1D: Diffusion Inversion

Preliminaries

We begin by defining the Lagrangian

$$\mathcal{L} = \int_0^T \langle \mu \cdot [\partial_t u - \nu \partial_x^2 u] \rangle dt + \frac{1}{2} \langle (u(T) - U(T))^2 \rangle$$

where $U(T)$ is the target end state, whose initial condition $U(0)$ we wish to recover. We continue by deriving the forward and adjoint problems

$$0 = \frac{\delta \mathcal{L}}{\delta \mu} = \int_0^T \langle \partial_t u - \nu \partial_x^2 u \rangle dt \quad \longrightarrow \quad \partial_t u = \nu \partial_x^2 u$$

Before deriving the adjoint system, we note that

$$\begin{aligned}
\mathcal{L} &= \int_0^T \langle \mu \partial_t u - \nu \mu \partial_x^2 u \rangle dt + \frac{1}{2} \langle (u(T) - U(T))^2 \rangle \\
&= \int_0^T \langle \partial_t (u\mu) - u \partial_t \mu - \nu \mu \partial_x^2 u \rangle dt + \frac{1}{2} \langle (u(T) - U(T))^2 \rangle \\
&= \left\langle \int_0^T \partial_t (u\mu) dt \right\rangle + \int_0^T \langle -u \partial_t \mu - \nu \mu \partial_x^2 u \rangle dt + \frac{1}{2} \langle (u(T) - U(T))^2 \rangle \\
&= \langle u(T)\mu(T) - u(0)\mu(0) \rangle + \int_0^T \langle -u \partial_t \mu - \nu \mu \partial_x^2 u \rangle dt + \frac{1}{2} \langle (u(T) - U(T))^2 \rangle
\end{aligned}$$

We continue by deriving the adjoint system. For periodic (among other) boundary conditions

$$0 = \frac{\delta \mathcal{L}}{\delta u} = -\int_0^T \langle \partial_t \mu + \nu \partial_x^2 \mu \rangle dt \quad \longrightarrow \quad \partial_t \mu = -\nu \partial_x^2 \mu$$

To initialize the adjoint system at $t = T$, we require

$$0 = \frac{\delta \mathcal{L}}{\delta u(T)} = \langle \mu(T) + u(T) - U(T) \rangle \quad \longrightarrow \quad \mu(T) = U(T) - u(T)$$

To determine a trajectory which is coincident with $U(0)$, we take

$$\frac{\delta \mathcal{L}}{\delta u(0)} = -\mu(0)$$

Linearity Cancellation

Here we demonstrate that the target state $U(0)$ manifests itself as a constant shift in the objective function wrt $u(0)$. Let $\tilde{f}$ represent a function f which has undergone diffusion with diffusivity ν over a time period T, i.e.

$$u(0) = f \quad \longrightarrow \quad u(T) = \tilde{f}$$

Suppose our guess's deviation from the target initial state is given by $-w$, i.e.

$$u(0) = U(0) - w$$

Then

$$u(T) = U(T) - \tilde{w} \quad \longrightarrow \quad \frac{1}{2}\langle (u(T) - U(T))^2 \rangle = \frac{1}{2}\langle w^2 \rangle$$

which is independent of the target state. So the objective only depends on our deviation from the target state, rather than the target state itself.

Objective Level Sets

Given that the target state doesn't matter for the following analysis, we take $U(0) = U(T) = 0$ for all x

Suppose

$$u(0) = a_1 e^{ik_1 x} + a_2 e^{ik_2 x} + ...$$

We wish to characterize the objective's dependence on the initial guess in this n-dimensional phase space $\boldsymbol{a} = (a_1, a_2, ..., a_n) \in \mathbb{C}^n$. We can solve for $u(T)$ analytically

$$\begin{aligned} u(T) &= a_1 e^{ik_1 x - \nu k_1^2 T} + a_2 e^{ik_2 x - \nu k_2^2 T} + ... \\ &= a_1 e^{-\nu k_1^2 T} e^{ik_1 x} + a_2 e^{-\nu k_2^2 T} e^{ik_2 x} + ... \end{aligned}$$

The objective

$$\begin{aligned} \frac{1}{2}\langle |u(T)|^2 \rangle &= \frac{1}{2}\langle \left(a_1 e^{-\nu k_1^2 T} e^{ik_1 x} + a_2 e^{-\nu k_2^2 T} e^{ik_2 x} + ... \right)^2 \rangle \\ &= \frac{1}{2}\left[a_1^2 e^{-2\nu k_1^2 T} \langle \cos^2(k_1 x) + \sin^2(k_1 x) \rangle + a_2^2 e^{-2\nu k_2^2 T} \langle \cos^2(k_2 x) + \sin^2(k_2 x) \rangle + ... \right] \\ &= \left(\pi e^{-2\nu k_1^2 T} \right) a_1^2 + \left(\pi e^{-2\nu k_2^2 T} \right) a_2^2 + ... + \left(\pi e^{-2\nu k_n^2 T} \right) a_n^2 \end{aligned}$$

Therefore the objective's level sets are given by infinite-dimensional hyperellipses in $\boldsymbol{a}$ space.

Preconditioner

The appropriate preconditioner should be obvious: we need to undiffuse the adjoint initial condition twice! Doing so remaps the levelsets of hyperellipses into hyperspheres, thereby ensuring that the gradient is always in the direction of $u(0) - U(0)$

Suppose the fourier coefficients of the adjoint initial condition are given by $\boldsymbol{b} = (b_1, b_2, b_3, ..., b_n) \in \mathbb{C}^n$. Then it would be advantageous to initialize the adjoint system with fourier coefficients $\tilde{\boldsymbol{b}}$ where

$$\begin{bmatrix} \tilde{b_1} \\ \tilde{b_2} \\ \tilde{b_3} \\ \vdots \\ \tilde{b_n} \end{bmatrix} = \begin{bmatrix} e^{2\nu k_1^2 T} & 0 & 0 & \dots & 0 \\ 0 & e^{2\nu k_2^2 T} & 0 & \dots & 0 \\ 0 & 0 & e^{2\nu k_3^2 T} & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & e^{2\nu k_n^2 T} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix}$$