Learning non-canonical Hamiltonian dynamics

Clémentine Courtès[1,2] Emmanuel Franck[1,2] Michael Kraus[3]
Laurent Navoret[1,2] Léopold Trémant[4]

[1]Institut de Recherche Mathématique Avancée, Univ. Strasbourg et CNRS
[2]INRIA Nancy-Grand Est, MACARON Project, Strasbourg, France
[3]Max Planck für Plasma physik, Garching bei München, Germany
[4]Laboratoire de Mathématiques de Lens, Univ. d'Artois, Lens, France

NODYCON 2025, June 22-25, 2025

Outline

Introducing the problem

- Hamiltonian dynamics
- Hard-coding structure
- Long-time integration

“Fixing” VF learning

- Analyzing the numerical scheme
- Goal-informed vector-field learning

Scheme learning

- Training procedure
- Scheme, structure and exact solution
- Results

What are Hamiltonian dynamics?

Canonical problems

Derive the dynamics on the *position & momentum* from a Hamiltonian $(q, p) \mapsto H(q, p) \in \mathbb{R}$,

$$\dot{q} = \nabla_p H(q, p), \qquad \dot{p} = -\nabla_q H(q, p)$$

Or on the full coordinates $u = (q, p)$,

$$\underbrace{\dot{u} = J^{-1} \nabla H(u),}_{\text{Hamiltonian dynamics}} \qquad \underbrace{J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}}_{\text{symplectic form}}$$

Example: simple pendulum, $q = \theta$, $p = \omega$.

Non-canonical problems

Other symplectic form $z \mapsto W(z)$

$$\dot{z} = f(z) := W(z)^{-1} \nabla H(z)$$

which derives from a potential $z \mapsto A(z)$.

Example: Lotka-Volterra, guiding center...

Learn a symplectic potential and a Hamiltonian with neural networks

[Greydanus, Dzamba, and Yosinski 2019]
[Chen, Matsubara, and Yaguchi 2021]

Different architectures – Lotka-Volterra

$$\begin{cases} \dot{x} = x(1-y), \\ \dot{y} = y(x-2). \end{cases}$$

fit a symplectic potential and a Hamiltonian on $\dot{x}, \dot{y}$

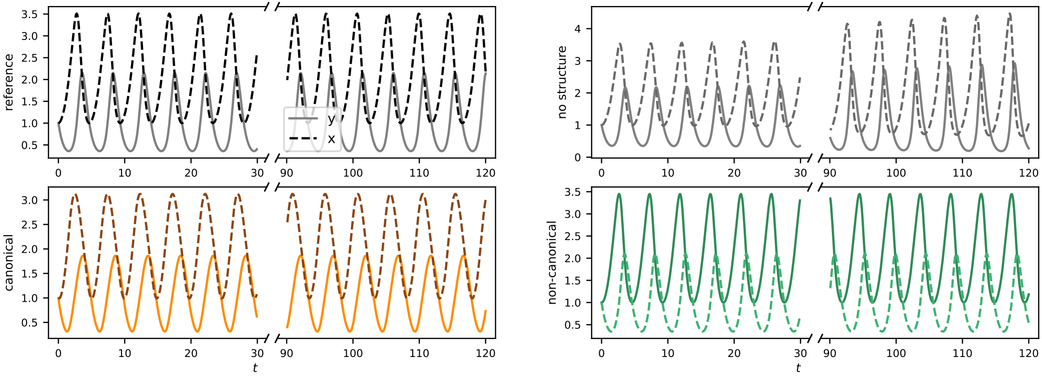
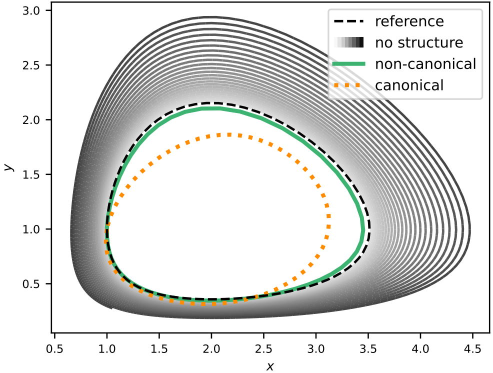


Figure: Comparison of the solutions for the different architectures, with initial conditions $x(0) = 1, y(0) = 1$.

The importance of structure – Lotka-Volterra

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & -xy \\ xy & 0 \end{pmatrix} \begin{pmatrix} 1 - 2/x \\ 1 - 1/y \end{pmatrix}$$

➜ fit a symplectic potential and a Hamiltonian on $\dot{x}, \dot{y}$



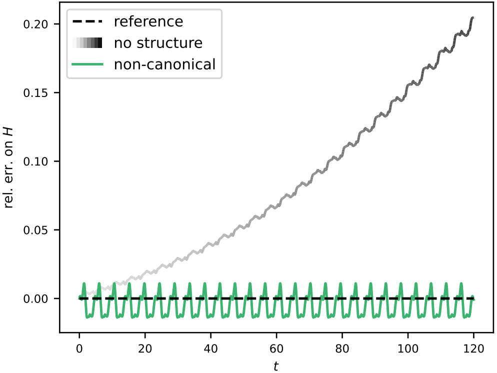


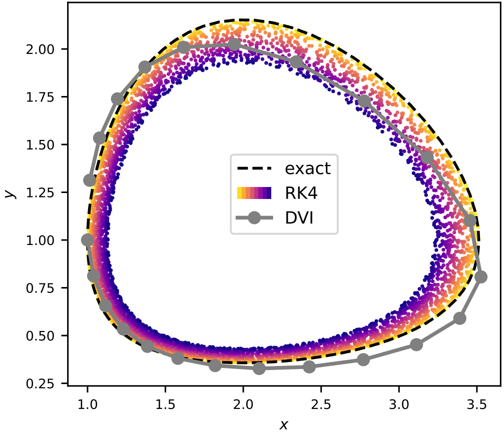
Figure: Comparison of the solutions for the different architectures, with initial conditions $x(0) = 1, y(0) = 1$.

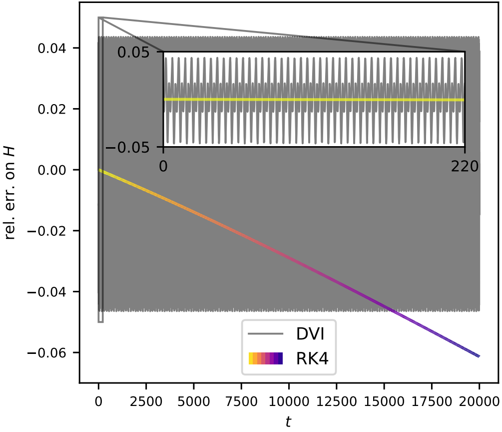
The importance of the scheme – Lotka-Volterra

Standard scheme — short time 👍 long time 👎

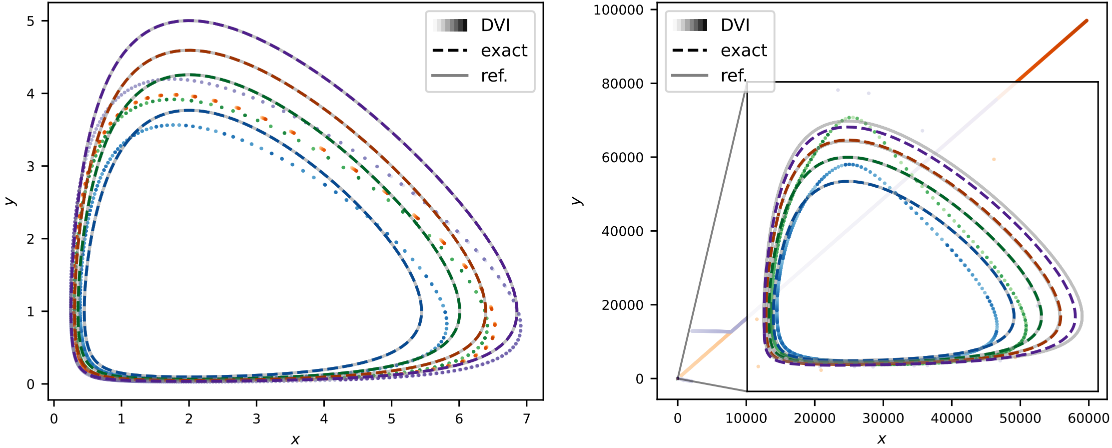
Geometric scheme — short time 👎 long time 👍

(Ellison et al. 2018)





Failure of the scheme for learnt models



Sample of solutions of the geometric scheme (DVI) for the reference model (left) and the learnt model (right) on the Lotka-Volterra problem, with a time-step $\Delta t = 0.1$.

Explaining the failure of the scheme

The symplectic form derives from a symplectic potential $A(z)$,

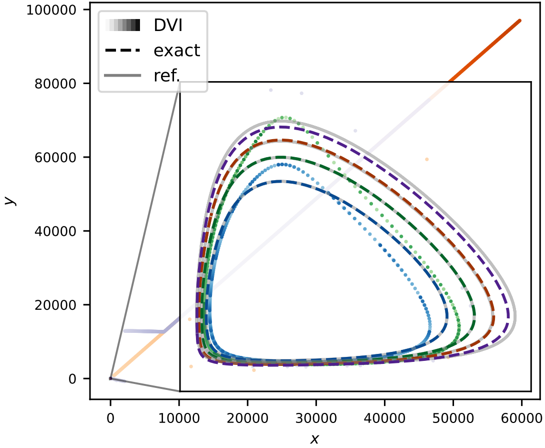
$$W(z) = D_z A(z)^T - D_z A(z).$$

and the dynamics used by the scheme are

$$\begin{cases} \dot{p} = D_z A(z)^T \dot{z} - \nabla_z H(z), \\ p = A(z). \end{cases}$$

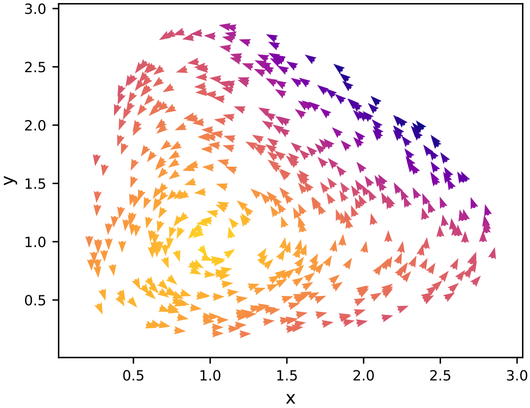
- the **symmetric** part of $D_z A$ (the gauge)
  - appears in the scheme error
  - does not appear in the loss

➔ "invisible" perturbations may appear in $z \mapsto A(z)$ during training



Goal-informed vector-field learning

dataset $\{(z, f(z))\}$



take into account the **scheme** that will be used by adding a penalization term to the loss,

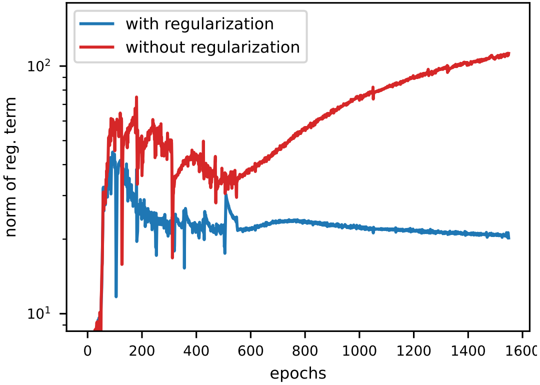
$$\mathcal{L}(\Theta) = \frac{1}{B}\sum_{b=1}^{B} \left\| f_\Theta(z^{(b)}) - f^{(b)} \right\|^2 + \varepsilon \left\| r_\Theta\left(z^{(b)}, f^{(b)}\right) \right\|^2$$

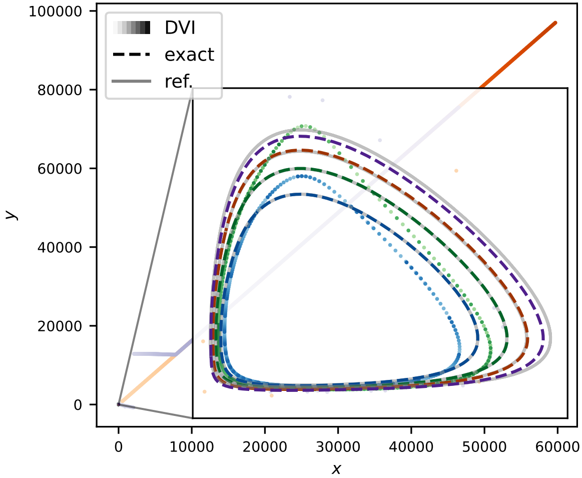
with the dominant error term $r(z, z_t)$,

$$\boxed{z(h) - z_h = h^2 r\left(z_0, f(z_0)\right) + \mathcal{O}(h^3)}$$

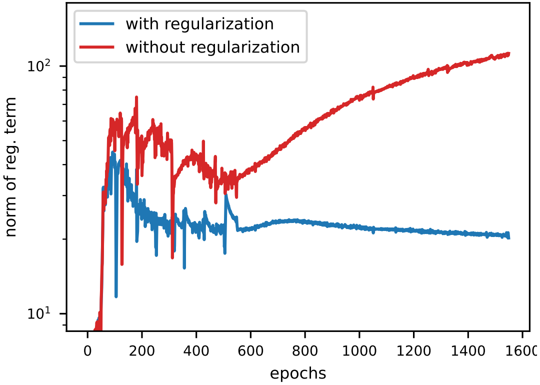
might not work with other schemes!

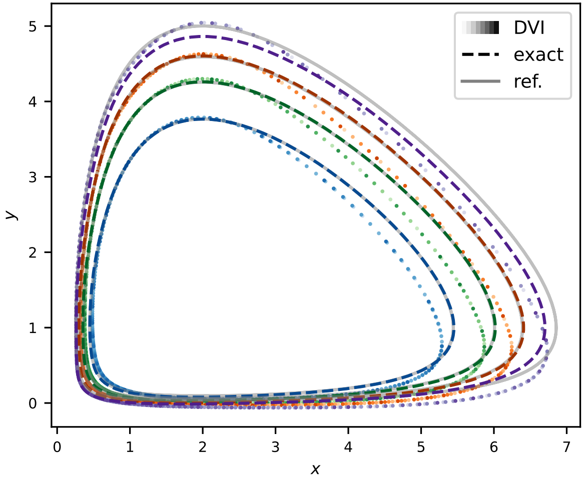
Effect of the regularization term





Effect of the regularization term

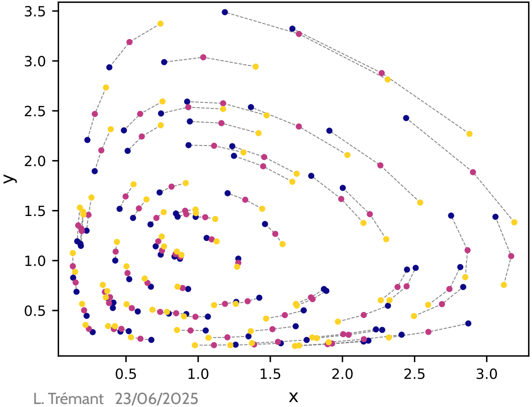




Scheme learning, broadly speaking

- learn a **discrete mapping** based on the num. scheme

dataset $\{(z(0), z(\Delta t))\}$



$$S_\Theta(z^n, z^{n+1}) \approx A_\Theta^{n+1} - A_\Theta^n - \left(D_z A_\Theta^n (z^{n+1} - z^n) - \Delta t\, \nabla_z H_\Theta^n\right)$$

- avoid the trivial $A_\Theta = 0$, $H = 0$ using $J_\Theta = D_z S_\Theta(z(0), z(\Delta t))$

$$\mathcal{L}(\Theta) = \frac{1}{B} \sum_{b=1}^{B} \left\| J_\Theta^{-1} S_\Theta\left(z^{(b)}(0), z^{(b)}(\Delta t)\right) \right\|^2 + \varepsilon \log_{10} \kappa(J_h)$$

Do A_Θ, H_Θ exist?

Backward error analysis in the linear case

the numerical solution is the exact solution of a perturbed vf

$$\begin{tikzcd} & f \arrow[dr, "\Phi_{\Delta t}"] \arrow[dl, dashed, bend right, "\text{b.e.a.}"'] & \\ \, \arrow[dr, dashed, bend right] & & u^n \\ & f_{\Delta t} \arrow[ur, "\varphi_{\Delta t}(\,\cdot\,;\Delta t)"'] & \end{tikzcd}$$

the exact solution is the numerical solution of a modified vf

$$\boxed{\begin{tikzcd} & \widetilde{f}_{\Delta t} \arrow[dr, "\widetilde{\Phi}_{\Delta t}"] & \\ \text{corrected b.e.a.} \arrow[ur, dashed, bend left] & & u(\Delta t) \\ & f \arrow[ul, dashed, bend left] \arrow[ur, "\varphi_{\Delta t}"'] & \end{tikzcd}}$$

For a linear problem $\dot{z} = Mz$,

Euler: $\Phi_{\Delta t} = \mathrm{id} + \Delta t\, M$,

$$\left.\begin{aligned} \varphi_{\Delta t} &= e^{\Delta t\, M} \\ \widetilde{\Phi}_{\Delta t} &= \mathrm{id} + \Delta t\, \widetilde{M}_{\Delta t} \end{aligned}\right\} \Rightarrow \widetilde{M}_{\Delta t} = \frac{1}{\Delta t}\left(e^{\Delta t\, M} - \mathrm{id}\right).$$

$\left(\begin{array}{c}\text{restrictions apply in the}\\ \text{non-linear case}\\ \text{with large time-steps}\end{array}\right)$

[Chartier, Hairer, and Vilmart 2010]

Error convergence without structure

$$\mathrm{err} = \left\| e^{TM} - \left(\mathrm{id} + \delta t\, \widetilde{M}_{\Delta t}\right)^{T/\delta t} \right\|, \qquad M = J^{-1} I_2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

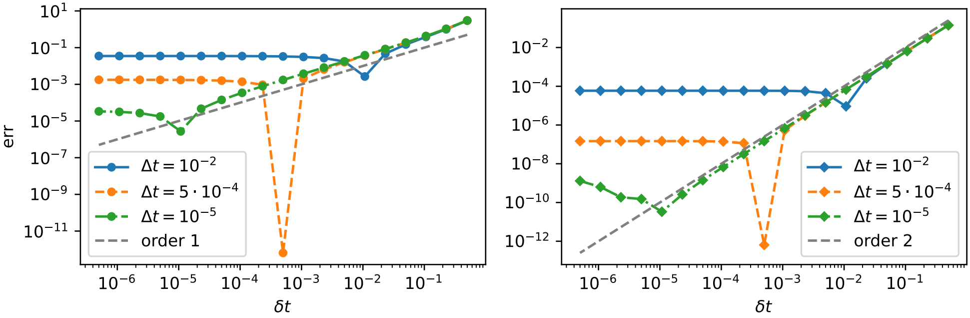


Figure: Error at a time $T = 5$ for different learning time-steps Δt and simulation time-steps δt using the explicit Euler method (left) and the midpoint method (right).

➡ simulate with the same time-step as for learning

Error convergence with structure

$$\text{err} = \left\| e^{TM} - \left(\text{id} + \delta t\, J^{-1}\widetilde{H}_{\Delta t}\right)^{T/\delta t} \right\|, \qquad \widetilde{H}_{\Delta t} = \frac{1}{2}\left(J\widetilde{M}_{\Delta t} + (J\widetilde{M}_{\Delta t})^{\mathsf{T}}\right)$$

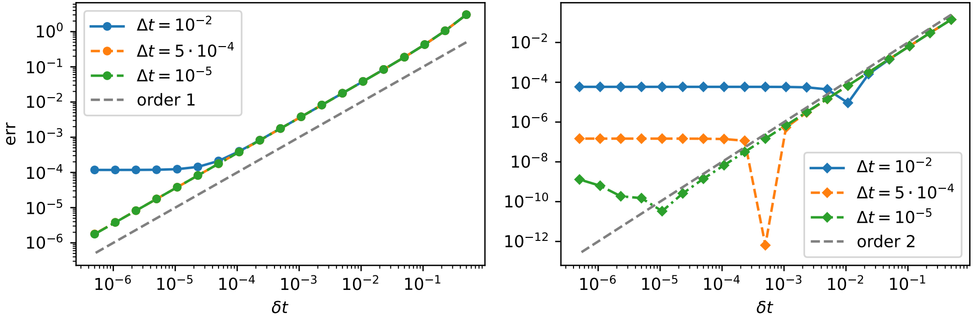
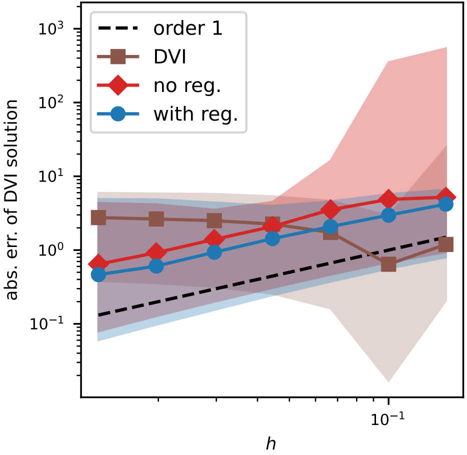


Figure: Error at a time $T = 5$ for different learning time-steps Δt and simulation time-steps δt using the explicit Euler method (left) and the midpoint method (right).

➔ use a scheme appropriate for the structure

Learning using the scheme



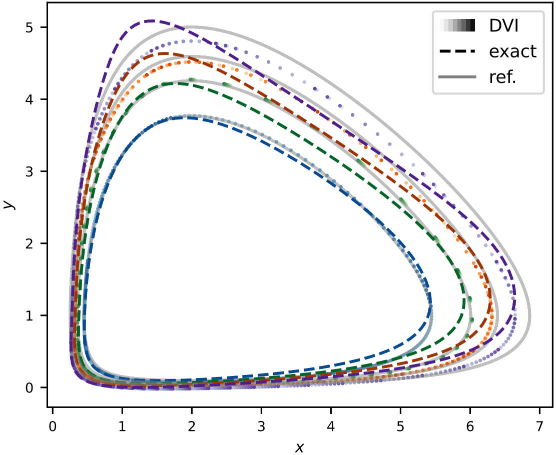


Figure: Results on the Lotka-Volterra problem, $\Delta t = 0.1$. Left: error on the numerical scheme in short time. Right: solution of the geometric scheme (dots) and the exact sol. of the learnt model.

Conclusion

For long-time simulations

- hard-code structure
- use an appropriate numerical scheme
- application to Lotka-Volterra and guiding center

Vector-field learning

- velocity dataset
- recover the underlying physics
- need care for long-time simulation

Scheme learning

- snapshot dataset
- allows larger time-steps
- hard-codes a specific scheme

Future work

- theoretical analysis of the numerical scheme
- other problems (time-dependent, reduced-order...)

Conclusion

For long-time simulations

- hard-code structure
- use an appropriate numerical scheme
- application to Lotka-Volterra and guiding center

Vector-field learning

- velocity dataset
- recover the underlying physics
- need care for long-time simulation

Scheme learning

- snapshot dataset
- allows larger time-steps
- hard-codes a specific scheme

Future work

- theoretical analysis of the numerical scheme
- other problems (time-dependent, reduced-order...)

Thanks for your attention!