Links between POD and balanced truncation

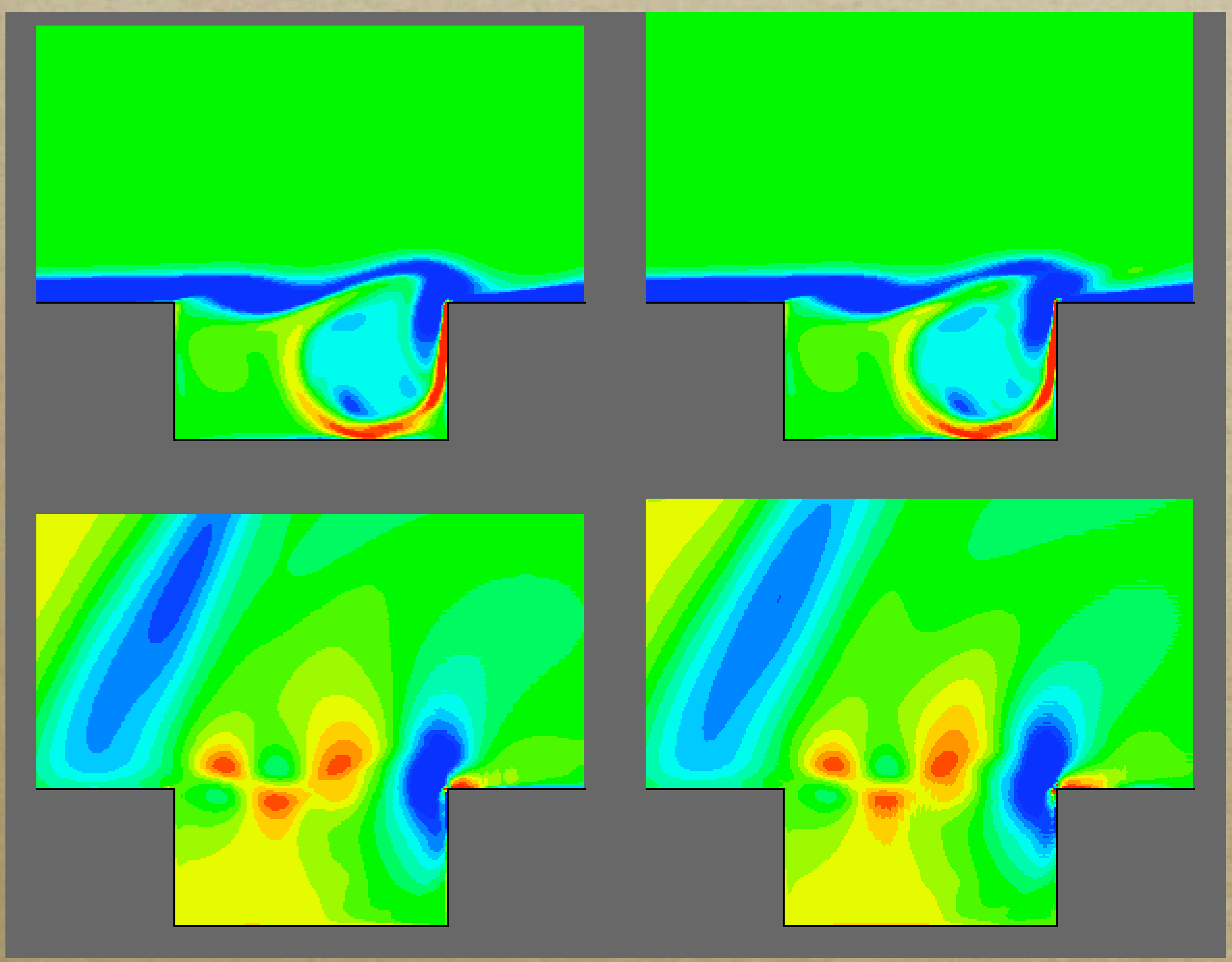
Clancy Rowley

Mechanical and Aerospace Engineering
Princeton University

CIMMS seminar, Caltech
August 21, 2003

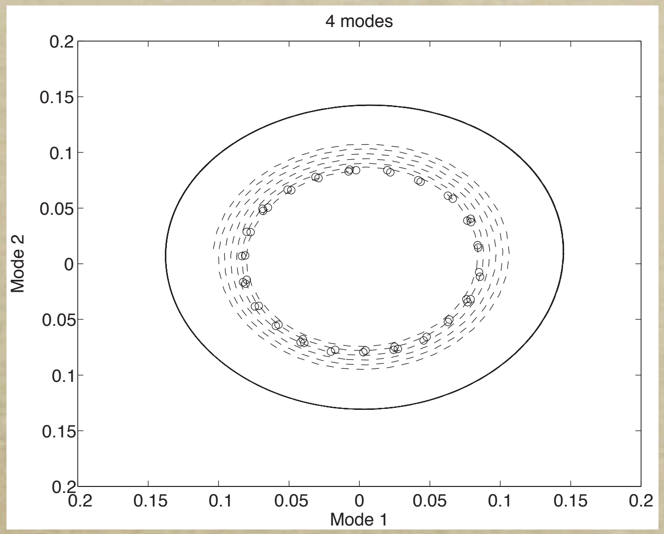
Motivation

Cavity flow oscillations

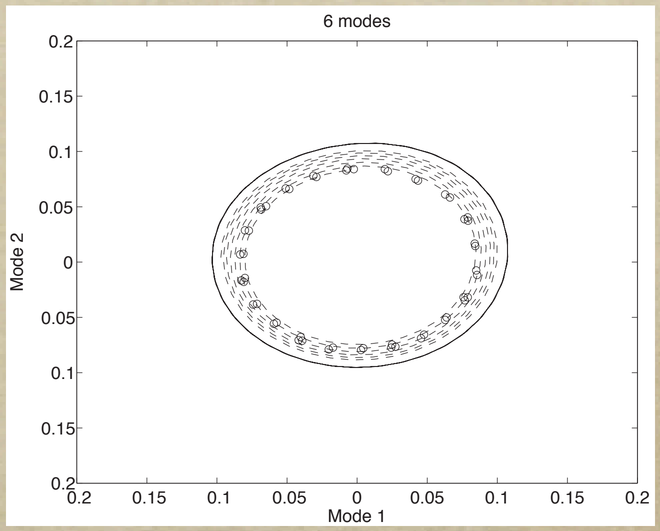


Full simulation *4-mode POD/Galerkin*

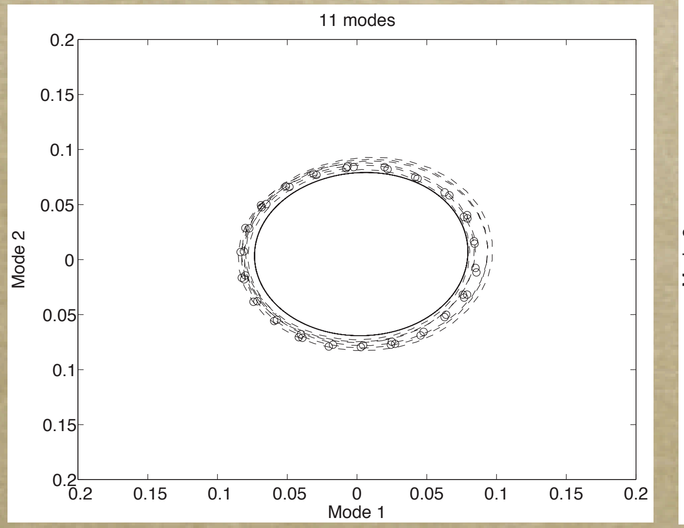
Phase portraits ($\varphi_1 \varphi_2$ plane)



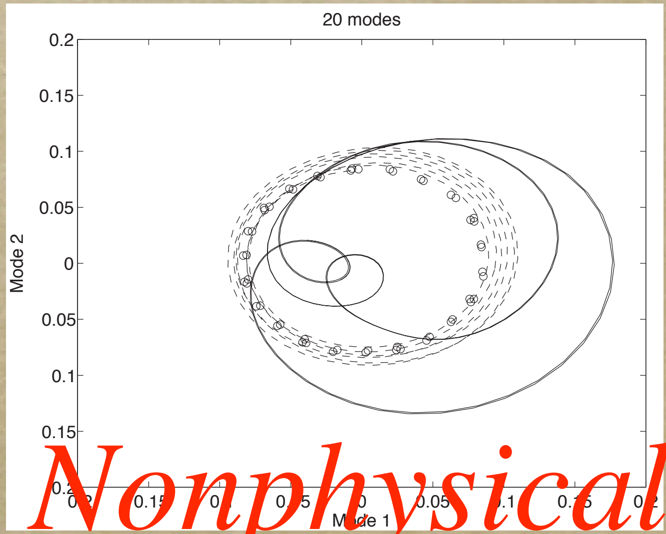
4 modes



6 modes



11 modes



20 modes

Nonphysical dynamics

POD vs. Balanced truncation

POD

- **Pro**: works for nonlinear systems
- **Pro**: tractable computation using eigenvalue solvers $(n > 10^6)$
- **Con**: often yields unpredictable results
  - *Most energetic modes are not necessarily most important to the dynamics*
  -

Balanced truncation

- **Pro**: Guaranteed error bounds
- **Con**: Works only for linear input-output systems
- **Con**: Computationally expensive, cannot compute for $n > 10,000$

Overview of POD/Galerkin

POD modes

Given a set of data $\{u(t) \in H \mid t \in I\}$, project onto orthonormal basis functions $\varphi_j \in H$:

$$\hat{u}(t) = Pu(t) = \sum_{j=1}^{n} a_j(t)\varphi_j, \qquad a_j(t) = \langle u(t), \varphi_j \rangle$$

Goal: Given $u(x,t)$, find *optimal* orthonormal functions $\varphi_k(x)$, called *POD modes*, which minimize the time average of $\|u - Pu\|$, for fixed n.

Solution: Eigenvalue problem

$$R\varphi_k = \lambda_k \varphi_k, \qquad R = \frac{1}{T}\int_0^T u(t) \otimes u^*(t)\, dt$$

Overview of POD/Galerkin

Standard approach

- Start with a nonlinear system $\dot{x} = f(x), \quad x \in H$
- Integrate solutions for one or more initial conditions $x_0 \in H$
- Compute POD modes from the ensemble of data
-
-
  $$\hat{x}(t) = \sum_{j=1}^{n} a_j(t)\varphi_j, \qquad \varphi_j \in H$$
-
- Galerkin projections are given by

$$\dot{a}_j = \langle f(\hat{x}), \varphi_j \rangle$$

Overview of balanced truncation

Standard approach

- Start with a stable, linear input-output system

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

- Compute controllability and observability Gramians

$$X = \int_0^\infty e^{At} BB^* e^{A^* t}\, dt \qquad Y = \int_0^\infty e^{A^* t} C^* C e^{At}\, dt$$

$$AX + XA^* + BB^* = 0 \qquad A^*Y + YA + C^*C = 0$$

- Find a transformation T that diagonalizes X and Y

$$x = Tz, \qquad T^{-1}X(T^{-1})^* = T^*YT = \Sigma$$

- Project onto first n columns of T.

Balanced truncation: properties

Error bounds

- Consider the transfer function $G(s) = C(sI - A)^{-1}B$
- Recall the L_2-induced norm

$$\|G\|_\infty = \max_\omega \sigma_1(G(i\omega)) = \max_u \frac{\|Gu\|_2}{\|u\|_2}$$

- Any reduction to *r* states must satisfy

$$\|G_r - G\|_\infty > \sigma_{r+1}$$

- Balanced truncation guarantees

$$\|G_r - G\|_\infty < 2 \sum_{j=r+1}^{n} \sigma_j$$

- Disclaimer: balanced truncation is *not* optimal. There are other methods for model reduction (e.g. Hankel norm reduction).

Main results

1. Balanced truncation may be viewed as POD/Galerkin with respect to an inner product defined by the observability Gramian
2. Computational procedure for computing approximate balanced truncations for very large systems, using the method of snapshots
3. Example: linearized channel flow

Can we compute balanced truncations without solving Lyapunov equations, or ever storing the full Gramians?

Empirical Gramians

Controllability Gramian

- Well known (e.g., Lall, Marsden, Glavaski, 2002)
- Construct an ensemble of solutions $\{x_1(t), x_2(t), \dots, x_p(t)\}$

$$\begin{aligned} x_1(0) &= B\hat{e}_1 \\ x_2(0) &= B\hat{e}_2 \\ &\vdots \\ x_p(0) &= B\hat{e}_p \end{aligned} \qquad u = 0 \qquad \Longrightarrow \qquad \begin{aligned} x_1(t) &= e^{At}B\hat{e}_1 \\ x_2(t) &= e^{At}B\hat{e}_2 \\ &\vdots \\ x_p(t) &= e^{At}B\hat{e}_p \end{aligned}$$

- Then

$$\begin{aligned} X &= \int_0^\infty e^{At}BB^*e^{A^*t}\,dt \\ &= \int_0^\infty (x_1x_1^* + x_2x_2^* + \cdots + x_px_p^*)\,dt \end{aligned}$$

- POD modes of this dataset are eigenvectors of X
  - *Standard POD procedure projects onto most controllable states, ignores observability*

Empirical Gramians

Notes

- Balancing modes are appropriately scaled eigenvectors of XY. Thus, balanced truncation is just POD with respect to an inner product defined by the observability Gramian:

$$\langle x_1, x_2 \rangle = x_1^* Y x_2$$

- Even if output is $y = x$, observability is still important—Y is solution to

$$A^* Y + Y A + I = 0$$

- If A is normal, A commutes with X and Y, so they have the same eigenvectors (Farrell & Ioannou 1993). In this case, POD modes are the same as balancing modes.

Empirical Gramians

Observability Gramian

- Consider the adjoint system $\dot{z} = A^* z$
- Construct an ensemble of solutions $\{z_1(t), \ldots, z_q(t)\}$

$$\begin{aligned} z_1(0) &= C^*\hat{e}_1 \\ z_2(0) &= C^*\hat{e}_2 \\ &\vdots \\ z_q(0) &= C^*\hat{e}_q \end{aligned} \quad \Longrightarrow \quad \begin{aligned} z_1(t) &= e^{A^*t}C^*\hat{e}_1 \\ z_2(t) &= e^{A^*t}C^*\hat{e}_2 \\ &\vdots \\ z_q(t) &= e^{A^*t}C^*\hat{e}_q \end{aligned}$$

- Empirical Gramian is

$$\begin{aligned} Y &= \int_0^\infty e^{A^*t}C^*Ce^{At}\,dt \\ &= \int_0^\infty (z_1z_1^* + z_2z_2^* + \cdots + z_qz_q^*)\,dt \end{aligned}$$

- Note that if C = Id, need to compute n different trajectories. (Doesn't scale well for very large n)

Approximate Gramians

Approximate observability Gramian

- Instead of the system
- $$\dot{x} = Ax + Bu$$
- $$y = x$$
-
- Consider the almost identical system
- $$\dot{x} = Ax + Bu$$
- $$y = Px$$
-
- $P = \varphi\varphi^*$ is a projection onto the first *r* POD modes (which are columns of φ)
- For this system, empirical observability Gramian is tractable for large *n*: compute *r* copies of adjoint system with initial conditions equal to each of the POD modes

Balanced truncation using snapshots

Procedure:

- Construct data matrices containing primal and dual snapshots

$$\rho = \underbrace{\begin{bmatrix} x(t_1) & \cdots & x(t_{n_p}) \end{bmatrix}}_{\textit{Primal}} \qquad \mu = \underbrace{\begin{bmatrix} z(t_1) & \cdots & z(t_{n_d}) \end{bmatrix}}_{\textit{Dual}}$$

- Form the singular value decomposition of $\mu^* \rho$

$$\mu^* \rho = \begin{bmatrix} z(t_1)^* x(t_1) & \cdots & z(t_1)^* x(t_{n_p}) \\ \vdots & \ddots & \vdots \\ z(t_{n_d}) * x(t_1) & \cdots & z(t_{n_d})^* x(t_{n_p}) \end{bmatrix} = U \Sigma V^*$$

- First *r* modes of balancing transformation (*r* is rank of $\sum$) are

$$T_1 = \rho V \Sigma^{-1/2}$$

Balanced truncation using snapshots

Theorem:

- Define $S_1 = \Sigma^{-1/2} U^* \mu^*$ $\quad T_1 = \rho V \Sigma^{-1/2}$
- If Σ has rank *n*, then $S_1 = (T_1)^{-1}$ and

$$S_1 X S_1^* = T_1^* Y T_1 = \Sigma$$

- If Σ has rank $r < n$, then there exist S_2 and T_2 such that

$$S = \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} \qquad T = \begin{bmatrix} T_1 & T_2 \end{bmatrix} \qquad S = T^{-1}$$

$$T^{-1} X Y T = \begin{bmatrix} \Sigma^2 & 0 \\ 0 & 0 \end{bmatrix}$$

$$S X S^* = \begin{bmatrix} \Sigma & 0 \\ 0 & X_2 \end{bmatrix} \qquad T^* Y T = \begin{bmatrix} \Sigma & 0 \\ 0 & Y_2 \end{bmatrix}$$

Galerkin projection

Original system

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

Galerkin projection (standard inner product)

$$\dot{z} = (T_1 T_1^*)^{-1} T_1^* A T_1 z + (T_1 T_1^*)^{-1} T_1^* B u$$
$$y = C T_1 z$$

Balanced truncation = Galerkin projection (Y inner product)

$$\dot{z} = S_1 A T_1 z + S_1 B u$$
$$y = C T_1 z$$

Stability

Galerkin projection using an “energy-based” inner product preserves stability of the origin

- Observability Gramian is a Lyapunov function, so stability of the origin is preserved under balanced truncation, or POD/Galerkin with Y as inner product

$$\dot{x} = Ax \qquad V(x) = x^* Q x$$

$$\dot{V}(x) = x^*(A^*Q + QA)x \leq 0$$

$$x = Tz$$

$$\dot{z} = (T^*QT)^{-1}T^*QATz \qquad V(z) = z^*T^*QTz$$

$$\begin{aligned} \dot{V}(z) &= \dot{z}^*T^*QTz + z^*T^*QT\dot{z} \\ &= z^*T^*(A^*Q + QA)Tz \leq 0 \end{aligned}$$

Also true for nonlinear systems
(Rowley, Murray, & Colonius, 2002), (Prajna, CDC 2003)

Summary of the method

1. Compute an ensemble of solutions to $\dot{x} = Ax$ with various relevant initial conditions (e.g. columns of *B*), and assemble these snapshots into a matrix ρ of dimension $n \times n_p$ (n_p snapshots)
2. Compute POD modes from this data (SVD of ρ)
3. If the first *r* modes capture a large fraction of energy, solve *r* copies of the adjoint system $\dot{z} = A^* z$ with initial conditions equal to the each of the first *r* POD modes, assembling this data into a matrix μ of dimension $n \times n_d$ (n_d snapshots)
4. Form the $n_d \times n_p$ matrix $\mu^* \rho$, and compute its SVD $\mu^* \rho = U \Sigma V^*$
5. The balancing modes are columns of the rectangular matrix

$$T_1 = \rho V \Sigma^{-1/2}$$

Largest matrix one has to store is
#snapshots by #states

Example: linearized channel flow

Plane channel flow

- Linearize Navier-Stokes about a parallel shear flow $(U(y), 0, 0)$
- In terms of wall-normal velocity v and wall-normal vorticity η pressure can be eliminated using continuity:

$$\left[\left(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x}\right)\nabla^2 - U''\frac{\partial}{\partial x} - \frac{1}{R}\nabla^4\right] v = 0$$

$$\left[\frac{\partial}{\partial t} + U\frac{\partial}{\partial x} - \frac{1}{R}\nabla^2\right]\eta + U'\frac{\partial v}{\partial z} = 0$$

$v = \partial v/\partial y = \eta = 0$ at solid walls

- Consider streamwise-constant perturbations $(\partial/\partial x = 0)$

$$\frac{\partial v}{\partial t} = \frac{1}{R}\nabla^2 v$$

$$\frac{\partial \eta}{\partial t} = \frac{1}{R}\nabla^2\eta - U'\frac{\partial v}{\partial z}$$

- Discretize with Chebyshev modes in *y*-direction, Fourier in *z*-direction

Computing modes

POD modes

- E-vectors of X: $AX + XA^* + BB^* = 0$

Balancing modes

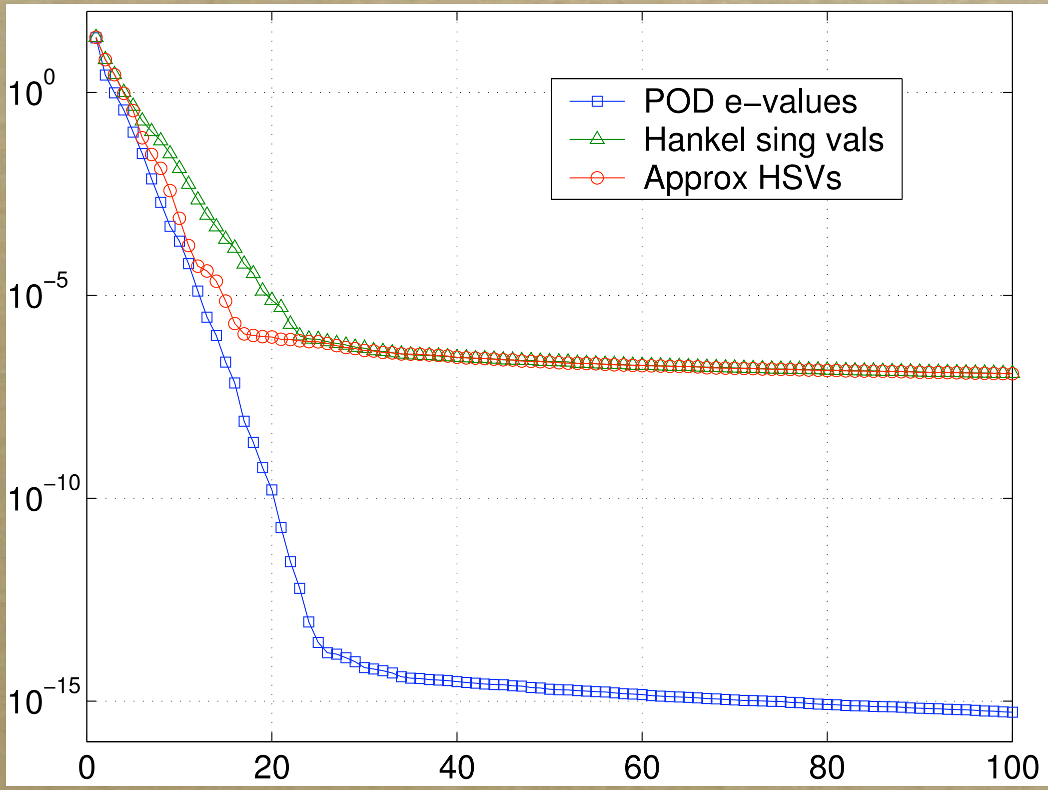
- Scaled e-vectors of XY: $A^*Y + YA + I = 0$

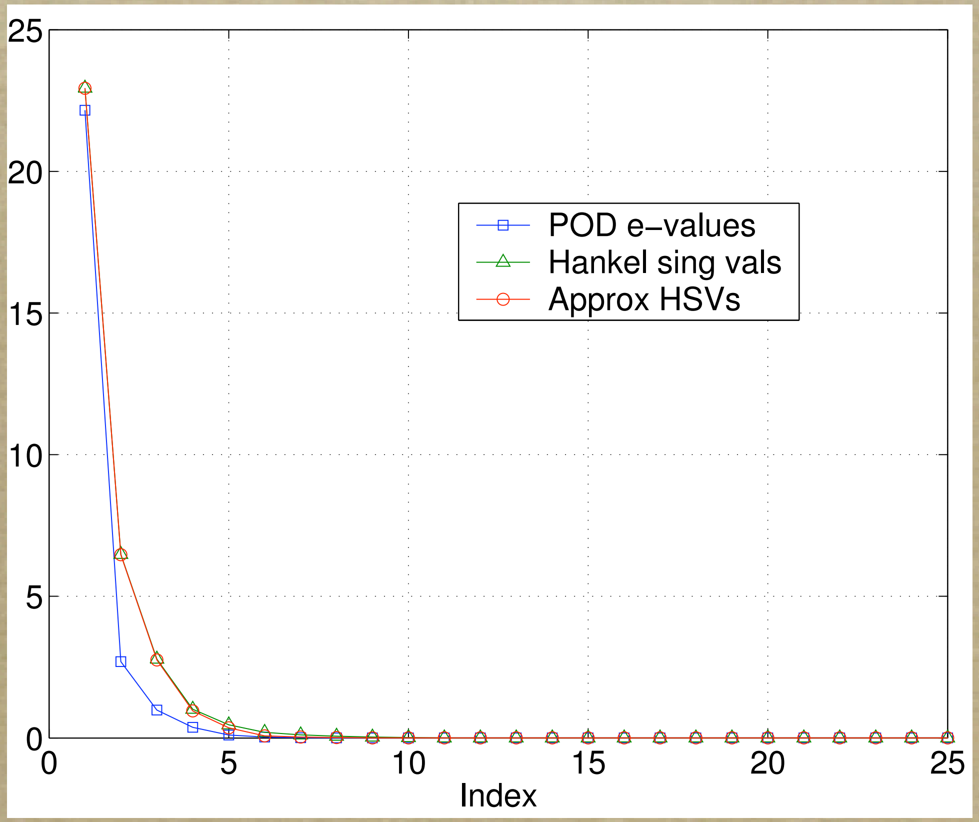
Approximate balancing modes

- Scaled e-vectors of XY: $A^*Y + YA + \varphi\varphi^* = 0$

Columns of φ are first 5 POD modes

Energy decay



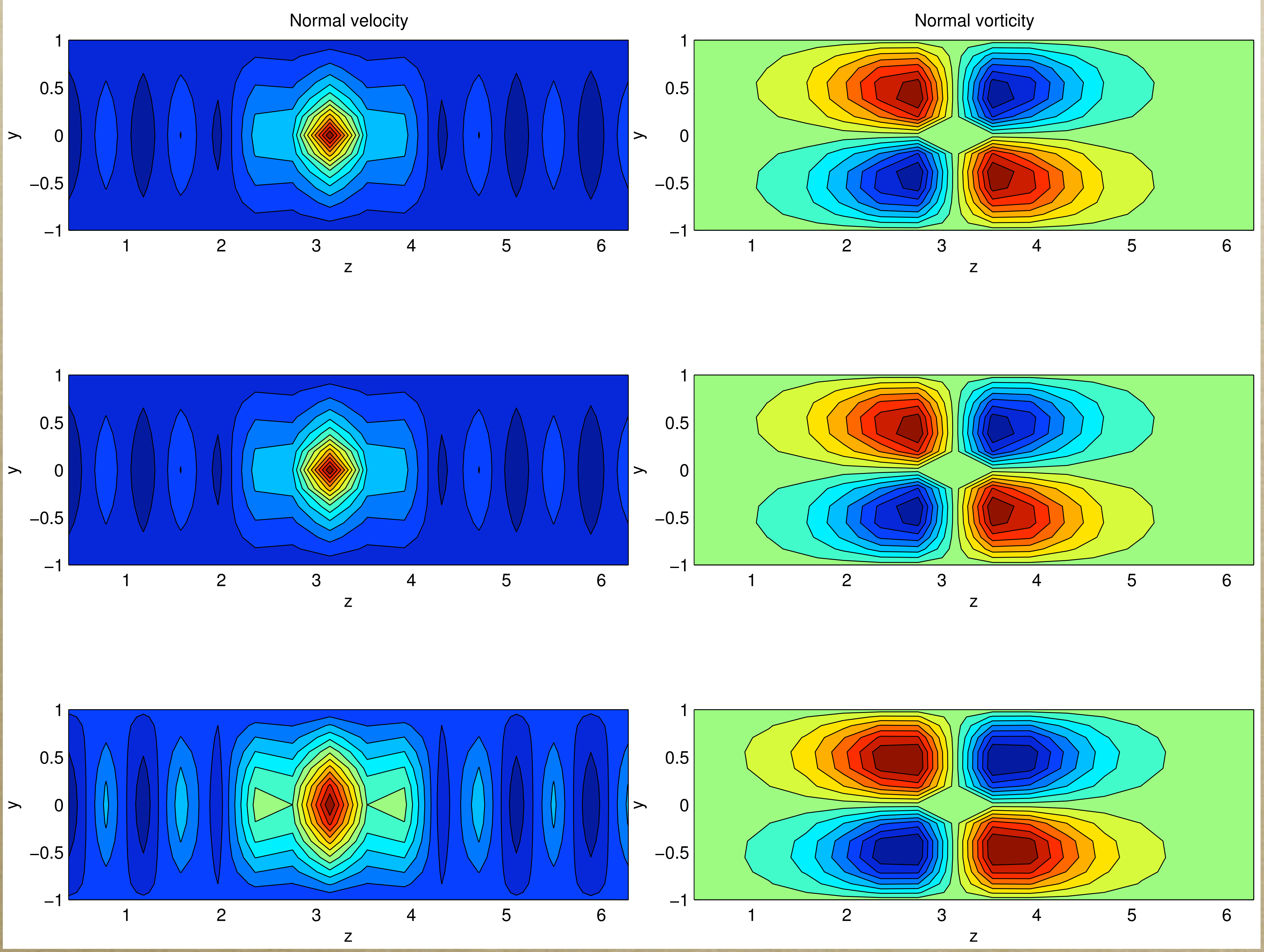


Mode 1

Balancing

Approximate balancing

POD

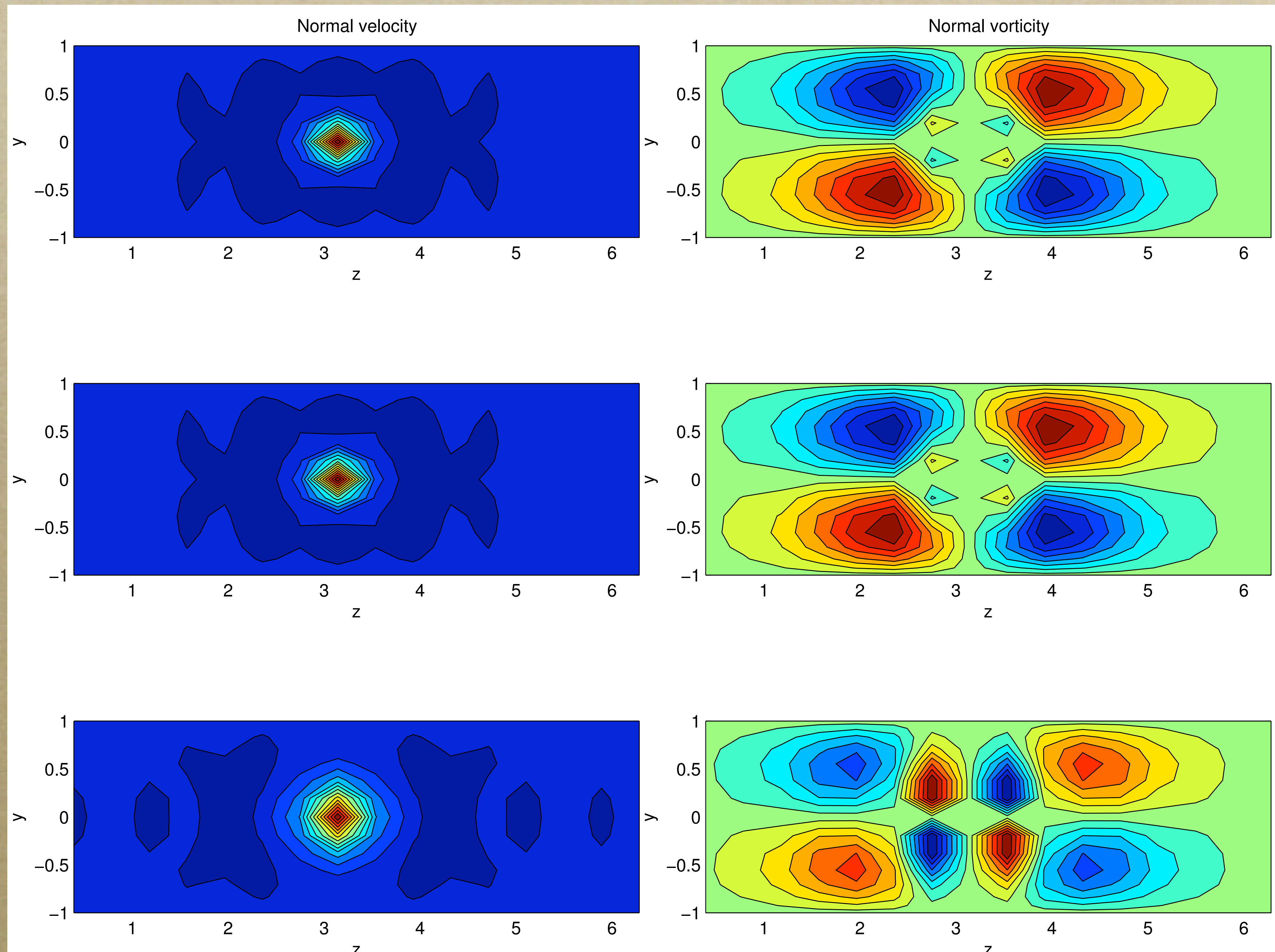


Mode 2

Balancing

Approximate balancing

POD

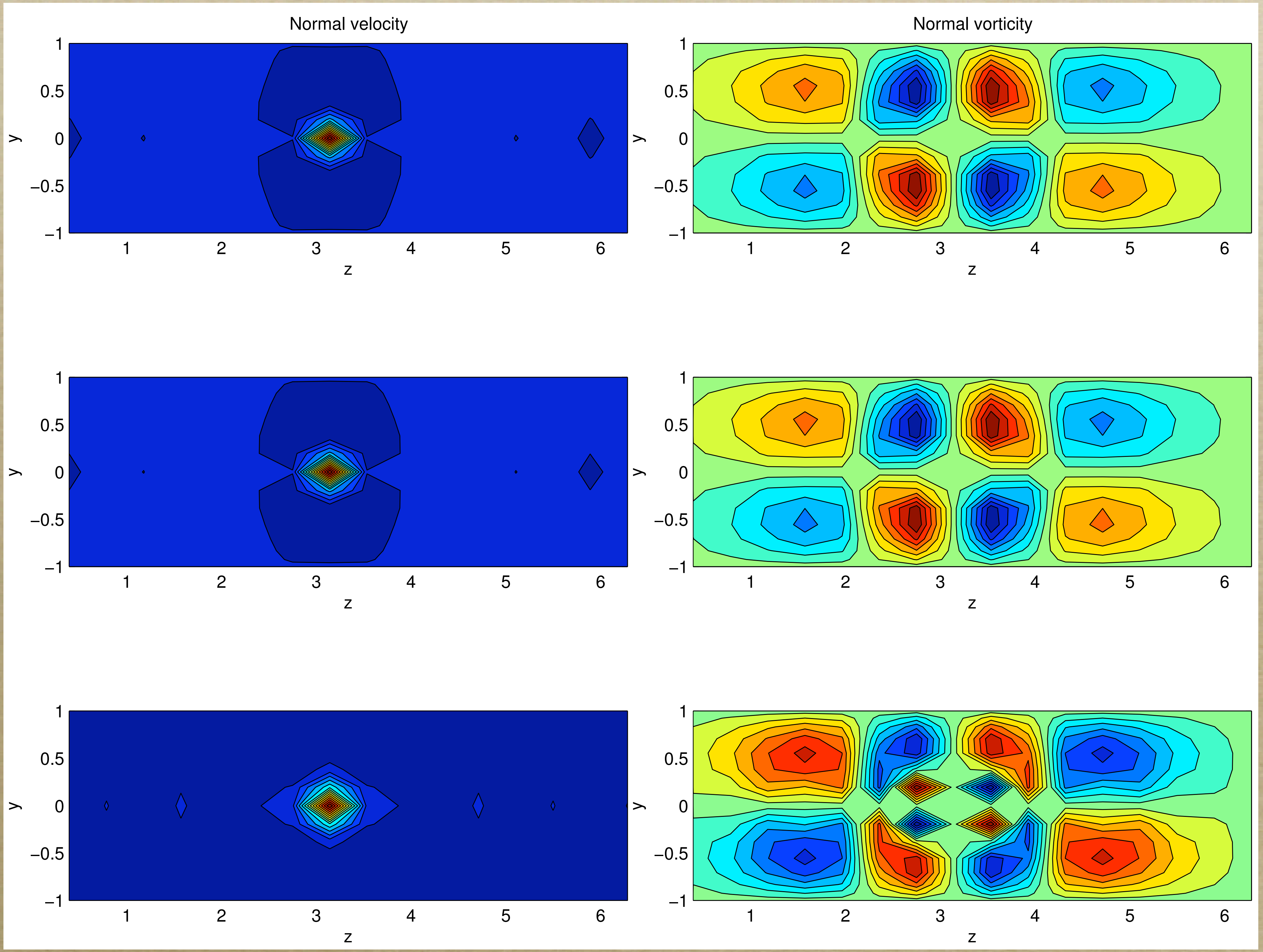


Mode 3

Balancing

Approximate balancing

POD

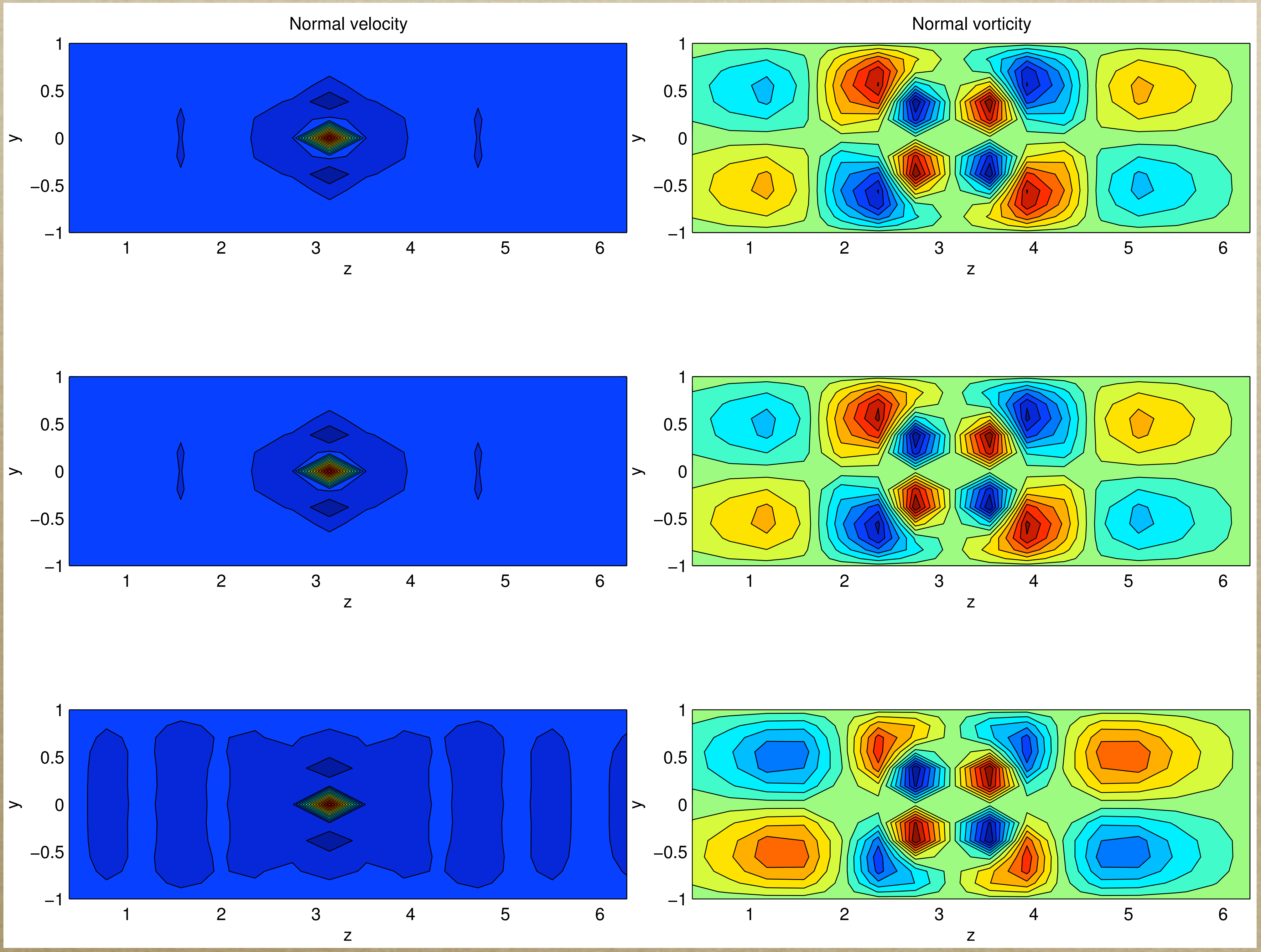


Mode 4

Balancing

Approximate balancing

POD



Error norms

Linear systems--have norms!

	BT	Approx BT	POD
$\|G_5 - G\|_\infty$	0.446	0.560	5.35
$\|G_5 - G\|_\infty / \|G\|_\infty$	1.18%	1.48%	14.1%

$$\varepsilon_1 < \|G_r - G\|_\infty < \varepsilon_2$$

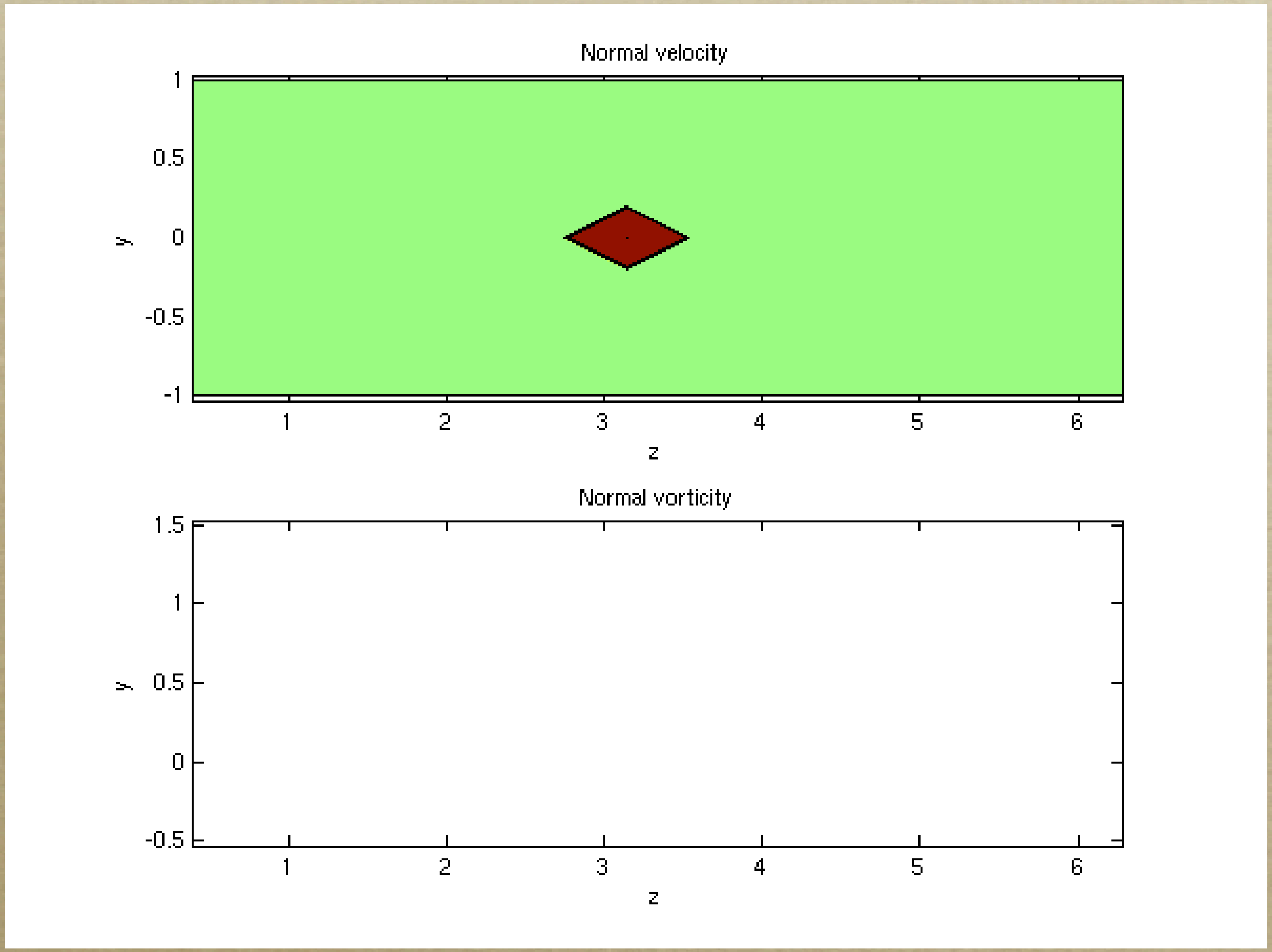
$$\varepsilon_1 = \sigma_6 = 0.2008 \qquad \varepsilon_2 = 2\sum_{j=6}^{n} \sigma_j = 0.8581$$

	BT, standard i.p.	POD, Y i.p.
$\|G_5 - G\|_\infty$	3.48	0.669
$\|G_5 - G\|_\infty / \|G\|_\infty$	9.19%	1.77%

Balanced truncation not optimal,
but works the best here

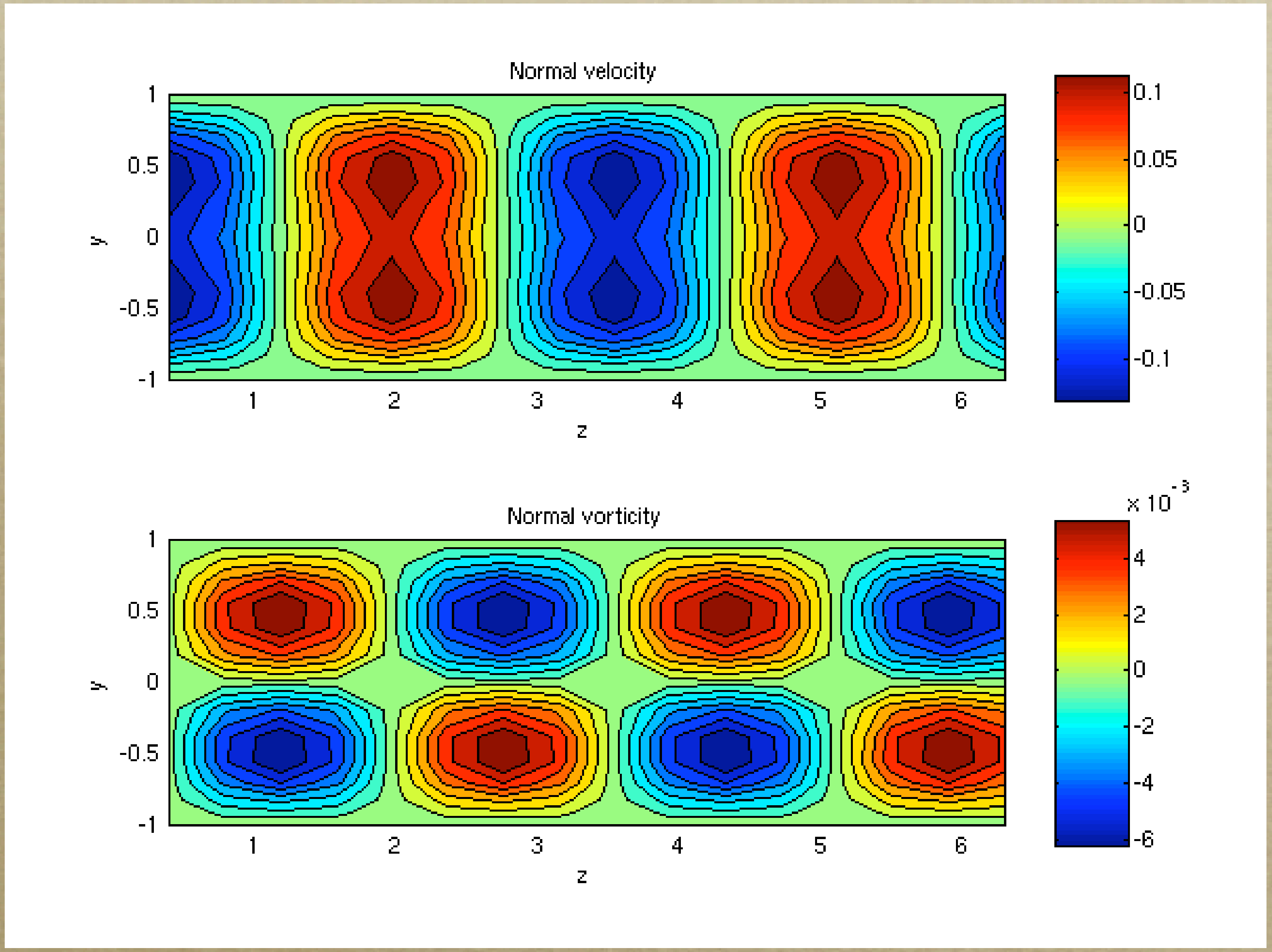
Movie

Full simulation, impulse response



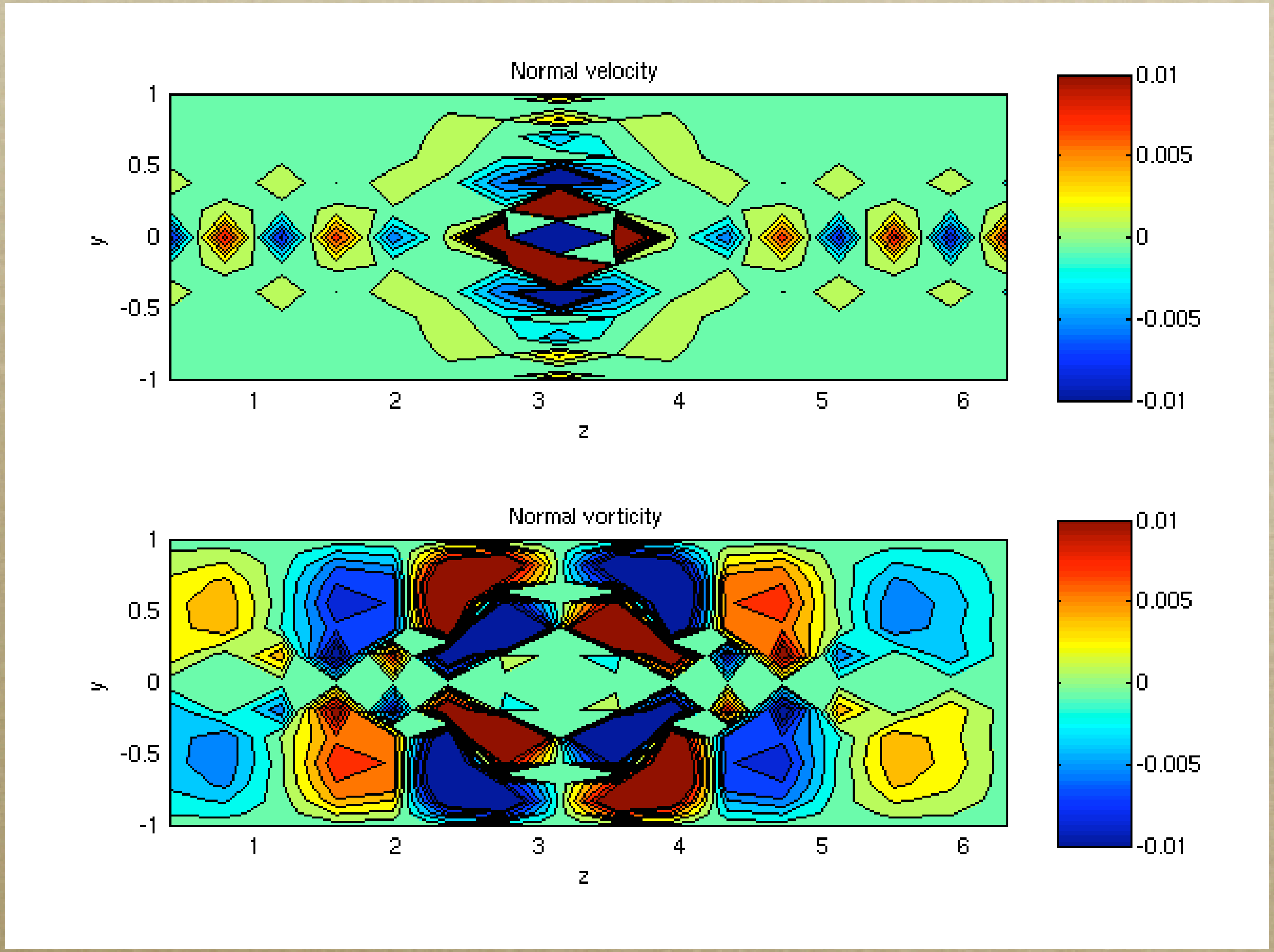
Movie

Error, 5-mode POD truncation



Movie

Error, 5-mode balanced truncation



Conclusions

POD modes of impulse response data represent most controllable modes

- Observability also important if A non-normal

Balanced truncation

- Same as POD/Galerkin of impulse response data, using inner product specified by observability Gramian Y
- Y is a good inner product: guarantees stability of Galerkin projections, gives good results in practice

Approximate empirical Gramians

- Involve several integrations of adjoint system
- Method of snapshots scales well for very large n
- 5-mode approximation agrees well with exact balanced truncation for streamwise-constant linearized channel flow.

The End