Least-Squares Time Reversal of PDEs

Liam O'Connor[1,2]

[1]*Department of Engineering Sciences and Applied Mathematics,*
Northwestern University, Evanston, IL 60208 USA and
[2]*Center for Interdisciplinary Exploration and Research in Astrophysics,*
Northwestern University, Evanston, IL, 60201 USA

abstract

I. INTRODUCTION

Rayleigh-Bénard convection plays a foundational role in astrophysical and geophysical settings. The resulting buoyancy-driven flows regulate heat transfer and generate large-scale vortices [1]. Turbulent convection, which is associated with large Rayleigh numbers Ra, is difficult to simulate. State of the art simulations performed by [2] have reached Ra $\sim 10^{14}$ but estimates for the sun's convective zone and earth's interior are Ra $\sim 10^{16} - 10^{20}$ and Ra $\sim 10^{20} - 10^{30}$ respectively [3, 4].

II.   GENERALIZED PROBLEM SETUP

We define a generalized solution state vector $\bar{X}$ and associated initial value problem

$$\mathcal{F}[\bar{X}] \equiv \partial_t \bar{X} + L.\bar{X} + N[\bar{X}, \bar{X}] = 0. \tag{1}$$

Given a target end state $\bar{X}(\boldsymbol{x}, T)$, our aim is to recover the associated initial condition $\bar{X}(\boldsymbol{x}, 0)$. This is realized by performing gradient-based minimization of the functional

$$H[X, \bar{X}] = \frac{1}{2}\langle |X(\boldsymbol{x}, T) - \bar{X}(\boldsymbol{x}, T)|^2 \rangle = \frac{1}{2}\langle X'^2(\boldsymbol{x}, T) \rangle, \tag{2}$$

with respect to $X(\boldsymbol{x}, 0)$, where $X(\boldsymbol{x}, t)$ obeys (1). The operator $\langle \cdot \rangle$ denotes the domain-integrated average and the quantity $X'(\boldsymbol{x}, t) \equiv X(\boldsymbol{x}, t) - \bar{X}(\boldsymbol{x}, t)$. For the remainder of this paper, symbols with overbars (such as $\bar{X}(\boldsymbol{x}, T)$) will be used to denote the target solution, while their corresponding unannotated symbols (such as $X(\boldsymbol{x}, T)$) play the role of optimization parameters.

The Lagrangian associated with our initial value problem is given by

$$\mathcal{L} = \int_0^T \langle \chi \cdot \mathcal{F}[X] \rangle dt + H[X, \bar{X}].$$

Where $\chi(\boldsymbol{x}, T)$ is the Lagrangian Multiplier corresponding to $X(\boldsymbol{x}, T)$. Equating

$$\frac{\delta \mathcal{L}}{\delta X} = 0 \quad \text{and} \quad \frac{\delta \mathcal{L}}{\delta \chi} = 0$$

yields the original and adjoint systems

$$\mathcal{F}[X] = 0 \quad \text{and} \quad \mathcal{F}^\dagger[\chi, X] = 0$$

respectively. At some particular guess $X^n(\boldsymbol{x}, 0)$, we compute the functional derivative

$$\frac{\delta \mathcal{L}}{\delta X^n(\boldsymbol{x}, 0)} = -\chi^n(\boldsymbol{x}, 0)$$

to determine the direction of steepest descent. Naive gradient-descent can then be performed by computing an updated guess

$$X^{n+1}(\boldsymbol{x}, 0) = X^n(\boldsymbol{x}, 0) + \varepsilon \chi^n(\boldsymbol{x}, 0)$$

where ε is the step size.

In addition to this algorithm, more advanced gradient-based optimization techniques such as conjugate gradient and BFGS have been employed in conjunction with adjoint-looping to improve performance and accelerate convergence. In the context of our time-inversion optimization problem, such techniques traverse the space of possible initial conditions $X(\boldsymbol{x}, 0)$ by relying on local information pertaining to the objective functional H and its functional derivative $\frac{\delta H}{\delta X(\boldsymbol{x}, 0)}$. For this reason, the focus of our investigation is largely centered around the objective functional's dependence on the initial condition. More specifically, we study the levelsets of the objective functional because these surfaces characterize the directions of steepest descent at any arbitrary point in the space of initial conditions. In doing so, we highlight the challenges and deficits associated with using the least squares functional (2) to recover the initial condition $\bar{X}(\boldsymbol{x}, 0)$ belonging to a target end state $\bar{X}(\boldsymbol{x}, T)$.

III.   PRELIMINARY ANALYSIS

A.   Linearity and Reversiblity

In this section, we study linear initial value problems which do not include the nonlinear term in (1), i.e. $N[X, X] = N[\bar{X}, \bar{X}] = 0$. Therefore $\mathcal{F}[X'] = \mathcal{F}[X] - \mathcal{F}[\bar{X}]$. In these special cases, the objective functional (2) depends only on the deviation's initial condition $X'(x, 0)$ about any arbitrary target $\bar{X}(x, 0)$. For convenience when dealing with linear problems, we take $U(0) = U(T) = 0$ for all x. In the following section, we demonstrate that the objective functionals of nonlinear systems with nonzero targets are directly influenced by one's choice of U.

Time-reversible wave equations of the form

$$\partial_t u = \sum_{n \geq 1} c_n \partial_x^{2n-1} u \tag{3}$$

are trivial cases of this optimization problem, because the quantity $\langle u^2(x, t) \rangle$ is conserved. This implies that the objective functional's levelsets form sphere's in u'-space, because its value is given by $\langle u'^2(T) \rangle = \langle u'^2(0) \rangle$. The spatial derivatives in (4) are anti-self-adjoint (i.e. $L[u] = -L^\dagger[u]$), so the path of steepest descent is always in the radial derivation

$$\frac{\delta H}{\delta u(x, 0)} = -u'(x, 0)$$

In contrast, diffusive equations of the form

$$\partial_t u = \sum_{n \geq 1} b_n \partial_x^{2n} u \tag{4}$$

do not conserve $\langle u^2(x, t) \rangle$, and consequently their objective functionals depend on $u'(x, 0)$'s *direction* along with its magnitude. Suppose initial deviation can be written as

$$u'(x, 0) = \sum_{k \geq 0} a_k e^{ikx}. \tag{5}$$

Let's just consider the diffusion equation ($n = 1$) with diffusivity b_1. The end state is given by

$$u'(x, T) = \sum_{k \geq 0} a_k \exp(ikx - Tb_1 k^2) \tag{6}$$

and the nominal objective functional is given by

$$\langle |u'(x, T)|^2 \rangle = \pi \sum_{k \geq 0} \left[a_k^2 \exp\left(-2T \sum_{n \geq 1} b_n k^{2n} \right) \right]. \tag{7}$$

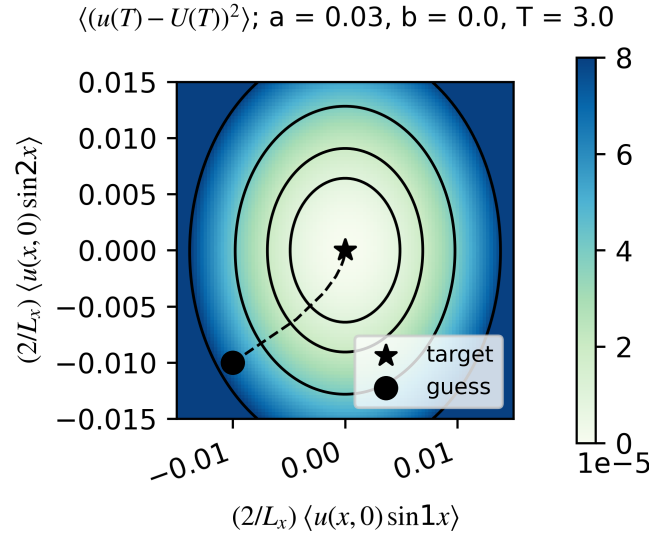Equation (7) gives us the elliptical heatmaps we observed previously, as shown in Figure 1.

$$\tag{8}$$

$\langle (u(T) - U(T))^2 \rangle$; a = 0.03, b = 0.0, T = 3.0

FIG. 1. Objective function and an example gradient descent path for the diffusion equation. The target simulation $U(t) = 0$ and the deviation is composed of two fourier modes, i.e. $u'(t) = a_1 \sin x + a_2 \sin 2x$. Because the system is linear self-adjoint, each subsequent guess in the gradient descent trajectory is constrained to lie in this 2D phase space.

Let's take a second-order spatial derivative of (6)

$$\partial_x^2 u'(x, T) = \sum_{k \geq 0} -k^2 a_k^2 \exp(ikx - Tb_1 k^2) \tag{9}$$

In general, applying a spatial derivative of order $2m$ to (6), yields

$$\partial_x^{2m} u'(x, T) = \sum_{k \geq 0} (-k^2)^m a_k^2 \exp(ikx - Tb_1 k^2) \tag{10}$$

We can leverage this fact to construct an objective functional whose levelsets are spherical.

Consider the following objective functional

$$H[u'(T)] = \langle |u'(x,T) - (Tb_1)\partial_x^2 u'(x,T) + \frac{(Tb_1)^2}{2!}\partial_x^4 u'(x,T) - \frac{(Tb_1)^3}{3!}\partial_x^6 u'(x,T)...|^2 \rangle$$

(11)

$$= \langle |\sum_{k\geq 0} a_k \exp(ikx - Tb_1 k^2) + \sum_{k\geq 0}(Tb_1 k^2)a_k\exp(ikx - Tb_1 k^2) + \sum_{k\geq 0}\frac{(Tb_1 k^2)^2}{2!}a_k\exp(ikx - Tb_1 k^2) + \sum_{k\geq 0}\frac{(Tb_1 k^2)^3}{3!}a_k\exp(ikx - Tb_1 k^2) + ...|^2 \rangle$$

$$= \langle |\sum_{k\geq 0}\left[a_k\exp(ikx - Tb_1 k^2)\left(1 + (Tb_1 k^2) + \frac{(Tb_1 k^2)^2}{2!} + \frac{(Tb_1 k^2)^3}{3!}...\right)\right]|^2 \rangle$$

$$= \langle |\sum_{k\geq 0}\left[a_k\exp(ikx - Tb_1 k^2)\left(\exp(Tb_1 k^2)\right)\right]|^2 \rangle$$

$$= \langle |\sum_{k\geq 0}\left[a_k\exp(ikx))\right]|^2 \rangle$$

$$= \pi \sum_{k\geq 0} a_k^2$$

That gives us spherical objective levelsets. Notice what happens as we increment the number of terms in the above Taylor Series

Taylor Series Terms: 5
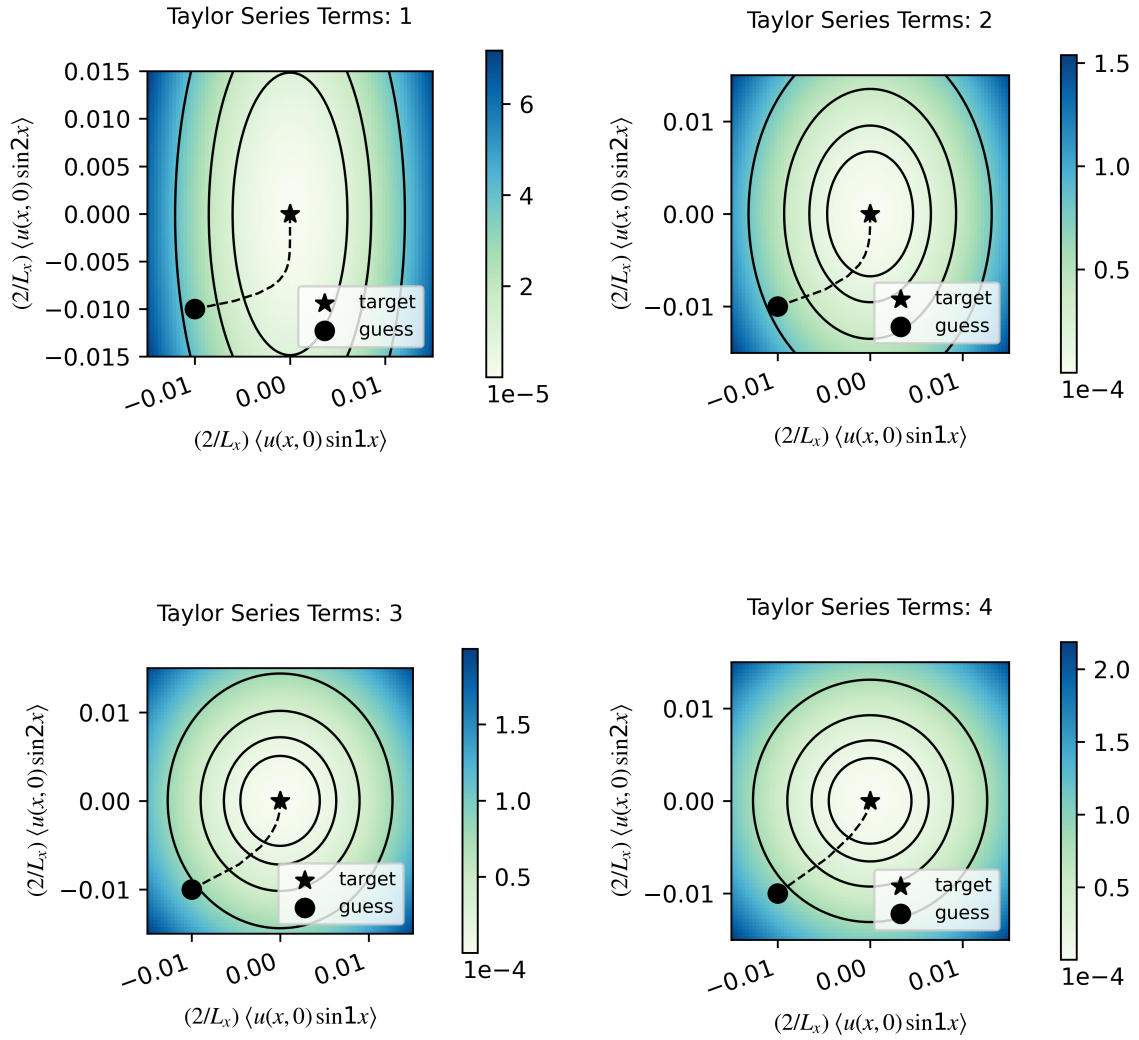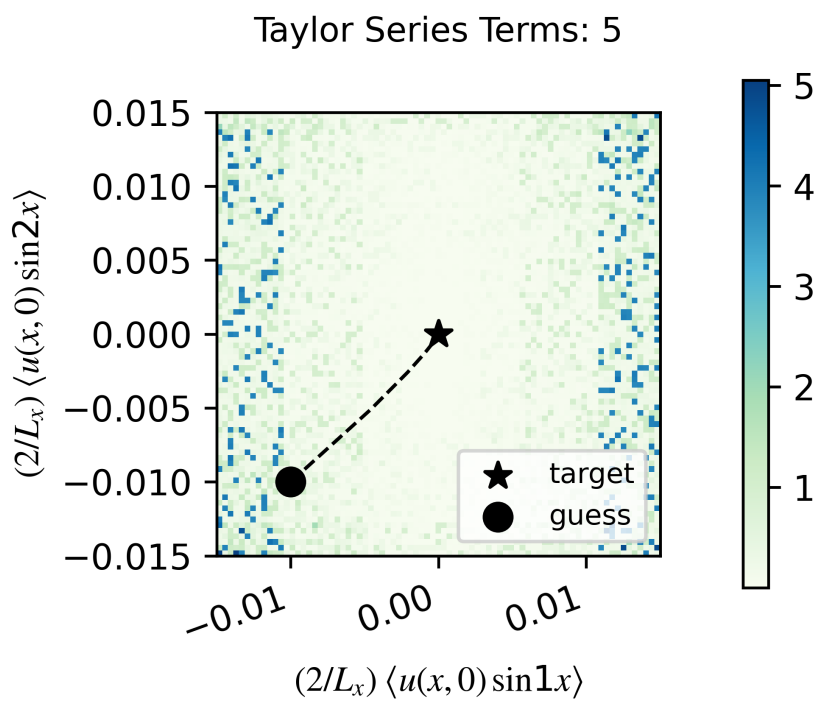
When the initial deviation $u'(x,0)$ is described using a basis of fourier coefficients, its levelsets form hyperellipses, whose semi-major (semi-minor) axes are aligned with $u'(x,0)$'s largest (smallest) fourier coefficient. Accordingly, the path of steepest descent is only aligned with $u'(x,0)$ when $u'(x,0)$ is composed of a single fourier mode. Such systems are self-adjoint, meaning $L[u] = L^{\dagger}[u]$

summary: Linear systems' objective functions' shapes do not depend on the target–the target manifests itself as a constant shift in ic space. Reversible wave equations, which conserve energy yield objective functions whose levelsets in ic space are spherical, so the gradient at any given guess always provides the shortest path to the target. In contrast, irreversible diffusive terms distort the spherical levelsets into hyperellipses, whose semi-minimum (maximum) axis is oriented in the direction of the lowest (highest) wavenumber fourier mode's coefficient. Thus, when traversing an arbitrary path of steepest descent toward the target IC, low wavenumber modes are matched with the target first, followed by increasingly higher wavenumber modes.

B. Nonlinear Reversible Problems

Here the target matters. We consider burger's equation

$$\partial_t u + u \partial_x u = b \, \partial_x^3 u$$

where energy is conserved

$$0 = \partial_t \langle \frac{1}{2} u^2 \rangle + \langle u^2 \partial_x u \rangle - b \langle u \partial_x^3 u \rangle$$
$$= \partial_t \langle \frac{1}{2} u^2 \rangle + \langle \frac{1}{3} \partial_x u^3 \rangle$$
$$= \partial_t \langle \frac{1}{2} u^2 \rangle$$

$$\partial_x^3 (u * u) = 2 \partial_x^2 [u \partial_x u]$$
$$= 2 \partial_x [(\partial_x u)^2 + u \partial_x^2 u]$$
$$= 2 \partial_x [(\partial_x u)^2] + \partial_x u \partial_x^2 u + u \partial_x^3 u$$
$$= \frac{5}{2} \partial_x [(\partial_x u)^2] + u \partial_x^3 u$$

Next we consider a guess with a known deviation from the target $u(x,0) = U(x,0) + u'(x,0)$. Our objective function is given by $\langle (u(T) - U(T))^2 \rangle = \langle u'^2(T) \rangle$. The forward evolution equation for the perturbation is given by

$$0 = \partial_t u' + u' \partial_x U + U \partial_x u' + u' \partial_x u' - b \partial_x^3 u' \tag{12}$$

then we have that

the evolution of the target is given by

$$0 = \frac{1}{2} \partial_t \langle u'^2 \rangle + \langle u' \partial_x (u'U) \rangle + \langle \frac{1}{3} \partial_x u'^3 \rangle$$
$$= \frac{1}{2} \partial_t \langle u'^2 \rangle + \langle u' \partial_x (u'U) \rangle$$
$$= \frac{1}{2} \partial_t \langle u'^2 \rangle - \frac{1}{2} \langle U \partial_x (u'^2) \rangle$$
$$\therefore$$
$$\langle u'^2(T) \rangle = \langle u'^2(0) \rangle + \int_0^T \langle U \partial_x (u'^2) \rangle dt \tag{13}$$

Numerical Examples

Suppose the target solution is $U(t) = 0$. Then the perturbation equation (12) is linear in u', and we revert back to the hyperspherical levelsets described for wave equations in the previous section, as shown in Fig 2.

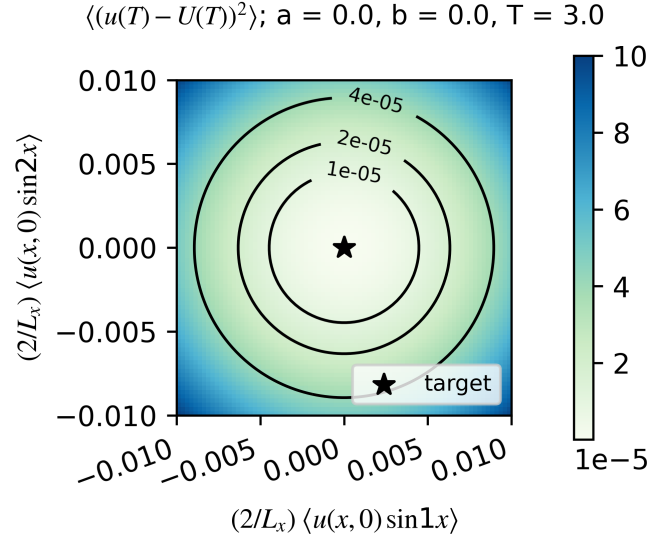$\langle (u(T) - U(T))^2 \rangle$; a = 0.0, b = 0.0, T = 3.0

FIG. 2. Objective functional the inviscid burger's equation. The target simulation $U(x,0) = 0$ and the deviation is composed of two fourier modes, i.e. $u'(t) = a_1 \sin x + a_5 \sin 2x$. Gradient descent paths will not deviate from this 2D phase space because the path of steepest descent is always in the radial direction.

Suppose instead that our target initial condition is given by $U(x,0) = 5\sin(x)$ and we are interested in the 2D phase space of initial deviations $u'(x,0) = a_1 \sin x + a_2 \sin 2x$, as shown in Fig 3..

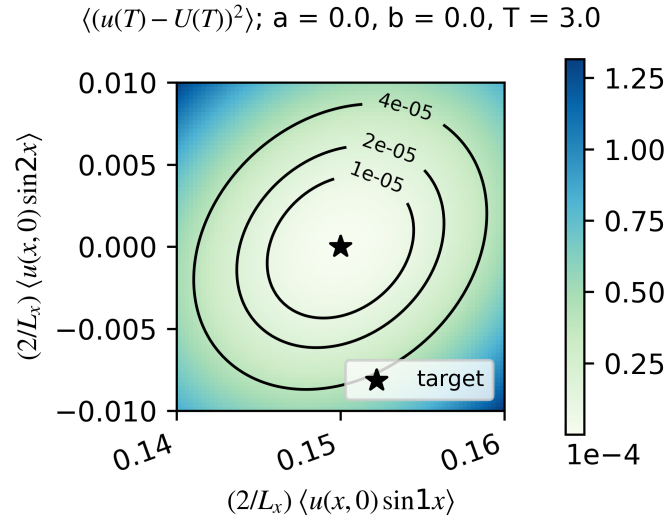$\langle (u(T) - U(T))^2 \rangle$; a = 0.0, b = 0.0, T = 3.0

FIG. 3. Objective functional the inviscid burger's equation. The target simulation $U(x,0) = 5\sin x$ and the deviation is composed of two fourier modes, i.e. $u'(t) = a_1 \sin x + a_5 \sin 2x$. Gradient descent paths will deviate from this 2D phase space because external fourier modes are populated by the nonlinearities in the forward and adjoint equations.

[1] L.-A. Couston, D. Lecoanet, B. Favier, and M. Le Bars, Shape and size of large-scale vortices: A generic fluid pattern in geophysical fluid dynamics, Phys. Rev. Research **2**, 023143 (2020).

[2] X. Zhu, V. Mathai, R. J. Stevens, R. Verzicco, and D. Lohse, Transition to the ultimate regime in two-dimensional rayleigh-bénard convection, Physical Review Letters **120**, 10.1103/physrevlett.120.144502 (2018).

[3] M. Ossendrijver, The solar dynamo, The Astronomy and astrophysics review **11**, 287 (2003).

[4] D. Gubbins, The rayleigh number for convection in the earth's core, Physics of the Earth and Planetary Interiors **128**, 3 (2001), dynamics and Magnetic Fields of the Earth's and Planetary Interiors.