ADJOINT–BASED OPTIMIZATION OF MULTISCALE PDE SYSTEMS

BARTOSZ PROTAS

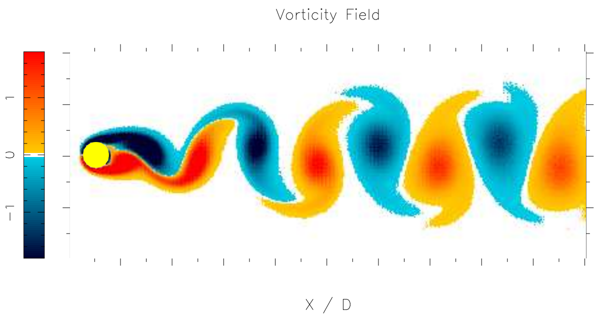
Department of Mathematics & Statistics
McMaster University
Hamilton, CANADA
http://www.math.mcmaster.ca/bprotas

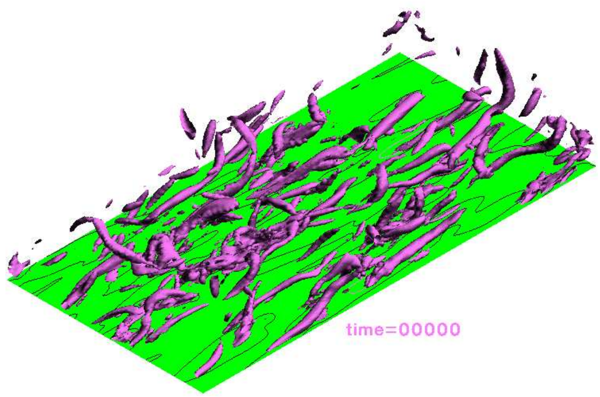
- COLLABORATORS:
  T. Bewley (UCSD)
- CURRENT SUPPORT:
  NSERC (Canada)
  MCAMM (Canada)
  General Motors (US / Canada)
  CNRS (France)

AGENDA

- Motivation
  - Origins of the multiscale behavior
  - Naive approach to PDE–constrained optimization
- Adjoint–based Analysis for Multi–scale Systems
  - The Big Picture and Available Regularization Options
  - Preconditioning Strategies
  - Results for the Kuramoto–Sivashinsky and the Navier–Stokes Systems
- Beyond linear preconditioning
  - Gradient extraction in Banach spaces
  - Nonlinear preconditioning
  - Computational results
- Conclusions & Future Directions

MOTIVATION





OBJECTIVES (inverse problems):

- Control fluid flow with the least amount of energy possible
- Estimate flow based on incomplete and/or noisy measurements

The Navier–Stokes system:

$$\begin{cases} \dfrac{\partial \mathbf{v}}{\partial t} + (\mathbf{v}\cdot\boldsymbol{\nabla})\mathbf{v} + \boldsymbol{\nabla} p - \mu\Delta\mathbf{v} = \boldsymbol{\phi}, & \text{in } \Omega\times(0,T) \\ \boldsymbol{\nabla}\cdot\mathbf{v} = 0, & \text{in } \Omega\times(0,T) \\ \text{Initial condition} & \text{on } \Gamma\times(0,T) \\ \text{Boundary condition} & \text{in } \Omega \text{ at } t=0 \end{cases}$$

ORIGINS OF MULTI–SCALE BEHAVIOR

Spectral (Fourier) representation of the Navier–Stokes equation:
$(v_p = \sum_{\mathbf{k}} \hat{u}_p(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}})$

$$\frac{\partial \hat{u}_p(\mathbf{k})}{\partial t} + \nu k^2 \hat{u}_p(\mathbf{k}) + \sum_{\mathbf{m}+\mathbf{l}=\mathbf{k}} \left(1 - \frac{k_p k_s}{k^2}\right) l_q \hat{u}_q(\mathbf{m}) \hat{u}_s(\mathbf{l}) = \hat{\phi}_p$$

- The triadic interactions represent an intrinsic mechanism responsible for transfer of energy across a wide range of time– and length–scales
- Regardless of the boundary conditions and the RHS forcing , motion exists at a very broad range of length– and time–scales
- In an accurate simulation, all length– and time–scales need to be resolved.

WAKE OPTIMIZATION PROBLEM [a] I

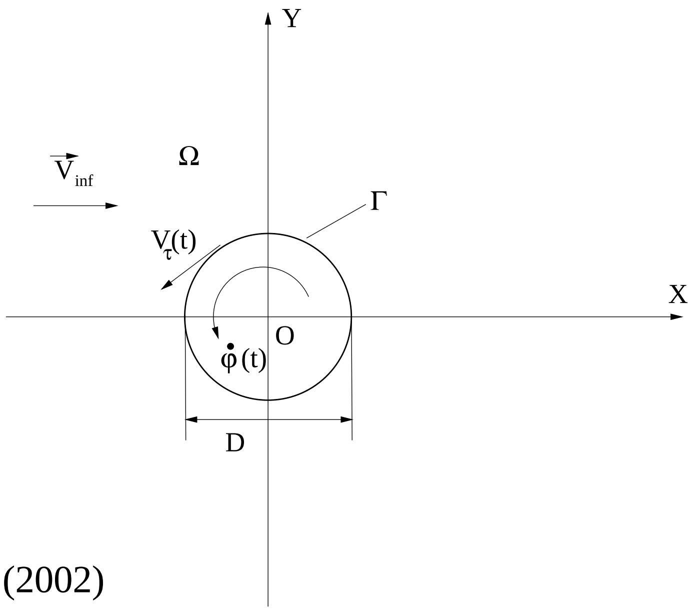
PRELIMINARIES

- Find $\dot{\varphi}_{opt} = \operatorname{argmin}_{\dot{\varphi}} \mathcal{J}(\dot{\varphi})$, where

$$\mathcal{J}(\dot{\varphi}) = \frac{1}{2}\int_0^T \left\{ \left[\begin{array}{c} \text{power related to} \\ \text{the drag force} \end{array} \right] + \left[\begin{array}{c} \text{power needed to} \\ \text{control the flow} \end{array} \right] \right\} dt$$

$$= \frac{1}{2}\int_0^T \oint_{\Gamma_0} \{[p(\dot{\varphi})\mathbf{n} - \mu\mathbf{n}\cdot\mathbf{D}(\mathbf{v}(\dot{\varphi}))]\cdot[\dot{\varphi}(\mathbf{e}_z \times \mathbf{r}) + \mathbf{v}_\infty]\}\, d\sigma dt,$$

- Assumptions:
  - viscous, incompressible flow
  - plane, infinite domain
  - $Re = 150$



[a] B. Protas & A. Styczek, *Physics of Fluids* **14**, 2073, (2002)

WAKE OPTIMIZATION PROBLEM II

THE COMPLETE OPTIMALITY SYSTEM FOR $\{\dot{\varphi}_{opt}, [\mathbf{v}_{opt}, p_{opt}], [\mathbf{v}*, p^*]\}$

$$
\left\{
\begin{array}{l}
\nabla \mathcal{J}(\dot{\varphi}_{opt}) = \dfrac{1}{2} \oint_{\Gamma} \{\mu R \mathbf{n} \cdot \mathbf{D}(\mathbf{v}^*) \cdot \tau + \mu \mathbf{n} \cdot \mathbf{D}(\mathbf{v}(\dot{\varphi}_{opt})) \cdot (\mathbf{e}_z \times \mathbf{r})\}\, d\sigma = 0 \\[2ex]
\left\{
\begin{array}{ll}
\begin{bmatrix} \frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \boldsymbol{\nabla})\mathbf{v} - \mu \Delta \mathbf{v} + \boldsymbol{\nabla} p \\ \boldsymbol{\nabla} \cdot \mathbf{v} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{in } \Omega \times (0, T), \\
\mathbf{v} = 0 & \text{at } t = 0, \\
\mathbf{v} = \dot{\varphi}_{opt} \tau & \text{on } \Gamma
\end{array}
\right. \\[2ex]
\left\{
\begin{array}{ll}
\mathcal{N}^* \begin{bmatrix} \mathbf{v}^* \\ p^* \end{bmatrix} = \begin{bmatrix} -\frac{\partial \mathbf{v}^*}{\partial t} - \mathbf{v} \cdot \left[\boldsymbol{\nabla} \mathbf{v}^* + (\boldsymbol{\nabla} \mathbf{v}^*)^T\right] - \mu \Delta \mathbf{v}^* + \boldsymbol{\nabla} p^* \\ -\boldsymbol{\nabla} \cdot \mathbf{v}^* \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{in } \Omega \times (0, T), \\
\mathbf{v}^* = 0 & \text{at } t = T, \\
\mathbf{v}^* = \mathbf{r} \times (\dot{\varphi}_{opt} \mathbf{e}_z) + \mathbf{v}_\infty & \text{on } \Gamma
\end{array}
\right.
\end{array}
\right.
$$

- A counterpart of the Euler–Lagrange equation
- Formidable computational task when solved at once
- Need an iterative procedure to decouple the different parts

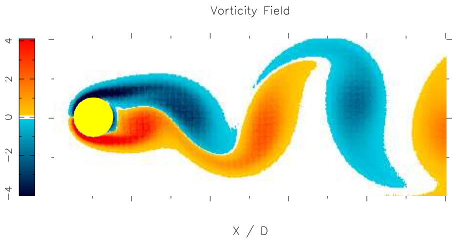
Wake Optimization Problem III

Primal and Adjoint Simulations for Cylinder Rotation as the Control

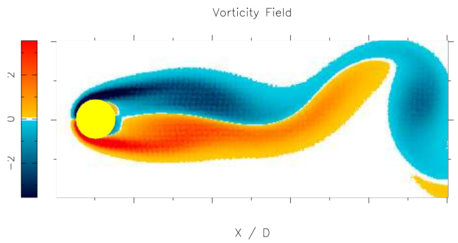
Animation

Wake Optimization Problem IV

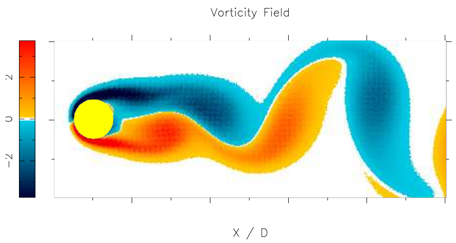
Modifications of Flow Patterns in the Controlled Flows $T = 6$



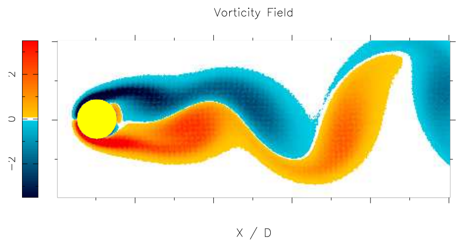
No Control



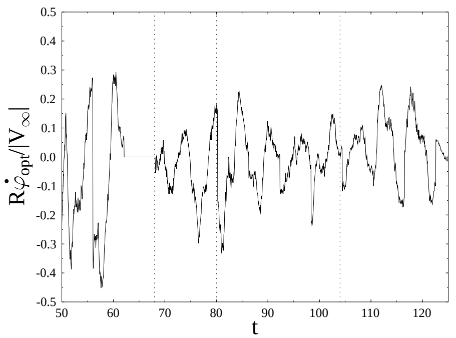
$t = 68$

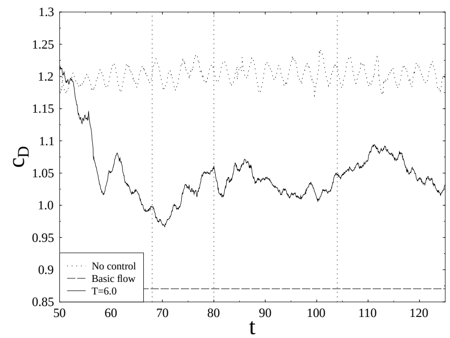


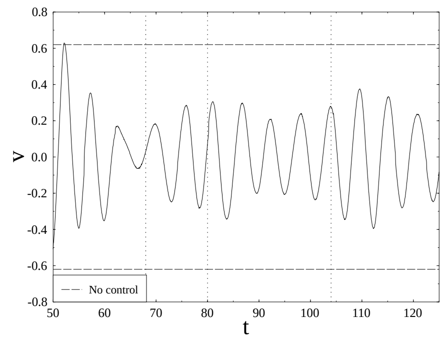
$t = 80$



$t = 104$







PDE–CONSTRAINED OPTIMIZATION

REDUCED HESSIANS [LEWIS (1997)]

- CONSTRAINED optimization problem

$$\begin{cases} \min\limits_{(u,a)} f(u,a) \\ S(u(a),a) = 0 \end{cases}$$

- Equivalent UNCONSTRAINED optimization problem

$$\min_a f(u(a),a) = \min_a F(a)$$

- CONDITIONING of the problem is determined by the REDUCED HESSIAN (of the unconstrained problem)

$$\nabla^2_{(a)} F = L_{aa} + L_{au}\left(\frac{du}{da}\right) + \left(\frac{du}{da}\right)^* L_{ua} + \left(\frac{du}{da}\right)^* L_{uu}\left(\frac{du}{da}\right)$$

where

$$L(u,a,\lambda) = f(u,a) + \langle \lambda, S(u(a),a)\rangle$$

$$\frac{du}{da} = -S_u^{-1} S_a, \qquad \text{(implicit function theorem)}$$

An Illuminating Example

Ritz–Galerkin Method for the Poisson equation

- Solve

$$\begin{cases} \Delta u = g, & u \in H_0^1(\Omega),\ g \in H^{-1}(\Omega) \\ u\big|_{Gamma} = 0, & \Delta\ :\ H_0^1(\Omega) \rightarrow H^{-1}(\Omega) \end{cases}$$

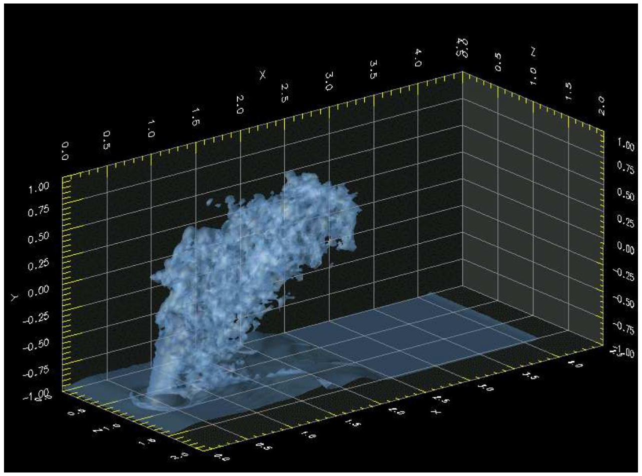
  by minimizing the functional $\mathcal{J}\ : H_0^1(\Omega) \rightarrow \mathbb{R}$, $\mathcal{J}(\Phi) = \int_\Omega [\frac{1}{2}(\boldsymbol{\nabla}\Phi)^2 + g\Phi]\, d\Omega$

- The Gâteaux differential (the optimality condition)

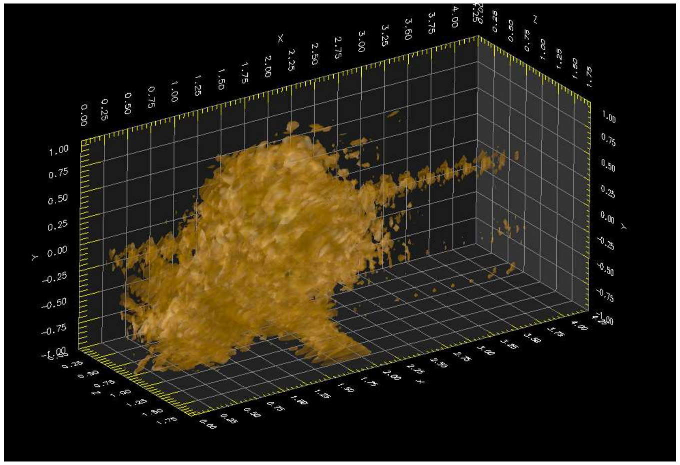
$$\mathcal{J}(\Phi;\Phi') = \int_\Omega [-\Delta\Phi + g]\Phi'\, d\Omega = 0$$

- Gradient in $L_2(\Omega)$: $\nabla^{L_2}\mathcal{J} = -\Delta\Phi + g \in L_2(\Omega)$
  - Hessian eigenvalues (Fourier space): $\{k_1^2, k_2^2, \dots, k_N^2\}$
  - The condition number: $\kappa = \frac{k_N^2}{k_1^2} \rightarrow \infty$ for $k_N \rightarrow \infty$
- Gradient in $H_0^1(\Omega)$: $\nabla^{L_2}\mathcal{J} = -\Delta_0^{-1}[\Delta\Phi - g \in L_2(\Omega)]$
  - Hessian eigenvalues (Fourier space): $\{1, 1, \dots, 1\}$
  - The condition number: $\kappa = 1$ independent of k_N

APPLICATIONS TO MULTI–SCALE SYSTEMS



ANIMATION — Courtesy of Dr. Peter Blossey



Wall–normal adjoint velocity

Issues:

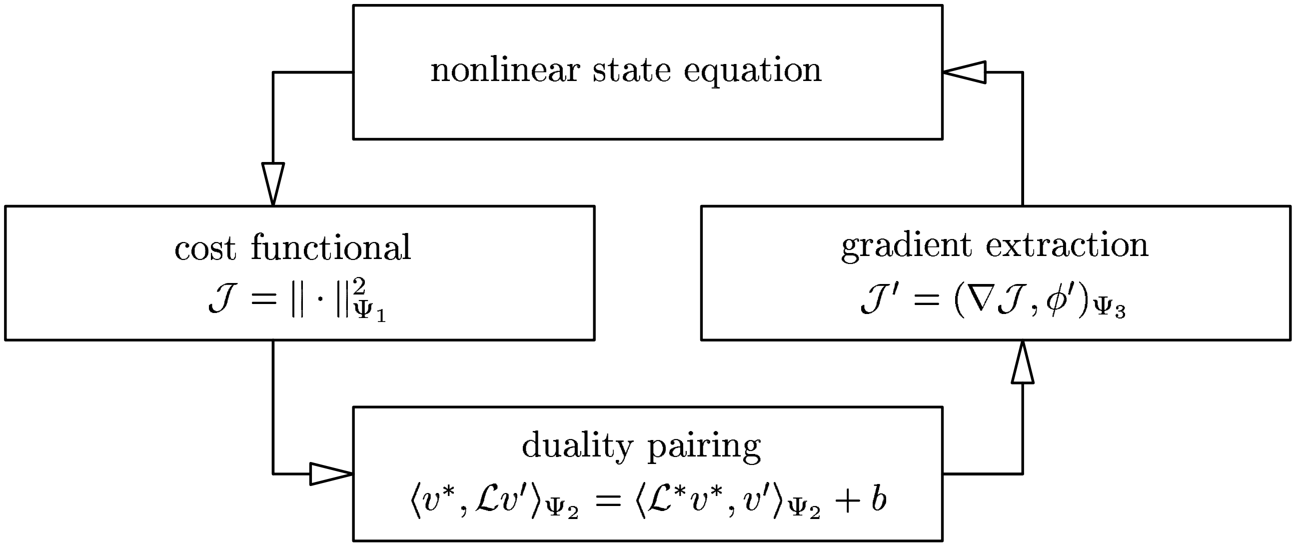
- Multiple local solutions
- Nonsmooth gradients
- Rapid growth of adjoint fields in regions with steep gradients

In order to better handle multi–scale systems we need REGULARIZATION to:

- focus optimization procedure on the physics of interest
- define numerically tractable adjoint operators
- extract properly preconditioned gradients

AVAILABLE OPTIONS – THE BIG PICTURE

(Protas et al. JCP 195)



- regularization of the state (evolution) equation itself
- different choices of the brackets $\| \cdot \|_{\Psi_1}$, $\langle \cdot, \cdot \rangle_{\Psi_2}$, and $\langle \cdot, \cdot \rangle_{\Psi_3}$
  - usually $L_2(0,T;L_2(\Omega_i))$, e.g. $\|u\|^2_{L_2(\Omega)} = \int_\Omega u^2 d\Omega$
  - now $H^p(0,T;H^q(\Omega_i))$, e.g. $\|u\|^2_{H^p(\Omega)} = \sum_{i=0}^p \int_\Omega (\nabla^{(i)} u)^2 d\Omega$

CONTINUOUS SETTING

KURAMOTO–SIVASHINSKY EQUATION

(nonlinear, multiscale, chaotic 1D model)

- The primitive form

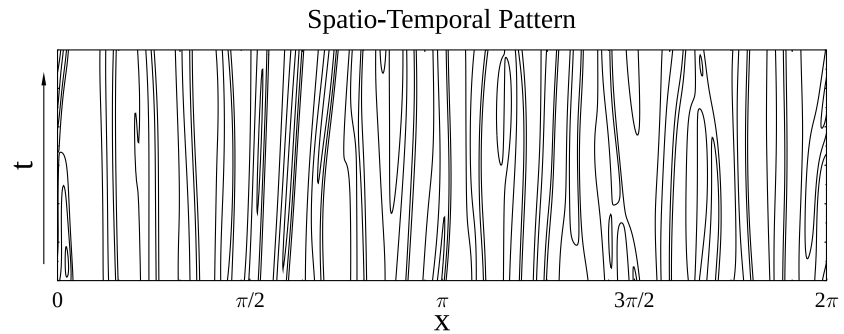
$$\begin{cases} \partial_t v + 4\partial_x^4 v + \kappa\left(\partial_x^2 v + v\partial_x v\right) = 0, \\ \partial_x^i v(0,t) = \partial_x^i v(2\pi,t),\ i = 0,\ldots,3, \\ v(x,0) = \varphi. \end{cases}$$

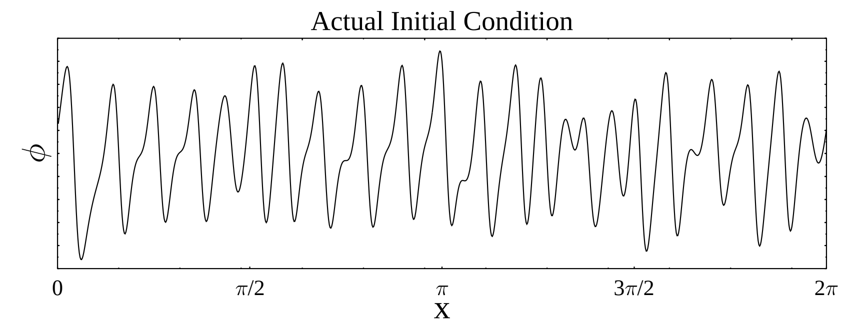
- The "streamfunction" form

$$\begin{cases} \partial_t u + 4\partial_x^4 u + \kappa\left[\partial_x^2 u + \frac{1}{2}\left(\partial_x u\right)^2\right] = 0, \\ \partial_x^i u(0,t) = \partial_x^i u(2\pi,t),\ i = 0,\ldots,3, \\ u(x,0) = \partial_x^{-1}\varphi \triangleq \psi. \end{cases}$$

- The "vorticity" form

$$\begin{cases} \partial_t w + 4\partial_x^4 w + \kappa\left(\partial_x^2 w + v\partial_x w + w^2\right) = 0, \\ w = \partial_x v \Rightarrow v = \partial_x^{-1} w, \\ \partial_x^i w(0,t) = \partial_x^i w(2\pi,t),\ i = 0,\ldots,3, \\ w(x,0) = \partial_x \varphi \triangleq \phi. \end{cases}$$





STATE ESTIMATION FOR THE KURAMOTO–SIVASHINSKY EQUATION

prototypical 4DVAR — Standard Derivation of the Adjoint Algorithm

$$\Psi_1 = \Psi_2 = \Psi_3 = L_2(0,T;L_2(\Omega_i))$$

- The optimization problem: $\min_{\varphi} \mathcal{J}(\varphi)$,

$$\mathcal{J}(\varphi) = \left\| (\mathcal{H}v - y) \right\|^2_{L_2(0,T;L_2(\Omega))} = \frac{1}{2}\int_0^T \int_0^{2\pi} [\mathcal{H}v - y]^2 \, dxdt$$

- Optimality condition – vanishing of the Gâteaux differential:

$$\mathcal{J}'(\varphi_{opt};\varphi') = \int_0^T \int_0^{2\pi} \mathcal{H}^*(\mathcal{H}v - y)\, v'(\varphi')\, dxdt = 0$$

- duality pairing (all boundary terms vanish due to periodicity):

$$\underbrace{\left\langle \mathcal{L}v', v^* \right\rangle_{L_2(0,T;L_2(\Omega))}}_{0} = \underbrace{\left\langle v', \mathcal{L}^* v^* \right\rangle_{L_2(0,T;L_2(\Omega))}}_{\mathcal{J}'(\varphi;\varphi')} - \int_0^{2\pi} \varphi' \, v^* \Big|_{t=0} dx.$$

- definition of adjoint system:

$$\begin{cases} \mathcal{L}^* v^* = -\partial_t v^* + 4\partial_x^4 v^* + \kappa\left(\partial_x^2 v^* + v\partial_x v^*\right) = \mathcal{H}^*(\mathcal{H}v - y) = f, \\ \frac{\partial^i v^*}{\partial x^i}(0,t) = \frac{\partial^i v^*}{\partial x^i}(2\pi,t), \; j = 0,\ldots,3, \\ v^*(x,T) = 0, \end{cases}$$

- cost functional gradient:

$$\mathcal{J}'(\varphi;\varphi') = \int_0^{2\pi} \varphi' v^* \Big|_{t=0} dx = (\nabla \mathcal{J}^{L_2}, \varphi')_{L_2} \Longrightarrow \nabla \mathcal{J}^{L_2} = v^* \Big|_{t=0}$$

OPTION #1 – DEFINITION OF ADJOINT EQUATION FOR ALTERNATIVE FORMS OF STATE EQUATION

- evolution governed by the vorticity equation $\Longrightarrow$
  the cost functional minimized w.r.t. the vorticity variable

$$\mathcal{J}'(\varphi;\varphi') = \mathcal{J}_\phi'(\partial_x\varphi;\partial_x\varphi') = -\int_0^T\int_0^{2\pi}\left[\partial_x^{-1}\mathcal{H}^*\left(\mathcal{H}v - y\right)\right]\partial_x w'(\varphi')\,dxdt.$$

- duality pairing for the vorticity equation:

$$\underbrace{\left\langle \mathcal{M}w', w^*\right\rangle_{L_2(0,T;L_2(\Omega))}}_{0} = \underbrace{\left\langle w', \mathcal{M}^*w^*\right\rangle_{L_2(0,T;L_2(\Omega))}}_{\mathcal{J}_\phi'(\partial_x\varphi;\partial_x\varphi')} - \int_0^{2\pi} w^*\Big|_{t=0}\partial_x h\,dx,$$

- "vorticity" adjoint equation:

$$\begin{cases} \mathcal{M}^*w^* = -\partial_t w^* + 4\partial_x^4 w^* + \kappa\left(\partial_x^2 w^* - \partial_x^{-1}(v\partial_x^2 w^*)\right) \\ \qquad\qquad = -\partial_x^{-1}\mathcal{H}^*\left(\mathcal{H}v - y\right), \\ \partial_x^i w^*(0,t) = \partial_x^i w^*(2\pi,t),\ i = 0,\dots,3, \\ w^*(x,T) = 0. \end{cases}$$

- cost functional gradient:

$$\nabla\mathcal{J}_\phi = w^*\Big|_{t=0}$$

DIFFERENTIATION AND TAKING ADJOINT DO NOT COMMUTE: $(\partial_x\mathcal{L})^* \neq \partial_x\mathcal{L}^*$

OPTION #2 – ALTERNATIVE DEFINITIONS OF THE COST FUNCTIONAL

(with the same set of available measurements)

In general: $\mathcal{J}^{H^k}(\varphi) = \left\| \mathcal{H} v(\varphi) - y \right\|^2_{L_2(0,T;H^k(\Omega))}$

- $k = 0$ — *standard formulation (discussed before)*
- $k = 1$ — differentiation of observations

$$\mathcal{J}'^{H^1}(\varphi;\varphi') = -\int_0^T \int_0^{2\pi} \partial_x^2 \mathcal{H}^* \left(\mathcal{H} v - y\right) v'(\varphi')\, dx\, d\tau$$

$$\Longrightarrow f_1 = -\partial_x^2 \mathcal{H}^* \left(\mathcal{H} v - y\right)$$

- $k = -1$ — integration of observations

$$\mathcal{J}'^{H^{-1}}(\varphi;\varphi') = -\int_0^T \int_0^{2\pi} \partial_x^{-2} \mathcal{H}^* \left(\mathcal{H} v - y\right) v'(\varphi')\, dx\, d\tau$$

$$\Longrightarrow f_{-1} = -\partial_x^{-2} \mathcal{H}^* \left(\mathcal{H} v - y\right)$$

DIFFERENT FUNCTIONALS RESULT IN DIFFERENT RHS FORCING TERMS IN THE ADJOINT SYSTEMS

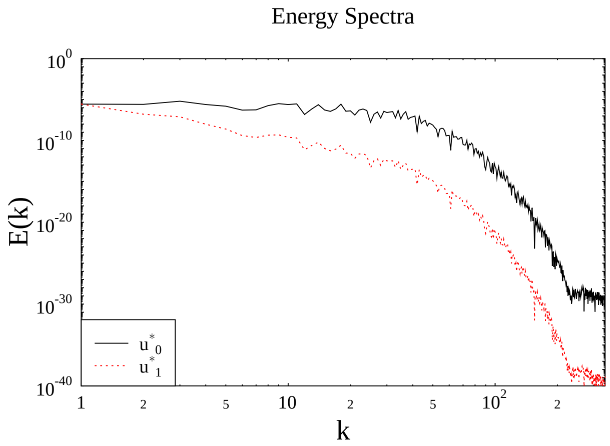
OPTION #3 – DEFINITION OF THE ADJOINT STATE WITH ALTERNATIVE DUALITY PAIRINGS

- definition of new duality pairing: $\left\langle z_1, z_2 \right\rangle_{L_2(0,T;H^1(\Omega))} \triangleq \int_0^T \int_0^{2\pi} (\partial_x z_1)(\partial_x z_2)\, dx\, dt$

$$\underbrace{\left\langle \mathcal{L}v', v^{*,H^1} \right\rangle_{L_2(0,T;H^1(\Omega))}}_{0} = \underbrace{\left\langle v', \mathcal{L}_1^* v^{*,H^1} \right\rangle_{L_2(0,T;H^1(\Omega))}}_{\mathcal{J}'_\varphi(\varphi;\varphi')} - \int_0^{2\pi} \partial_x v_1^* \Big|_{t=0} \partial_x \varphi'\, dx.$$

- new adjoint equation:

$$\begin{cases} \mathcal{L}^{*,H^1} v^{*,H^1} = -\partial_t v^{*,H^1} + 4\partial_x^4 v^{*,H^1} + \\ \qquad\qquad \kappa\left[\partial_x^2 v^{*,H1} + \partial_{xx}^{-1}(v\partial_{xxx} v^{*,H^1})\right] \\ \qquad\qquad = -\partial_x^{-2}\mathcal{H}^*(\mathcal{H}v - y), \\ \partial_x^i v^{*,H^1}(0,t) = \partial_x^i v^{*,H^1}(2\pi,t),\ j = 0,\dots,3, \\ v^{*,H^1}(x,T) = 0. \end{cases}$$



- new (smoother) gradient: $\nabla \mathcal{J}^{H^1} = v^{*,H^1}\big|_{t=0}$

OPTION #4 – ALTERNATIVE (PRECONDITIONED) DEFINITIONS OF THE GRADIENT

Riesz representation theorem: $\mathcal{J}'(\varphi;\varphi') = \int_0^{2\pi} \varphi' \, v^* \big|_{t=0} dx = (\nabla \mathcal{J}^X, \varphi')_X$, e.g.,

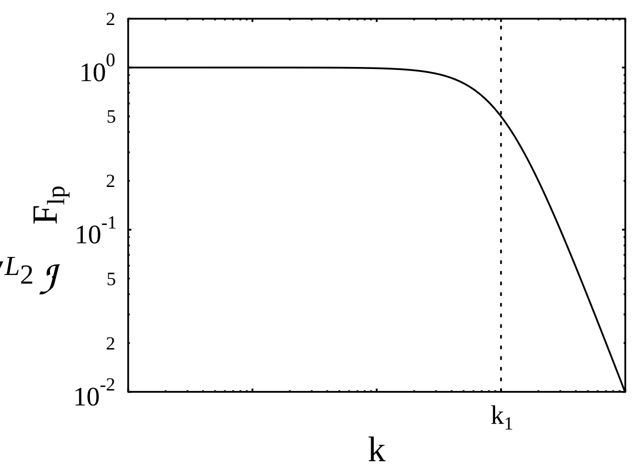
$$\left(z_1, z_2\right)_{W^{l_1,l_{-1}}} = \frac{l_{-1}^2}{(1+l_1^2)(1+l_{-1}^2)} \int_0^{2\pi} \left[z_1 z_2 + \frac{l_1^2 l_{-1}^2}{l_1^2 + l_{-1}^2} (\partial_x z_1)(\partial_x z_2) + \frac{1}{l_1^2 + l_{-1}^2} (\partial_x^{-1} z_1)(\partial_x^{-1} z_2) \right] dx$$

Assuming $l_{-1} \to \infty$, we obtain:

$$\left(\nabla^{W^{l_1,\infty}} \mathcal{J}, \varphi'\right)_{W^{l_1,\infty}} = \int_0^{2\pi} \varphi' \, v^* \Big|_{t=0} dx = \frac{1}{1+l_1^2} \int_0^{2\pi} \left(z_1 z_2 + l_1^2 \partial_x z_1 \, \partial_x z_2\right) dx \quad \Longrightarrow \quad \begin{cases} \frac{1}{1+l_1^2} \left[1 - l_1^2 \partial_x^2\right] \nabla^{W^{l_1,\infty}} \mathcal{J} = v^* \big|_{t=0} \\ \nabla^{W^{l_1,\infty}} \mathcal{J}(0) = \nabla^{W^{l_1,\infty}} \mathcal{J}(2\pi) \end{cases}$$

In Fourier space:

$$\widehat{\nabla^{W^{l_1,\infty}} \mathcal{J}} = \frac{1/l_1^2}{k^2 + 1/l_1^2} \, \hat{v}^* \Big|_{t=0}$$



Low–pass filter with the cut–off length l_1*:* $\lim_{l_1 \to 0} \nabla^{W^{l_1,\infty}} \mathcal{J} = \nabla^{L_2} \mathcal{J}$

Gradient extraction in Sobolev spaces
can be applied as post-processing

SUMMARY – ZOO OF ADJOINTS

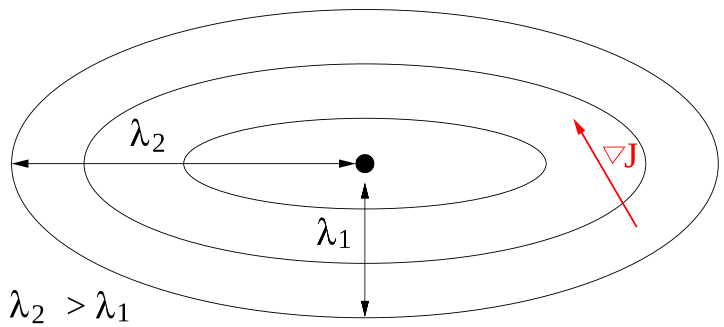
Perturbation system	Duality pairing	Adjoint system	$\nabla^{L_2}\mathcal{J}$
$\mathcal{L}v' = 0$ $v'(0) = \phi$	$\langle\cdot,\cdot\rangle_{L_2(0,T;L_2(\Omega))}$	$\mathcal{L}^* v^* = f = \mathcal{H}^*(\mathcal{H}v - y)$, $v^*(T) = 0$	$v^*\big\vert_{t=0}$
$\mathcal{M}w' = \partial_x \mathcal{L}v' = 0$, $w'(0) = \varphi = \partial_x \phi$	$\langle\cdot,\cdot\rangle_{L_2(0,T;L_2(\Omega))}$	$\mathcal{M}^* w^* =$ $\partial_x^{-1}\mathcal{L}^*(\partial_x w^*) = -\partial_x^{-1} f$, $w^*(T) = 0$	$-\partial_x w^*\big\vert_{t=0}$
$\mathcal{K}u' = \partial_x^{-1}\mathcal{L}v' = 0$, $u'(0) = \psi = \partial_x^{-1}\phi$	$\langle\cdot,\cdot\rangle_{L_2(0,T;L_2(\Omega))}$	$\mathcal{K}^* u^* =$ $\partial_x \mathcal{L}^*(\partial_x^{-1} u^*) = -\partial_x f$, $u^*(T) = 0$	$-\partial_x^{-1} u^*\big\vert_{t=0}$
$\mathcal{L}v' = 0$, $v'(0) = \phi$	$\langle\cdot,\cdot\rangle_{L_2(0,T;H^1(\Omega))}$	$\mathcal{L}^{*,H^1} v^{*,H^1} =$ $\partial_x^{-2}\mathcal{L}^*(\partial_x^2 v^{*,H^1}) = \partial_x^{-2} f$, $v^{*,H^1}(T) = 0$	$-\partial_x^2 v^{*,H^1}\big\vert_{t=0}$
$\mathcal{L}v' = 0$, $v'(0) = \phi$	$\langle\cdot,\cdot\rangle_{L_2(0,T;H^{-1}(\Omega))}$	$\mathcal{L}^{*,H^{-1}} v^{*,H^{-1}} =$ $\partial_x^2\mathcal{L}^*(\partial_x^{-2} v^{*,H^{-1}}) = \partial_x^2 f$, $v^{*,H^{-1}}(T) = 0$	$-\partial_x^{-2} v^{*,H^{-1}}\big\vert_{t=0}$

EXACT PRECONDITIONING

AN ANALYTICAL EXAMPLE

- assume:
  - $\mathcal{H} = Id$ (complete observations) ,
  - $v \equiv 0$ (first iteration)
- gradient extracted using an L_2 inner product

$$\left(\widehat{\nabla^{L_2}\mathcal{J}}\right)_k = \frac{\hat{\varphi}_k^{act}}{2(4k^4 - \kappa k^2)}\left(1 - e^{-2(4k^4-\kappa k^2)T}\right) = \hat{F}_k \hat{\varphi}_k^{act}$$



Exact Preconditioning

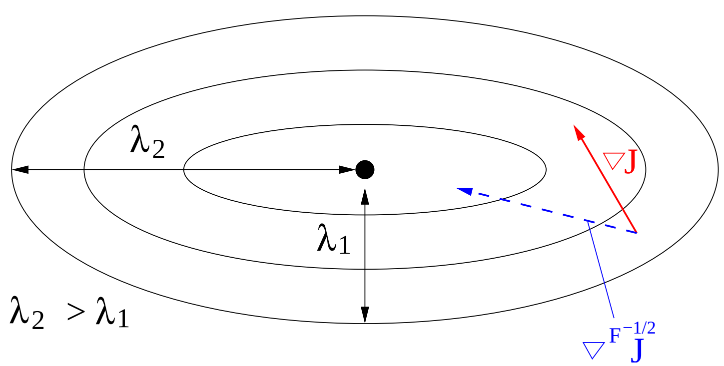
an Analytical Example

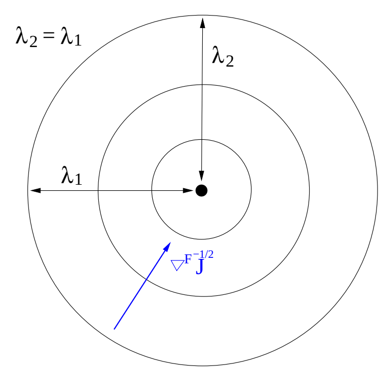
- assume:
  - $\mathcal{H} = Id$ (complete observations) ,
  - $v \equiv 0$ (first iteration)
- gradient extracted using an L_2 inner product

$$\left(\widehat{\nabla^{L_2} \mathcal{J}}\right)_k = \frac{\hat{\varphi}_k^{act}}{2(4k^4 - \kappa k^2)} \left(1 - e^{-2(4k^4 - \kappa k^2)T}\right) = \hat{F}_k \hat{\varphi}_k^{act}$$

- gradient extracted using the inner product $(v_1, v_2)_{F^{-1/2}}$, where $\hat{F}_k = \sqrt{\frac{\left(1 - e^{-2(4k^4 - \kappa k^2)T}\right)}{2(4k^4 - \kappa k^2)}}$:

$$\left(\widehat{\nabla^{F^{-1/2}} \mathcal{J}}\right)_k = \hat{\varphi}_k^{act}$$



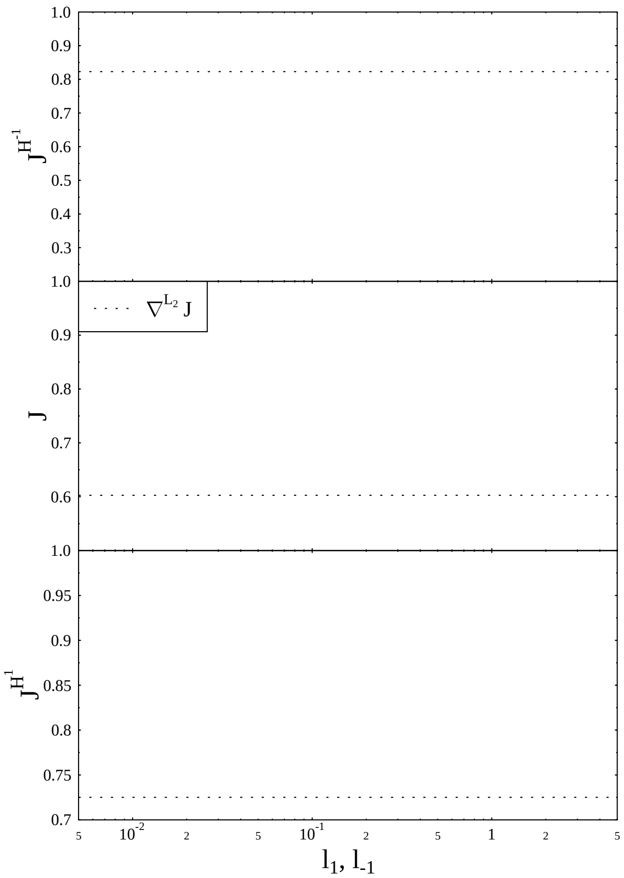


Computational results I

Minimizing $\mathcal{J}^{L_2}(\varphi) = \left\| \mathcal{H} v(\varphi) - y \right\|^2_{L_2(0,T;L_2(\Omega))}$

first iteration, zero initial guess

$$\mathcal{J}'(\varphi;\varphi') = (\nabla \mathcal{J}^{L_2}, \varphi')_{L_2}$$
$$= \int_0^{2\pi} (z_1\, z_2)\, dx$$



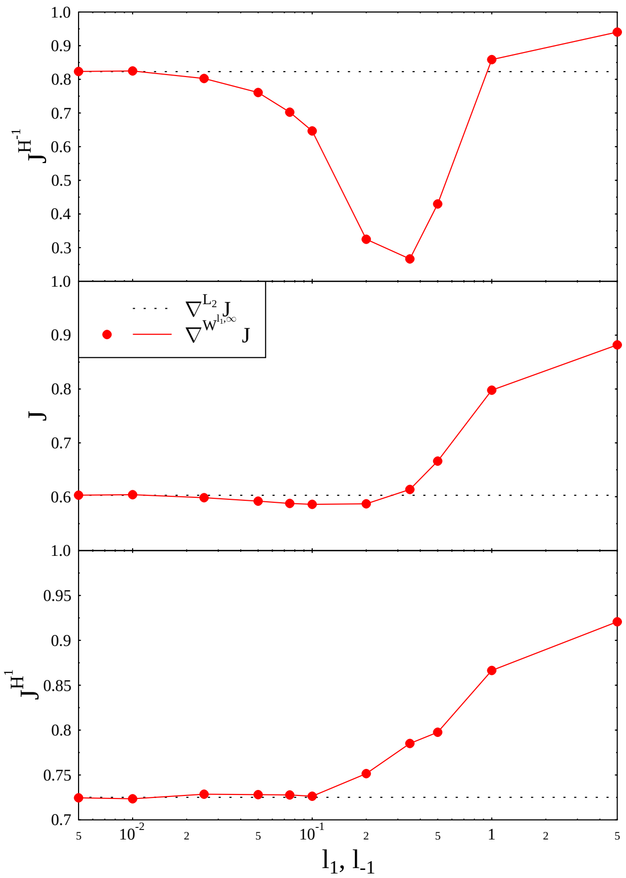
Computational results I

Minimizing $\mathcal{J}^{L_2}(\varphi) = \left\|\mathcal{H}v(\varphi) - y\right\|^2_{L_2(0,T;L_2(\Omega))}$

first iteration, zero initial guess

$$\mathcal{J}'(\varphi;\varphi') = (\nabla\mathcal{J}^{L_2},\varphi')_{L_2} = \int_0^{2\pi} (z_1\, z_2)\, dx$$

$$\mathcal{J}'(\varphi;\varphi') = (\nabla^{W^{l_1,\infty}}\mathcal{J},\varphi')_{W^{l_1,\infty}} = \frac{1}{1+l_1^2}\int_0^{2\pi} \left(z_1 z_2 + l_1^2 \partial_x z_1\, \partial_x z_2\right)\, dx$$

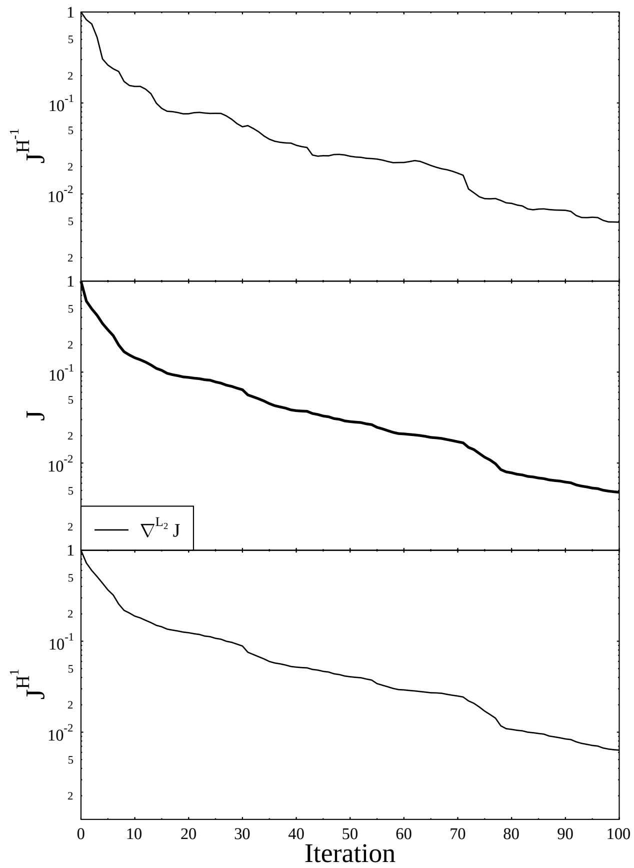


COMPUTATIONAL RESULTS II

Minimizing $\mathcal{J}^{L_2}(\varphi) = \left\| \mathcal{H} v(\varphi) - y \right\|^2_{L_2(0,T;L_2(\Omega))}$

many iterations, zero initial guess

- $\nabla^{L_2} \mathcal{J}$
-
-
-

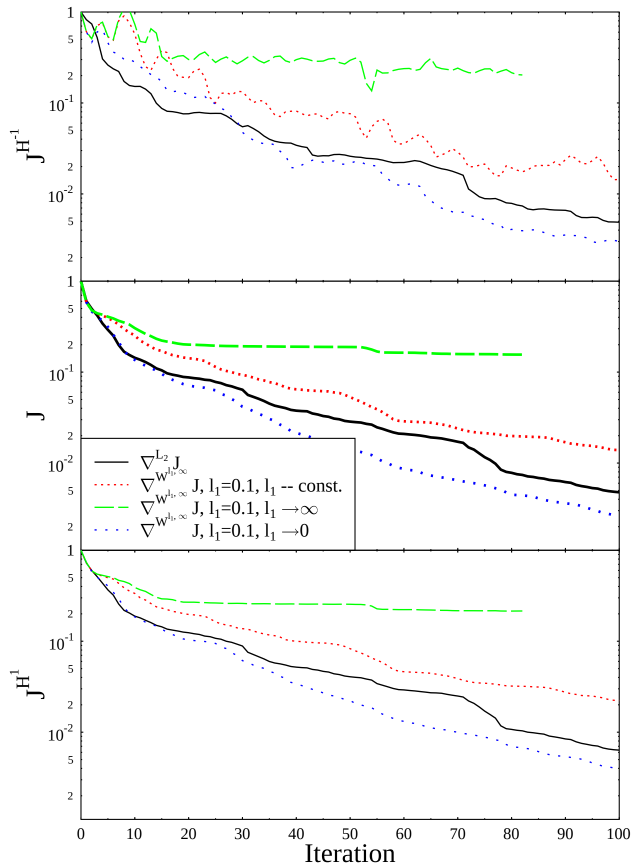


COMPUTATIONAL RESULTS II

Minimizing $\mathcal{J}^{L_2}(\varphi) = \left\| \mathcal{H} v(\varphi) - y \right\|^2_{L_2(0,T;L_2(\Omega))}$

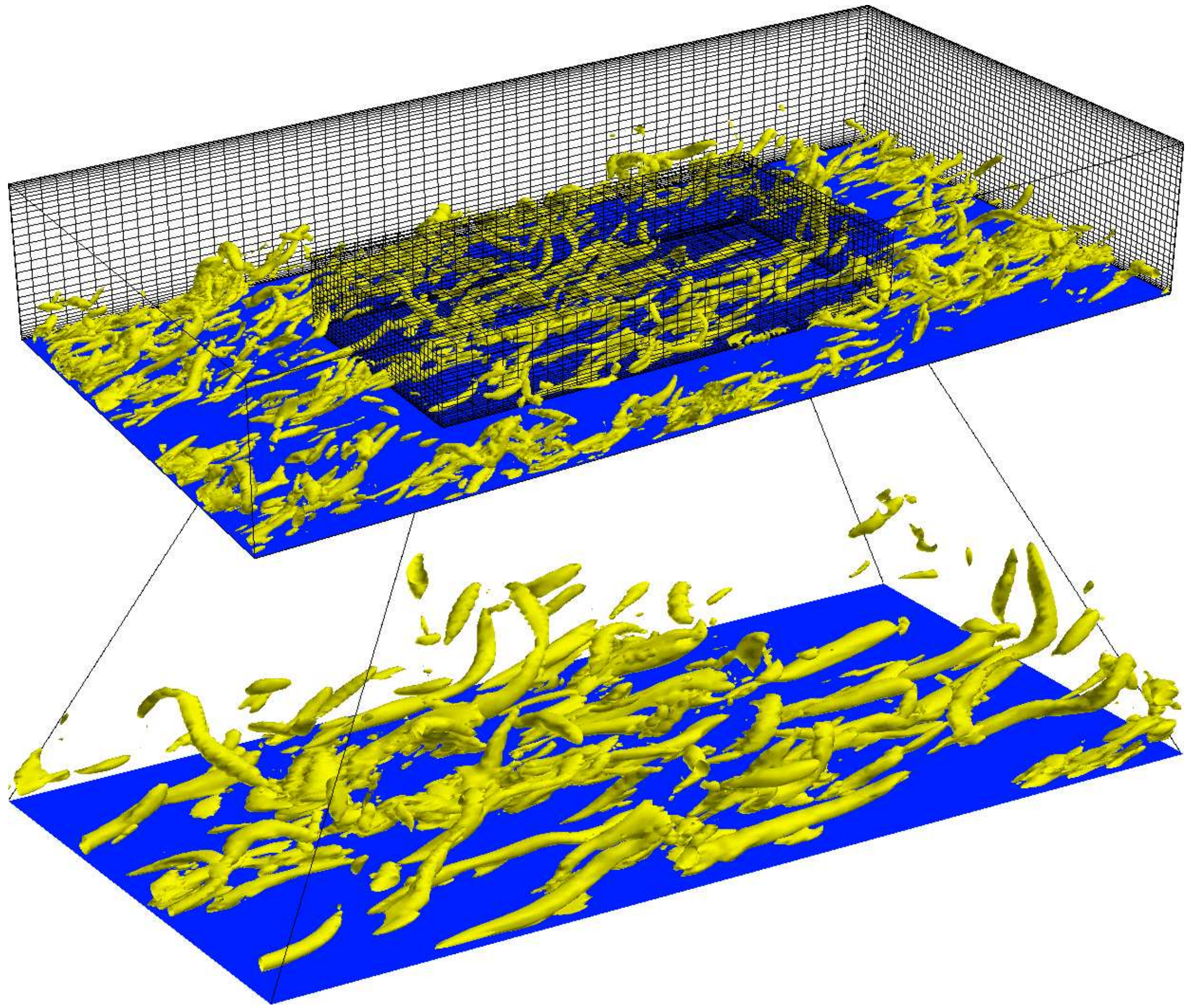
many iterations, zero initial guess

- $\nabla^{L_2} \mathcal{J}$
- $\nabla^{W^{l_1,\infty}} \mathcal{J}$, l_1 — constant
- $\nabla^{W^{l_1,\infty}} \mathcal{J}$, $l_1 \to \infty$ with n
- $\nabla^{W^{l_1,\infty}} \mathcal{J}$, $l_1 \to 0$ with n

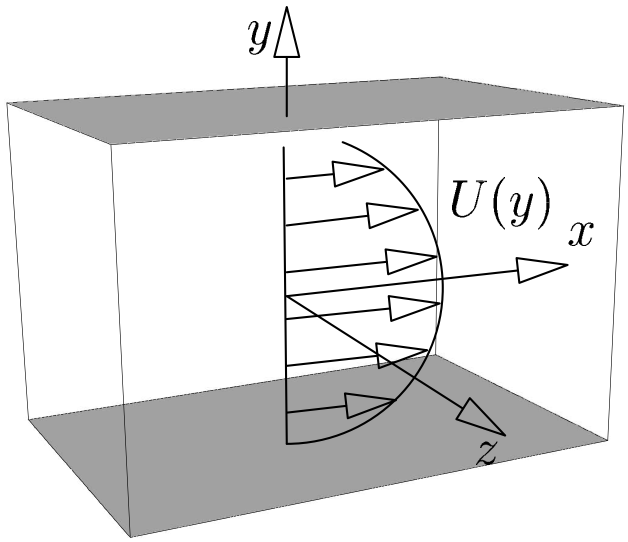


RECONSTRUCTION OF A TURBULENT CHANNEL FLOW I

WALL SHEAR — THE "FOOTPRINT" OF STREAKY STRUCTURES IN THE BOUNDARY LAYER



Reconstruction of a Turbulent Channel Flow II



Navier–Stokes system:

$$\begin{cases} \mathcal{N}(\mathbf{q}) = \begin{pmatrix} \frac{\partial u_j}{\partial x_j} \\ \frac{\partial u_i}{\partial t} + \frac{\partial u_j u_i}{\partial x_j} - \nu \frac{\partial u_i}{\partial x_j} + \frac{\partial p}{\partial x_i} \end{pmatrix} = 0 \\ \mathbf{u}|_{t=0} = \Phi \text{ in } \Omega, \\ \mathbf{u}(0,y,z) = \mathbf{u}(2\pi L_x, y, z); \\ \mathbf{u}(x,y,0) = \mathbf{u}(x,y,2\pi L_z) \\ \mathbf{u}(x,\pm 1, x) = 0 \end{cases}$$

- Solver: finite–differences/spectral in space, hybrid implicit/explicit in time
- wall shear and wall pressure measurements

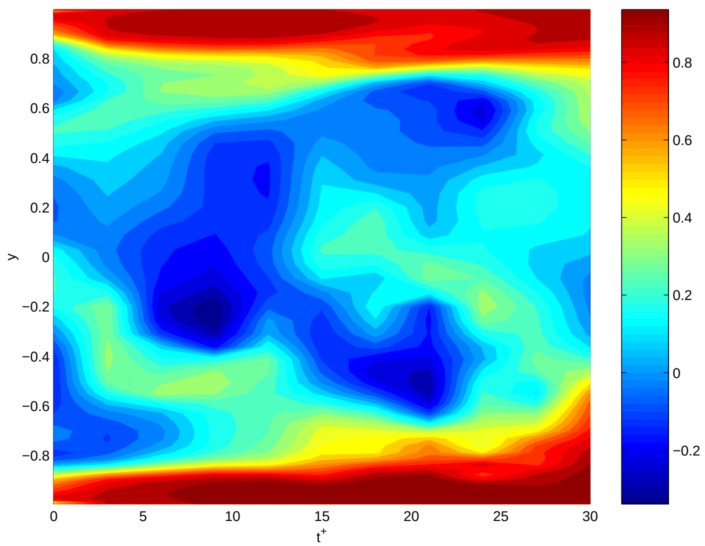
- constant mass flux
- turbulent flow at $Re_\tau = 100$

$$\mathcal{J}(\Phi) = \frac{1}{2}\int_0^T \left[\alpha_1 \left\| \frac{\partial u_1}{\partial x_2} - m_1 \right\|^2_{\Gamma_2^\pm} + \alpha_2 \left\| p - m_2 \right\|^2_{\Gamma_2^\pm} + \alpha_3 \left\| \frac{\partial u_3}{\partial x_2} - m_3 \right\|^2_{\Gamma_2^\pm} \right] dt,$$

Reconstruction of a Turbulent Channel Flow III

Different strategies for gradient extraction (Option # 4)

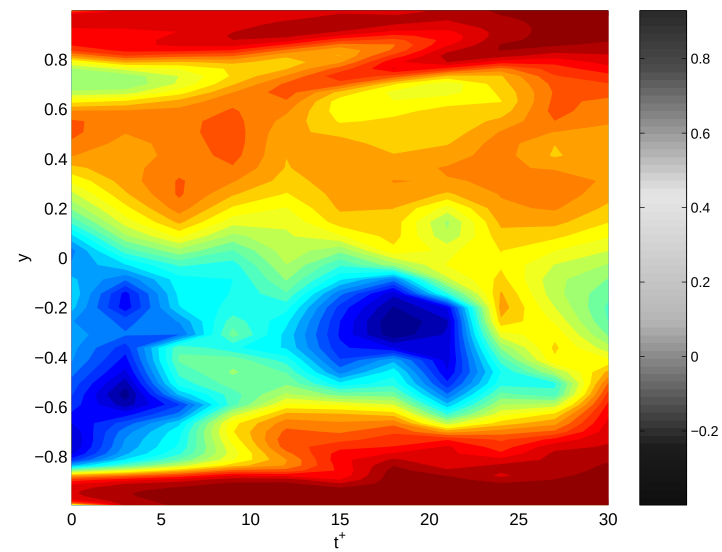
L_2 Gradient extraction



$$\mathcal{J}'(\Phi;\mathbf{h}) = -\int_\Omega \mathbf{u}^*\Big|_{t=0} \cdot \mathbf{h}\, d\Omega = \left(\boldsymbol{\nabla}\mathcal{J}^{L_2(\Omega)}, \mathbf{h}\right)_{L_2(\Omega)}$$

$$\Longrightarrow \boldsymbol{\nabla}\mathcal{J}^{L_2(\Omega)} = -\mathbf{u}^*\Big|_{t=0}$$

H^1 Gradient extraction



$$\mathcal{J}'(\Phi;\mathbf{h}) = \frac{1}{1+l_1^2}\int_\Omega \left(\boldsymbol{\nabla}\mathcal{J}^{H^1}\cdot\mathbf{h} + l_1^2\,\partial_{\mathbf{x}}\boldsymbol{\nabla}\mathcal{J}^{H^1}\cdot\partial_{\mathbf{x}}\mathbf{h}\right) d\Omega$$

$$\Longrightarrow \begin{cases} \overbrace{\frac{1}{1+l_1^2}\left[1+l_1^2\,\Delta\right]}^{\text{Helmholtz operator}}\boldsymbol{\nabla}\mathcal{J}^{H^1} = -\mathbf{u}^*\Big|_{t=0} \\ \boldsymbol{\nabla}\mathcal{J}^{H^1}\Big|_\Gamma = 0 \end{cases}$$

H^1 gradients lead to better flow reconstruction than L_2 gradients

GRADIENT EXTRACTION IN BANACH SPACES — THEORY (I)

- Consider a Banach space $\mathbf{X}$ (without Hilbert structure!)

$$\mathcal{J}'(\varphi;\varphi') = \int_0^{2\pi} \varphi' \, v^* \big|_{t=0} \, dx = \langle \nabla^{\mathbf{X}} \mathcal{J}, \varphi' \rangle_{\mathbf{X}^* \times \mathbf{X}}$$

  Note that $\mathbf{X}^*$ (the dual space) is usually "bigger" than $\mathbf{X}$

  Hence $\nabla^{\mathbf{X}} \mathcal{J} \notin \mathbf{X}$ is not an acceptable descent direction !!!

  No Riesz Theorem in Banach spaces ...

- Alternative definition of the descent direction g ensuring that $g \in \mathbf{X}$

$$G = \operatorname{argmax}_{\|\varphi'\|_{\mathbf{X}}=1} \langle \nabla^{\mathbf{X}} \mathcal{J}, \varphi' \rangle_{\mathbf{X}^* \times \mathbf{X}} = \max \left[\int_0^{2\pi} \varphi' \, v^* \big|_{t=0} \, dx + \mu \|\varphi'\|_{\mathbf{X}} \right]$$

- For instance, when $\mathbf{X} = W^{p,q}(\Omega)$ with $\|z\|_{W^{p,q}} = \int_0^{2\pi} |z|^q + l^{qp} |\partial_x^p u|^q \, dx$,

$$\begin{cases} p|u|^{(p-2)} u + p \partial_x^q \left(|\partial_x^q u|^{(p-2)} \partial_x^q u \right) = -v^* \big|_{t=0} \\ \partial_x^m u\big|_{x=0} = \partial_x^m u\big|_{x=2\pi} = 0 \end{cases} \qquad p\text{–LAPLACIAN equation}$$

- The spaces $W^{p,q}(\Omega)$ are the natural setting for many nonlinear PDEs

Gradient Extraction in Banach Spaces — Theory (II)

- Nonlinear equations with p–Laplacian
  - p–Laplacian as a nonlinear generalization of the Laplacian
  - for a positive–definite RHS, a unique positive-definite solution exists
  - solution readily obtained with nonlinear FEM
- Special cases:
  - $p=0,\ q=1 \quad \Rightarrow \quad W^{p,q}(\Omega)=L_1(\Omega)$ (not reflexive) ,

    The steepest descent doe not exist ...
  - $p=0,\ q>1 \quad \Rightarrow \quad W^{p,q}(\Omega)=L_q(\Omega)$,

    the gradient $G=-\operatorname{sgn}(w^*)\sqrt[q-1]{\frac{1}{\mu}\operatorname{sgn}w^* w^*}$
  - $p=0,\ q=\infty \quad \Rightarrow \quad W^{p,q}(\Omega)=L_\infty(\Omega)$,

    the gradient $G=-\operatorname{sgn}(w^*)$
  - $p=1,\ q=2 \quad \Rightarrow \quad W^{p,q}(\Omega)=H^1(\Omega)$,

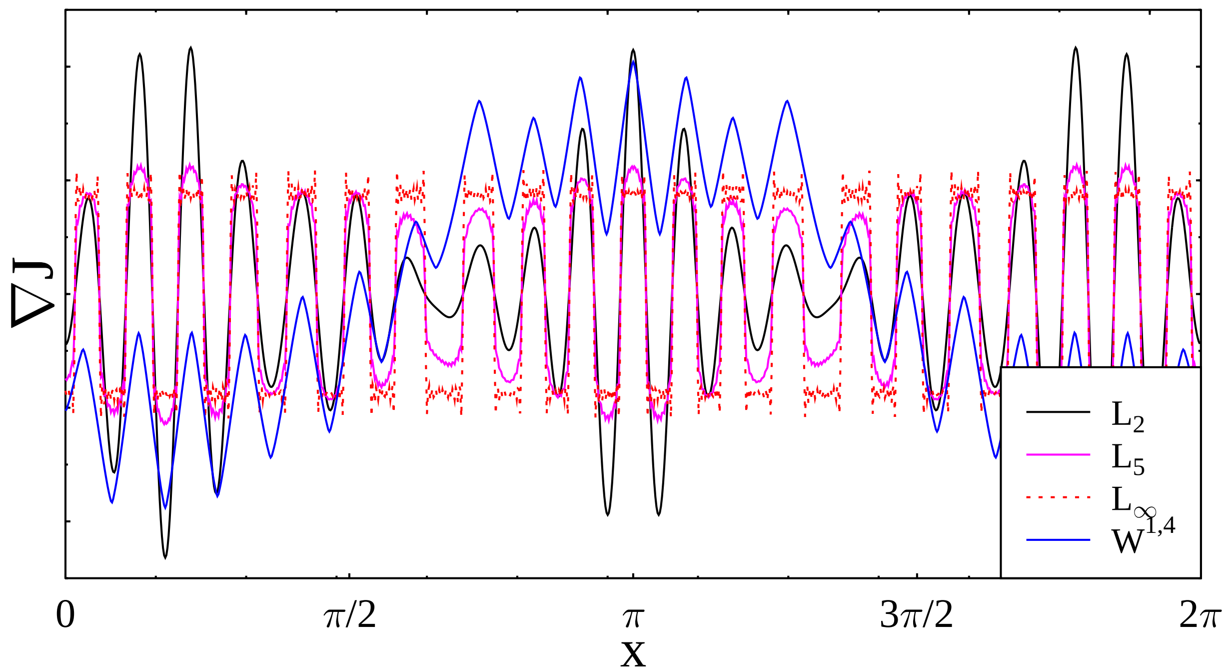
    the gradient $(I-l^2\Delta)G=-w^*,\ G|_0=G|_{2\pi}=0$

Gradient Extraction in Banach Spaces

Results for the Kuramoto–Sivashinsky Equation (I)

Comparison of the different (nonlinear) gradients

(the first iteration)



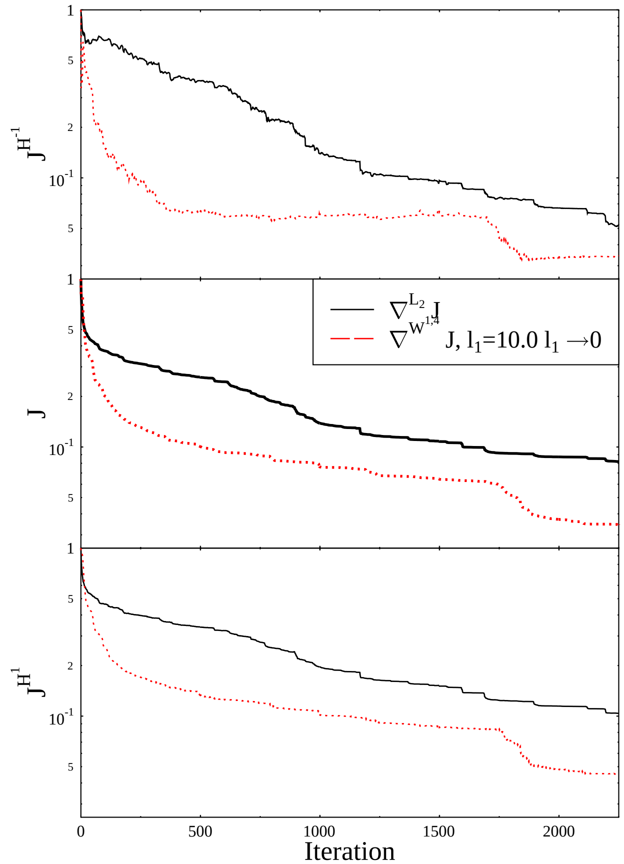
Gradient Extraction in Banach Spaces

Results for the Kuramoto–Sivashinsky Equation (II)

Minimizing $\mathcal{J}^{L_2}(\varphi) = \left\| \mathcal{H} v(\varphi) - y \right\|^2_{L_2(0,T;L_2(\Omega))}$

Very long optimization horizon (tough problem)

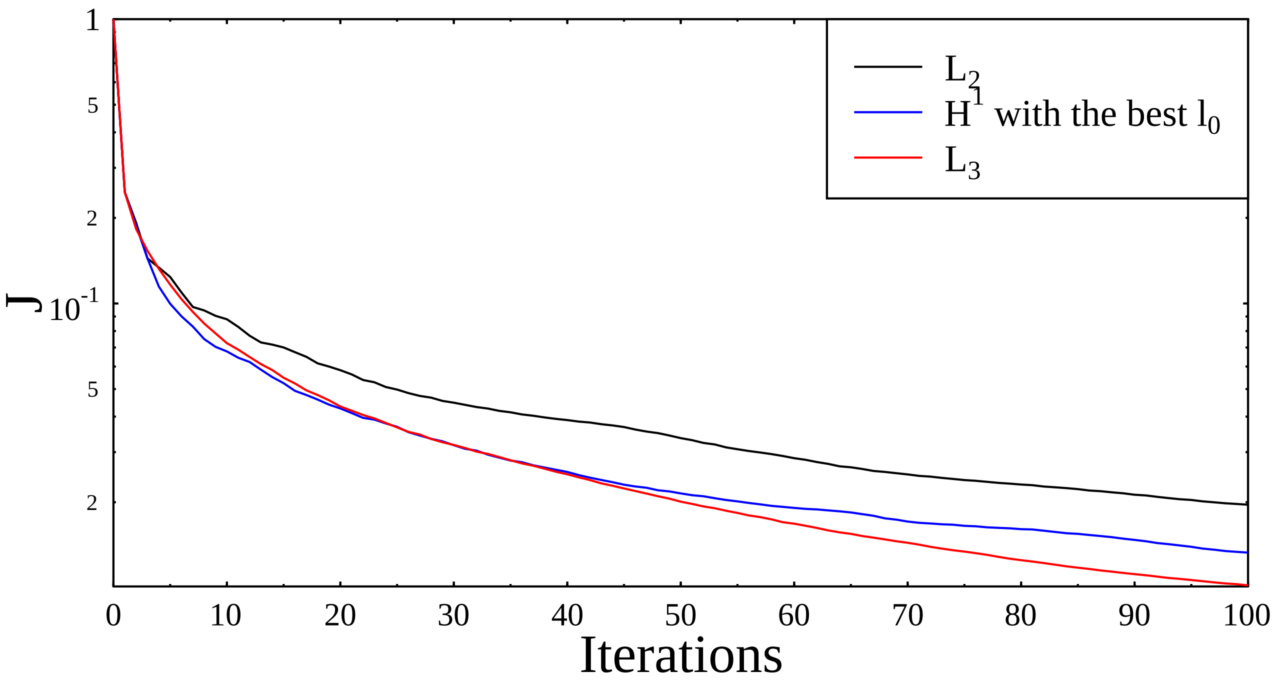
- $\nabla^{L_2} \mathcal{J}$
- $\nabla^{W^{1,4}} \mathcal{J},\ l_1 \to 0$ with n



Gradient Extraction in Banach Spaces

Reconstruction of a Turbulent Channel Flow (III)

Convergence of iterations with different gradients



Summary & Future Directions

- Extended formulation of an adjoint–based analysis allowing for a better treatment of multiscale systems;
  Examples:
  - derivation of numerically tractable adjoint operators
    $\Longrightarrow$ stability of sensitivity calculation
  - extraction of better preconditioned gradients
    $\Longrightarrow$ accelerated convergence of iterations
    $\Longrightarrow$ advantages of spaces without Hilbert structure (nonlinear preconditioning)
- Future directions:
  - Solution of Optimization Problems in non–standard spaces (e.g., Besov $B_s^{p,q}$, ...)
    $\Longrightarrow$ exploit adaptivity of gradient preconditioning
    $\Longrightarrow$ understand why! (numerical analysis)

More information at: http://math.mcmaster.ca/bprotas/research