

END-TO-END LEARNING OF DYNAMICAL SYSTEMS WITH THE MORI-ZWANZIG FORMALISM

T. MONSEL ¹, O. SEMERARO ¹, L. MATHELIN ¹, G. CHARPIAT ^{1,2}

¹LISN, Université de Paris-Saclay, 91400 Orsay, France

²INRIA, 91400 Saclay, France

March 6, 2023



- ❶ Introduction
- ❷ The Mori-Zwanzig (MZ) formalism
 - Derivation of the MZ equation
 - Modelling possibilities
 - Examples
- ❸ Applying Delay Differential Equations to MZ

Introduction

Modelling dynamical systems into the age of DL

- Experimental data is often noisy, corrupted and does not account for the system's full state.
- Generally, partially observed states cannot reconstruct its own dynamics but ...
- ... With historical values of such a partial state, it can be done (Takens theorem).

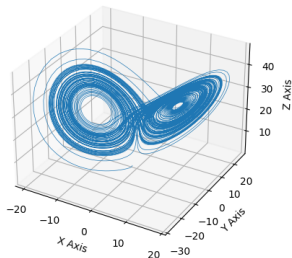


Figure: Lorenz system

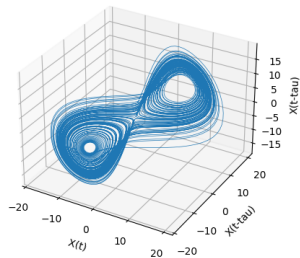


Figure: Delay coordinate embedding $h_t = [x_t, x_{t-\tau}, x_{t-2\tau}]$

Introduction

Modelling partially observed dynamical systems into the age of Deep Learning (DL)

- Latent variables enhance model expressivity but ...
- these black box models (RNN, LSTM ...) lack interpretability, robustness guarantees and insights into the structure of the dynamics.

$$h_{t+1} = f_{\theta_t}(h_t, x_t)$$

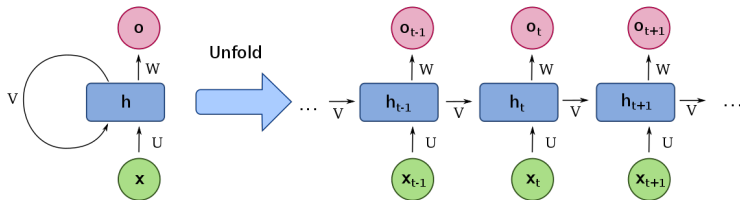


Figure: Example of RNN

Introduction

Modelling dynamical systems into the age of Deep Learning (DL)

Neural Ordinary Differential Equations (NODEs)¹ is a **natural way** of modelling a dynamical system since NODE is a continuous equivalent ResNets (to be precise Leaky RNN):

$$\mathbf{x}_{t+1} = \mathbf{x}_t + f_{\theta}(\mathbf{x}_t, t)$$
$$\frac{d\mathbf{x}(t)}{dt} = f_{\theta}(\mathbf{x}(t), t)$$

More often than not :

- Experiments are done with the full state $\mathbf{x}(t)$.
- Partially observed states coupled with non physical latent variables to reconstruct the dynamics (i.e. Augmented NODEs², LSTM ...).

Ideally what we want :

- a DL framework with a principled sound structure.

Mori-Zwanzig (MZ) formalism provides a rigorously grounded theory for our needs.

¹R Chen et al. *Neural ordinary differential equations*. 2018.

²E. Dupont et al. *Augmented Neural ODEs*. 2019.

The Mori-Zwanzig Formalism

Coping with partially observed systems

- It allows to derive **exact** evolution equations for **coarse-grained quantities** of interest from high-dimensional systems.
- If we consider a nonlinear system evolving on a smooth manifold $\mathcal{S} = \mathbb{R}^n$.

$$\frac{dx}{dt} = F(x), \quad x(0) = x_0$$

- The system is seen through the lens of the observables; $\forall i, g_i : \mathcal{S} \rightarrow \mathbb{C}$, that obey³:

$$\mathcal{L}g_i(x) = F(x) \cdot \nabla g_i(x)$$

where \mathcal{L} is the Liouville operator

³Liouville operator Linked to the Koopman operator $\mathcal{K}(t,s) = e^{(t-s)\mathcal{L}}$

The Mori-Zwanzig Equation

Coping with partially observed systems

By concatenating the observables g_i , $\mathbf{g} = [g_0, \dots, g_n]$, we get the MZ equation^{4,5,6}:

$$\frac{\partial}{\partial t} \mathbf{g}(x, t) = \underbrace{M(\mathbf{g}(x, t))}_{\text{Markov term}} + \underbrace{F(x, t)}_{\text{Noise term}} - \underbrace{\int_0^t K(\mathbf{g}(x, t-s), s) ds}_{\text{Memory term}}$$

where

$$M(\mathbf{g}(x, t)) = e^{t\mathcal{L}} \mathcal{P} \mathcal{L} \mathbf{g}(x)$$

$$F(x, t) = e^{t\mathcal{Q}\mathcal{L}} \mathcal{Q} \mathcal{L} \mathbf{g}(x)$$

$$K(\mathbf{g}(x, t-s), s) = -e^{(t-s)\mathcal{L}} \mathcal{P} \mathcal{L} e^{s\mathcal{Q}\mathcal{L}} \mathcal{Q} \mathcal{L} \mathbf{g}(x)$$

with \mathcal{P} a projector operator and its orthogonal $\mathcal{Q} = \mathcal{I} - \mathcal{P}$

⁴Y. Zhu. Mori-Zwanzig equation: Theory and Applications. 2019.

⁵H. Mori. A Continued-Fraction Representation of the Time-Correlation Functions. 1965.

⁶R. Zwanzig. Approximate Eigenfunctions of the Liouville Operator in Classical Many-Body Systems. 1966.

The Mori-Zwanzig Equation

Coping with partially observed systems

Several ways are possible to cancel the noise term $F(x, t)$ in the MZ equation. With $F(x, t) = 0$, we get :

$$\frac{\partial}{\partial t} \mathbf{g}(x, t) = \underbrace{M(\mathbf{g}(x, t))}_{\text{Markov term}} - \underbrace{\int_0^t K(\mathbf{g}(x, t-s), s) ds}_{\mathbf{y}(x, t) = \text{Memory term}}$$

Additionally, we can rewrite the MZ equation as coupled PDEs⁷:

$$\left\{ \begin{array}{l} \frac{\partial}{\partial t} \mathbf{g}(x, t) = M(\mathbf{g}(x, t)) + \mathbf{y}(x, t) \\ \frac{\partial}{\partial t} \mathbf{y}(x, t) = K(\mathbf{g}(x, t), 0) - \int_0^t \partial_2 K(\mathbf{g}(x, t-s), s) ds \end{array} \right.$$

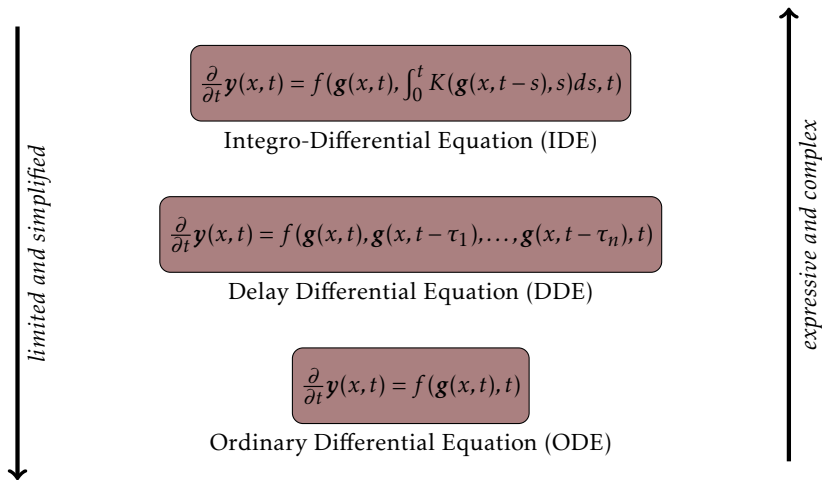
On-going theoretical link of MZ with Deep Learning theory⁸.

⁷ ∂_2 refers to the partial derivative with respect to the second argument

⁸D. Venturi et al. *The Mori-Zwanzig formulation of deep learning*. 2022.

Modelling possibilities with MZ

An overview



Practical example of MZ

Linear systems

$$X'(t) = AX(t)$$

$$A = \begin{bmatrix} A_{rr} & A_{ru} \\ A_{ur} & A_{uu} \end{bmatrix}, \quad X = \begin{bmatrix} X_r \\ X_u \end{bmatrix}$$

$$\mathcal{P}A = A_{rr}, \quad \mathcal{Q}A = A_{uu}$$

The MZ equation applied to the resolved variables $X_r(t)$ gives

$$X_r'(t) = A_{rr}X_r(t) + A_{ru} \int_0^t e^{(t-s)A_{uu}} A_{ur} X_r(s) ds + \underbrace{\cancel{A_{ru} e^{tA_{uu}} X_u(0)}}_{X_u(0) := \mathbf{0}}$$

Practical example of MZ

Harmonic Oscillator

An harmonic oscillator dynamic is specified by

$$\ddot{x} + \frac{1}{\tau} \dot{x} + \omega_0^2 x = 0$$

than can be rewritten in matrix form as

$$A = \begin{bmatrix} 0 & 1 \\ -\omega_0^2 & -\frac{1}{\tau} \end{bmatrix} \quad X = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}$$

If only the velocity $g = \dot{x}$ is observed then we get :

$$\begin{cases} \frac{dg(t)}{dt} = -\frac{1}{\tau}g(t) + y(t) \\ \frac{dy(t)}{dt} = -\omega_0^2 g(t) \end{cases}$$

Analytical

$$\begin{cases} \frac{dg(t)}{dt} = f_{\theta_1}(g(t), t) + y(t) \\ \frac{dy(t)}{dt} = f_{\theta_2}(g(t), t) \end{cases}$$

DL-based

Practical example of MZ

Harmonic Oscillator

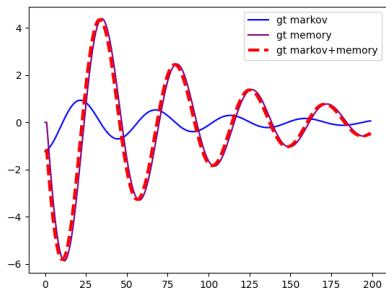


Figure: Analytical derivation of quantities of interest \dot{x}

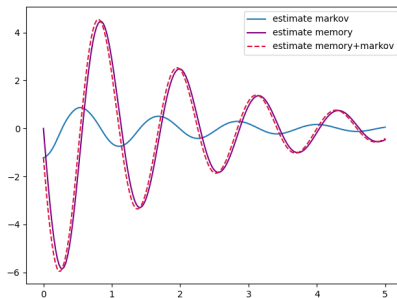


Figure: NODEs estimation of Markov and memory terms

Neural Network based DDEs

Applied to MZ equation

- Modelling partially observed dynamics with DDEs is a compromise between an ODE and IDE model for the memory term's dynamic $y(x, t)$.
- Therefore, our MZ equation modelling is :

$$\begin{cases} \frac{\partial}{\partial t} g(x, t) = M_{\theta}(g(x, t)) + y(x, t) \\ \frac{\partial}{\partial t} y(x, t) = f_{\theta}(g(x, t), g(x, t - \tau_1), \dots, g(x, t - \tau_n), t) \end{cases}$$

- A DDE approach can also be justified by Takens theorem.
- **So far, our work focuses on neural-network based DDE implementation since no tool is available. DDE solver was implemented with autodifferentiation software JAX.**

Neural Network based DDEs

Validating the DDE approach

Mackey-Glass system ($\beta = 0.25$, $n = 10$, $\tau = 10$ and $\gamma = 0.1$):

$$\dot{x}(t) = \beta \frac{x(t-\tau)}{1 + x(t-\tau)^n} - \gamma x(t)$$

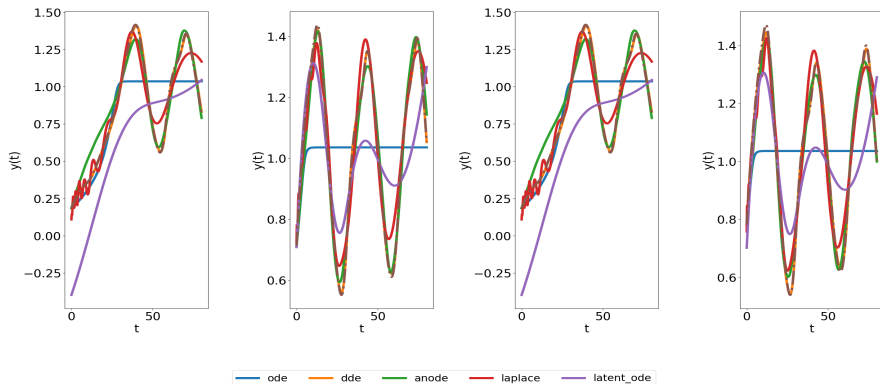


Figure: DL model $\dot{x}(t) = f_{\theta}(x(t), x(t-\tau))$ testset predictions

Neural Network based DDEs

Practical Example : Van der Pol oscillator

- Stiff Van der Pol oscillator ($\mu = 2$) : $\ddot{x} + \mu(1 - x^2)\dot{x} + x = 0$
- If only the velocity $g = \dot{x}$ is observed then no obvious analytical form is given. Moreover, the current NODE approach for linear system fails

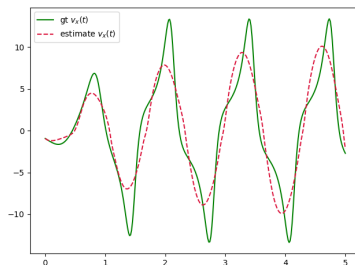


Figure: Results for the same MZ structure used in linear systems for Van der Pol

$$\begin{cases} \frac{dg(t)}{dt} = f_{\theta_1}(g(t), t) + y(t) \\ \frac{dy(t)}{dt} = f_{\theta_2}(g(t), t) \end{cases}$$

DL-based MZ formalism is underway for modelling partially observed systems because:

- Deep Learning based theory for MZ.
- Theoretically and empirically sound results for linear systems.
- Neural networks based DDEs are now possible with known delays.
- Tool developed with be used for MZ modelling.

References I

- [1] R Chen et al. *Neural ordinary differential equations*. 2018.
- [2] E. Dupont et al. *Augmented Neural ODEs*. 2019.
- [3] Y. Zhu. *Mori-Zwanzig equation: Theory and Applications*. 2019.
- [4] H. Mori. *A Continued-Fraction Representation of the Time-Correlation Functions*. 1965.
- [5] R. Zwanzig. *Approximate Eigenfunctions of the Liouville Operator in Classical Many-Body Systems*. 1966.
- [6] D. Venturi et al. *The Mori-Zwanzig formulation of deep learning*. 2022.