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
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, $u \in \mathbb{R}^N$, with linear dynamics and an energy preserving quadratic part

$$\frac{du}{dt} = \mathcal{L}[u] = (L + D)u + B(u, u) + F(t) + \sigma_k(t) \dot{W}_k(t; \omega),$$

subject to $L^* = -L; D \leq 0; u \cdot B(u, u) \equiv 0$. 

---

Modeling Stages

**Model Selection**

- Ergodic Theory
- Statistical Measures in Equilibrium

**Model Calibration**

- Empirical Information Theory
- Linear Response Theory
- Total Statistical Energy Equations

**Model Prediction**

- Accurate and Efficient Schemes
- Numerical Stability Analysis

---

**Math. & Computational Tools**

- Rigorous Math Theories
- Computational Methods
Related Works and Papers

- **Recent new developments**

- **Statistical theories**

- **Improving imperfect model skill**
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

\[
\frac{\partial q_\psi}{\partial t} + J(\psi, q_\psi) + J(\tau, q_\tau) + \beta \frac{\partial}{\partial x} \left( \frac{\partial \Delta \tau}{\partial x} \right) = -\frac{\kappa}{2} \Delta (\psi - \tau) - \nu \Delta^s q_\psi + \mathcal{F}_\psi,
\]

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

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:

- \( p_\psi,k = \frac{q_\psi,k}{|k|} = -|k| \psi_k \),
- \( p_\tau,k = \frac{q_\tau,k}{\sqrt{|k|^2 + k_d^2 \tau_k}} = -\sqrt{|k|^2 + k_d^2} \tau_k \).
Flow in high-latitude homogeneous regimes

<table>
<thead>
<tr>
<th>regime</th>
<th>$N$</th>
<th>$\beta$</th>
<th>$k_d$</th>
<th>$U$</th>
<th>$\kappa$</th>
<th>$(k_{\text{min}}, k_{\text{max}})$</th>
<th>$\sigma_{\text{max}}$</th>
<th>$(k_x, k_y)_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ocean, high lat.</td>
<td>256</td>
<td>10</td>
<td>10</td>
<td>1</td>
<td>9</td>
<td>(2.25, 14.61)</td>
<td>0.411</td>
<td>(4, 0)</td>
</tr>
<tr>
<td>atmosphere, high lat.</td>
<td>256</td>
<td>1</td>
<td>4</td>
<td>0.2</td>
<td>0.2</td>
<td>(1.58, 6.78)</td>
<td>0.099</td>
<td>(2, 0)</td>
</tr>
</tbody>
</table>

(a) high-latitude ocean regime

(b) high-latitude atmosphere regime
Flow in low/mid-latitude regimes with zonal jets

Fig. 5.4: Time-averaged statistics (in radial average) in mean and second-order moments in low/mid-latitude regime. The first row compares the statistical mean states. The following two rows show the variances, and statistical energy, in barotropic and baroclinic modes, as well as the potential energy.

Fig. 5.5: Autocorrelation functions and the probability distribution functions in low/mid-latitude ocean and atmosphere regime. The first three most energetic baroclinic modes are displayed. In the autocorrelations, the solid lines show the real part while the dashed lines are the imaginary part of the functions. In the pdfs, the corresponding Gaussian distributions with the same variance are also plotted in dashed black lines.
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
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 \( u \in \mathbb{R}^N \)

\[
\frac{du}{dt} = \mathcal{L} [u(t; \omega); \omega] = (L + D)u + B(u, u) + F(t) + \sigma_k(t) \dot{W}_k(t; \omega), \quad (2)
\]

\[
u(t_0; \omega) = u_0(\omega). \quad (3)
\]

- \( L \) being a **skew-symmetric linear operator** \( L^* = -L \), representing the \( \beta \)-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.
- \( B(u, u) \) being a **quadratic operator** which conserves the energy by itself so that it satisfies \( B(u, u) \cdot u = 0 \).
- \( F(t) + \sigma_k(t) \dot{W}_k(t; \omega) \) being the **effects of external forcing**, i.e. solar forcing, seasonal cycle, which can be split into a mean component \( F(t) \) and a stochastic component with white noise characteristics.
Exact statistical moment equations
Statistical mean and covariance dynamics, \( \mathbf{u} = \bar{\mathbf{u}} + Z_i \mathbf{v}_i, R_{ij} = \left\langle Z_i Z_j^* \right\rangle, \)
\[
\frac{d\bar{\mathbf{u}}}{dt} = (L + D) \bar{\mathbf{u}} + B (\bar{\mathbf{u}}, \bar{\mathbf{u}}) + R_{ij} B (\mathbf{v}_i, \mathbf{v}_j) + \mathbf{F} (t),
\]
\[
\frac{dR}{dt} = L \mathbf{v} R + R L^*_\mathbf{v} + Q_F + Q_\sigma.
\]

- the linear dynamics operator \( L \mathbf{v} \) expressing energy transfers between the \textit{mean field and the stochastic modes} (\( B \)), as well as energy \textit{dissipation} (\( D \)), and \textit{non-normal dynamics} (\( L \))

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

- the positive definite operator \( Q_\sigma \) expressing energy transfer due to external stochastic forcing

\[
\{Q_\sigma\}_{ij} = \mathbf{v}_i^* \sigma_k^* \sigma_k \mathbf{v}_j.
\]

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

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

\textit{Note} that energy is still conserved in this nonlinear interaction part

\[
\text{Tr} [Q_F] = 2 \left\langle Z_m Z_n Z_i \right\rangle B (\mathbf{v}_m, \mathbf{v}_n) \cdot \mathbf{v}_i = 2 \left\langle B (Z_m \mathbf{v}_m, Z_n \mathbf{v}_n) \cdot Z_i \mathbf{v}_i \right\rangle = 2 \left\langle B (\mathbf{u}', \mathbf{u}') \cdot \mathbf{u}' \right\rangle = 0.
\]
Reduced-Order Statistical Energy Closure

The true statistical model

\[
\frac{d\bar{u}}{dt} = (L + D)\bar{u} + B(\bar{u}, \bar{u}) + R_{ij} B(v_i, v_j) + F(t), \quad \bar{u} \in \mathbb{R}^N,
\]

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

\[
Q_{F,ij} = \langle Z_m Z_n Z_j \rangle B(v_m, v_n) \cdot v_i + \langle Z_m Z_n Z_i \rangle B(v_m, v_n) \cdot v_j
\]
Reduced-Order Statistical Energy Closure

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

$$\frac{d\tilde{u}_M}{dt} = (L + D)\tilde{u}_M + B(\tilde{u}_M, \tilde{u}_M) + R_{M,ij}B(v_i, v_j) + F,$$

$$\frac{dR_M}{dt} = L_vR_M + R_ML_v^* + Q^M_F + Q^\sigma,$$

A preferred approach for the nonlinear flux $Q^M_F$ combining both the detailed model energy mechanism and control over model sensitivity is proposed

$$Q^M_F = Q^{M,-}_F + Q^{M,+}_F = f_1(E) \left[ -(N_{M,eq} + d_M I_N) R_M \right] + f_2(E) \left[ Q^+_{F,eq} + \Sigma_M \right].$$
Reduced-Order Statistical Energy Closure

A preferred approach for the nonlinear flux $Q^M_F$ combining both the detailed model energy mechanism and control over model sensitivity is proposed

\[
Q^M_F = Q^{M,-}_F + Q^{M,+}_F = f_1(E) \left[ - \left( N_{M,eq} + d_M I N \right) R_M \right] + f_2(E) \left[ Q^+_{F,eq} + \Sigma_M \right].
\]

- **Higher-order corrections from equilibrium statistics:**

  \[
  Q_{F,eq} = Q^{-}_{F,eq} + Q^{+}_{F,eq} = -L_v(\bar{u}_{eq}) R_{eq} - R_{eq} L^*_v(\bar{u}_{eq}) - Q_\sigma, \quad N_{M,eq} = \frac{1}{2} Q^{-}_{F,eq} R_{eq}^{-1}.
  \]

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

  \[
  Q^{\text{add}}_M = -d_M R_M + \Sigma_M.
  \]

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

  \[
  f_1(E) = \left( \frac{E}{E_{eq}} \right)^{1/2}, \quad f_2(E) = \left( \frac{E}{E_{eq}} \right)^{3/2}.
  \]
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, $Q_M \to Q_{F,eq}$, when no external perturbation is added

$$\frac{dR_{M,eq}}{dt} = 0 = L_v(\tilde{u}_{eq}) R_{M,eq} + R_{M,eq} L^*(\tilde{u}_{eq}) + Q_{M,eq} + Q_{\sigma} \to R_{M,eq} = R_{eq}.$$

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

$$\Sigma_M = \frac{1}{2} d_M R_{eq}, \quad f_1 (E_{eq}) = 1, \ f_2 (E_{eq}) = 1.$$
**Statistical energy conservation principle**

**Theorem**

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

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

where $R$ satisfies the exact covariance equation.

**Corollary**

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

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

---

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 $R_A(t)$

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

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

$$\delta E'_A = \int_0^t R_A(t-s) \delta f'(s) \, ds.$$

*without the need of applying any perturbations.*

**kicked response:** For $\delta$ small enough, the linear response operator $R_A(t)$ can be calculated by solving the unperturbed system with a perturbed initial distribution

$$\pi|_{t=0} = \pi_{eq}(u - \delta u) = \pi_{eq} - \delta u \cdot \nabla \pi_{eq} + O(\delta^2).$$

$$\delta R_A(t) \equiv \delta u \cdot R_A = \int A(u) \delta \pi' + O(\delta^2).$$
Link between equilibrium fidelity and forecasting skill

Given the optimal model for the unperturbed climate $\pi_{\text{eq}}$, how can we assess the error in the climate change prediction

$$\mathcal{P}(\pi, \pi^M) = \int \pi \ln \frac{\pi}{\pi^M},$$

based on the unperturbed climate?

Under assumptions with diagonal covariance matrices $R = \text{diag}(R_k)$ and equilibrium model fidelity $\mathcal{P}(\pi_G, \pi^M_G) = 0$

$$\mathcal{P}(\pi_\delta, \pi^M_\delta) = \mathcal{I}(\pi_G, \delta) - \mathcal{I}(\pi_\delta)$$

$$+ \frac{1}{2} \sum_k (\delta \bar{u}_k - \delta \bar{u}_{M,k}) R_k^{-1} (\delta \bar{u}_k - \delta \bar{u}_{M,k})$$

$$+ \frac{1}{4} \sum_k R_k^{-2} (\delta R_k - \delta R_{M,k})^2 + O(\delta^3).$$

$R_k$ is the equilibrium variance in $k$-th component, and $\delta \bar{u}_k$ and $\delta R_k$ are the linear response operators for the mean and variance in $k$-th component.

(Majda & Gershgorin, PNAS, 2011)
Optimization in the training phase

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

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

- $p_\delta^M$ can be achieved through a kicked response procedure in the training phase compared with the actual observed data $p_\delta$ in nature;
- the information distance between perturbed responses $\mathcal{P} \left( p_\delta, p_\delta^M \right)$ can be calculated through the expansion formula;
- the information distance $\mathcal{P} \left( p_\delta(t), p_\delta^M(t) \right)$ 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

\[
\frac{dp_k}{dt} = B_k(p_k, p_k) + (\mathcal{L}_k - \mathcal{D}_k)p_k + \mathcal{F}_k, \quad p_k = (p_{\psi,k}, p_{\tau,k})^T, \quad \sum_k p_k \cdot B_k(p_k, p_k) \equiv 0,
\]

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

\[
\mathcal{L}_k = \begin{bmatrix}
\frac{ik_x \beta}{|k|^2} & -\frac{ik_x U}{\sqrt{1+(k_d/|k|)^2}} \\
-i k_x U \frac{1-(k_d/|k|)^2}{\sqrt{1+(k_d/|k|)^2}} & \frac{ik_x \beta}{|k|^2+k_d^2}
\end{bmatrix}, \quad \mathcal{D}_k = \frac{k}{2} \begin{bmatrix}
-1 & 0 \\
0 & \frac{1}{\sqrt{1+(k_d/|k|)^2}} \\
0 & \frac{1}{1+(k_d/|k|)^2}
\end{bmatrix},
\]

\[
B_k(p_k, p_k) = \begin{bmatrix}
B_{\psi,k} \\
B_{\tau,k}
\end{bmatrix} = \begin{bmatrix}
\sum_{m+n=k} \frac{m \cdot n}{|k|} \left( \frac{|n|}{|m|} p_{\psi,m} p_{\psi,n} + \sqrt{\frac{|n|^2+k_d^2}{|m|^2+k_d^2}} p_{\tau,m} p_{\tau,n} \right) \\
\sum_{m+n=k} \frac{m \cdot n}{\sqrt{|k|^2+k_d^2}} \left( \sqrt{\frac{|n|^2+k_d^2}{|m|^2+k_d^2}} p_{\psi,m} p_{\tau,n} + \frac{|n|}{\sqrt{|m|^2+k_d^2}} p_{\tau,m} p_{\psi,n} \right)
\end{bmatrix}.
\]
Exact statistical moment equations

Statistical energy in each spectral mode

\[ R_k = \mathbf{p}_k^* \mathbf{p}_k = \begin{bmatrix} |p_{\psi,k}|^2 & p_{\psi,k}^* p_{\tau,k} \\ p_{\psi,k} p_{\tau,k}^* & |p_{\tau,k}|^2 \end{bmatrix}, \quad p_{1,k}^* p_{2,k} = \bar{p}_{1,k}^* \bar{p}_{2,k} + p_{1,k}' p_{2,k}'. \]

\( R_k \) combines the variability in both mean and variance. The true statistical dynamical equations form a \( 2 \times 2 \) system about \( R_k \in \mathbb{C}^{2 \times 2} \)

\[ \frac{dR_k}{dt} = (\mathcal{L}_k - \mathcal{D}_k) R_k + Q_{F,k} + Q_{\sigma,k} + c.c., \quad |k| \leq N, \]

\[ Q_{F,k} = \mathbf{p}_k^* B_k (\mathbf{p}_k, \mathbf{p}_k) = \begin{bmatrix} p_{\psi,k}^* B_{\psi,k} & p_{\psi,k}^* B_{\tau,k} \\ p_{\tau,k}^* B_{\psi,k} & p_{\tau,k}^* B_{\tau,k} \end{bmatrix}, \quad \sum_k \text{tr}Q_{F,k} \equiv 0. \]
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.

\[
E = \frac{1}{2} \sum_{1 \leq |k| \leq N} |k|^2 |\psi_k|^2 + \left( |k|^2 + k_d^2 \right) |\tau_k|^2 = \frac{1}{2} \sum_{1 \leq |k| \leq N} |p_{\psi,k}|^2 + |p_{\tau,k}|^2.
\]

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.

\[
H_f = k_d^2 U \int \psi_x \tau = k_d^2 U \sum i k_x \bar{\psi_k^* \tau_k}, \quad F = \sum k_d^2 |\tau_k|^2 + 2 |k|^2 \Re \bar{\psi_k^* \tau_k}.
\]
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 $\sigma_k \hat{W}_k$ is introduced according to the equilibrium energy so that

$$
\sigma_{\psi,k}^2 = \delta \sigma_0^2 |q_{\psi,k}|_{eq}^2, \; \sigma_{\tau,k}^2 = \delta \sigma_0^2 |q_{\tau,k}|_{eq}^2.
$$

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

$$
\delta f_{\psi,k} = \delta U ik_x \left(-|\mathbf{k}|^2\right) \tau_k, \; \delta f_{\tau,k} = \delta U ik_x \left(-|\mathbf{k}|^2 + k_d^2\right) \psi_k.
$$
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

<table>
<thead>
<tr>
<th>regime</th>
<th>$N$</th>
<th>$\beta$</th>
<th>$k_d$</th>
<th>$U$</th>
<th>$\kappa$</th>
<th>$(k_{\text{min}}, k_{\text{max}})$</th>
<th>$\sigma_{\text{max}}$</th>
<th>$(k_x, k_y)_{\text{max}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ocean, high lat.</td>
<td>256</td>
<td>100</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>$(7.14, 15.63)$</td>
<td>0.104</td>
<td>(2, 8)</td>
</tr>
<tr>
<td>atmosphere, high lat.</td>
<td>256</td>
<td>2.5</td>
<td>4</td>
<td>0.2</td>
<td>0.05</td>
<td>$(2.51, 7.06)$</td>
<td>0.053</td>
<td>(3, 0)</td>
</tr>
</tbody>
</table>

[Diagram of zonal averaged flow field, low/mid-lat ocean and atmosphere]
Low-latitude: Stochastic perturbation with $\delta \sigma_0^2 = 0.2$ autocorrelation functions and probability density functions

(a) ocean

(b) atmosphere

(c) ocean regime

(d) atmosphere regime
Current and future work

- Reduced-order stochastic modeling strategies to capture passive scalar intermittency

- 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

• Statistical theories

• Improving imperfect model skill