled set of equations of (1) can be summarized in the abstract form

$$\frac{d\mathbf{p_k}}{dt} = B_{\mathbf{k}}(\mathbf{p_k},\mathbf{p_k}) + (\mathscr{L}_{\mathbf{k}} - \mathscr{D}_{\mathbf{k}})\mathbf{p_k} + \mathscr{F}_{\mathbf{k}}, \quad \mathbf{p_k} = (p_{\psi,\mathbf{k}}, p_{\tau,\mathbf{k}})^T, \quad \sum_{\mathbf{k}} \mathbf{p_k} \cdot B_{\mathbf{k}}(\mathbf{p_k},\mathbf{p_k}) \equiv 0,$$

where the normalized state variable $\mathbf{p_k} = (p_{\psi,\mathbf{k}}, p_{\tau,\mathbf{k}})^T$ is in barotropic and baroclinic mode, the linear operator is decomposed into non-symmetric part $\mathscr{L}_{\mathbf{k}}$ involving β-effect and shear flow U and dissipation part $\mathscr{D}_{\mathbf{k}}$, together with the forcing $\mathscr{F}_{\mathbf{k}}$ combining deterministic component and stochastic component.

$$\mathscr{L}_{\mathbf{k}} = \begin{bmatrix} \frac{ik_x\beta}{|\mathbf{k}|^2} & -\frac{ik_xU}{\sqrt{1+(k_d/|\mathbf{k}|)^2}} \\ -ik_xU\frac{1-(k_d/|\mathbf{k}|)^2}{\sqrt{1+(k_d/|\mathbf{k}|)^2}} & \frac{ik_x\beta}{|\mathbf{k}|^2+k_d^2} \end{bmatrix}, \quad \mathscr{D}_{\mathbf{k}} = \frac{\kappa}{2}\begin{bmatrix} -1 & \frac{1}{\sqrt{1+(k_d/|\mathbf{k}|)^2}} \\ \frac{1}{\sqrt{1+(k_d/|\mathbf{k}|)^2}} & -\frac{1}{1+(k_d/|\mathbf{k}|)^2} \end{bmatrix},$$

$$B_{\mathbf{k}}(\mathbf{p_k},\mathbf{p_k}) = \begin{bmatrix} B_{\psi,\mathbf{k}} \\ B_{\tau,\mathbf{k}} \end{bmatrix} = \begin{bmatrix} \sum_{\mathbf{m}+\mathbf{n}=\mathbf{k}} \frac{\mathbf{m}^\perp\cdot\mathbf{n}}{|\mathbf{k}|}\left(\frac{|\mathbf{n}|}{|\mathbf{m}|}p_{\psi,\mathbf{m}}p_{\psi,\mathbf{n}} + \sqrt{\frac{|\mathbf{n}|^2+k_d^2}{|\mathbf{m}|^2+k_d^2}}p_{\tau,\mathbf{m}}p_{\tau,\mathbf{n}}\right) \\ \sum_{\mathbf{m}+\mathbf{n}=\mathbf{k}} \frac{\mathbf{m}^\perp\cdot\mathbf{n}}{\sqrt{|\mathbf{k}|^2+k_d^2}}\left(\frac{\sqrt{|\mathbf{n}|^2+k_d^2}}{|\mathbf{m}|}p_{\psi,\mathbf{m}}p_{\tau,\mathbf{n}} + \frac{|\mathbf{n}|}{\sqrt{|\mathbf{m}|^2+k_d^2}}p_{\tau,\mathbf{m}}p_{\psi,\mathbf{n}}\right) \end{bmatrix}.$$

Exact statistical moment equations

Statistical energy in each spectral mode

$$R_{\mathbf{k}} = \overline{\mathbf{p_k}^* \mathbf{p_k}} = \begin{bmatrix} \overline{\left|p_{\psi,\mathbf{k}}\right|^2} & \overline{p^*_{\psi,\mathbf{k}} p_{\tau,\mathbf{k}}} \\ \overline{p_{\psi,\mathbf{k}} p^*_{\tau,\mathbf{k}}} & \overline{\left|p_{\tau,\mathbf{k}}\right|^2} \end{bmatrix}, \quad \overline{p^*_{1,\mathbf{k}} p_{2,\mathbf{k}}} = \bar{p}^*_{1,\mathbf{k}} \bar{p}_{2,\mathbf{k}} + \overline{p'^*_{1,\mathbf{k}} p'_{2,\mathbf{k}}}.$$

$R_{\mathbf{k}}$ combines the variability in both *mean* and *variance*. The true statistical dynamical equations form a 2×2 system about $R_{\mathbf{k}} \in \mathbb{C}^{2\times 2}$

$$\boxed{\frac{dR_{\mathbf{k}}}{dt} = \left(\mathscr{L}_{\mathbf{k}} - \mathscr{D}_{\mathbf{k}}\right) R_{\mathbf{k}} + Q_{F,\mathbf{k}} + Q_{\sigma,\mathbf{k}} + c.c., \quad |\mathbf{k}| \leq N,}$$

$$Q_{F,\mathbf{k}} = \overline{\mathbf{p_k}^* B_{\mathbf{k}}\left(\mathbf{p_k}, \mathbf{p_k}\right)} = \begin{bmatrix} \overline{\mathbf{p}^*_{\psi,\mathbf{k}} B_{\psi,\mathbf{k}}} & \overline{\mathbf{p}^*_{\psi,\mathbf{k}} B_{\tau,\mathbf{k}}} \\ \overline{\mathbf{p}^*_{\tau,\mathbf{k}} B_{\psi,\mathbf{k}}} & \overline{\mathbf{p}^*_{\tau,\mathbf{k}} B_{\tau,\mathbf{k}}} \end{bmatrix}, \quad \sum_{\mathbf{k}} \mathrm{tr} Q_{F,\mathrm{k}} \equiv 0.$$

Statistical energy conservation principle

The total statistical energy dynamical equation concerns the evolution of the *total variability in mean and variance* in response to external perturbations

$$E=\frac{1}{2}\sum_{1\leq|\mathbf{k}|\leq N}|\mathbf{k}|^2\overline{|\psi_{\mathbf{k}}|^2}+\left(|\mathbf{k}|^2+k_d^2\right)\overline{|\tau_{\mathbf{k}}|^2}=\frac{1}{2}\sum_{1\leq|\mathbf{k}|\leq N}\overline{\left|p_{\psi,\mathbf{k}}\right|^2}+\overline{\left|p_{\tau,\mathbf{k}}\right|^2}.$$

The exact dynamics for the statistical energy can be derived as

$$\frac{dE}{dt}+H_f=-\kappa E+\frac{\kappa}{2}F-\nu H+Q_\sigma.$$

H_f is the meridional heat flux due to baroclinic instability, F is the additional damping effects due to the non-symmetry in Ekman drag only applied on the bottom layer

$$H_f=k_d^2U\int\overline{\psi_x\tau}=k_d^2U\sum ik_x\overline{\psi_{\mathbf{k}}^*\tau_{\mathbf{k}}},\quad F=\sum k_d^2\overline{|\tau_{\mathbf{k}}|^2}+2\,|\mathbf{k}|^2\,\mathfrak{Re}\,\overline{\psi_{\mathbf{k}}^*\tau_{\mathbf{k}}}.$$

Set-up for the numerical problem

- The true statistics is calculated by a pseudo-spectra code with 128 spectral modes zonally and meridionally, corresponding to $256 \times 256 \times 2$ grid points in total.
- In the reduced-order methods, only the large-scale modes $|\mathbf{k}| \leq 10$ are resolved, which is about 0.15% of the full model resolution.

External forcing in stochastic and deterministic component:

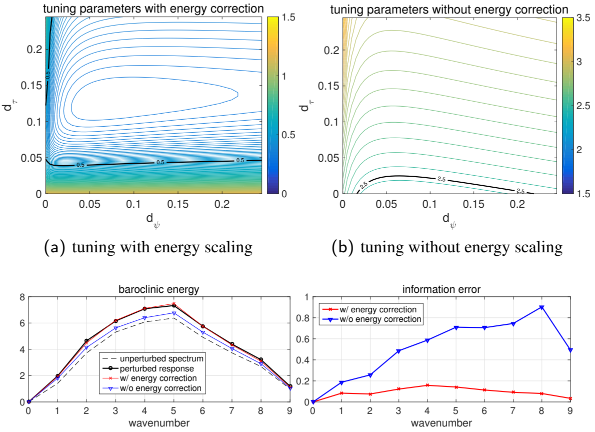
- The amplitude of the *stochastic forcing* $\sigma_k \dot{W}_k$ is introduced according to the equilibrium energy so that

$$\sigma^2_{\psi,\mathbf{k}} = \delta\sigma_0^2 \overline{\left|q_{\psi,\mathbf{k}}\right|^2}_{\mathrm{eq}}, \quad \sigma^2_{\tau,\mathbf{k}} = \delta\sigma_0^2 \overline{\left|q_{\tau,\mathbf{k}}\right|^2}_{\mathrm{eq}}.$$

- The *deterministic forcing* is introduced through a perturbation in the background shear $U_\delta = U + \delta U$

$$\delta f_{\psi,\mathbf{k}} = \delta U i k_x \left(-\left|\mathbf{k}\right|^2\right) \tau_{\mathbf{k}}, \quad \delta f_{\tau,\mathbf{k}} = \delta U i k_x \left(-\left|\mathbf{k}\right|^2 + k_d^2\right) \psi_{\mathbf{k}}.$$

Tuning parameters in the training phase



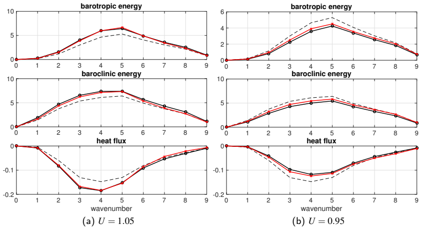
(a) tuning with energy scaling (b) tuning without energy scaling

(c) prediction and info. errors

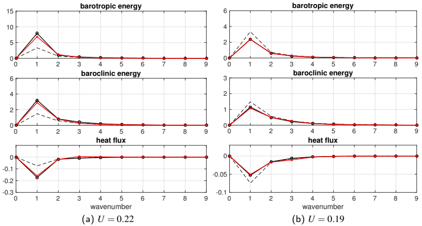
Figure: Tuning imperfect model parameters in the training phase. The information errors with varying model parameters, $d_M = (d_\psi, d_\tau)$, are plotted for stochastic barotropic perturbation case.

High-latitude: Mean shear flow perturbation

- The model is perturbed by changing the background zonal flow strength U;
- The entire spectral is perturbed due to the mean flow advection in each spectral mode.



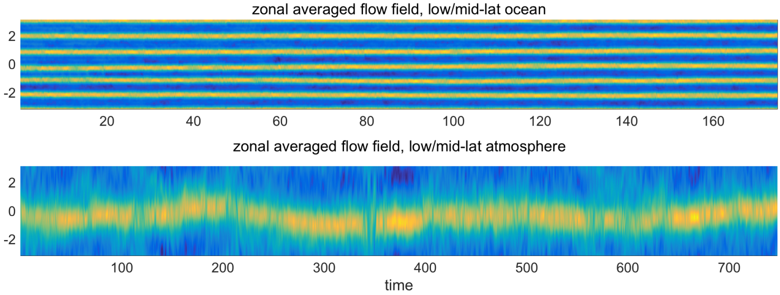
(a) ocean regime $\delta U = \pm 0.05$, $U_0 = 1$



(b) atmos. regime $\delta U = \pm 0.02$, $U_0 = 0.2$

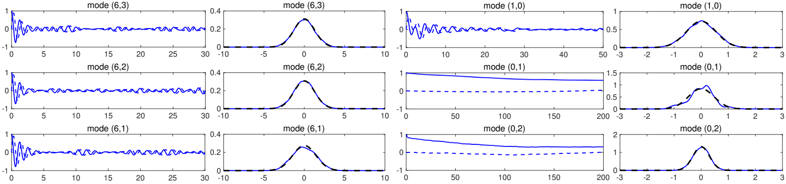
Flow in low-latitude regimes with zonal jets

regime	N	β	k_d	U	κ	$(k_{\min}, k_{\max})$	$\sigma_{\max}$	$(k_x, k_y)_{\max}$
ocean, high lat.	256	100	10	1	1	$(7.14, 15.63)$	0.104	(2, 8)
atmosphere, high lat.	256	2.5	4	0.2	0.05	$(2.51, 7.06)$	0.053	(3, 0)

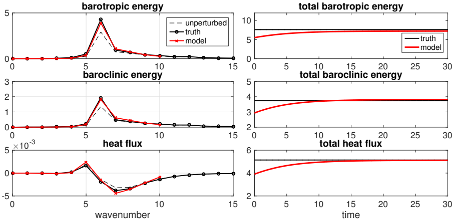


Low-latitude: Stochastic perturbation with $\delta\sigma_0^2 = 0.2$

autocorrelation functions and probability density functions

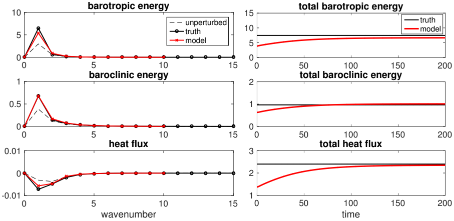


(a) ocean (b) atmosphere



(a) spectra for the energy and heat flux (b) time-series for total responses

(c) ocean regime



(a) spectra for the energy and heat flux (b) time-series for total responses

(d) atmosphere regime

Current and future work

- Reduced-order stochastic modeling strategies to capture passive scalar intermittency
  - Majda & Tong, *Intermittency in turbulent diffusion models with a mean gradient*, Nonlinearity, 2015
  - Majda & Gershgorin, *Elementary models for turbulent diffusion with complex physical features: eddy diffusivity, spectrum and intermittency*, Phil. Trans. R. Soc. A, 2013
  - Qi & Majda, *Predicting fat-tailed intermittent probability distributions in passive scalar turbulence with imperfect models through empirical information theory*, Comm. Math. Sci, 2015.
  - Qi & Majda, *Predicting Extreme Events for Passive Scalar Turbulence in Two-Layer Baroclinic Flows through Reduced-Order Stochastic Models*, Comm. Math. Sci, 2017.
- Design of a mitigation control strategy by using novel low-order statistical models;
  - Majda & Qi, *Effective control of complex turbulent dynamical systems through statistical functionals,* PNAS, 2017
  - Majda & Qi, *Using Statistical Functionals for Effective Control of Inhomogeneous Complex Turbulent Dynamical Systems*, submitted to Physica D.
- Rigorous statistical UQ for turbulent geophysical flows
  - Majda & Qi, Rigorous statistical uncertainty quantification for one-layer turbulent geophysical flows, in preparation.

Related Works and Papers

- Recent new developments
  - Majda, *Introduction to turbulent dynamical systems in complex systems*, Springer, 2016.
  - Majda and Qi, *New strategies for reduced-order models for predicting the statistical responses and uncertainty quantification in complex turbulent dynamical systems*, SIAM Review, 2017.
- Statistical theories
  - Majda, *Statistical energy conservation principle for inhomogeneous turbulent dynamical systems*. PNAS, 2015.
  - Majda and Gershgorin, *Link between statistical equilibrium fidelity and forecasting skill for complex systems with model error*. PNAS, 2011.
  - Majda and Wang, *Linear response theory for statistical ensembles in complex systems with time-periodic forcing*. CMS, 2010.
- Improving imperfect model skill
  - Majda and Qi, *Improving prediction skill of imperfect turbulent models through statistical response and information theory*, Journal of Nonlinear Science, 2015.
  - Qi and Majda, *Low-dimensional reduced-order models for statistical response and uncertainty quantification: two-layer baroclinic turbulence,* JAS, 2016.
  - Qi and Majda, *Low-dimensional reduced-order models for statistical response and uncertainty quantification: barotropic turbulence with topography,* Physica D, 2016.