

# DSTLS SINDy Optimizer and regression of Hodgkin-Huxley model onto FitzHugh-Nagumo model

Štěpán Zapadlo (email: stepan.zapadlo@mail.muni.cz), Lenka Příbylová (email: pribylova@math.muni.cz)

Department of Mathematics and Statistics, Faculty of Science, Masaryk University, Brno, Czech Republic

## DSTLS SINDy Optimizer

In the current day and age, we have an abundance of data about many different systems. This data can be mined for useful information and an ODE formulation of the original systems is among the most important. While a myriad of techniques emerged, *Sparse Identification of Nonlinear Dynamics* (SINDy) [1] is one of the most popular, which uses  $l$ -dimensional regularized linear regression

$$\min_{\Xi} \frac{1}{2} \|\dot{\mathbf{X}} - \Theta(\mathbf{X})\Xi\|_2^2$$

s.t. sparsity inducing constraint.

Although it can be solved using different optimization techniques, we propose *Dynamic Sequentially Thresholded Least Squares* (DSTLS); a modification of *Sequentially Thresholded Least Squares* (STLS).

### DSTLS optimization problem

For a  $k$ -th variable of our system, we can estimate its derivative using a DSTLS optimization problem:

$$\min_{\xi_k} \frac{1}{2} \|\dot{\mathbf{X}}_{\cdot,k} - \Theta(\mathbf{X})\xi_k\|_2^2$$

s.t.  $\forall i \in \{1, \dots, p\} : |\xi_{ki}| \geq \tau \cdot \max |\xi_k|$

$\mathbf{X}$  measurements of our systems

$\dot{\mathbf{X}}$  derivatives of our systems

$\Theta$  library of candidate functions

$\xi_k$   $k$ -th column of coefficient matrix  $\Xi$ , corresponding to the equation for the derivative of  $k$ -th variable

$p$  number of candidate functions

$\tau$  sparsity threshold

Such modification is motivated by the *FitzHugh-Nagumo* (FHN) model of neuron

$$\begin{aligned} \dot{V} &= V - \frac{V^3}{3} - W + i_e, \\ \dot{W} &= a \cdot (bV - cW + d), \end{aligned}$$

with the following parameters and initial conditions

$$\begin{aligned} a &= 0.08, b = 1, c = 0.8, d = 0.7, i_e = 0.8, \\ V(0) &= 3.3, W(0) = -2. \end{aligned}$$

The disparity in magnitudes of parameters between equations causes sensitivity to the value of threshold  $\tau$  with the ordinary STLS method. Choosing  $\tau$  too big, only a constant zero solution will be discovered for  $\dot{W}$ . For  $\tau$  small enough unnecessary terms will be identified for  $\dot{V}$ . Scaling the threshold by the largest absolute value of estimated parameters aims to address this issue.

### Applications and results

For the illustration of the proposed method, let us assume the derivatives are unknown and our data are corrupted by a *additive white Gaussian noise* (AWGN) with its variance equal to 5 percent of the data's variance. The derivative is estimated with total variation regularized numerical differentiation, which is computed on raw noisy data, and smooth the data using total variation. At last, polynomials up to 4th order were used as a candidate library. DSTLS optimizer correctly chooses the appropriate candidate functions, unlike STLS.

### Comparison of discovered models

$$\begin{aligned} \dot{V} &= 0.71 + 0.95 \cdot V - 0.32 \cdot V^3 - 0.91 \cdot W, \\ \dot{W} &= 0.05 + 0.08 \cdot V - 0.06 \cdot W \end{aligned}$$

DSTLS

$$\begin{aligned} \dot{V} &= 0.85 + 0.81 \cdot V + 0.27 \cdot W \cdot V^2 + 0.28 \cdot V \cdot W^2 \\ &\quad - 1.02 \cdot W - 0.16 \cdot V^2 - 0.22 \cdot V^3 - 0.44 \cdot V \cdot W \\ &\quad - 0.09 \cdot V^2 \cdot W^2, \\ \dot{W} &= 0.08 \cdot V \end{aligned}$$

STLS

Furthermore, a comparison of STLS, DSTLS, and SR3 [2] optimizers was performed with respect to the variance of AWGN.

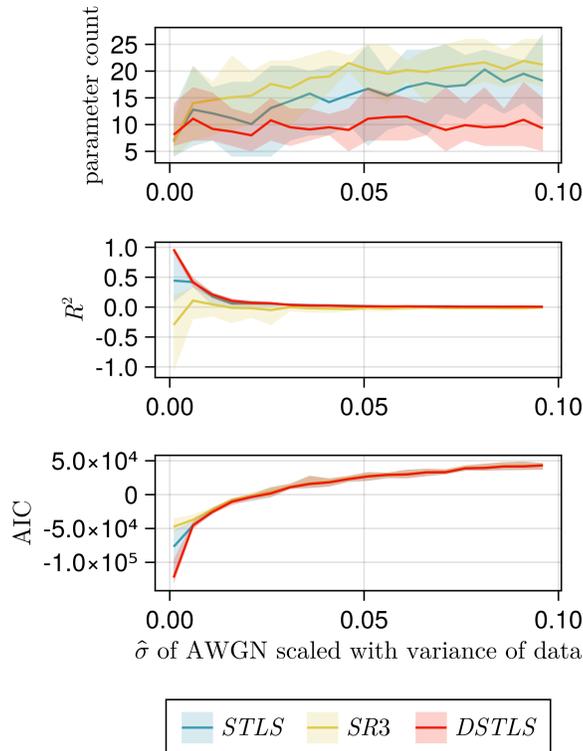


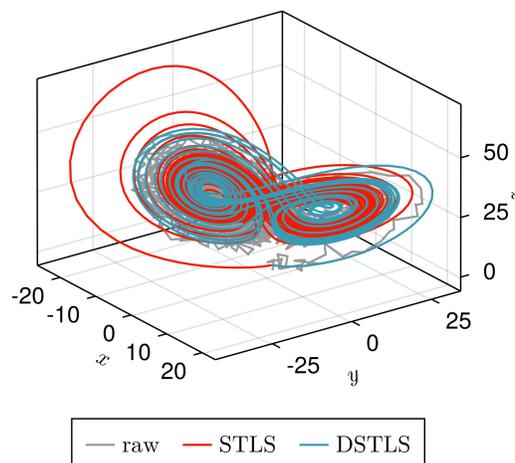
Figure 1: The effect of the magnitude of applied noise and used optimizer on the behavior of SINDy learning the FHN model. The DSTLS optimizer best retains sought qualities of low parameter count, high  $R^2$  and low AIC with a rising variance of noise in the FitzHugh-Nagumo model.

### Effects on Lorenz system

Although DSTLS was designed to be used on models with differences in magnitudes of parameters between equations, it still reliably works in other cases. As an example, we show its behavior on the well-known Lorenz model:

$$\begin{aligned} \dot{x} &= \sigma(y - x), \\ \dot{y} &= x(\rho - z) - y, \\ \dot{z} &= xy - \beta z \end{aligned}$$

with the standard choice of chaos inducing parameters  $\sigma = 10, \rho = 28, \beta = \frac{8}{3}$ . The derivatives are unknown and trajectory data is affected by AWGN with variance equal to 1 percent of the data's variance. Polynomials up to 3rd order were used as a candidate library. Due to the chaotic nature of the said system, prediction of the exact same trajectory with the discovered model cannot be expected, but the goal is to identify the attractor reasonably well. Only trajectories lying on the attractor were used for training.



### Acknowledgements

This work is supported by GAMU Interdisciplinary project # MUNI/G/1213/2022 and Mathematical and Statistical Modeling project # MUNI/A/1132/2022.

## Transformations of HH onto FHN model with Optim.jl

The highly nonlinear 4D *Hodgkin-Huxley* (HH) system is a physiological neuron model [3], unlike its conceptual simplification in the form of the 2D polynomial FHN model. As such, we strive to fit these two models to obtain physiological parameters for the FHN model as well.

### Transformation optimization problem

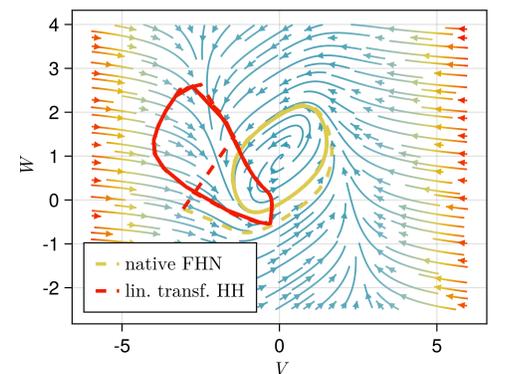
The trajectory  $\mathbf{H}$  of the HH model is presumed to be the truth and we strive to find both the transformation  $\mathbf{\Lambda} \in \mathbb{R}^{4 \times 2}$  and the parameters  $\mathbf{p}$  for FHN model with a trajectory  $\mathbf{u}(T; \mathbf{p})$  as a solution to the problem

$$\min_{\mathbf{\Lambda}, \mathbf{p}} \text{loss}(\mathbf{u}(T; \mathbf{p}), \mathbf{H}\mathbf{\Lambda}), \quad (1)$$

such that for  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^{k \times 2}$  the loss function reads

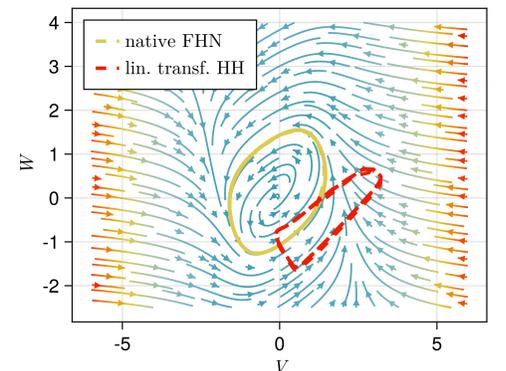
$$\begin{aligned} \text{loss}(\mathbf{x}, \mathbf{y}) &= \frac{1}{k} \|\mathbf{x} - \mathbf{y}\|_2^2 + \alpha \text{osc}(\mathbf{x}, \mathbf{y}) \\ &\quad + \beta \text{osc}(\mathbf{x}_{1:\lfloor \frac{k}{2} \rfloor}, \mathbf{x}_{\lfloor \frac{k}{2} \rfloor + 1 \dots k}), \\ \text{osc}(\mathbf{a}, \mathbf{b}) &= \left| 1 - \frac{\text{Var}(\mathbf{a})}{\text{Var}(\mathbf{b})} \right| \end{aligned}$$

The optimization problem (1) has a non-convex loss function and due to the inclusion of the simulated trajectory  $\mathbf{u}(T; \mathbf{p})$  of the FHN model, computation of the gradient would be expensive. Therefore we employed the Nelder-Mead method from the Optim.jl package, which does not require the gradient.



### Enhancing fit with Metaheuristics.jl

While the Nelder-Mead method works reasonably well, in our case it happens to be very sensitive to the initial condition, i.e. initial transformation matrix  $\mathbf{\Lambda}$  and parameters  $\mathbf{p}$ . To overcome this issue, we utilized the *Whale Optimization Algorithm* (WOA) from the Metaheuristics.jl package, which can be thought of as using an ensemble of initial conditions.



### References

- BRUNTON, Steven L.; PROCTOR, Joshua L.; KUTZ, J. Nathan. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*. 2016, vol. 113, no. 15, pp. 3932–3937.
- ZHENG, Peng; ASKHAM, Travis; BRUNTON, Steven L.; KUTZ, J. Nathan; ARAVKIN, Aleksandr Y. *A Unified Framework for Sparse Regularized Regression: SR3*. 2018.
- HODGKIN, A. L.; HUXLEY, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of Physiology*. 1952, vol. 117, no. 4, pp. 500–544.