COB-2021-0316
THE EFFECT OF DATA SELECTION ON THE SPARSE IDENTIFICATION OF DYNAMICAL SYSTEMS

Davi Saadi de Almeida Lettieri
Leonardo Santos de Brito Alves
Departamento de Engenharia Mecânica, Universidade Federal Fluminense, Niterói, RJ 24210-240, Brasil
davilettieri@id.uff.br
lsbalves@id.uff.br

Abstract. *Machine learning has continuously provided new and more efficient techniques for the discovery of governing equations. A well known branch of such techniques is known as symbolic regression. Although it had been left aside for a long time due to its high costs, it has been recently brought back through the development of SINDy, i.e. Sparse Identification of Nonlinear Dynamics. This technique is analyzed here when applied to the identification of partial differential equations, more specifically the unsteady, one-dimensional and viscous Burgers' equation. The analysis is performed by choosing different temporal data sets generated at different spatial locations during different time spans in order to investigate SINDy's ability to rediscover the Burgers' equation while employing different libraries of candidate functions. The results indicate a strong sensitivity to the spatial location and time period chosen, even when the library contains only the original terms of this equation.*

Keywords: *Machine Learning, Symbolic Regression, SINDy, Burgers' Equation, Data Selection.*

1. INTRODUCTION

Most times when studying a physical phenomena of interest, an available mathematical model able to describe its behavior makes possible a deeper and richer analysis. In this sense, for centuries scientists searched for best system identification techniques including the data-based modeling today known as symbolic regression, which can be defined as a symbolic analysis that searches for the best model that fits available data set using a given space of candidate functions. However, the need of high computational power alongside a series of required assumptions on the form of the model resulting in linear dynamics made it left aside for a long time. Recently, (Schmidt and Lipson, 2009) used symbolic regression to successfully develop a new approach to identify nonlinear dynamical system, bringing back this topic although the high cost, bad scaling to larger systems and over-fitting were still a problem. Moreover, (Brunton *et al.*, 2016), taking advantage of sparse regression (Tibshirani, 1996) and compressive sensing (Donoho, 2006), developed the Sparse Identification of Nonlinear Dynamics commonly known as SINDy which reduces significantly the symbolic regression cost. Furthermore, another recent work (Alves, 2020) observed at the time most SINDy's application were considering only low order polynomial nonlinearities in the candidate functions space, and showed that as the library of candidate functions matrix assumes a Vandermonde form increasing the size of matrix would not only turn it more ill-conditioned but also increases the error propagation preventing SINDy from identifying the correct nonlinear dynamical model and explaining why the LASSO regularization fails as the maximum nonlinearity order becomes larger.

In the present work we analyse the behavior of SINDy applied to different data sets in different spatial points and time periods generated by a solution of an partial differential equation, the unsteady, one-dimensional and Burgers' equation. The analysis indicates the existence of a region in the spatial domain with higher probability towards convergence associated with lower condition number.

2. METHODOLOGY

Since the main focus in this paper is the analysis of SINDy's behavior as we select distinct spatial points for a time period we may consider a system of partial differential equation in the general form

$$\frac{\partial \boldsymbol{\xi}(\boldsymbol{x},t)}{\partial t} = \boldsymbol{f}(\Theta^{d}(\boldsymbol{\xi}(\boldsymbol{x},t))), \tag{1}$$

where the state vector and function can be written as

$$\boldsymbol{\xi}(\boldsymbol{x},t) = \{\xi_1(\boldsymbol{x},t), \xi_2(\boldsymbol{x},t), \ldots, \xi_n(\boldsymbol{x},t)\}^T, \tag{2}$$

and

$$\boldsymbol{f}(\boldsymbol{\xi}(\boldsymbol{x},t)) = \{f_1(\boldsymbol{\xi}(\boldsymbol{x},t)), f_2(\boldsymbol{\xi}(\boldsymbol{x},t)), \ldots, f_n(\boldsymbol{\xi}(\boldsymbol{x},t))\}^T, \tag{3}$$

respectively, in which $\boldsymbol{x}$ representing the space coordinates. Moreover, the operator $\boldsymbol{\Theta}^d$ represents all maximum d^{th} order partial derivatives combinations possible.

2.1 SINDy

Some few assumptions must be made about Eq. (1), Eq. (2) and Eq. (3), in order to properly use SINDy. They are:

1. The state size n is arbitrary but small;
2. The state vector time history $\boldsymbol{\xi}(\boldsymbol{x},t)$ is avaiable from data;
3. The state function $\boldsymbol{f}(\boldsymbol{\xi}(\boldsymbol{x},t))$ dependence on the state vector $\boldsymbol{\xi}(\boldsymbol{x},t)$ is unknown but sparse.

In order to accomplish the second assumption the data is colected, associated with a sampling rate selection m and a period $\tau = t_m - t_1$,

$$\boldsymbol{\Xi} = \begin{pmatrix} \boldsymbol{\xi}(\boldsymbol{x},t_1)^T \\ \boldsymbol{\xi}(\boldsymbol{x},t_2)^T \\ \vdots \\ \boldsymbol{\xi}(\boldsymbol{x},t_m)^T \end{pmatrix} = \begin{pmatrix} \xi_1(\boldsymbol{x},t_1) & \xi_2(\boldsymbol{x},t_1) & \ldots & \xi_n(\boldsymbol{x},t_1) \\ \xi_1(\boldsymbol{x},t_2) & \xi_2(\boldsymbol{x},t_2) & \ldots & \xi_n(\boldsymbol{x},t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \xi_1(\boldsymbol{x},t_m) & \xi_1(\boldsymbol{x},t_m) & \ldots & \xi_n(\boldsymbol{x},t_m) \end{pmatrix}_{m\times n} \tag{4}$$

and

$$\dot{\boldsymbol{\Xi}} = \begin{pmatrix} \dot{\boldsymbol{\xi}}(\boldsymbol{x},t_1)^T \\ \dot{\boldsymbol{\xi}}(\boldsymbol{x},t_2)^T \\ \vdots \\ \dot{\boldsymbol{\xi}}(\boldsymbol{x},t_m)^T \end{pmatrix} = \begin{pmatrix} \dot{\xi}_1(\boldsymbol{x},t_1) & \dot{\xi}_2(\boldsymbol{x},t_1) & \ldots & \dot{\xi}_n(\boldsymbol{x},t_1) \\ \dot{\xi}_1(\boldsymbol{x},t_2) & \dot{\xi}_2(\boldsymbol{x},t_2) & \ldots & \dot{\xi}_n(\boldsymbol{x},t_2) \\ \vdots & \vdots & \ddots & \vdots \\ \dot{\xi}_1(\boldsymbol{x},t_m) & \dot{\xi}_1(\boldsymbol{x},t_m) & \ldots & \dot{\xi}_n(\boldsymbol{x},t_m) \end{pmatrix}_{m\times n}, \tag{5}$$

then the first time derivative state vector can be constructed by direct measure or numerically approximate. After that, according with the third assumption we must define a library of candidate nonlinear functions so we can estimate the unknown dependence of $\boldsymbol{f}(\boldsymbol{\Theta}^d(\boldsymbol{\xi}(\boldsymbol{x},t)))$ on $\boldsymbol{\xi}(\boldsymbol{x},t)$. Hence, we come up with a matrix of candidate functions $\boldsymbol{F}$

$$\boldsymbol{F}(\boldsymbol{\Xi}) = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \boldsymbol{\Xi} & \boldsymbol{\Xi}^2 & \boldsymbol{\Xi}^3 & \ldots & \boldsymbol{\Xi}^p \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}_{m\times Q}, \tag{6}$$

where Q is the total number of different functions in the library and the term $\boldsymbol{\Xi}^p$ indicates the maximum nonlinearity order p^{th} for higher polynomials. For example, $\boldsymbol{\Xi}^2$ denotes a submatrix in which its elements presents only quadratic nonlinearities, including the partial derivatives of the state vector,

$$\boldsymbol{F}(\boldsymbol{\Xi}) = \begin{pmatrix} \frac{\partial^{d_1+d_2}\boldsymbol{\xi}(\boldsymbol{x},t)}{\partial \boldsymbol{x}^{d_1}\partial t^{d_2}} \frac{\partial^{d_1+d_2}\boldsymbol{\xi}(\boldsymbol{x},t)}{\partial \boldsymbol{x}^{d_1}\partial t^{d_2}} \end{pmatrix}_{m\times q}, \tag{7}$$

where q is the number of different functions in this matrix and $d_1 = \sum_{\delta=1}^{n} \delta$, is the sum of all derivative orders $\delta \geq 0$ associated with the spatial coordinates. Furthermore, $d_1 = 0, 1, \ldots, d$ and $d_2 = 0, 1, \ldots, d$, but $0 \leq d_1 + d_2 \leq d$. The only combination that satisfies these conditions but is not allowed is $d_1 = 0$ e $d_2 = 1$ when $p = 1$. Finally, defining the matrix of unknown linear coefficients $\boldsymbol{C}$ as

$$\boldsymbol{C} = \begin{pmatrix} c_{1,1} & c_{1,2} & \ldots & c_{1,n} \\ c_{2,1} & c_{2,2} & \ldots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{Q,1} & c_{Q,2} & \ldots & c_{Q,n} \end{pmatrix}_{Q\times n}, \tag{8}$$

means we can estimate the functional form of each $f_j(\boldsymbol{\xi}(\boldsymbol{x},t))$ from its respective data set $\boldsymbol{\Xi}_j$ by solving

$$\dot{\boldsymbol{\Xi}}_j^T = \boldsymbol{F} \cdot \boldsymbol{C}_j^T, \tag{9}$$

instead of Eq. (1) for each column $j = 1, 2, \dots, n$.
It is important to note that $m \gg Q$ in most applications, i.e., $\boldsymbol{F}$ is a rectangular (low rank) matrix. Hence, Eq. (9) is over-determined and, in fact, represents a linear optimization problem. In order to take advantage of the third assumption,, the objective function $\boldsymbol{v}_j$ to be minimized is usually defined as

$$\boldsymbol{v}_j = \|\dot{\boldsymbol{\Xi}}_j^T - \boldsymbol{F} \cdot \boldsymbol{C}_j^T\|_2 + \lambda \|\boldsymbol{C}_j^T\|_1, \tag{10}$$

where λ is the sparsity identification LASSO regularization parameter.

2.2 The test case

The analysis made here to evaluate SINDy's performance took as reference the one-dimensional Burgers equation (Burgers, 1939), one of the simplest mathematical formulation illustrating the competition between diffusion and convection, which has the form

$$\frac{\partial u}{\partial t} = \nu \frac{\partial^2 u}{\partial x^2} - u \frac{\partial u}{\partial x}, \tag{11}$$

where u is the fluid velocity and ν represents the kinematic viscosity. Furthermore, the generated data was based on two different solutions from (Benton and Platzman, 1972), where they used nondimensional variables after some substitutions and mathematical transformations such as

$$x/L \to x, \nu t/L^2 \to t \text{ and } uL/\nu \to u \tag{12}$$

where L is the domain's total length, which had the effect of $\nu = 1$ in Eq. (11), even when we are able to modify this parameter. Thus, the solutions chosen were

1. $$u(x,t) = \frac{x/t}{1 + e^{x^2/4t}\sqrt{t}}, \tag{13}$$

2. $$u(x,t) = \frac{2}{\sqrt{\pi t}} \frac{e^{-z^2}}{\alpha + \text{erfc}(z)}, \quad \text{where } z = x/(2\sqrt{t}) \tag{14}$$

the first one illustrates the decay of a solitary pair of equal compression and expansion pulses, while the other represents a solitary compression pulse, and a similarity solution, here the parameter α has a relation with the Reynolds number R, given by $R = 2\ln(1 + 2\alpha^{-1})$, thus, for small values of R diffusion dominates, otherwise, convection prevails.

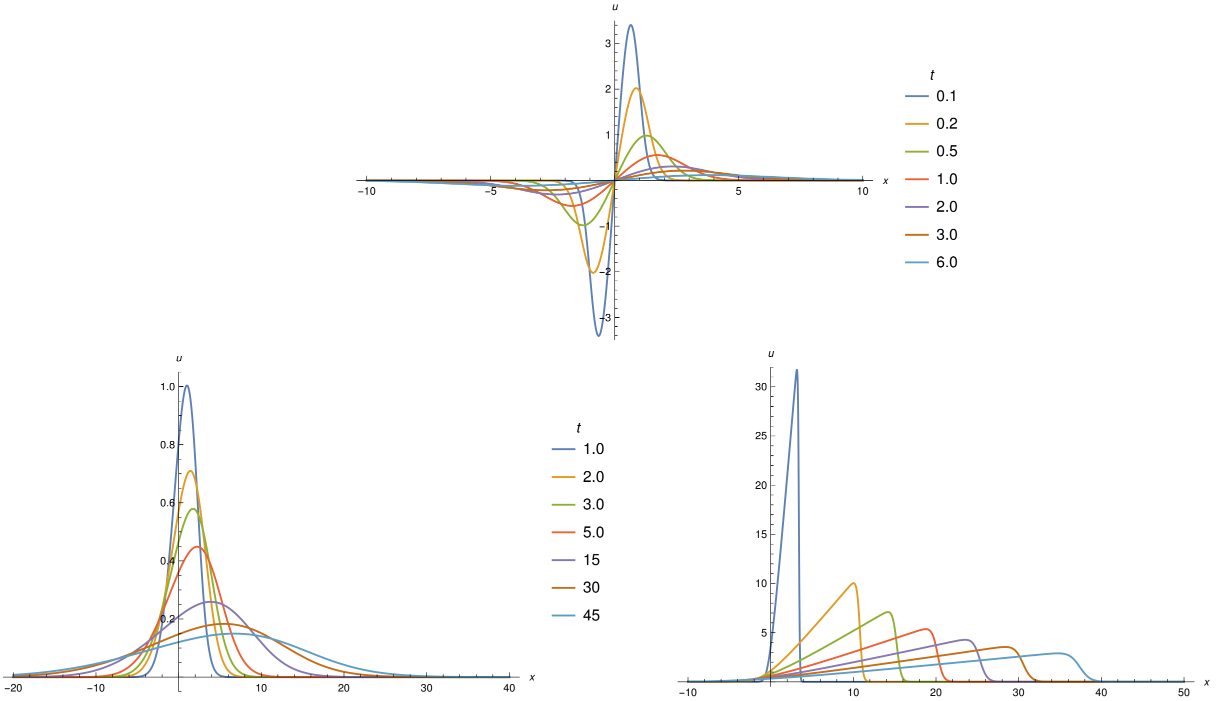


Figure 1: Solution 1 (above) and solution 2 (below) for different $R = 3.6$ (left) and $R = 64$ (right).

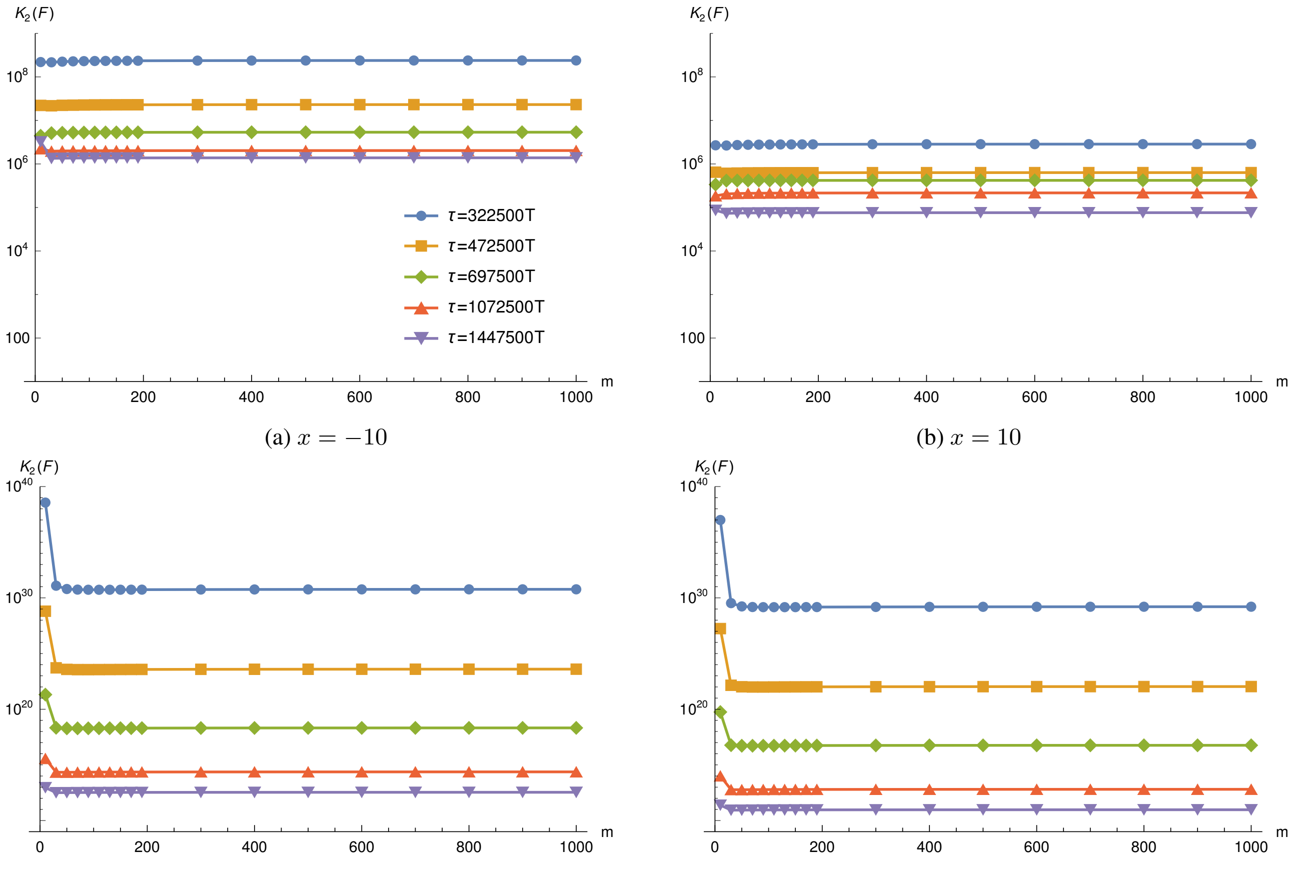
3. RESULTS

In order to investigate the behavior of SINDy when dealing with a partial differential equation we follow the same approach from (Alves, 2020) which assumed several inverse problems scenarios instead of a machine learning one, where we already know the goal model. Hence, for the Burgers' equation, the two linear coefficients are unknown. Furthermore the same work showed that a good SINDy's performance is usually associated with a low condition number. In this sense, as we are dealing with an one-dimensional partial differential equation we look forward to evaluate if the described trend is also followed by distinct regions on the spatial domain, in other words, we are searching for regions that may present better performance as we observe the behavior of its matrix condition number.

Thus, first we evaluate the condition number on a space domain for different parameters in order to choose the optimal final time and sampling rate. The condition number of a matrix $\mathbf{A}$ is defined as $k_i(\mathbf{A}) = \|\mathbf{A}\|_i \|\mathbf{A}^{-1}\|_i$, where $\|\cdot\|_i$ represents an arbitrary i norm and $\mathbf{A}^{-1}$ is the inverse of $\mathbf{A}$. Furthermore, as we are dealing with a rectangular linear system a few assumptions must be made in order to validate the condition number effect: i) a pseudo-norm inverse is employed instead of a regular inverse and ii) either a maximum or a Euclidean norm is employed instead of an arbitrary norm (Demko, 1986). Again, taking as reference (Alves, 2020) we use $k_2(\mathbf{F}) = \|\mathbf{F}\|_2 \|\mathbf{F}^\dagger\|_2$ where $\mathbf{F}^\dagger$ is the Moore-Penrose inverse of $\mathbf{F}$ and $\|\cdot\|_2$ is the Euclidean norm.

Hence, if we analyze the domain $x \in \mathbb{R} \mid -60 \leq x \leq 60$, taking Eq. (14) as reference, for $R = 3.6$, the characteristic time can be set as $T = 2 \times 10^{-5}$ so we can investigate the condition number behavior for a few different points in the domain space for distinct time periods as function of the sampling rate m.

Figure 2 clearly shows that there is a maximum sampling rate beyond which the condition number does not decrease. However a maximum time period is only achievable near to $x = 0$, and taking a point further away from it weakens this trend indicating that it will only be true setting even longer time periods which can be confirmed in Fig. 3. Beyond that, it is also important to note a small the convective effect impacting the condition number across the domain, as the $x - axis$ negative side presents a lower value compared with the opposite region. This phenomena can indicate a association between the SINDy's performance and the solution used to generate the sample data, in other words, maybe there is a available region in space evaluated at certain time period with higher probability of convergence as it presents a lower condition number.

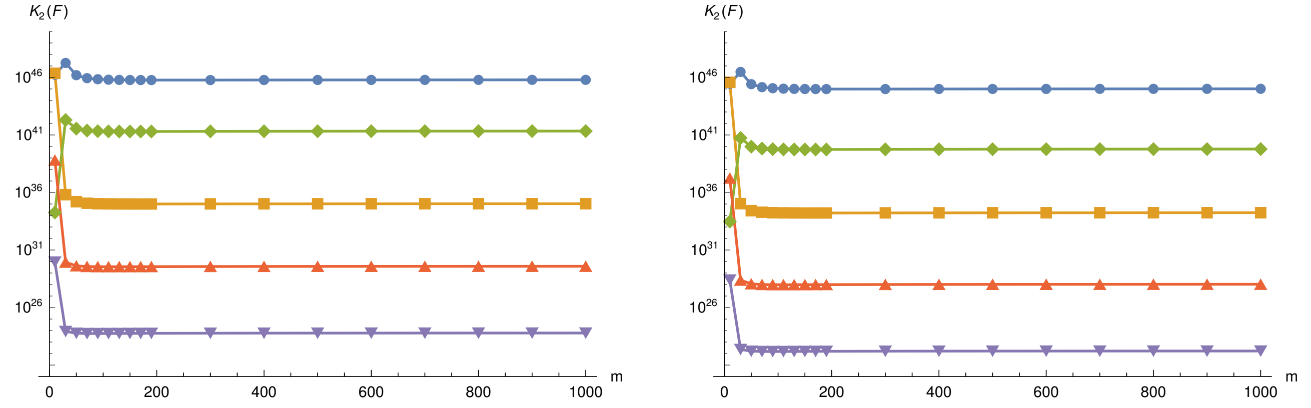


(a) $x = -10$

(b) $x = 10$

(c) $x = -30$

(d) $x = 30$



(e) $x = -50$

(f) $x = 50$

Figure 2: Condition number of the library matrix versus sampling rate m for Eq. (14) when $R = 3.6$ with $p = 1$, $d = 2$, $t_1 = 1.05$ and $T = 2 \times 10^{-5}$ for different sampling periods evaluated in distinct spatial points.

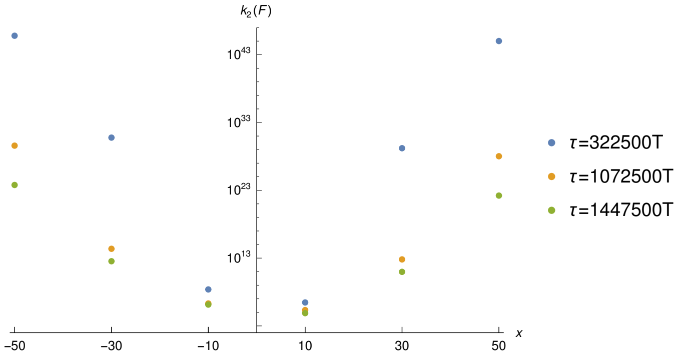


Figure 3: Optimal condition number across the spatial domain for different sampling periods for Eq. (14) when $R = 3.6$ with $p = 1$, $d = 2$, $t_1 = 1.05$, $T = 2 \times 10^{-5}$ and $m = 1000$.

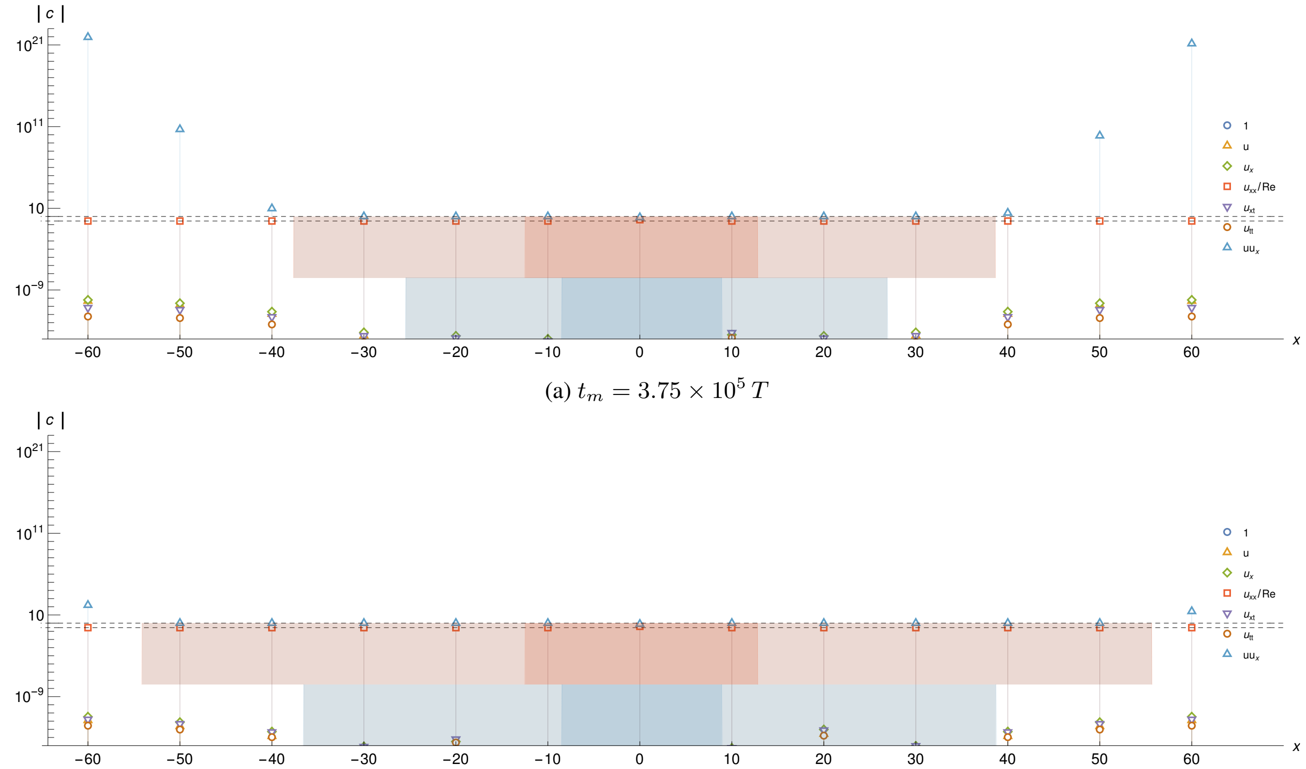
Thus, for better observation, we evaluate the values of the linear coefficients considered unknown for the points highlighted in Fig. 2 and Fig. 3. In this case, we only expect a better SINDy performance far from $x = 0$ if we employ longer time periods as a low condition number indicates a higher probability of good convergence towards the data-generated model. Table 1 illustrates that exact behavior as it include accurate values of both coefficients near $x = 0$ while the nonlinear one acquires extremely high magnitude away from it.

Table 1: Linear coefficients for each Burgers' equation term, for Eq. (14) when $R = 3.6$, obtained by SINDy with $p = 1$, $d = 2$, $t_1 = 1.05$ and $m = 1000$. For each term the first line considers $t_m = 10.5$ while the last $t_m = 22.5$.

Term	$x = -50$	$x = -30$	$x = -10$	$x = 10$	$x = 30$	$x = 50$
u_{xx}	0.2777777 0.2777777	0.2777777 0.2777777	0.2777777 0.2777777	0.2777777 0.2777777	0.2777777 0.2777777	0.2777777 0.2777777
uu_x	4.6717836×10^{10} -0.99938914	-1.0000000 -1.0000000	-1.0000000 -1.0000000	1.0000000 1.0000000	-1.0000000 -1.0000000	-7.7224063×10^{9} -1.0001010

Figure 4 is another way to exhibit this behavior across all chosen spatial domain also including the coefficient values of the other candidate functions for the specific library. It becomes even more clear the already mentioned association between SINDy's performance and the solution used to generate the data. As we keep the initial time, increasing t_m enable more spatial points far from $x = 0$ to achieve a satisfactory convergence. Again, although not very strong as the convective effect is not really significative, is possible to observe that the right side of the spatial domain is favored as the convective effect leads to higher values faster than the opposite side.

The region plotted is an attempt to estimate an area where the probability of convergence towards the correct coefficients of the Burger's equation occurs. It represents the region where the respective term coefficient value based on the solution taken as reference is greater than a fixed tolerance tol for both the initial and final time period. In this work,the value 10^{-15} was employed because we consider 15 digits of precision for the generated data. Moreover, it is clear that the nonlinear term coefficient is much more sensitive to this area, as it is necessary to the point be at least inside the linear coefficient region to achieve an accurate convergence. For this particular case, the only exception point inside the area



(a) $t_m = 3.75 \times 10^5\, T$

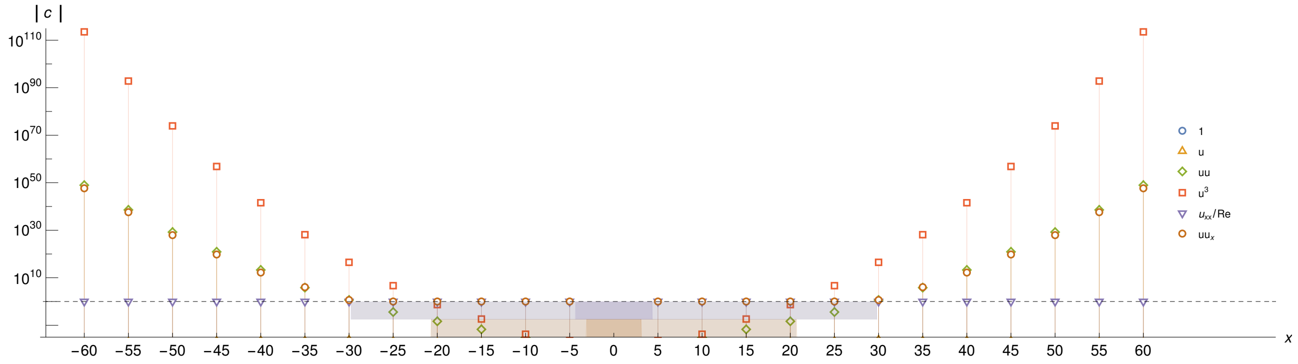
(b) $t_m = 1.125 \times 10^6\, T$

Figure 4: Linear coefficients for the Burgers' equation obtained by SINDy using Eq. (14) when $R = 3.6$ with $p = 1$, $d = 2$, $t_1 = 1.05$, $T = 2 \times 10^{-5}$ and $m = 1000$. Each region is associated with the correspondent color of the reference equation term and shows the range where their respective values are larger than 10^{-15}, for t_1 (darker) and t_m (lighter).

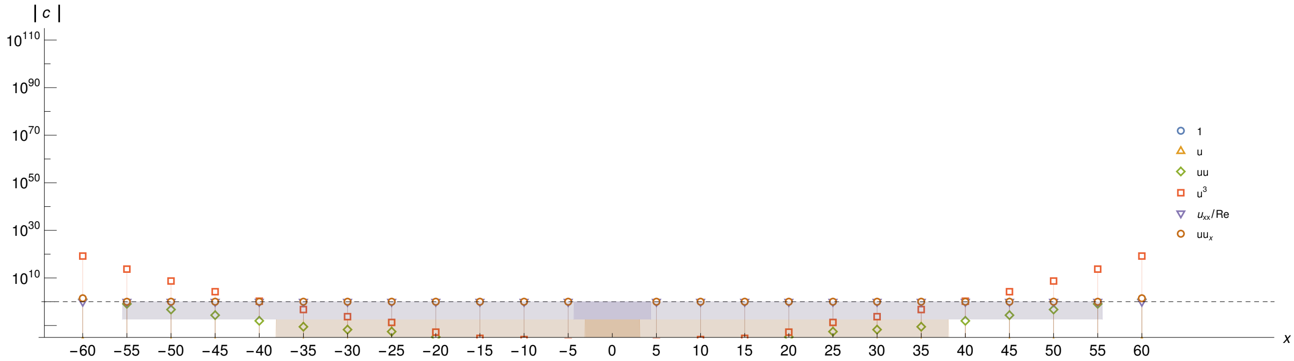
that does not vary for different parameters neither show an accuracy as good as the others is when $x = 0$, where the coefficients for the linear and nonlinear terms are 0.4005760 and -0.7653720, respectively.

The same behavior can be observed when the database is generated by Eq. (13) or Eq. (14) when $R = 64$ in Fig. 5 and Fig. 6, respectively. In the first one, again, only the points inside the auxiliary region will exhibit a good convergence for the nonlinear term coefficient, assuming a symmetric shape about a vertical axis on $x = 0$ as the absolute solution is used as reference. Moreover, Fig. 6 shows the already mentioned convective effect towards the positive values of x when $R = 3.6$, but now with a clearer impact as it becomes stronger for greater values of R. Interestingly, when $x = 0$, the same values already mentioned were achieved even now employing $R = 64$ and using a new library matrix.

Furthermore, another topic that is relevant when dealing with SINDy is the amount of unphysical terms that appear in the analysis. In Fig. 4 particular case, these terms seem to be irrelevant in most spatial points along the colored region, but is not clear that the values increase as we move away from it. However, increasing the nonlinearity or the number of different candidate function has a heavy impact in this matter. Figures 5 and 6 shows exactly the latter, where a good SINDy's performance without significative values of meaningless terms is only probably to occur in points with very slow condition number.

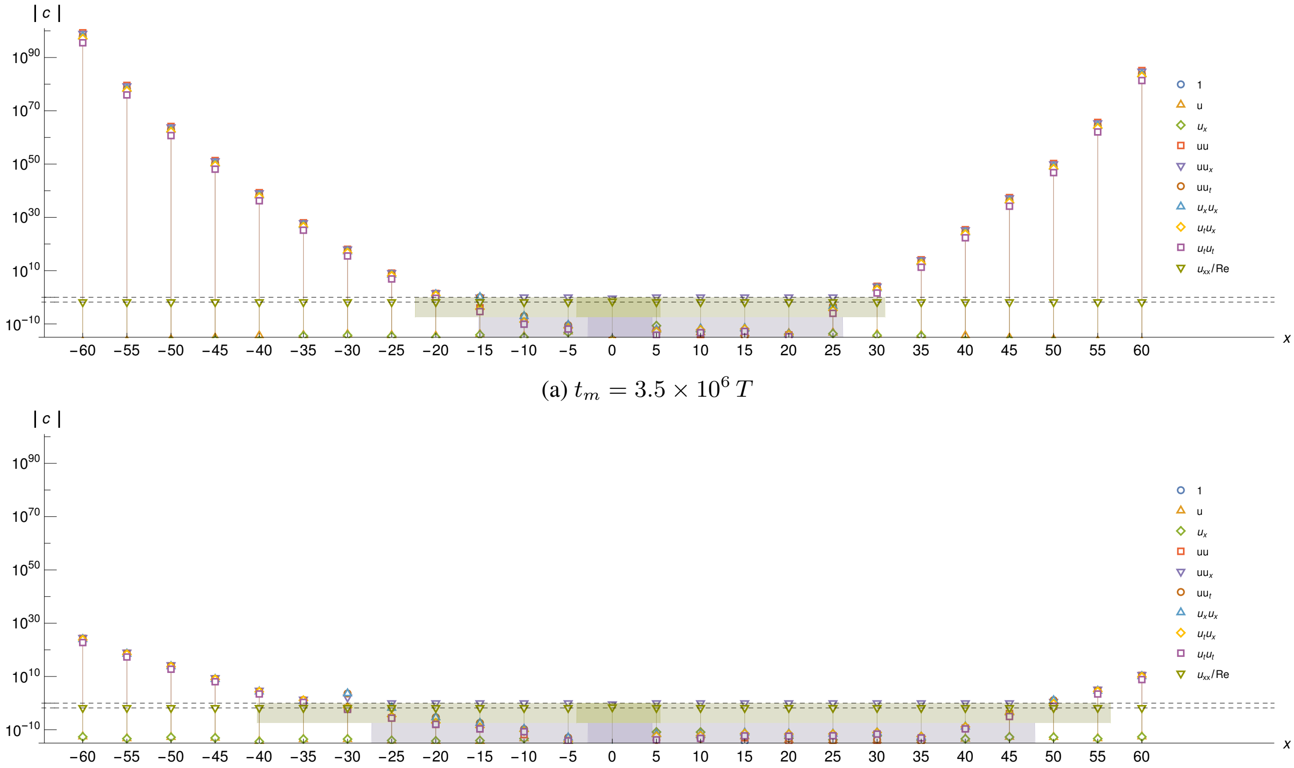


(a) $t_m = 6 \times 10^4\, T$



(b) $t_m = 2.25 \times 10^5\,T$

Figure 5: Linear coefficients for the Burgers' equation obtained by SINDy using Eq. (13) with $p = 3$, $d = 0$, $t_1 = 0.1$, $T = 1 \times 10^{-4}$ and $m = 1000$, $tol = 10^{-15}$.

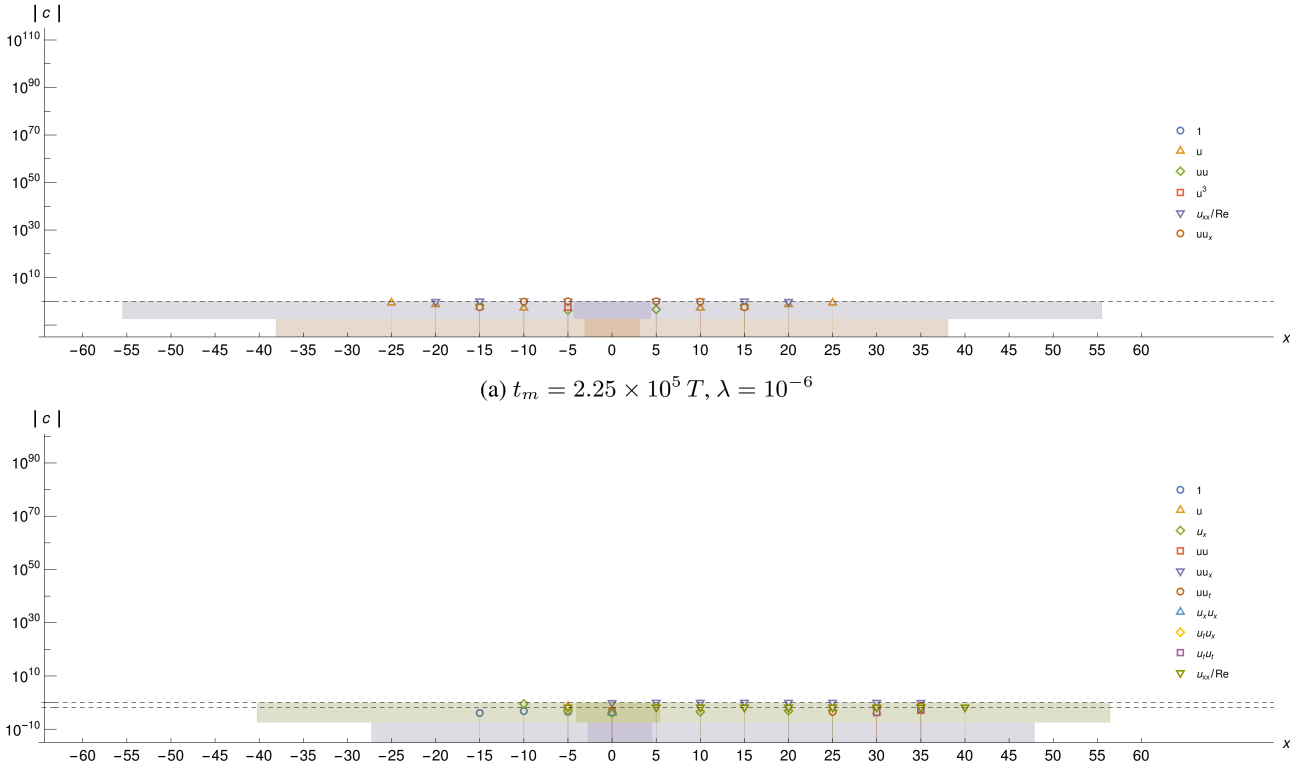


(a) $t_m = 3.5 \times 10^6\,T$

(b) $t_m = 12 \times 10^6\,T$

Figure 6: Linear coefficients for the Burgers' equation obtained by SINDy using Eq. (14) when $R = 64$ with $p = 2$, $d = 1$, $t_1 = 0.1$, $T = 1 \times 10^{-6}$ and $m = 1000$, $tol = 10^{-15}$.

Unfortunately using the LASSO regularization for several λ did not help, in fact most of the times even the coefficient values of the Burger's equation terms did not converge. Only for a limited set of parameters and spatial points where we could already distinguish the real physical terms, based their coefficients magnitude, it has a non negative impact, in this case they just approximate the lowest coefficients to zero. However, it was observed that other regions where SINDy without regularization performed well produced a worse estimated model. In fact, we could argue that the best set of regularization parameters only shrank the plotted area used to predict where a good result can be found. Figure 7 (below) illustrate this behavior, as we can see less points leading towards Burgers equation coefficients compared with Fig. 6 (below) where we employed the same parameters without LASSO. Choosing $\lambda < 10^{-3}$ do not have a significant impact. Moreover that is only true if we employ low order polynomials in the library matrix, otherwise, as Fig. 7 (above) shows, no points selected presented a satisfactory result and for $\lambda < 10^{-6}$ similar models were obtained. However, its clear that where SINDy without regularization showed poor performance, in the region where the solution tends to zero, for almost every case the coefficients found were null. This indicates that the regularization does not work well when dealing with very small magnitude values of data.

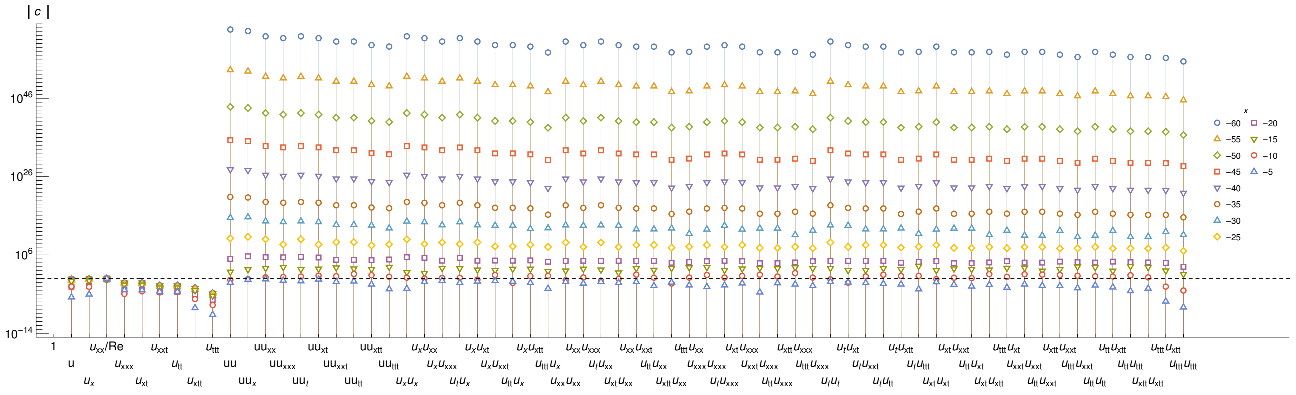


(a) $t_m = 2.25 \times 10^5\, T$, $\lambda = 10^{-6}$

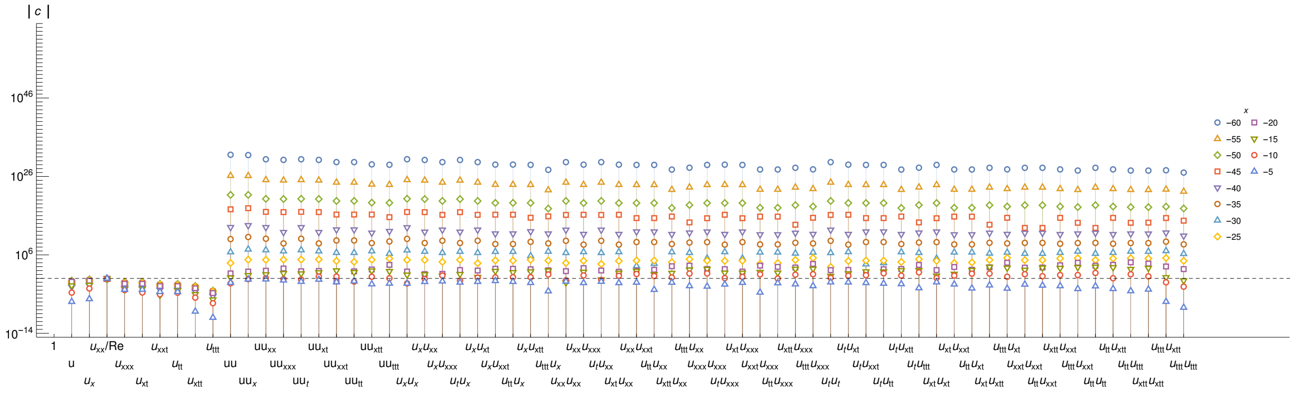
(b) $t_m = 12 \times 10^6\, T$, $\lambda = 10^{-3}$

Figure 7: Linear coefficients for the Burgers' equation obtained by SINDy using Eq. (13) with $p = 3$, $d = 0$, $t_1 = 0.1$, $T = 1 \times 10^{-4}$ (above) and using Eq. (14) when $R = 64$ with $p = 2$, $d = 1$, $t_1 = 0.1$, $T = 1 \times 10^{-6}$ (below), $m = 1000$, using a LASSO regularization.

We applied SINDy for several different library matrices, they varied from including only the exact terms of the original reference equation up to fourth order nonlinearity also modifying the number of different functions based on derivative orders following each solution used, however the increase of nonlinearity's order also limits the number of new functions in the matrix, i.e., most of the times we were not able to achieve a satisfactory convergence when employing a candidate function matrix with both high nonlinearity and maximum derivative order as illustrated in Fig. 8. Furthermore, as expected the increase of the matrix of candidate functions size by including new derivatives or increasing the nonlinearity order has two bad effects: i) a good converge for the coefficients of Burger's terms only occurs if we take points nearer to $x = 0$, this effect is more relevant for the Burger's nonlinear term; ii) moreover the increase of the unphysical terms values is significative, in fact, we probably would not find the physical coefficients in a real scenario. Figure 8 also shows the latter, where we can see that the nonlinear terms far from the auxiliary region even for very long times do not decrease as desired.



(a) $t_m = 6 \times 10^4\, T$



(b) $t_m = 12 \times 10^4\,T$

Figure 8: Linear coefficients for the Burgers' equation obtained by SINDy using Eq. (13) with $p = 2$, $d = 3$, $t_1 = 0.1$, $T = 1 \times 10^{-4}$ and $m = 1000$.

4. CONCLUSION

The results showed above indicates that when applying SINDy in order to modeling a dimensional physical phenomena is always interesting to evaluate the condition number across the spatial domain as it clearly may be used to determine which points has higher probability of converge towards the more appropriate dynamical system.

5. ACKNOWLEDGEMENTS

The authors appreciate the financial support of CNPq.

6. REFERENCES

Alves, L.S.d.B., 2020. "Addressing overfitting issues in the sparse identification of nonlinear dynamical systems". In *Proceedings of the 18th Brazilian Congress of Thermal Sciences and Engineering - ENCIT 2020*. Online.

Benton, E.R. and Platzman, G.W., 1972. "A table of solutions of the one-dimensional burgers equation". *Quarterly of Applied Mathematics*, Vol. 30, No. 2, pp. 195–212.

Brunton, S.L., Proctor, J.L. and Kutz, J.N., 2016. "Discovering governing equations from data by sparse identification of nonlinear dynamical systems". *Proceedings of the national academy of sciences*, Vol. 113, No. 15, pp. 3932–3937.

Burgers, J.M., 1939. "Mathematical examples illustrating relations occuring in the theory of turbulent fluid motion". *Trans. Roy. Neth. Acad. Sci. Amsterdam*, Vol. 17, pp. 1–53.

Demko, S., 1986. "Condition numbers of rectangular systems and bounds for generalized inverses". *Linear Algebra and its Applications*, Vol. 78, pp. 199–206.

Donoho, D.L., 2006. "Compressed sensing". *IEEE Transactions on information theory*, Vol. 52, No. 4, pp. 1289–1306.

Schmidt, M. and Lipson, H., 2009. "Distilling free-form natural laws from experimental data". *science*, Vol. 324, No. 5923, pp. 81–85.

Tibshirani, R., 1996. "Regression shrinkage and selection via the lasso". *Journal of the Royal Statistical Society: Series B (Methodological)*, Vol. 58, No. 1, pp. 267–288.

7. RESPONSIBILITY NOTICE

The authors are solely responsible for the printed material included in this paper.