

Computer Assisted Discovery of Integrability via SILO: Sparse Identification of Lax Operators*

Jimmie Adriazola[†], Wei Zhu[‡], Panayotis G. Kevrekidis[§], and Alejandro Aceves[¶]

Abstract. We formulate the discovery of Lax integrability of Hamiltonian dynamical systems as a symbolic regression problem, which, loosely speaking, seeks to maximize the compatibility between a pair of Lax operators and the known Hamiltonian of the dynamical system. Our approach is first tested on the simple harmonic oscillator. We then move on to the Henon–Heiles system, i.e., a two-degree-of-freedom system of nonlinear oscillators. The integrability of the Henon–Heiles system is critically dependent on a set of three parameters within its Hamiltonian, a fact that we leverage to assess the robustness of our approach in detecting the integrability of this system with respect to the parameter dependence of the Hamiltonian. We then adapt our method to canonical examples of Hamiltonian partial differential equations, including the Korteweg–de Vries and cubic nonlinear Schrödinger equations, again testing robustness against nonintegrable perturbations of their respective Hamiltonians. In all examples, our approach reliably confirms or denies the integrability of the equations of interest. Moreover, by appropriately adjusting the loss function and applying thresholded l^0 regularization to enforce sparsity in the operator weights, we successfully recover accurate forms of the Lax pairs despite wide initial hypotheses on the operators. Some of the relevant Lax pairs, notably for the Henon–Heiles system and the Korteweg–de Vries equation, are distinct from the ones that are typically reported in the literature. The Lax pairs that our methodology discovers warrant further mathematical and computational investigation, and we discuss extensively the opportunities for further improvement of SILO as a viable tool for interpretable exploration of integrable Hamiltonian dynamical systems.

Key words. lax pairs, sparse regression, integrable Hamiltonian systems, statistical learning

MSC codes. 37K10, 65P99

DOI. 10.1137/25M1741716

*Received by the editors March 14, 2025; accepted for publication (in revised form) by V. Zharnitsky August 14, 2025; published electronically January 2, 2026.

<https://doi.org/10.1137/25M1741716>

Funding: The work of the first author was supported by the U.S. National Science Foundation under award PHY-2316622. The work of the second author was supported by the Air Force Office of Scientific Research (AFOSR) under grant FA9550-25-1-0079. The work of the third author was supported by the U.S. National Science Foundation under awards PHY-2110030, PHY-2408988 and DMS-2204702. The work of the fourth author was supported by the U.S. National Science Foundation under award DMS-2502900. The work of the first and third authors was also supported by the Society for Industrial and Applied Mathematics (SIAM) postdoctoral support program, established by Martin Golubitsky and Barbara Keyfitz.

[†]Corresponding author. School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ 85281 USA (jimmie.adriazola@asu.edu).

[‡]School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332 USA (weizhu@gatech.edu).

[§]Department of Mathematics and Statistics, University of Massachusetts, Amherst, MA 01003 USA; Department of Physics, University of Massachusetts, Amherst 01003 USA (kevrekid@umass.edu).

[¶]Department of Mathematics, Southern Methodist University, Dallas, TX 75205 USA (aaceves@smu.edu).

1. Introduction. The theory of integrable dynamical systems occupies a central place in the historical development of mathematical physics [43, 2, 41, 1], further inspiring branches of modern mathematical fields in differential geometry [24, 7], algebra [17, 41], and functional analysis [15, 53]. The classical sense of the term integrability refers to the exact solvability of a differential equation up to the calculation of a suitable number of integrals [36, 45]. More precisely, according to the Liouville–Arnold theorem, if an n degree-of-freedom Hamiltonian dynamical system has n algebraically independent integrals of motion, whose level sets are compact, then there exists a local coordinate transformation to a so-called action-angle torus [3]. To resolve the dynamics, one only needs to perform the necessary quadratures in this action-angle coordinate system [21]. Moreover, since the dynamics of an integrable system is constrained to take place on such tori, the dynamics cannot exhibit chaos and is thus regular and indeed generically quasi-periodic in its nature [44]. Of course, there are extended notions of integrability in various different mathematical contexts; see, for a recent discussion, e.g., [38]. Yet, what we have just described is the Liouville sense of integrability that we will often adopt throughout this paper when we refer to the integrability of a dynamical system.

If a Hamiltonian system is integrable, then Hamilton–Jacobi theory [28] or, notably, inverse scattering transform (IST) methods [2, 41, 1] may be employed to construct these action-angle coordinates and ultimately solve the physics problem. The IST amounts effectively to a nonlinear form of the Fourier transform in which suitable scattering data are propagated forward in time and eventually the solution of an inverse problem reconstructs the solution of the PDE at future times.

However, one does not know a priori if a given Hamiltonian dynamical system is integrable, a feature that would be highly desirable to recognize. Newton knew that the two-body gravitational problem is integrable; he solved the problem exactly [21]. Yet, while attacking the three-body problem in the late 17th century, Newton, when pressed by his colleague Edmund Halley, asserted that the problem [50] “made his head ache, and kept him awake so often, that he would think of it no more.” This problem later also challenged Poincaré, as has been eloquently described, e.g., in [16]. In the end, he correctly asserted the nonintegrability of the three-body problem. A mathematical and historical perspective of Poincaré’s work can be found in [4].

Thus, the efforts of Newton and all those who attempted to solve the three-body problem exactly were always in vain. Notwithstanding the beautiful mathematics that resulted from such efforts, had nonintegrability been confirmed at the time, it would have spared the celestial mechanics community throughout the scientific revolution great pain. Considering that there will always exist a vast number of physical models whose integrability is left to be determined by mathematicians and physicists, a tool for the automated detection of integrability retains its value even in the present day.

Founded on this principle, our aim in this paper is to develop and demonstrate a broad-scope detector of integrability for systems with a known Hamiltonian. The key technology we use in the detection of integrability is a symbolic regression [49] over a space of hypothesized Lax pairs [37]. With this tool that we call the Sparse Identification of Lax Operators (SILO), we expect to catalyze the further future study of Hamiltonian systems where the integrability of the dynamical system is unknown. We note in passing here that there have been substantial efforts during the past few years towards identifying that Hamiltonian directly from data

[6, 12, 22, 38, 48]. Accordingly, the discovery of integrability directly from data (rather than from equations) is a natural next step along the vein presented herein.

To keep this high-level and introductory discussion simple, let us consider a mechanical system with finite degrees of freedom n . Lax pairs in this setting are time-dependent matrices $L(t)$ and $P(t)$ in $\mathbb{R}^{n \times n}$ that satisfy the matrix equation

$$(1.1) \quad \frac{dL}{dt} = [L, P],$$

where $[L, P] = LP - PL$ denotes the matrix commutator. It is a fact that any finite-dimensional L , where each component $L_{ij} \in C^1([0, T])$ with $t \in [0, T]$, that satisfies (1.1) is isospectral [37, 35].

Now, recall that a Hamiltonian system in canonical coordinates, with a smooth Hamiltonian function $H : \mathbb{R}^{2n} \times [0, T] \rightarrow \mathbb{R}$, is given by

$$(1.2) \quad \begin{aligned} \dot{q}_i &= \frac{\partial H}{\partial p_i}, & i &= 1, \dots, n, \\ \dot{p}_i &= -\frac{\partial H}{\partial q_i}, & i &= 1, \dots, n. \end{aligned}$$

If equations (1.2) further arise as the compatibility conditions of (1.1), then this Hamiltonian dynamical system is said to be *Lax* integrable, with integrals of motion that can be constructed from the associated Lax operator L [35], assuming that the resulting integrals are in involution. If it can be shown that this Lax pair reproduces all conserved quantities or satisfies certain geometric conditions, then the Lax integrability implies Liouville integrability [23]. Nevertheless, the presence of a Lax pair underlying the integrable dynamics enables us to reformulate the detection of integrability in the present work. This will be formulated and accomplished as a symbolic regression problem that seeks to maximize the compatibility (in the precise sense of (2.7) discussed below) between a hypothesized library of Lax pairs and a known Hamiltonian of the dynamics. This is the essence of our work.

We are aware of two previous attempts that use Lax pairs for numerical detection of integrability, as well as a third, different approach towards algorithmic discovery of integrability. The first is due to Krippendorf, Lüst, and Syvaeri [34], where neural networks are employed to find parametrized Lax pairs that are as compatible with the known Hamiltonian dynamical system as possible. Although this work was successful in its approach and, indeed, paved the way toward such data-driven methodologies for detecting integrability, its effectiveness was somewhat limited. Due to the use of neural networks, the numerical precision and interpretability of their results could, arguably, be further improved. Additionally, the sampling scheme used for their PDE problems assumes that one knows the soliton solution of the model. De Koster and Wahls took a different, more sophisticated data-driven approach to construct the spectral operator for members of the Ablowitz–Kaup–Newell–Segur (AKNS) hierarchy, even discovering the governing equations from noisy data [14]. However, this approach is inherently restricted, as it is only applied to AKNS systems (and thus uses rather limited libraries of functions). Moreover, this approach uses an error function that requires a priori knowledge (and necessitates the subsequent in time preservation) of the conserved quantities of the system.

A third, very recent approach, developed by Kantamneni, Liu, and Tegmark [29], attempts to search over the space of possible PDEs to maximize the number of conserved quantities. However, this approach is limited in that it can only discover a few conserved quantities at most (instead of an algorithmic sequence of infinitely many such quantities). Thus, it subsequently relies on human judgment to determine whether or not the hierarchy of conserved quantities can be continued onward towards an infinite number of such (which, moreover, need to be in involution).

We find that reinterpreting the problem of detecting integrability via Lax pairs as a sparse regression problem yields highly promising results. First, this allows us to obtain high-precision detection of integrability against nonintegrable Hamiltonian perturbations. Second, by using sparse regression techniques, we are able to rediscover accurate structural forms of the Lax pairs in numerous examples. These include ones that are accurate even though they were not—and in some cases, to the best of our knowledge, are not—the standard ones used in the literature. We note that the idea of searching over a wide library of hypothesized operators while maintaining a sparsification is similar in spirit to the Sparse Identification of Nonlinear Dynamics (SINDy) methodology that discovers governing equations from data [8]. A major departure from SINDy here is largely in the execution of the computations, since discovery of sparse Lax pairs from data involves an entirely different loss function than what is used in equation discovery.

Consequently, our work demonstrates that SILO advances us in the direction of building general, robust, interpretable, and reliable tools for precise mathematical discoveries in the context of integrable Hamiltonian systems. We also aim to be as transparent as possible about our methodology in that we discuss at length the limitations, computational choices, and many of the salient implementation details when using SILO. We reiterate here what we consider to be a major advantage of framing integrability detection via a sparse regression on Lax pairs, as opposed to a numerical search for models with a high, yet finite number of conserved quantities. That is, once Lax integrability is established, work by Gui-zhang [23] demonstrates that one can, in principle, elevate this to Liouville integrability by checking certain geometric conditions on the discovered Lax pair. Thus, the discovery of *Liouville* integrability from Lax integrability remains, in principle, automated with the use of computer algebra systems. Such an examination goes beyond the scope of our paper, but we make note of this here to communicate that such studies can be undertaken if desired.

The paper is organized as follows. In section 2, we build and motivate the loss function, used in the regression, for a system with one degree of freedom. This loss function, as we will show, scales to scenarios with many degrees of freedom. In this section, we also take the opportunity to detail our sparsification strategy, employed throughout this work, as much as possible. Since autonomous Hamiltonian systems with one degree of freedom are always integrable, we proceed to study a two-degree-of-freedom system, the Hénon–Heiles model, in section 3. Because the integrability of the Hénon–Heiles model depends on parameters defining its Hamiltonian, this case allows us to assess the accuracy of our method for detecting integrability. In sections 4 and 5, we extend and test our method for Hamiltonian nonlinear partial differential equations (PDEs), including the Korteweg–de Vries equation and the nonlinear Schrödinger equation, respectively. In all sections, we employ a sparsification technique for the regression to facilitate the interpretation of the Lax pair discovered by our numerics.

We also comment on the relevant findings in comparison to the “standard” Lax pair settings that exist in the corresponding theoretical literature.

2. Warm-up problem: The simple harmonic oscillator. In this section, we introduce the basic methodology we use to detect integrability for Hamiltonian systems and interpret the resulting Lax pairs. As an instructive warm-up, we first develop the method for the simple harmonic oscillator. We find that with appropriate adjustments, the techniques discussed in this section can be scaled to the discovery of Lax pairs for Hamiltonian PDEs, i.e., systems with infinitely many degrees of freedom. We emphasize again that while certain aspects of our methodology share similarities with those of Krippendorf, Lüst, and Syvaeri [34], our method differs significantly by adopting a more fundamental interpretation of the optimization problem formulated. Specifically, the approach presented here uses an expanded basis and a more robust sampling approach compared to the earlier work of [34]. Moreover, the subsequent sparsification process leads to the interpretable Lax pairs identified throughout this work. Furthermore, our work only uses the known Hamiltonian and is, in principle, agnostic to knowledge of conserved quantities or mathematical form of the Lax pair, such as what is assumed in [14].

To begin, recall that the dynamics of the simple harmonic oscillator are given by the following system:

$$(2.1) \quad \begin{aligned} \dot{q} &= \frac{\partial H}{\partial p} = p/m, \\ \dot{p} &= -\frac{\partial H}{\partial q} = -kq, \end{aligned}$$

where the Hamiltonian $H = (p^2/m + m\omega^2 q^2)/2$ and the spring constant k is defined in terms of the oscillator’s mass $m > 0$ and angular frequency $\omega \in \mathbb{R}$, namely $k = m\omega^2$.

A Lax pair reported in the literature for this system is given in [35]:

$$(2.2) \quad L = \begin{pmatrix} p/m & \omega q \\ \omega q & -p/m \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} 0 & \omega/2 \\ -\omega/2 & 0 \end{pmatrix}.$$

It is important to note that the Lax pair (2.2) corresponding to (2.1) is not unique; indeed, Krippendorf, Lüst, and Syvaeri [34] correctly point out that there exist two two-parameter families of Lax pairs

$$(2.3a) \quad L_1 = a \begin{pmatrix} p/m & b\omega q \\ \omega q/b & -p/m \end{pmatrix}, \quad P_1 = \begin{pmatrix} 0 & b\omega/2 \\ -\omega/2b & 0 \end{pmatrix},$$

$$(2.3b) \quad L_2 = a \begin{pmatrix} q & p/b\omega \\ bp/\omega & -q \end{pmatrix}, \quad P_2 = \begin{pmatrix} 0 & -\omega/2b \\ 2\omega/b & 0 \end{pmatrix},$$

where $a, b \in \mathbb{R}/\{0\}$. Also note that the only conserved quantity, the energy of the system E , is related to the eigenvalues of the operator L appearing in (2.2) and can be written as

$$(2.4) \quad E = \frac{1}{2} \left(\frac{p^2}{m} + m\omega^2 q^2 \right) = -\frac{m}{2} \det L = \frac{m}{4} \operatorname{tr} L^2.$$

Already at this elementary stage, the nonuniqueness of the Lax pairs prevents us from ever making the claim that we can reconstruct *the* Lax pair for a Hamiltonian dynamical system. Nevertheless, we will attempt to construct one such pair. We hypothesize that the operators \tilde{L} and \tilde{P} , satisfying Lax's equation we would like construct are of the form

$$\begin{aligned}\tilde{L}(t) &= \sum_{k=1}^{N_\xi} \xi_k \Theta_k^{(L)}(t), \\ \tilde{P}(t) &= \sum_{k=1}^{N_\zeta} \zeta_k \Theta_k^{(P)}(t),\end{aligned}$$

where the symbol Θ represents a user-specified library of matrices with parameters $\xi \in \mathbb{R}^{N_\xi}$ and $\zeta \in \mathbb{R}^{N_\zeta}$. In the case of the simple harmonic oscillator, we hypothesize that the 2×2 matrix operators contain terms that are at most linear in both q and p . This gives $N_\xi = N_\zeta = 2 \times 2 \times 3 = 12$. We make no other assumptions on the form of the operators \tilde{L} and \tilde{P} . Note that despite limiting each entry to a linear polynomial, we already have 24 parameters $\eta := [\xi \ \zeta]$ for which to search.

It is obvious that the size of this parameter space N_η is much too large, in fact by at least a factor of 6. Indeed, if we make the following assumptions,

- the operator \tilde{L} does not contain any constant terms,
- the operator \tilde{L} is a symmetric matrix with zero trace, and
- the operator \tilde{P} is skew-symmetric,

then there are only four parameters to find. To show this, we write the hypotheses explicitly:

$$\tilde{L} = \begin{pmatrix} \xi_1 q + \xi_3 p & \xi_2 q + \xi_4 p \\ \xi_2 q + \xi_4 p & -\xi_1 q - \xi_3 p \end{pmatrix}$$

and

$$\tilde{P} = \begin{pmatrix} 0 & \zeta_1 q + \zeta_2 p + \zeta_3 \\ -\zeta_1 q - \zeta_2 p - \zeta_3 & 0 \end{pmatrix}.$$

With this hypothesis, the governing equations are given by

$$(2.5) \quad \begin{aligned}\dot{q} &= 2\zeta_3 \left(\frac{\xi_2}{\xi_4} - \frac{\xi_1}{\xi_3} \right)^{-1} \left(\left(\frac{\xi_1}{\xi_4} + \frac{\xi_2}{\xi_3} \right) q + \left(\frac{\xi_3}{\xi_4} + \frac{\xi_4}{\xi_3} \right) p \right), \\ \dot{p} &= 2\zeta_3 \left(\frac{\xi_4}{\xi_2} - \frac{\xi_3}{\xi_1} \right)^{-1} \left(\left(\frac{\xi_1}{\xi_2} + \frac{\xi_2}{\xi_1} \right) q + \left(\frac{\xi_3}{\xi_2} + \frac{\xi_4}{\xi_1} \right) p \right),\end{aligned}$$

while

$$(2.6) \quad \text{tr}(L^2) = 2((\xi_1^2 + \xi_2^2)q^2 + (\xi_3^2 + \xi_4^2)p^2 + (\xi_1\xi_3 + \xi_2\xi_4)qp).$$

We see that any combination of coefficients satisfying $(\xi_1\xi_3 + \xi_2\xi_4) = 0$ eliminates the undesirable cross-terms appearing in (2.6) and the equations of motion (2.5). Moreover, the

parameters ζ_1 and ζ_2 are redundant, since they do not appear anywhere in the dynamics. These five remaining parameters along with the condition $(\xi_1\xi_3 + \xi_2\xi_4) = 0$ leave four total parameters.

Of course, having the hindsight to make such a judicious choice of assumptions on the Lax pairs is a luxury. When no user-specified knowledge is available, it is standard in data-driven science and statistics to sparsify the problem using some variant of the $l^p(\mathbb{R}^{N_\eta})$ norm [8, 25, 26]; this is an approach that we adopt in this work and will discuss in greater depth shortly.

We now motivate the loss function. The Lax equation includes a time derivative of the operator L . Numerically working with time derivatives naturally introduces a time discretization, leading to numerical errors. To circumvent this issue, we observe via the chain rule that

$$\begin{aligned}\frac{dL}{dt} &= \frac{\partial L}{\partial q} \dot{q} + \frac{\partial L}{\partial p} \dot{p} \\ &= \frac{\partial L}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial L}{\partial p} \frac{\partial H}{\partial q} \\ &:= \{L, H\},\end{aligned}$$

where the last line defines the Poisson bracket of the operator L with the Hamiltonian H . Evaluating this Poisson bracket $\{L, H\}$ is straightforward to do analytically by hand. Also, since $\frac{dL}{dt} = [L, P]$, it follows that

$$(2.7) \quad \{L, H\} - [L, P] = 0$$

should hold for any $q, p \in \mathbb{R}$. However, this last equation admits the trivial solution $L \equiv 0$, so care should be taken to penalize away from $\xi \equiv 0$.

With these pieces in hand, we formulate the sparse identification of a Lax pair associated with a given Hamiltonian dynamical system as the following optimization problem:

$$(2.8) \quad \min_{\eta \in \mathbb{R}^{N_\eta}} J[\eta] = \min_{\eta \in \mathbb{R}^{N_\eta}} \frac{1-r}{N_{\text{samples}}} \sum_{k=1}^{N_{\text{samples}}} \left[\sum_{i,j} \frac{(\{\tilde{L}, H\} - [\tilde{L}, \tilde{P}])_{i,j}^2}{\{\tilde{L}, H\}_{i,j}^2} \right]_{\Omega} + r\mathcal{R}(\eta),$$

where $r \in [0, 1)$ controls the amount of desired sparsification in the search. The first contribution to the loss function J rewards the discovery of the pointwise relationship (2.7), in phase space, that the Lax pair needs to satisfy. Since the numerator is divided by the Poisson bracket that acts on \tilde{L} , this formulation discourages trivial solutions where $\tilde{L} \equiv 0$. Since it is infeasible to access the continuous nature of the phase space, we restrict the evaluation to the set Ω , a subset of the phase space \mathbb{R}^2 , and N_{samples} is the number of times we sample from Ω . The sum over i, j ensures that we consider every component of the 2×2 matrices. In summary, we search for a pair of matrices that on average over some subset of the phase space minimizes the relative residual of the Lax equation, compatible with the Hamiltonian of the system, while maintaining parsimony through the sparsification function \mathcal{R} .

Typical choices for the sparsification function include $\mathcal{R}_0 = \|\eta\|_{l^0(\mathbb{R}^{N_\eta})}$ or $\mathcal{R}_1 = \|\eta\|_{l^1(\mathbb{R}^{N_\eta})}$.

We find that neither of these choices works quite as well as we could hope. After several numerical experiments, we found that

$$(2.9) \quad \mathcal{R}^*(\eta) := \frac{\|\tilde{\eta}\|_{l^0(\mathbb{R}^{N_\eta})}}{\|\eta\|_{l^0(\mathbb{R}^{N_\eta})}}$$

consistently yielded the best results. Here, $\|\tilde{\eta}\|_{l^0(\mathbb{R}^{N_\eta})}$ is the number of parameters that have an absolute value greater than some user-specified tolerance τ . In short, the regularization $\mathcal{R}^*(\eta)$ counts the proportion of components that exist outside of the threshold level τ . Thus, we always have $0 < \mathcal{R}^* < 1$ independent of the size N^η of the search space.

Especially for the larger-dimensional problems that we tackle later in this work, the optimization problem equipped with $\mathcal{R}^*(\eta)$ is not enough to reconstruct a sparse set of Lax operators. The steps we take to perform sparsification are precisely as follows:

- (1) Solve the optimization problem (2.8) with some fixed $r > 0$ and threshold level $\tau > 0$.
- (2) Solve the problem again over the subspace of nonzero parameters, using a line search method with the initial iterate set to the output of step (1), but with no further sparsification nor thresholding. That is, set $r = \tau = 0$.
- (3) Solve the problem yet again with $r = 0$ using the output of step (2) as the initial guess for a line search, but introduce a smaller threshold level to remove any small components of η that may still remain.

The rationale behind this approach is as follows. Having fixed the parameter r , step (1) executes the sparse search without any guesswork. We used a 24-core machine to execute the search on 24 thresholding values τ in parallel. In principle, this only costs us the real-time cost of one search. In step (2), we remove all bias that sparsification introduces into the problem by setting $r = \tau = 0$. However, when we remove the threshold, this introduces the likelihood that thresholded components may be resurrected. To avoid resurrection and to reintroduce smoothness into the problem, we remove thresholding and ensure that the search is executed only in the subspace of nonzero parameter values that survived from step (1). Now, since the search performed in step (2) may have some remaining components that are small but nonzero, we reintroduce a threshold of $\tau/10$ for each value of τ and execute a last parallel search maintaining $r = 0$ so that the sparsification bias is kept to a minimum. This strategy may not be the most elegant, but it is algorithmic and has proven successful in all the examples of this work. Future work with sparsification certainly involves attempts to find how to successfully execute a more standard l^1 penalization within our overall computational program.

Some further comments on sparsification are warranted. For all the numerical examples in this work, we fix $r = 1/2$. Varying this parameter, of course, interpolates between the two clear extremes of the optimization problem given by (2.8). We found, through numerical experimentation, that fixing r and varying τ are sufficient for our purposes. Additionally, the more commonly used alternatives to \mathcal{R}^* , such as the previously discussed \mathcal{R}_0 and \mathcal{R}_1 sparsification functions, could perhaps be made viable. One could also imagine that using \mathcal{R}_p , $0 < p < 1$, may also be deployable. When testing these different strategies, we found that for \mathcal{R}_0 to be reliable, we required different values of r to be used in the different examples in this paper, while \mathcal{R}_1 overrelaxed sparsification, and we could not discover the sparsest possible Lax pairs this way. Our sparsification strategy, as will be demonstrated throughout this work,

achieves the desired result of finding sparse Lax pairs as best we could hope. For brevity, we omit detailed comparisons among different \mathcal{R}_p , $0 < p < 1$, and let the numerical results speak for themselves.

Returning to the simple harmonic oscillator for a numerical demonstration, we set $k = 5$ and $m = 2$ —chosen arbitrarily for this proof of concept. The phase space is sampled using $N_{\text{samples}} = 100$ points from a uniform distribution over $[-1, 1] \times [-1, 1]$. To numerically solve the optimization problem (2.8), we use a nonconvex search method called the cross-entropy method [47], implemented in the CEopt MATLAB package [13]. We then feed the result of cross-entropy minimization into a BFGS method [42], as implemented by the `fminunc` function of MATLAB, which refines the result and quickly converges to the nearest local minimum. In this way, this hybrid optimization approach overcomes the nonconvexity of the optimization landscape while remaining computationally efficient. Without any sparsification, we reliably minimize the loss function to a value on the order of 10^{-15} .

For sparsification, we performed our parallel search over the 24 uniformly spaced values of $\tau \in [0.1, 2.4]$. Despite the bias in the training, we are still able to achieve a loss on the order of 10^{-15} (evaluated when $r = 0$) and correctly identify the exact structure of the known Lax pairs given by (2.3a) and (2.3b), that is, correctly identify the six parameters needed in the 24-parameter set η . The reason why six parameters should be identified, instead of the four as argued previously, is due to the free parameters a and b present in (2.3a) and (2.3b).

Precisely, the algorithm discovers

$$\tilde{L}_1 = \begin{pmatrix} \eta_1 q & \eta_4 p \\ \eta_6 p & \eta_7 q \end{pmatrix}, \quad \tilde{L}_2 = \begin{pmatrix} \eta_2 p & \eta_3 q \\ \eta_5 q & \eta_8 p \end{pmatrix}$$

and

$$\tilde{P} = \begin{pmatrix} 0 & \eta_{22} \\ \eta_{23} & 0 \end{pmatrix},$$

with all other 18 entries of η equal to zero. The recovered Lax pairs reproduce the equations of motion with seven-digit precision. At this stage, it is quite encouraging that our methodology correctly identifies *both* known Lax pairs. Figure 1 shows two examples of the values of η before and after thresholding. We report that our algorithm did not find any other family of Lax pairs for the simple harmonic oscillator. We therefore report with confidence that no other such family of Lax pairs exists within the operator hypothesis used for this study.

Thus, integrability detection and the recovery of a computationally meaningful Lax pair for the simple harmonic oscillator is achieved. But, of course, one-degree-of-freedom autonomous Hamiltonian systems are always integrable because the Hamiltonian itself is conserved. Moreover, a Lax pair can always be identified with the dynamics since for any Hamiltonian of the form $H = \frac{1}{2}p^2 + V(q)$, with $V(q) \in C^1(\mathbb{R})$, the compatibility of the operators

$$L = \begin{pmatrix} p & \sqrt{2V(q)} \\ \sqrt{2V(q)} & -p \end{pmatrix},$$

$$P = \begin{pmatrix} 0 & \partial_q \sqrt{2V(q)} \\ -\partial_q \sqrt{2V(q)} & 0 \end{pmatrix}$$

gives rise to Hamilton's equations of motion $\dot{q} = \frac{\partial H}{\partial p}$, $\dot{p} = -\frac{\partial H}{\partial q}$. For this reason, we move on to a two-degree-of-freedom system, carrying forward the lessons learned from this test case.

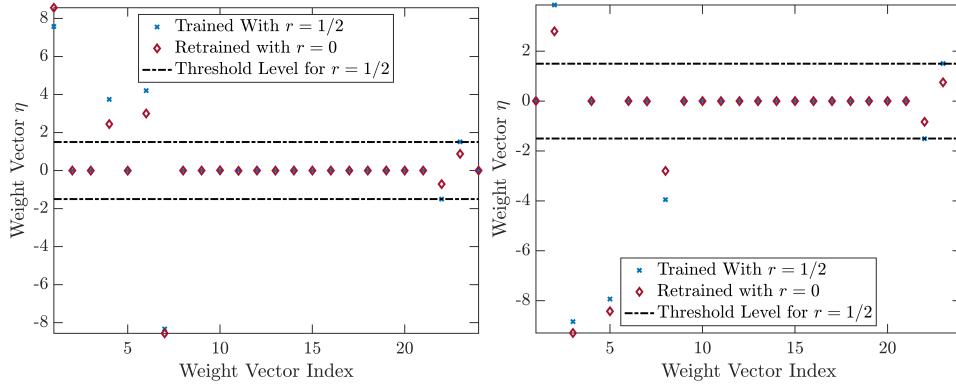


Figure 1. Two numerical results from studying problem (2.8) for the simple harmonic oscillator (2.1). We show the result of optimizing before and after thresholding. SILO correctly identifies that there should only be six parameters in the 24-parameter operator hypothesis. SILO also finds both families of Lax pairs as given by (2.3a) and (2.3b).

3. A two-degree-of-freedom scenario: Hénon–Heiles. Now, with a viable approach in hand, we go on to study a slightly more complicated example: the famous Hénon–Heiles (HH) system. This is a system with two degrees of freedom defined by the Hamiltonian

$$(3.1) \quad H = \frac{1}{2} (p_x^2 + p_y^2) + \frac{1}{2} (Ax^2 + By^2) + x^2y + \varepsilon y^3,$$

where the parameters A , B , and ε are arbitrary [27, 19]. This system is known to be integrable for three different cases of these parameters [46].

We will focus on one such case: $A = B$ and $\varepsilon = 1/3$. In this case, a Lax pair for the HH system can be written as [46]

$$(3.2) \quad L = \begin{pmatrix} L_1 & 0 \\ 0 & L_2 \end{pmatrix}, \quad P = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix},$$

where

$$\begin{aligned} L_1 &= \begin{pmatrix} \frac{p_y - p_x}{12} & \frac{y - x - k}{p_x - p_y} \\ \frac{2k^2 + k[3A + 2(y - x)] + 3A(y - x) + 2(y - x)^2}{12} & \end{pmatrix}, \\ P_1 &= \begin{pmatrix} 0 & 1 \\ -\frac{y - x}{3} - \frac{3A + 2k}{12} & 0 \end{pmatrix}, \\ L_2 &= \begin{pmatrix} \frac{p_y + p_x}{12} & \frac{y + x - k}{-p_x - p_y} \\ \frac{2k^2 + k[3A + 2(y + x)] + 3A(y + x) + 2(y + x)^2}{12} & \end{pmatrix}, \\ P_2 &= \begin{pmatrix} 0 & 1 \\ -\frac{y + x}{3} - \frac{3A + 2k}{12} & 0 \end{pmatrix}, \end{aligned}$$

and k is a free parameter.

We are interested in testing our methodology as a proof of concept. For this reason, we use a mix of basic assumptions about the structure of the Lax pair along with some sparsification. Specifically, we exploit the block-diagonal structure of the Lax pairs and inform our search so that we only seek to achieve a small loss over one of the blocks. Our hypothesis for the Lax operators is

$$(3.3) \quad \tilde{L} = \begin{pmatrix} \xi_1 p_x + \xi_2 p_y & \xi_8 x + \xi_9 y + \xi_{10} xy + \xi_{11} x^2 + \xi_{12} y^2 \\ \xi_3 x + \xi_4 y + \xi_5 xy + \xi_6 x^2 + \xi_7 y^2 & -\xi_1 p_x - \xi_2 p_y \end{pmatrix}$$

and

$$(3.4) \quad \tilde{P} = \begin{pmatrix} 0 & \zeta_4 + \zeta_5 x + \zeta_6 y \\ \zeta_1 + \zeta_2 x + \zeta_3 y & 0 \end{pmatrix}.$$

Note that the most general quadratic hypothesis in all variables and without assuming a block diagonal structure would result in an optimization dimension of $N_\eta = 288$. Although this remains computationally feasible, such a study would obfuscate the essence of how our methodology performs. Therefore, we instead study this much smaller problem of dimension $N_\eta = 18$, leaving higher-dimensional studies for future work.

Only two tasks remain. First, we need the Poisson bracket $\{L, H\}$ again:

$$(3.5) \quad \begin{aligned} \frac{dL}{dt} &= \frac{\partial L}{\partial x} \dot{x} + \frac{\partial L}{\partial y} \dot{y} + \frac{\partial L}{\partial p_x} \dot{p}_x + \frac{\partial L}{\partial p_y} \dot{p}_y \\ &= \frac{\partial L}{\partial x} \frac{\partial H}{\partial p_x} + \frac{\partial L}{\partial y} \frac{\partial H}{\partial p_y} - \frac{\partial L}{\partial p_x} \frac{\partial H}{\partial x} - \frac{\partial L}{\partial p_y} \frac{\partial H}{\partial y} \\ &= \{L, H\}, \end{aligned}$$

which is straightforward to evaluate by hand. The other is to generalize the sampling of the phase space to four dimensions, that is, $(x_k, p_{x_k}, y_k, p_{y_k}) \sim U([-1, 1]^4)$.

We train again on $N_{\text{samples}} = 100$ samples from the phase space. As illustrated in Figure 2, the optimization correctly identifies that only 13 parameters are necessary in the Lax pair. The thresholds, difficult to distinguish from zero in the scaling of the figure, were found to be

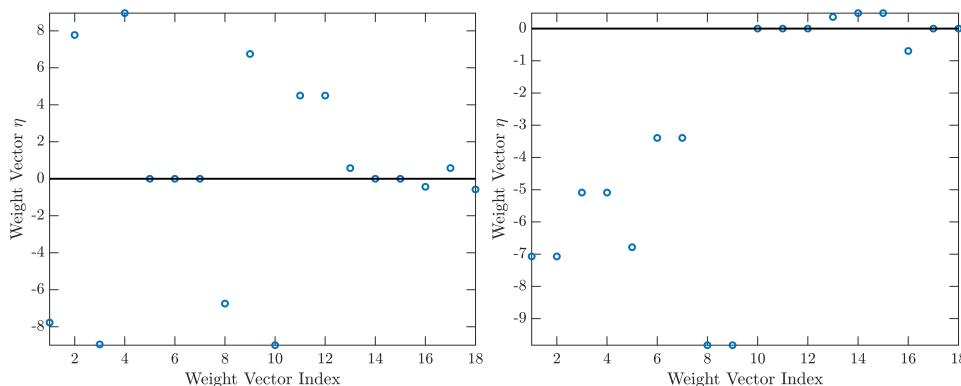


Figure 2. The correct identification, to six digits of precision, of a Lax pair for the integrable case of the HH system given by Hamiltonian (3.1). The panel on the left reproduces the expected structure of the known Lax pair. The panel on the right discovers that, up to a sign, the transpose of the known Lax pair is also valid for reproducing the HH system. A small thresholding of 10^{-4} was used but not visualized here.

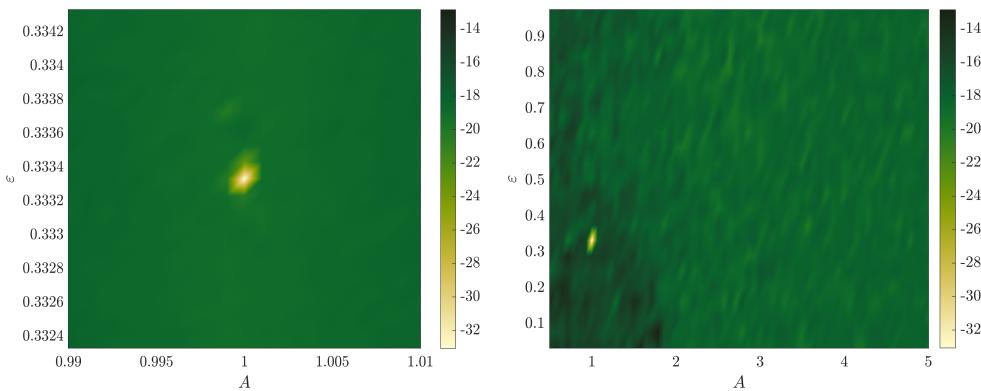


Figure 3. A broad parameter search (with B fixed to 1) for integrability detection in the HH system given by Hamiltonian (3.1). The optimization of the loss function, shown on a logarithmic scale, identifies a distinct position at $(A, \varepsilon) = (1, 1/3)$ where integrability is meaningfully detected, differing by several orders of magnitude from background loss values. Cubic spline interpolation of the landscape is used for visualization.

on the order of 10^{-3} after some experimentation. We believe that such small threshold levels were needed because so much sparsification had been performed manually through the choice of our operator hypothesis.

We report that the discovered Lax pairs reproduce the dynamical system to six digits of precision with a loss on the order of 10^{-15} . In fact, we recognize that one Lax pair has not been previously reported. Some simple algebra indeed shows that up to a sign, the transpose of the known Lax pair also reproduces the Henon–Heiles system. Although this finding is only mildly interesting, it once again provides further confidence that SILO, with high likelihood, indeed identifies all possible sparse options for compatible Lax pairs.

Our best results through sparsification find that the loss is a number on the order of 10^{-15} in the integrable case of $A = B = 1$ and $\varepsilon = 1/3$. We now ask whether SILO is sensitive enough to detect when a system is nonintegrable. Our findings suggest that it is. We show the numerical results of the search for Lax pairs, without sparsification, as A and ε are varied, with $B = 1$ fixed, in Figure 3. These results show a significant increase in the computed loss—by several orders of magnitude—across the parameter space. Thus, restricted to the operator hypothesis used, we can determine where the integrability lies in the parameter space (A, ε) for fixed B . We note that sparsification is not used for this type of study because the focus is not on the interpretation of the discovered Lax pair but instead on whether compatible Lax pairs could be identified at all.

4. The Korteweg–de Vries equation. We now extend our methodology to Hamiltonian PDEs beginning with the Korteweg–de Vries (KdV) equation

$$(4.1) \quad \partial_t u - 6u\partial_x u + \partial_x^3 u = 0.$$

We first review the Hamiltonian form of the KdV equation [51] which is given by

$$u_t = Q \frac{\delta H}{\delta u},$$

where

$$(4.2) \quad H = \int_{-\infty}^{\infty} \left(u^3 - \frac{1}{2} u u_{xx} \right) dx := \int_{-\infty}^{\infty} h(u, u_{xx}) dx.$$

The KdV equation is bi-Hamiltonian; two choices of Q (with corresponding choices of H) yield the equation of motion. For the choice of H used here, the relevant operator is $Q = \partial_x$. Specifically, since the Fréchet derivative of the functional H is given by

$$\frac{\delta H}{\delta h} = \frac{\partial h}{\partial u} - \partial_x \frac{\partial h}{\partial u_x} + \partial_x^2 \frac{\partial h}{\partial u_{xx}},$$

we see that

$$\partial_x \frac{\delta H}{\delta u} = 6u \partial_x u - \partial_x^3 u = \partial_t u$$

reproduces the KdV equation (4.1).

The KdV equation on the real line constitutes a prototypical example of a Hamiltonian PDE that is integrable. A Lax pair is known [20, 2]:

$$(4.3) \quad \begin{aligned} L &= -\partial_x^2 + u, \\ P &= 4\partial_x^3 - 6u\partial_x - 3u_x. \end{aligned}$$

The KdV equation can thus be viewed as the compatibility between these differential operators; that is, the equation

$$\partial_t L = [L, P]$$

reproduces (4.1).

To yet again build a loss function that does not involve the explicit use of time derivatives, we use the chain rule. Observe that

$$\begin{aligned} \partial_t L &= \frac{\partial L}{\partial u} \frac{\partial u}{\partial t} \\ &= \frac{\partial L}{\partial u} Q \frac{\delta H}{\delta u} = [L, P]. \end{aligned}$$

Thus, the expression $\frac{\partial \tilde{L}}{\partial u} Q \frac{\delta H}{\delta u}$ plays the role of the Poisson bracket in the design of the loss functions from previous sections. By direct analogy with problem (2.8), our optimization problem for this Hamiltonian setting is given by

$$(4.4) \quad \min_{\eta \in \mathbb{R}^{N_\eta}} J[\eta] = \min_{\eta \in \mathbb{R}^{N_\eta}} \frac{1-r}{N_{\text{samples}}} \sum_{j=1}^{N_{\text{samples}}} \frac{\int \left| \frac{\partial \tilde{L}}{\partial u} Q \frac{\delta H}{\delta u} u_j - [\tilde{L}, \tilde{P}] u_j \right|^2 dx}{\int \left| \frac{\partial \tilde{L}}{\partial u} Q \frac{\delta H}{\delta u} u_j \right|^2 dx} \Bigg|_{u_j \in \Omega} + r \mathcal{R}^*(\eta),$$

where Ω is a subset of the KdV phase space and \mathcal{R}^* is the sparsification function given by (2.9) earlier. Notice that by the large vertical bar on the right of the ratio on the first term of the right-hand side, we mean that the expression is evaluated on the functions u_j .

In this PDE setting, there are only two additional modifications we need to make to the numerical framework from previous sections. First, we must carefully consider how we sample the phase space and how we construct the operator hypotheses. This amounts to sampling from the function space $L^p(\mathbb{R})$, $p > 1$. To this end, we construct random samples from the overcomplete basis

$$(4.5) \quad u^{\text{rand}}(x) = N \sum_j e^{-a_j(x-b_j)^2} \sum_k \frac{A_{jk}}{k^3} \sin \frac{k\pi x}{l},$$

where all parameters a_j , b_j , A_{jk} are appropriately sampled from uniform distributions, l is the length of the truncated spatial domain, and the coefficient N ensures a unit norm in the space $L^1(\mathbb{R})$. In this way, we try to verify the compatibility of our operators with functions sampled from the function space $C^3(\mathbb{R}) \cap L^p(\mathbb{R})$, $p > 1$, and with equivalent masses in the sense of $L^1(\mathbb{R})$. The smoothness of our samples is ensured by the decay of the Fourier coefficients $A_{j,k}/k^3$ while also being regularized by the fact that we only take the sum over k to be finite. We used $k = 10$ throughout our study. This smoothness is used for computational tractability in the evaluation of differential operators in the loss. We will comment on how this choice of sampling, as opposed to sampling from any other function space, affects numerical results shortly.

The second modification is the operator hypothesis. In general, Lax pairs are linear differential operators (in x) with coefficients dependent on x , u , and derivatives of u . A fairly wide class of operators is given by

$$\begin{aligned} \tilde{L} &= \xi_1 u + \xi_2 \partial_x + \xi_3 \partial_x^2, \\ \tilde{P} &= \sum_l \sum_j \sum_k \sum_m \zeta_{l j k m} x^{l-1} u^{j-1} \left(\partial_x^{k-1} u \right) \partial_x^{m-1}, \end{aligned}$$

where the indices in the sum all start at one. It may seem like the hypothesis on \tilde{L} is overly restrictive. However, we must keep in mind that we seek to reproduce an evolution equation involving $\partial_t \tilde{L}$. Should any higher powers of u enter into the hypothesis of \tilde{L} , then the likelihood of finding an *explicit* equation of the form $u_t = F(u, u_x, u_{xx}, \dots)$ decreases substantially. In this sense, we seek to preserve the semilinear (in time) nature of the PDE of interest.

We also comment that the dependence on polynomial coefficients x is redundant when it comes to Lax pairs for the KdV equation. Therefore, to keep our methodology simple, we drop the dependence on the monomial x^l and its associated index on the tensor coefficients ξ and ζ . We also assume that P has at most third-order coefficients. For clarity on these choices, reproducing the KdV Lax pair amounts to setting $\xi_1 = 1$, $\xi_3 = -1$, $\zeta_{4,1,1} = 4$, $\zeta_{2,2,1} = -6$, $\zeta_{1,1,2} = -3$ with the remaining 34 coefficients set to zero. Last, to compute the spatial derivatives in the operator hypothesis, we use the fast Fourier transform. That is, we use the formula

$$\partial_x^j u = \mathcal{F}^{-1} \{ (ik)^j \mathcal{F}\{u\} \}$$

to compute the derivatives spectrally, where k denotes the grid-dependent wavenumbers.

Since we are computing derivatives assuming the use of periodic boundary conditions, we sample from Ω , ensuring that all samples have compact support to double precision in the

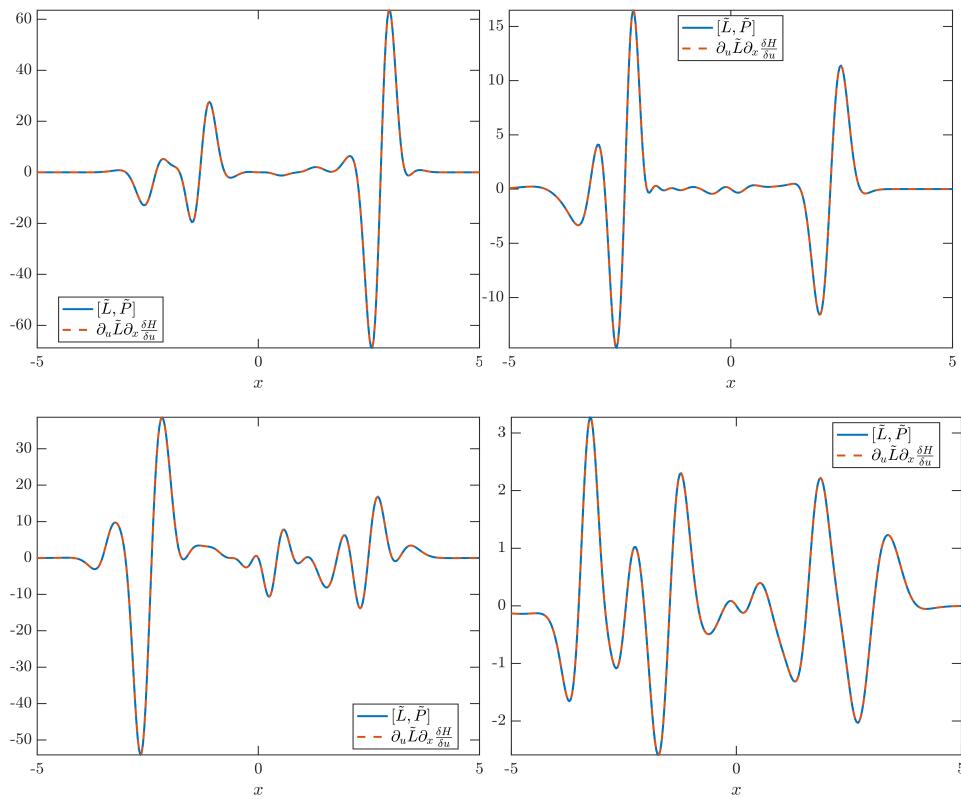


Figure 4. A numerical result of solving Problem (4.4) without sparsification. Visualized here is a cross-validation study displaying the generalized Poisson brackets and commutators evaluated at the optimal point η^* and on four samples from the function space Ω that were unseen during training. For all four cases, the loss is on the order of 10^{-11} .

interval $[-20\pi, 20\pi]$. Despite using just 2^{11} points, we verify that the loss in problem (4.4), evaluated at $r = 0$ and at the values of η that give the known KdV Lax pair, is zero to double precision over 100 such samples from Ω .

In our training, we find that only using 20 samples is sufficient in our cross-validation. After training on these samples, we find that the loss, on average and without sparsification, is on the order of 10^{-11} when evaluated on unseen samples from Ω . We show, in Figure 4, a visual comparison between the action of $\frac{\partial \tilde{L}}{\partial u} Q \frac{\delta H}{\delta u}$ and $[\tilde{L}, \tilde{P}]$ on four unseen sample functions. Note that despite the unit norms of the samples in $L^1(\mathbb{R})$, the action of the generalized Poisson bracket and matrix commutators yield arbitrary norms, which account for the variance of the scales across these images.

The following remark is critical both mathematically and computationally. The careful reader will notice that we have not made a distinction between the functions $u(x)$ used in our operator hypotheses and the function $u(x)$ on which the operators act. That is because, computationally, we do not make such a distinction. This is a *mathematically restricted* interpretation of how Lax pairs work since, in principle, the relevant action of the operators in the Lax equation is applicable to *arbitrary* functions; that is, this is valid as an operator equation.

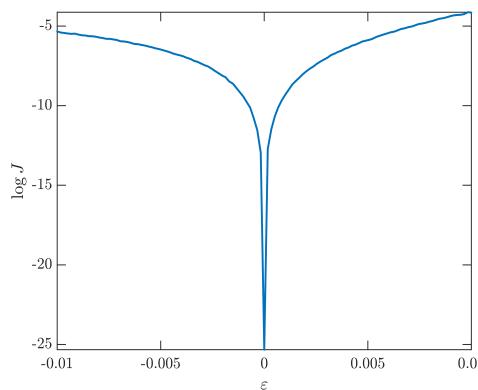


Figure 5. A perturbation study using the perturbed Hamiltonian density $h_1 = \frac{\varepsilon_1}{2}(\partial_x^2 u)^2$. Shown here is the numerical solution of problem (4.4), without sparsification, with density $h + \varepsilon h_1$. We observe a near-smooth dependence on ε with a clearly discernible “special” point associated with the detection of integrability at $\varepsilon = 0$.

Nevertheless, despite its limitations, we will show that this *computational interpretation* of the problem still achieves many of the desired objectives of detecting integrability and discovering sparse Lax pairs.

A natural question arises: Why not adopt the interpretation

$$(\partial_t L(u) - [L(u), P(u)]) w = 0 \quad \forall w \in \mathcal{F},$$

where \mathcal{F} is the relevant function space? We did pursue this. We used the simple scenario of fixing u as a single function, specifically $u(x) = e^{-x^2}$. Then we randomly sample 20 functions w using our sampling scheme. We report that our optimization methodology employed up until now absolutely fails to find any Lax pair that minimizes the unbiased loss past values smaller than 10^{-4} . This, as we found, turns out not to be an acceptable order of magnitude to resolve the detection of integrability against smooth nonintegrable perturbations; see Figure 5 to observe the degree of precision required to detect integrability. Of course, this is a point that merits further investigation both mathematically and through computational experiments, especially since this order of magnitude for the loss is often deemed acceptable when using neural networks; see, for example, [38, Table II] or [54].

It is rather remarkable that this mathematically restricted interpretation of the optimization problem still yields numerical outputs that are mathematically correct when interpreted. We leveraged substantial computational resources to solve the unrestricted mathematical interpretation of the optimization problem, to no avail (at least for the prescribed numerical precision). For this reason, we believe that our interpretation, relative to our current computational framework, is crucial to ensuring the desired degree of tractability for this problem.

To demonstrate the validity of our approach, we show that our framework is sensitive enough to detect the known integrability of the KdV equation. Consider the nonintegrable perturbation $h_1 = \frac{1}{2}(\partial_x^2 u)^2$. We solve problem (4.4), again without sparsification, for Hamiltonian densities $h + \varepsilon h_1$, where h is defined in (4.2) and $\varepsilon \in [-.01, .01]$. Figure 5 shows that the loss has a nearly smooth dependence on the parameter ε with a minimum at the expected integrable point $\varepsilon = 0$. Again, the integrability point is privileged, as the loss is several orders of magnitude smaller than the nearby values of ε .

We now discuss the interpretability of building Lax pairs from solving problem (4.4). Without sparsification, it is not surprising that all 39 coefficients in our operators are activated. Therefore, even with computer algebra systems such as *Mathematica*, we have no chance to interpret the PDE that the Lax pair is producing. Before discussing what we discover through sparsification, we make the following basic observations.

Suppose that we were fortunate enough to choose the perfect operator hypothesis

$$\begin{aligned} L &= a\partial_x^2 + bu, \\ P &= c\partial_x^3 + du\partial_x + e\partial_x u. \end{aligned}$$

We can determine these coefficients by hand and show that they match the operators in (4.3), up to a constant factor in the operator L . This implies that if our numerical method successfully identifies this Lax pair, it should do so uniquely, modulo a scaling factor. Because of our restricted approach of taking the coefficients in the Lax pair to be the same functions on which we evaluate, this introduces some indeterminacy. A calculation by hand shows that the compatibility of these operators, with the interpretation $\partial_t L(u)u = [L(u), P(u)]u$, yields

$$\begin{aligned} ae - cb + b &= 0, \\ -bd - 6b &= 0, \\ 3ad + 2ae - 6cb &= 0. \end{aligned}$$

This system is, of course, underdetermined, featuring 3 equations for 5 unknowns. If we solve problem (4.4) with this perfect hypothesis, we find that the combination of parameters identified is such that these equations are satisfied to $\mathcal{O}(10^{-6})$. Meanwhile, the loss is $\mathcal{O}(10^{-12})$.

We also make the observation that, in principle, we have the freedom to rescale by space and amplitude so that we can remove two parameters, that is,

$$\begin{aligned} L &= \partial_x^2 - u, \\ P &= c\partial_x^3 + du\partial_x + eu_x, \end{aligned}$$

with the three remaining parameters having been appropriately rescaled. Upon solving the optimization problem, we practically recover the known coefficients. We find that $c = 4$, $d = 6$, and $e = -3$ to six digits of precision with a loss that evaluates to $\mathcal{O}(10^{-13})$. Thus, with all of this fortuitous hindsight, we can rediscover the expected KdV Lax pair with a small loss and with no ambiguity in the coefficients. These observations give us a target and a level of precision to aim for as we employ our sparsification strategy.

Once again, we use the same sparsification from the finite-degree-of-freedom setting to aid us in our interpretation of our discovered Lax pairs. Interestingly, our numerics have a lot of difficulty rediscovering the known KdV Lax pair just discussed and commonly reported in the literature because, in fact, *this is not the most parsimonious Lax pair that exists*. Our numerics discover, to the best of our knowledge, two entirely new families of Lax pairs, each containing more parsimonious four-term Lax pairs—instead of the five-term commonly reported one in (4.3)—as special cases. Before discussing these more parsimonious Lax pairs, we state the discovered Lax pair in its full generality.

Theorem 4.1 (a new KdV Lax pair). *For every $u \in C^1([0, T]; C^3(\mathbb{R}))$, there exists a parameter $v \in \mathbb{R}$ such that the pair of operators*

$$\begin{aligned} L &= \alpha u + \beta \partial_x, \\ P &= \gamma u + \delta u^2 + \epsilon u_{xx} + \kappa \partial_x, \end{aligned}$$

satisfying Lax's equation $\partial_t L = [L, P]$, understood as acting on the function space $C^3(\mathbb{R})$, reproduces the KdV equation

$$u_t = \frac{2\beta\delta}{\alpha} u u_x + \frac{\beta\epsilon}{\alpha} u_{xxx}$$

in the co-traveling reference frame $x \rightarrow x - vt$.

Proof. By direct calculation, observe that for every $w \in C^3(\mathbb{R})$, we have

$$[L, P]w = LPw - PLw = ((\beta\gamma - \alpha\kappa)u_x + 2\beta\delta u u_x + \beta\epsilon u_{xxx})w.$$

Since $\partial_t L = \alpha u_t$, we see that the operator equation $\partial_t L = [L, P]$ reproduces the desired KdV equation after the appropriate Galilean boost of velocity $v = \frac{\beta\gamma - \alpha\kappa}{\alpha}$. ■

Before moving forward with interpreting this result, we comment that we are aware that the function spaces for which we state our results are classical. In fact, it is well known that the Cauchy problem for the KdV equation is well-posed for initial conditions in the Sobolev space $H^s(\mathbb{R})$, $s > 3/4$ [30] and even more recently it was discovered that, remarkably, the problem is well-posed in $H^{-1}(\mathbb{R})$ [32]. For these reasons, it is obvious that we could employ standard embedding techniques [18] to relax the regularity assumptions that we make. We find this unnecessary at this stage, as this would detract from the essence of this finding. Therefore, we proceed with classical regularity assumptions that are compatible with the discovered Lax pairs.

In light of Theorem 4.1, the most parsimonious Lax pair has four, not five, terms. Specifically, this is the valid choice of $\gamma = \kappa = 0$. In this degenerate case, the operator P is not a differential operator, as it is purely multiplicative. This degenerate example is consistently found by SILO. We delay a discussion of the modification made to find the Lax pair of Theorem 4.1 in full generality until the end of this section. This is because the second type of Lax pair that SILO consistently finds is one that only makes sense under an integral with functions that vanish on the boundary, which is what is used in the evaluation of the loss function. The following theorem expresses our discovery of these *weak* Lax pairs precisely.

Theorem 4.2 (existence of a weak Lax pair). *Denote the Schwartz space of distributions by $\mathcal{S}(\mathbb{R})$. Then, for every $u \in H^2([0, T]; H^4(\mathbb{R}))$, the Lax pair*

$$\begin{aligned} L &= \alpha u + \beta \partial_x^2, \\ P &= \delta u_x + \varepsilon u \partial_x, \end{aligned}$$

satisfying

$$\int_{\mathbb{R}} (\partial_t L - [L, P]) \varphi(x) dx = 0 \quad \forall \varphi \in \mathcal{S}(\mathbb{R}),$$

reproduces the equation

$$(4.6) \quad \alpha u_t + \beta(\varepsilon - \delta)u_{xxx} + \alpha\varepsilon uu_x = 0$$

Lebesgue almost everywhere.

Proof. By direct calculation, we compute the left-hand side of the equation $\langle (\partial_t L - [L, P])\varphi, 1 \rangle_{L^2(\mathbb{R})} = 0$, in the sense of distributions, and find that

$$(4.7) \quad \langle \alpha u_t - \beta\delta u_{xxx} + \alpha\epsilon uu_x, \varphi \rangle_{L^2(\mathbb{R})} - \beta(2\delta + \varepsilon)\langle u_{xx}, \varphi_x \rangle_{L^2(\mathbb{R})} - 2\beta\epsilon\langle u_x, \varphi_{xx} \rangle_{L^2(\mathbb{R})} = 0.$$

Integrating by parts once in the second inner product and integrating by parts twice in the third reproduces the weak form of (4.6). Inferring the strong form of the KdV equation follows directly from the continuous embeddings of $H^2([0, T]) \hookrightarrow C^1([0, T])$ and $H^4(\mathbb{R}) \hookrightarrow C^3(\mathbb{R})$ [18]. \blacksquare

We note that despite how the Lax pair just discussed in Theorem 4.2 reproduces the KdV equation in the sense of distributions, one can consider altering the function space so that the Lax pair is a strong one. Suppose that instead of the wide function space \mathcal{S} , we operate on functions $\varphi \in C^3(\mathbb{R})$ such that $L\varphi = \lambda\varphi$ for all $u \in C^1([0, T]; C^3(\mathbb{R}))$. Additionally, if $2\delta + \varepsilon = 0$, then the strong Lax equation $\partial_t L = [L, P]$, with this Lax pair being the one stated in Theorem 4.2, reproduces the KdV equation

$$\alpha u_t - \beta\delta u_{xxx} + 3\alpha\varepsilon uu_x - 2\varepsilon\lambda u_x = 0.$$

As was done in the proof of Theorem 4.1, we may once again exploit the Galilean invariance of the KdV equation to frame boost away the parameter λ through $x \rightarrow x + \frac{2\lambda\varepsilon}{\alpha}t$. Thus, we strengthen the sense in which the Lax pairs of Theorem 4.2 produce an unambiguous KdV equation without the need to integrate by parts. Note, however, that this sense of Lax pair is not fully mathematically general, as it operates only on a subspace of $C^3(\mathbb{R})$ and not the whole space of functions for which classical Lax pairs are typically thought to operate on. Therefore, we do not claim that this Lax pair with this interpretation has the same mathematical relevance as the typical Lax pair given by (4.3) and leave this as a curious observation in passing.

Shifting back to our interpretation of the numerics, we see that SILO indeed serves as a computationally aided proof technique in discovering these new Lax pairs since our conclusions are indeed mathematical and only the means of discovery were computational. Although not entirely insightful, we still report the coefficients we discovered during optimization in Figure 6. There, we show coefficients that achieved unbiased losses of $\mathcal{O}(10^{-15})$. The strong Lax pair reproduces the KdV equation consistent with the Hamiltonian (4.2) to eight digits of precision. The weak Lax pair, however, does not produce the KdV equation expected by Theorem 4.2. This is because the loss function (4.4) involves squaring inside the integral, and integration by parts cannot be carried out as cleanly as was done in the proof of Theorem 4.2. A tedious calculation reveals that the numerical compatibility of the coefficients is captured by the equations $\varepsilon = -6$ and $5\alpha + 3\beta(\delta - \varepsilon) = 0$, where these symbols are consistent with Theorem 4.2. Indeed, our numerically discovered coefficients satisfy these equations to eight digits of precision.

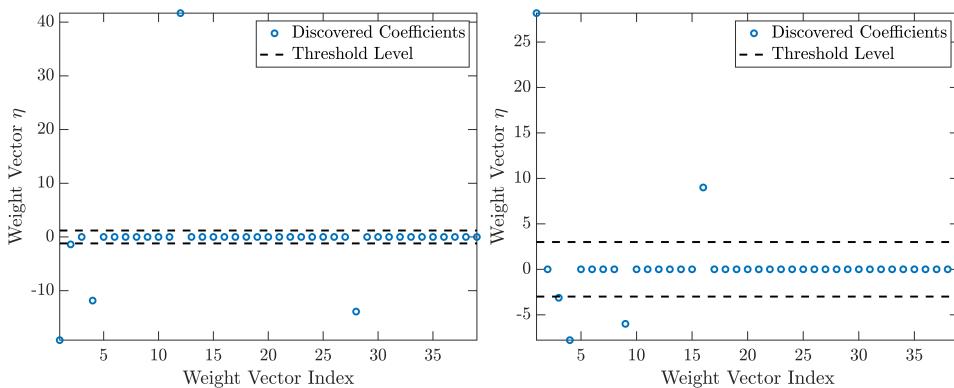


Figure 6. Coefficients discovered during the sparse KdV Lax pair regression as given by problem (4.4). Coefficients shown in the left panel correspond to the degenerate ($\gamma = \kappa = 0$) family of Lax pairs discussed in Theorem 4.1, while coefficients in the right panel correspond to the family of weak Lax pairs discussed in Theorem 4.2. As discussed in the text, only four coefficients matter since η_4 in both panels does not contribute in either case to the equation compatible with the computed Lax pairs. In both cases, the loss with $r = 0$ evaluates to $\mathcal{O}(10^{-15})$.

Our findings, shown in Figure 6, show that our numerics discover five terms. However, each of our new KdV Lax pairs is a family with only four terms. This is because the coefficients η_4 shown in both panels correspond to the constant operator appearing in \tilde{P} , that is, where all the indices produce $\zeta_{1,1,1}$. The calculations in the proofs of Theorems 4.1 and 4.2 with the additional constant γ added to the operators \tilde{P} would show that the reproduced KdV equations have no terms involving γ . Therefore, the numerically discovered Lax pairs actually only have four meaningful coefficients.

It is interesting to note that our numerics did not discover the known Lax pair. This is perhaps best explained by the fact that by searching for the sparsest possible Lax pairs, we missed the classical Lax pair with five terms. Indeed, the unbiased loss for the four-term Lax pair is at least two orders of magnitude smaller than the loss when we searched for the best Lax pairs assuming the five-term hypothesis without sparsification.

For this reason, we searched once again, sampling from the space of periodic functions on $x \in (0, 2\pi)$ that do not necessarily vanish at the boundary. We did this to attempt to circumvent the repeated discovery of weak and degenerate Lax pairs. Only then did we discover the fully general six-term Lax pair of Theorem 4.1. Furthermore, we were able to recover the well-known Lax pair when the search was restricted to operators of the form $\tilde{L} = \alpha u + \beta \partial_x^2$. Moreover, there exist alternative formulations of the Lax pair compatibility, such as, e.g., that of [11].

Therefore, increasing SILO's thoroughness (or potentially modifying its setup) poses some intriguing challenges for further efforts in this direction. Indeed, it does not escape us that there still remain several options for sampling function spaces and evaluating operators, and it is unclear how these may affect discoveries of Lax pairs. We simply report here how our choices explored thus far enable us, as computational users, to discover suitable Lax pairs for the problem of interest. We leave discussions about the mathematical implications of widening the space of test functions to accommodate the existence of weak Lax pairs to section 6.

5. The cubic nonlinear Schrödinger equation. Our last example involves the focusing nonlinear Schrödinger (NLS) equation

$$(5.1) \quad i\partial_t\psi = -\partial_x^2\psi + \frac{2}{p^2-1}|\psi|^2\psi,$$

where $0 < p < 1$ is a free parameter and $x \in \mathbb{R}$. In principle, the parameter p can be absorbed by rescaling the equation's variables, yet we maintain its presence so that the discussed Lax pairs are historically aligned with the literature [52]. We make the arbitrary choice of $p = 1/\sqrt{2}$ when executing computations.

The NLS is a Hamiltonian system in the sense that $i\partial_t\psi = \delta H/\delta\psi^*$ with Hamiltonian

$$(5.2) \quad H = \int_{-\infty}^{\infty} \left(|\partial_x\psi|^2 + \frac{1}{1-p^2}|\psi|^4 \right) dx := \int_{-\infty}^{\infty} h(\psi, \psi^*, \partial_x\psi, \partial_x\psi^*) dx.$$

In this way, ψ^* plays the role of the conjugate variable. Using the Hamiltonian, we build the Poisson bracket once more:

$$\begin{aligned} \frac{\partial L}{\partial t} &= \frac{\partial L}{\partial \psi} \frac{\partial \psi}{\partial t} + \frac{\partial L}{\partial \psi^*} \frac{\partial \psi^*}{\partial t} \\ &= \frac{\partial L}{\partial \psi} \left(-i \frac{\delta H}{\delta \psi^*} \right) + \frac{\partial L}{\partial \psi^*} \left(i \frac{\delta H}{\delta \psi} \right) \\ &= i \left(\frac{\partial L}{\partial \psi^*} \frac{\delta H}{\delta \psi} - \frac{\partial L}{\partial \psi} \frac{\delta H}{\delta \psi^*} \right) \\ &= i\{L, H\} = i[L, P], \end{aligned}$$

where the Fréchet derivatives are given by

$$\begin{aligned} \frac{\delta H}{\delta \psi} &= -\partial_x^2\psi^* - \frac{2}{1-p^2}|\psi|^2\psi^*, \\ \frac{\delta H}{\delta \psi^*} &= -\partial_x^2\psi + \frac{2}{1-p^2}|\psi|^2\psi. \end{aligned}$$

To build the operator hypothesis, motivated in part by the well-known Lax pair of the NLS due to Zakharov and Shabat [52], we assume that the operators in L are at most linear (in the field ψ) and the operators in P are at most quadratic with constant and complex matrix coefficients in 2×2 . That is, the hypothesis is

$$\tilde{L} = \sigma_1\partial_x + \sigma_2\psi + \sigma_3\psi^*,$$

$$\tilde{P} = \sigma_4\partial_x^2 + \sigma_5\partial_x + \sigma_6\partial_x\psi + \sigma_7\partial_x\psi^* + \sigma_8|\psi|^2 + \sigma_9\psi^2 + \sigma_{10}\psi^{*2} + \sigma_{11}\psi + \sigma_{12}\psi^*.$$

This amounts to the discovery of 12 matrices $\sigma_j \in \mathbb{C}^{2 \times 2}$, or 96 real parameters, in the optimization problem

$$(5.3) \quad \min_{\eta \in \mathbb{R}^{N\eta}} J[\eta] = \min_{\eta \in \mathbb{R}^{N\eta}} (1-r) \left[\frac{\int_{\mathbb{R}} |\{\tilde{L}, H\} - [\tilde{L}, \tilde{P}]\|^2 dx}{\int_{\mathbb{R}} |\{\tilde{L}, H\}|^2 dx} \right]_{\Omega} + r\mathcal{R}^*(\eta),$$

where the evaluation over Ω is understood in the same sense of (4.4) but over two component functions drawn from the NLS phase space. To sample from this phase space, we similarly use the overcomplete basis given by (4.5), replacing the sine function with a complex exponential. We make the computational choice to compute the evaluation of the operators in the sense that they act on the column vector $w(x) = [u(x) \ u(x)^*]^\top$. We, once again, emphasize that this is not the most general mathematical choice but has been empirically found to suffice for our purposes.

We note that, due to Zakharov and Shabat, the Lax pair for the NLS equation is available in the form [52]

$$(5.4) \quad \begin{aligned} L &= i \begin{pmatrix} 1+p & 0 \\ 0 & 1-p \end{pmatrix} \partial_x + \begin{pmatrix} 0 & \psi^* \\ \psi & 0 \end{pmatrix}, \\ P &= -p \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \partial_x^2 + \begin{pmatrix} |\psi|^2 & i\partial_x\psi^* \\ \frac{1}{1+p} & -i\partial_x\psi \end{pmatrix}. \end{aligned}$$

Therefore, we know that our operator hypothesis is wide enough to enclose the Zakharov–Shabat Lax pair. However, despite our best efforts, we were only able to achieve a loss on the order of 10^{-7} . We believe that with 96 parameters, the loss landscape of this optimization problem becomes highly nonconvex, with numerous local minima preventing convergence to the global solution that coincides with the Zakharov–Shabat solution. To achieve a higher precision detection of integrability, we discard the matrices $\sigma_9 - \sigma_{12}$ because, as can be seen, these matrices are not relevant to the resulting equation of motion. Therefore, the optimization problem has a more manageable 64 parameters to find. We defer the exploration of more sophisticated numerical optimization tools for higher-dimensional computations to future studies.

With 64 parameters, we are able to minimize the loss to $\mathcal{O}(10^{-10})$. We perform another cross-validation study and visualize the results in Figure 7. Since NLS generalized Poisson brackets and commutators have four components, we display the real and imaginary parts of the first row and the first column components for different unseen samples in Figure 7. We again see that our numerical optimization generalizes to unseen samples.

To investigate again the sensitivity of the integrability detection, we introduce nonintegrable Hamiltonians $h_1 = \frac{1}{3}|\psi|^6$ and $h_2 = \frac{1}{2}|\partial_x^2\psi|^2$. We solve problem (5.3), without sparsification, for Hamiltonian densities $h + \varepsilon_1 h_1 + \varepsilon_2 h_2$, where h is defined in (5.2) and $\varepsilon_1, \varepsilon_2 \in [-.01, .01]$. We show in Figure 8 that the loss has distinguished minima at the integrable points $\varepsilon_1 = 0$ (for fixed $\varepsilon_2 = 0$) and $\varepsilon_2 = 0$ (for fixed $\varepsilon_1 = 0$). Once again, we conclude that our methodology is precise enough to detect the integrability of the PDE under study.

Just as in previous sections, we aim to interpret our numerical results by introducing sparsification. We omit the numerical values and simply report the mathematical findings that we infer from our computation. These computations typically lead to an unbiased loss of $\mathcal{O}(10^{-15})$. That is, we report that with the same sparsification strategies employed earlier, we once again discover a new family of weak Lax pairs. However, this time SILO led us to a Lax pair that simultaneously produces a coupled system of linear and nonlinear Schrödinger equations.

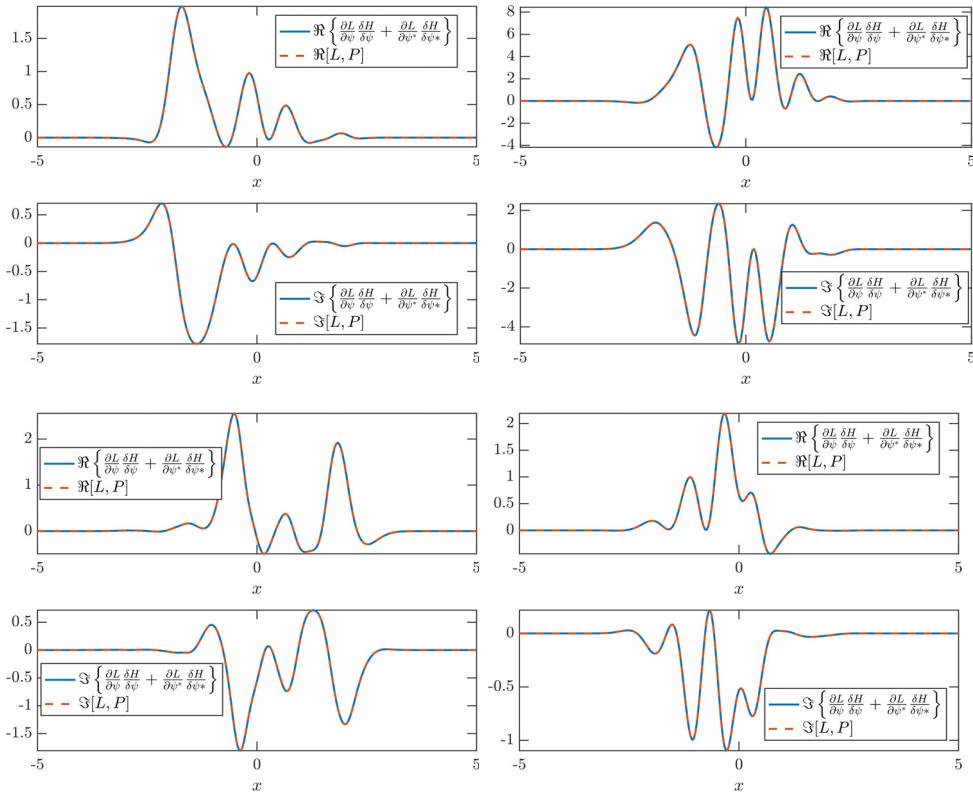


Figure 7. A numerical result of solving problem (5.3) without sparsification and with the 64-parameter hypothesis discussed in the text. Visualized here is a cross-validation study displaying the Poisson brackets and commutators evaluated at the optimal point η^* and on four samples from the function space Ω that were unseen during training. For all four cases, the loss is on the order of 10^{-10} . Displayed are the real and imaginary parts of the first components of the vectors resulting from the evaluation of the Lax pairs.

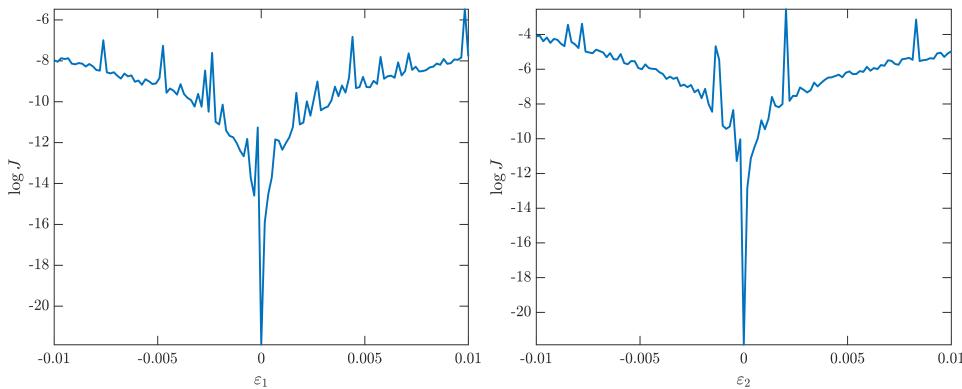


Figure 8. A perturbation study on problem (5.3). We solve problem (5.3), without sparsification, for Hamiltonian densities $h + \varepsilon_1 h_1 + \varepsilon_2 h_2$, where h is defined in (5.2) with the Hamiltonians $h_1 = \frac{1}{3}|\psi|^6$ and $h_2 = \frac{1}{2}|\partial_x^2\psi|^2$. We see that the loss has a distinguished minimum at the integrable points $\varepsilon_1 = 0$ (for fixed $\varepsilon_2 = 0$) and $\varepsilon_2 = 0$ (for fixed $\varepsilon_1 = 0$).

Theorem 5.1 (weak Lax pairs for the linear and nonlinear Schrödinger equations). Let $w(x) = [u(x) \ v(x)]^\top$, where each component is a complex function in the Schwartz space $\mathcal{S}(\mathbb{R})$. Then, for every $\psi, \varphi \in H^1([0, T]; H^3(\mathbb{R}))$, the Lax pair

$$L = \begin{pmatrix} a_1\varphi^* & a_2\psi \\ a_3\psi^* & a_4\varphi \end{pmatrix}, \quad P = \begin{pmatrix} b_1\partial_x^2 - b_2|\psi|^2 & 0 \\ 0 & -b_1\partial_x^2 + b_2|\psi|^2 \end{pmatrix}$$

produces the decoupled pair of equations

$$\begin{aligned} i\varphi_t &= b_1\varphi_{xx}, \\ i\psi_t &= b_1\psi_{xx} - 2b_2|\psi|^2\psi \end{aligned}$$

almost everywhere together with their complex conjugates.

Proof. By direct analogy with the KdV case, an evaluation on $w \in \mathcal{S}(\mathbb{R})^2$ along with integration by parts and function space embeddings wherever necessary reproduces the desired linear and nonlinear Schrödinger equations (almost everywhere). ■

It is interesting to note that this Lax pair not only produces the NLS but its linear counterpart, too. The free space linear Schrödinger equation is, after all, exactly solvable. We note that in the classical case, the Zakharov–Shabat Lax pair produces the NLS and its complex conjugate counterpart; that is, it does not reproduce a linear Schrödinger model.

Our numerical findings of sparse Lax pairs helped us in the discovery of the Lax pair in Theorem 5.1. During the manual process of generalizing our numerical results to a mathematical statement about Lax pairs, we recognized a degenerate case where the spectrum of L is zero for all time. This is indeed the sparsest possible weak Lax pair that produces the NLS equation, as we show now.

Theorem 5.2 (sparsest weak Lax pair for the NLS equation). Let $w(x) = [u(x) \ v(x)]^\top$, where each component is a function in the Schwartz space $\mathcal{S}(\mathbb{R})$. Then, for every $\psi \in H^2([0, T]; H^3(\mathbb{R}))$, the Lax pair

$$L = \begin{pmatrix} 0 & 0 \\ \alpha\psi^* & 0 \end{pmatrix}, \quad P = \begin{pmatrix} \beta\partial_x^2 + \gamma|\psi|^2 & 0 \\ 0 & 0 \end{pmatrix}$$

is the sparsest weak Lax pair that reproduces the NLS equation

$$i\psi_t = -\beta\psi_{xx} - \gamma|\psi|^2\psi$$

almost everywhere.

Proof. In this case, verifying the Lax pair is a short calculation, so we provide the details here. Observe that

$$\int_{\mathbb{R}} [L, P] w dx = \begin{pmatrix} 0 \\ \int_{\mathbb{R}} (\alpha\beta\psi^* u_{xx} + \alpha\gamma|\psi|^2\psi^* u) dx \end{pmatrix}$$

while

$$\int_{\mathbb{R}} \partial_t L w dx = \begin{pmatrix} 0 \\ \alpha \int_{\mathbb{R}} \psi_t^* u dx \end{pmatrix}.$$

Setting $\int_{\mathbb{R}} (\partial_t L - i[L, P]) w dx = 0$, dividing through by α , and integrating by parts twice, we see that the second component is now

$$\langle -i\psi_t^* - \beta\psi_{xx}^* - \gamma|\psi|^2\psi^*, u \rangle_{L^2(\mathbb{R})} = 0,$$

which is the desired weak form of the NLS equation. The fact that this is the sparsest Lax pair naturally follows by contradiction. ■

In a certain sense, the results of SILO in the present example provided us with only a trivial weak Lax pair that produces the NLS equation. One can argue that this is perhaps not (sufficiently) mathematically useful, but with respect to our computational framework, this is the best we could hope for. It is therefore not surprising that our numerics, aimed at finding the sparsest possible Lax pairs, correctly discover the Lax pairs shown in Theorems 5.1 and 5.2, all of which, under an integral, are sparser than the Zakharov–Shabat one.

6. Conclusion and outlook. In many ways, we (hope that we) have demonstrated that SILO is a step in the right direction for automated discovery of integrability in Hamiltonian dynamical systems. Broadly speaking, SILO achieves two major goals: high-precision detection of integrability and the discovery of interpretable, sparse Lax pairs (potentially weak ones through a relevant definition that was made precise earlier in the text). Nevertheless, it is important to highlight that for the former one of these goals (the detection of integrability), there already exist adequate, accurate, alternative methods, such as notably the detection of the Lyapunov spectrum [5], which has been used widely in recent studies [39, 40]. In the Hamiltonian setting, the vanishing of this spectrum is tantamount to integrability (for any finite-dimensional Hamiltonian system). Rather, here, the principal scope of our work is, once integrability is suspected, to seek to identify a (potentially approximate) Lax pair that optimizes the representation of the Lax equation and can guide the efforts for the analytical consideration of the solutions of such a putative integrable model.

Despite its successes, SILO is certainly not without flaws. SILO consistently finds the six-term lower-order Lax pair of Theorem 4.1 instead of the typical five-term one reported in the literature. It was only with some guidance that we could reproduce the well-known pair. Additionally, despite attempts to filter out the discovery of weak Lax pairs by sampling from periodic function spaces, SILO, in its current form, did not succeed in identifying a strong Lax pair for the NLS equation. This warrants future study on how to tune our framework toward a more reliable discovery of *all* Lax pairs that may exist. Additionally, both the function spaces within which functions are selected and the very choice of examining the action of the operator equation on the solution (rather than on arbitrary functions) are aspects worth revisiting and improving, as may be possible, in future efforts. Indeed, accounting in some suitably generalized way for the operator nature of the Lax equation is a significant direction that can potentially be targeted for future efforts.

There are numerous additional veins for improvement and continued exploration. One such study is to construct a wide operator hypothesis to attempt to thoroughly investigate the integrability of the Henon–Heiles system. Henon–Heiles indeed has two other known integrable points, yet the Lax pairs known to exist at these points are functionally different, and not just by the placement of the p 's and q 's, than the one we investigated in this work. Searching over such wide hypotheses is a computational challenge that should be addressed because this

is a clear path to the potential discovery of unknown points of integrability over the (A, B, ε) landscape. There are also opportunities to improve our computational framework to support large-scale nonconvex numerical optimization methods to attack large lattice-based systems such as Ablowitz–Ladik and Toda-type systems. There, potentially, further complications may arise, such as, e.g., the nonstandard Poisson brackets of the Ablowitz–Ladik setting [31]. Another dimension worth exploring is that of settings where nonalgebraic integrability is involved (and how to formulate relevant function libraries therein).

We also leave behind interesting mathematical questions in the context of the Lax pairs discovered in the PDE setting. Are weak Lax pairs of potential mathematical use, or are they simply weeds in the landscape of viable operator pairs? Does the strong Lax pair of Theorem 4.1 have an impact on the inverse scattering theory for the KdV equation? [A concern in the latter vein is the non-Hermitian nature of the resulting L operator, especially since the isospectrality of this operator and the nature of its eigenvalues are crucial to the entire inverse scattering machinery [1].] Is this strong pair an example of a “fake” Lax pair [9], that is, one which reproduces the Hamiltonian system yet fails to reproduce the correct conserved quantities? Or does this Lax pair imply Liouville integrability; that is, does it satisfy the geometric conditions of Gui-zhang [23]?

We believe that sparse symbolic regression techniques will continue to lead to breakthroughs in automated discovery within mathematical physics. Indeed, our intention in the present work is to plant the seed towards further potential classical and quantum, low-dimensional and field theoretic data-driven approaches. Beyond integrability, there is also the question of (maximal) superintegrability on an N -degree-of-freedom Hamiltonian system with $m > N$ integrals up to $m = 2N - 1$ for maximal superintegrability. Perhaps the best known model, both in the classical and the quantum settings, is the Calogero–Moser model [10]. Although much is known in the classical setting, in the quantum case, superintegrability is an open question for $N > 2$.

Another excellent testbed for adapting the essence of the SILO framework is the realm of exactly solvable models of statistical and quantum field theories. In fact, in that context, one could think of detecting exact solvability by exploiting the Yang–Baxter relation [33]. The 1D and 2D Ising, eight-vertex, quantum Heisenberg, Lieb–Liniger, and Hubbard models are all examples of field theories that are consistent with regard to the Yang–Baxter equation. We ask the following: Could the sparse identification of Yang–Baxter scattering matrices be the key to automated discovery of integrable quantum field theories as well?

Acknowledgments. The first author acknowledges helpful discussions with Roy Goodman, who suggested studying the Henon–Heiles system, Peter Miller, who pointed out the possibility of fake Lax pairs, and Sébastien Motsch, who suggested studying the effects of using \mathcal{R}_p for $0 < p < 1$ in sparsification. The authors also acknowledge Nicholas Bagley for his initial efforts on this project.

REFERENCES

- [1] M. J. ABLOWITZ, *Nonlinear Dispersive Waves: Asymptotic Analysis and Solitons*, Cambridge University Press, 2011, <https://doi.org/10.1017/CBO9780511998324>.
- [2] M. J. ABLOWITZ AND H. SEGUR, *Solitons and the Inverse Scattering Transform*, SIAM, 1981, <https://doi.org/10.1137/1.9781611970883>.

- [3] V. I. ARNOL'D, *Mathematical Methods of Classical Mechanics*, Grad. Texts in Math. 60, Springer Science & Business Media, 2013.
- [4] J. BARROW-GREEN, *Poincaré and the Three Body Problem*, Hist. Math. 11, American Mathematical Society, 1997.
- [5] G. BENETTIN, L. GALGANI, A. GIORGILLI, AND J.-M. STRELCYN, *Lyapunov characteristic exponents for smooth dynamical systems and for Hamiltonian systems; a method for computing all of them. Part 1: Theory*, Meccanica, 15 (1980), pp. 9–20, <https://doi.org/10.1007/BF02128236>.
- [6] T. BERTALAN, F. DIETRICH, I. MEZIĆ, AND I. G. KEVREKIDIS, *On learning Hamiltonian systems from data*, Chaos, 29 (2019), 121107, <https://doi.org/10.1063/1.5128231>.
- [7] A. V. BOLSINOV AND A. T. FOMENKO, *Integrable Hamiltonian Systems: Geometry, Topology, Classification*, CRC Press, 2004.
- [8] S. L. BRUNTON, J. L. PROCTOR, AND J. N. KUTZ, *Discovering governing equations from data by sparse identification of nonlinear dynamical systems*, Proc. Natl. Acad. Sci. USA, 113 (2016), pp. 3932–3937, <https://doi.org/10.1073/pnas.1517384113>.
- [9] S. BUTLER AND M. HAY, *Two definitions of fake lax pairs*, in AIP Conference Proceedings, Vol. 1648, AIP Publishing, 2015.
- [10] F. CALOGERO, *Solution of the one-dimensional n-body problems with quadratic and/or inversely quadratic pair potentials*, J. Math. Phys., 12 (1971), pp. 419–436, <https://doi.org/10.1063/1.1665604>.
- [11] F. CALOGERO AND M. C. NUCCI, *Lax pairs galore*, J. Math. Phys., 32 (1991), pp. 72–74, <https://doi.org/10.1063/1.529096>.
- [12] S. CHEN, P. G. KEVREKIDIS, H.-K. ZHANG, AND W. ZHU, *Data-driven discovery of conservation laws from trajectories via neural deflation*, Commun. Nonlinear Sci., 143 (2025), 108563, <https://doi.org/10.1016/j.cnsns.2024.108563>.
- [13] A. CUNHA, JR., M. V. ISSA, J. C. BASILIO, AND J. G. T. RIBEIRO, *CEopt: A MATLAB Package for Non-convex Optimization with the Cross-Entropy Method*, preprint, [arXiv:2409.00013](https://arxiv.org/abs/2409.00013), 2024.
- [14] P. DE KOSTER AND S. WAHLS, *Data-driven identification of the spectral operator in AKNS Lax pairs using conserved quantities*, Wave Motion, 127 (2024), 103273, <https://doi.org/10.1016/j.wavemoti.2024.103273>.
- [15] P. DEIFT AND X. ZHOU, *A steepest descent method for oscillatory Riemann–Hilbert problems. Asymptotics for the MKdV equation*, Ann. of Math. (2), 137 (1993), pp. 295–368, <https://doi.org/10.2307/2946540>.
- [16] F. DIACU AND P. HOLMES, *Celestial Encounters: The Origins of Chaos in Classical and Quantum Mechanics*, Princeton University Press, 1996.
- [17] V. G. DRINFEL'D, *Quantum groups*, J. Soviet Math., 41 (1988), pp. 898–915, <https://doi.org/10.1007/BF01247086>.
- [18] L. C. EVANS, *Partial Differential Equations*, Grad. Stud. Math. 19, American Mathematical Society, 2022.
- [19] A. P. FORDY, *The Hénon-Heiles system revisited*, Phys. D, 52 (1991), pp. 204–210, [https://doi.org/10.1016/0167-2789\(91\)90122-P](https://doi.org/10.1016/0167-2789(91)90122-P).
- [20] C. S. GARDNER, J. M. GREENE, M. D. KRUSKAL, AND R. M. MIURA, *Method for solving the Korteweg-deVries equation*, Phys. Rev. Lett., 19 (1967), pp. 1095–1097, <https://doi.org/10.1103/PhysRevLett.19.1095>.
- [21] H. GOLDSTEIN, C. POOLE, AND J. SAFKO, *Classical Mechanics*, 3rd ed., Addison- Wesley, 2002.
- [22] S. GREYDANUS, M. DZAMBA, AND J. YOSINSKI, *Hamiltonian neural networks*, Adv. Neural Inf. Process. Syst., 32 (2019).
- [23] T. GUI-ZHANG, *On Liouville integrability of zero-curvature equations and the Yang hierarchy*, J. Phys. A: Math. General, 22 (1989), pp. 2375–2392, <https://doi.org/10.1088/0305-4470/22/13/031>.
- [24] V. GUILLEMIN AND S. STERNBERG, *Symplectic Techniques in Physics*, Cambridge University Press, 1990.
- [25] T. HASTIE, R. TIBSHIRANI, AND J. FRIEDMAN, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd ed., Springer, 2009.
- [26] T. HASTIE, R. TIBSHIRANI, AND M. WAINWRIGHT, *Statistical Learning with Sparsity: The Lasso and Generalizations*, CRC Press, 2015.
- [27] M. HENON AND C. HEILES, *The applicability of the third integral of motion: Some numerical experiments*, Astronomical J., 69 (1964), 73, <https://doi.org/10.1086/109234>.

- [28] J. V. JOSÉ AND E. J. SALETAN, *Classical Dynamics: A Contemporary Approach*, Cambridge University Press, 1998.
- [29] S. KANTAMneni, Z. LIU, AND M. TEGMARK, *OptPDE: Discovering novel integrable systems via AI-human collaboration*, Phys. Rev. E., 111 (2025), 025303.
- [30] C. E. KENIG, G. PONCE, AND L. VEGA, *Well-posedness of the initial value problem for the Korteweg-de Vries equation*, J. Amer. Math. Soc., 4 (1991), pp. 323–347, <https://doi.org/10.1090/S0894-0347-1991-1086966-0>.
- [31] P. G. KEVREKIDIS, *The Discrete Nonlinear Schrödinger Equation: Mathematical Analysis, Numerical Computations and Physical Perspectives*, Springer Tracts Modern Phys. 232, Springer, 2009, <https://doi.org/10.1007/978-3-540-89199-4>.
- [32] R. KILLIP AND M. VIŞAN, *KdV is well-posed in H^1* , Ann. of Math. (2), 190 (2019), pp. 249–305, <https://doi.org/10.4007/annals.2019.190.1.4>.
- [33] V. E. KOREPIN, V. E. KOREPIN, N. BOGOLIUBOV, AND A. IZERGIN, *Quantum Inverse Scattering Method and Correlation Functions*, Vol. 3, Cambridge University Press, 1997.
- [34] S. KRIPPENDORF, D. LÜST, AND M. SYVAERI, *Integrability ex machina*, Fortschr. Physik, 69 (2021), 2100057, <https://doi.org/10.1002/prop.202100057>.
- [35] G. S. KRISHNASWAMI AND T. VISHNU, *An Introduction to Lax Pairs and the Zero Curvature Representation*, preprint, [arXiv:2004.05791](https://arxiv.org/abs/2004.05791), 2020.
- [36] J. L. LAGRANGE, *Mécanique analytique*, Vol. 1, Mallet-Bachelier, 1853.
- [37] P. D. LAX, *Integrals of nonlinear equations of evolution and solitary waves*, Comm. Pure Appl. Math., 21 (1968), pp. 467–490, <https://doi.org/10.1002/cpa.3160210503>.
- [38] Z. LIU, V. MADHAVAN, AND M. TEGMARK, *Machine learning conservation laws from differential equations*, Phys. Rev. E (3), 106 (2022), 045307, <https://doi.org/10.1103/PhysRevE.106.045307>.
- [39] M. MALISHAVA AND S. FLACH, *Lyapunov spectrum scaling for classical many-body dynamics close to integrability*, Phys. Rev. Lett., 128 (2022), 134102, <https://doi.org/10.1103/PhysRevLett.128.134102>.
- [40] T. MITHUN, A. MALUCKOV, A. MANČIĆ, A. KHARE, AND P. G. KEVREKIDIS, *How close are integrable and nonintegrable models: A parametric case study based on the Salerno model*, Phys. Rev. E (3), 107 (2023), 024202, <https://doi.org/10.1103/PhysRevE.107.024202>.
- [41] A. C. NEWELL, *Solitons in Mathematics and Physics*, SIAM, 1985, <https://doi.org/10.1137/1.9781611970227>.
- [42] J. NOCEDAL AND S. J. WRIGHT, *Numerical Optimization*, Springer, 1999.
- [43] D. Ó'MATHÚNA, *Integrable Systems in Celestial Mechanics*, Prog. Math. Phys. 51, Birkhäuser Boston, 2008.
- [44] E. OTT, *Chaos in Dynamical Systems*, Cambridge University Press, 2002.
- [45] H. POINCARÉ, *New Methods of Celestial Mechanics*, Vol. 450-451, National Aeronautics and Space Administration, 1967.
- [46] V. RAVOSON, L. GAVRILOV, AND R. CABOZ, *Separability and Lax pairs for Hénon–Heiles system*, J. Math. Phys., 34 (1993), pp. 2385–2393, <https://doi.org/10.1063/1.530123>.
- [47] R. RUBINSTEIN, *A stochastic minimum cross-entropy method for combinatorial optimization and rare-event estimation*, Methodol. Comput. Appl. Probab., 7 (2005), pp. 5–50, <https://doi.org/10.1007/s11009-005-6653-7>.
- [48] P. TOTH, D. J. REZENDE, A. JAEGLE, S. RACANIÈRE, A. BOTEV, AND I. HIGGINS, *Hamiltonian generative networks*, in Proceedings of the International Conference on Learning Representations, 2020.
- [49] E. J. VLADISLAVLEVA, *Symbolic Regression Methods: Integrating Domain Knowledge into Machine Learning*, CRC Press, 2020.
- [50] R. S. WESTFALL, *Never at Rest: A Biography of Isaac Newton*, Cambridge University Press, 1980.
- [51] V. E. ZAKHAROV AND L. D. FADDEEV, *Korteweg–de Vries equation: A completely integrable Hamiltonian system*, Funktsional'nyi Analiz i ego Prilozheniya, 5 (1971), pp. 18–27.
- [52] V. E. ZAKHAROV AND A. B. SHABAT, *Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media*, Soviet Phys. JETP, 34 (1972), pp. 62–69.

- [53] X. ZHOU, *The Riemann–Hilbert problem and inverse scattering*, SIAM J. Math. Anal., 20 (1989), pp. 966–986, <https://doi.org/10.1137/0520065>.
- [54] W. ZHU, H.-K. ZHANG, AND P. KEVREKIDIS, *Machine learning of independent conservation laws through neural deflation*, Phys. Rev. E (3), 108 (2023), L022301, <https://doi.org/10.1103/PhysRevE.108.L022301>.