On Numerical Integration in Neural ODE

Aiqing Zhu

joint with Pengzhan Jin, Beibei Zhu & Yifa Tang

LSEC, ICMSEC, AMSS

July 13, 2022

Neural ODE

Consider autonomous systems of first-order ordinary differential equations

$$\frac{d}{dt}y(t) = f(y(t)), \quad y(0) = x, \tag{1}$$

where $y(t) \in \mathbb{R}^D$, $f : \mathbb{R}^D \to \mathbb{R}^D$ is smooth and x is the initial value. For fixed t, $y(t)$ can be regarded as a function of its initial value x. We denote

$$\phi_t(x) := y(t) = x + \int_0^t f(y(\tau))d\tau,$$

which is known as the time-t flow map of dynamical system (1). In general, we chose a numerical integrator Φ_h that approaches ϕ_h and compose it to obtain the numerical solution. In order to emphasize specific differential equation, we will add the subscript f and denote ϕ_t as $\phi_{t,f}$ and Φ_h as $\Phi_{h,f}$.

Neural ODE

In this work, we consider the empirical risk optimization problem

$$L = \frac{1}{N}\sum_{n=1}^{N} l(\phi_{T,f_\theta}(x_n), z_n),$$

where ϕ_{T,f_θ} is a Neural ODE model with a trainable neural network f_θ

Dividing T in S equally-spaced intervals, the ϕ_{T,f_θ} can be approximated by S compositions of a predetermined one-step numerical integrator Φ_h,

$$\phi_{T,f_\theta} \approx \underbrace{\Phi_{h,f_\theta} \circ \cdots \circ \Phi_{h,f_\theta}}_{S \text{ compositions}}(x) = \left(\Phi_{h,f_\theta}\right)^S(x),$$

where $h = T/S$ is the discrete step. Therefore, the practical input of loss function is given by the predetermined ODE solver, i.e.,

$$L = \frac{1}{N}\sum_{n=1}^{N} l\left(\left(\Phi_{h,f_\theta}\right)^S(x_n), z_n\right).$$

Numerical integration in Neural ODE

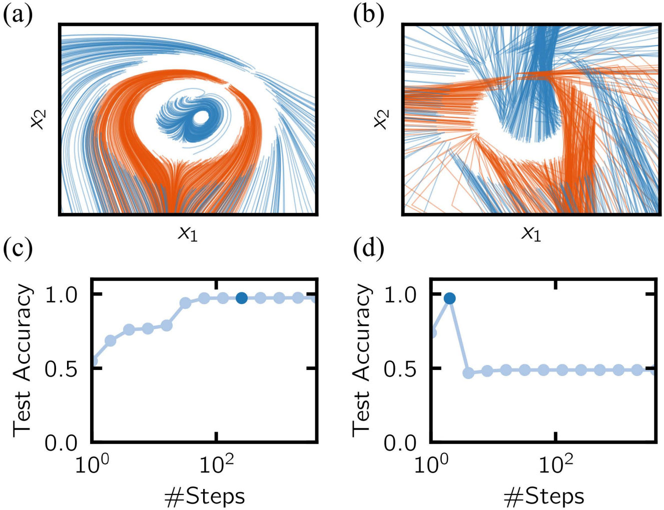


Figure: From Ott et al.

Inverse Modified Differential Equations (IMDE)

We aim to find a perturbed differential equation (named IMDE)

$$\frac{d}{dt}\tilde{y}(t) = f_h(\tilde{y}(t)) = f_0(\tilde{y}) + hf_1(\tilde{y}) + h^2 f_2(\tilde{y}) + \cdots,$$

such that $\Phi_{h,f_h}(x) = \phi_{h,f}(x)$ formally.

1. Expand $\phi_{h,f}(x)$ into a Taylor series around $h = 0$,

$$\phi_{h,f}(x) = x + hf(x) + \frac{h^2}{2} f' f(x) + \frac{h^3}{6}(f''(f,f)(x) + f' f' f(x)) + \cdots.$$

2. Expand numerical solution as

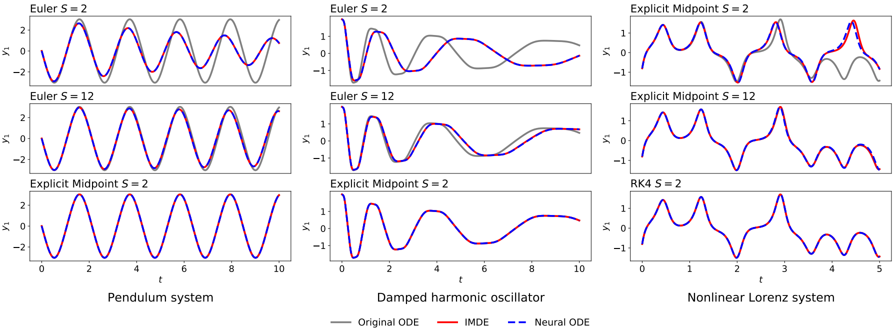
$$\Phi_{h,f_h}(x) = x + hd_{1,f_h}(x) + h^2 d_{2,f_h}(x) + \cdots.$$

3. Compare the coefficients of equal powers of h, obtain f_k in a recursive manner.

Main Results

The trained Neural ODE model is a close approximation of the IMDE, i.e., the difference between the learned Neural ODE model and the truncation of the IMDE is bounded by the sum of the learning loss and a discrepancy which can be made sub exponentially small,

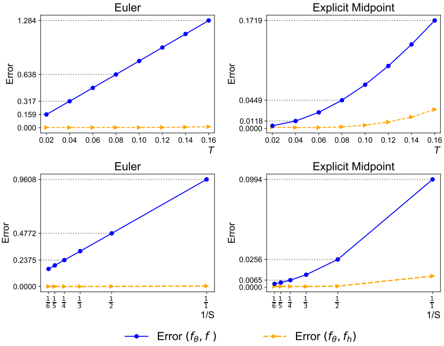
$$\left\| f_\theta(x) - f_h^K(x) \right\| \leq c_1 m e^{-\gamma/h^{1/q}} + \frac{e}{e-1}\mathcal{L},$$



Main Results

The difference between the learned Neural ODE model and the true hidden system is bounded by the sum of the discretization error Ch^p and the learning loss, where h is the discrete step and p is the order of the numerical integrator, i.e.,

$$\|f_\theta(x) - f(x)\| \leq c_2 m h^p + \frac{e}{e-1}\mathcal{L}.$$



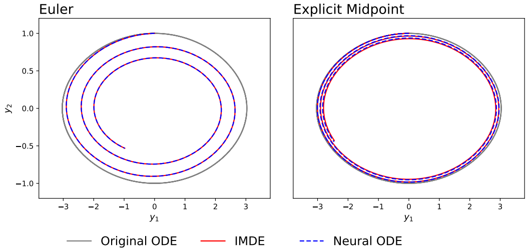
Main Results

A Hamiltonian system is formulated as

$$\frac{d}{dt}y = J^{-1}\nabla H(y), \quad J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

where I is $D/2$-by-$D/2$ identity matrix.

Neural ODE using non-symplectic numerical integration fail to learn conservation laws theoretically, since the IMDE of non-symplectic integrator is not a Hamiltonian system.



Thanks for your attention!