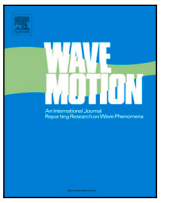
Wave Motion 127 (2024) 103273

Contents lists available at ScienceDirect

Wave Motion

journal homepage: www.elsevier.com/locate/wamot



Data-driven identification of the spectral operator in AKNS Lax pairs using conserved quantities

Pascal de Koster [a,*], Sander Wahls [b]

[a] Delft University of Technology, Delft Center for Systems and Control, Mekelweg 2, Delft, 2628 CD, Netherlands
[b] Karlsruhe Institute of Technology, Institute of Industrial Information Technology, Hertzstr. 16, Karlsruhe, 76187, Germany

ARTICLE INFO

Keywords:
AKNS
Identification
Forward scattering transform
Nonlinear Fourier transform

ABSTRACT

Lax-integrable partial differential equations (PDEs) can by definition be described through a compatibility condition between two linear operators. These operators are said to form a Lax pair for the PDE, which itself is usually nonlinear. Lax pairs are a very useful tool, but unfortunately finding them is a difficult problem in practice. In this paper, we propose a method that determines the spectral operator of an AKNS-type Lax pair such that the corresponding PDE fits given measurement data as well as possible. The spectral operator then enables practitioners to solve or analyze the underlying PDE using the induced nonlinear Fourier transform. The underlying PDE only has to be approximately Lax-integrable; the method will find the spectral operator that explains the data best. Together with the dispersion relation, the spectral operator of AKNS type completely determines an integrable PDE that approximates the true underlying PDE. We identify the most suitable spectral operator by matching PDE-dependent quantities that should be conserved during evolution. The method is automatic and only requires recordings of solutions at two different values of the evolution variable, which do not have to be close.

1. Introduction

Many nonlinear partial differential equations (PDEs) can be solved with the help of Lax pairs, which consist of a linear spectral operator L and a linear propagation operator A, and the inverse scattering method (ISM) [1,2]. Well-known examples include the Korteweg–de Vries equation (KdV) [3], the modified Korteweg–de Vries equation (MKdV) [4] and the nonlinear Schrödinger equation (NLSE) [5]. The ISM for the solution of such so-called Lax-integrable systems proceeds as follows. First, the initial condition is used as a potential function for the linear spectral operator L. The spectrum of this operator together with the corresponding (generalized) eigenfunctions leads to an equivalent spectral representation of the initial condition. The evolution of this spectral representation can then be performed in closed form with the help of the linear propagation operator A. Finally, a suitable spectral theorem is used to recover the solution in the original domain from the evolved spectral representation. The spectral representation can be interpreted as a nonlinear Fourier transform (NFT; a.k.a. forward scattering transform) that moves a signal from its original domain to a PDE-specific spectral domain in which the propagation becomes trivial [2]. The spectral representation furthermore has physical interpretations. Most importantly, it is able to reveal hidden solitons, which are localized particle-like waveforms [6]. NFTs have therefore found practical applications in areas such as ocean wave data analysis (e.g. [7–12]) and fiber-optical communication (e.g. [13–22]). In these areas, many dynamics are known to be well-approximated by the most common Lax-integrable PDEs.

However, the situation is different in areas where the notion of Lax-integrability is not yet well-known. When a system is suspected to be Lax-integrable, e.g. because it admits soliton solutions, the question arises if the system can be described by a

* Corresponding author.
E-mail addresses: p.b.j.dekoster@tudelft.nl (P. de Koster), wahls@kit.edu (S. Wahls).

https://doi.org/10.1016/j.wavemoti.2024.103273
Received 17 September 2023; Received in revised form 11 January 2024; Accepted 18 January 2024
Available online 1 February 2024

Lax-integrable PDE, and if so, which PDE. In practice, it is often very difficult to determine which Lax-integrable PDE approximates a given system best. From a practical point of view, it is often preferable to approach this question in a data-driven way as the analytical derivation of Lax-integrable approximations requires very specific human expertise, time and sometimes also some luck. Data-driven approaches instead try to identify the system directly from measurements of the system. Of course, they come with their own challenges. For the identification of PDEs, one of the biggest practical problems is that the data can typically only be measured for a few sparse values of the evolution variable, and thus the derivative in the direction of propagation cannot be determined. The sparseness of measurement points occurs naturally e.g. in wave flume experiments [12,23], where wave gauges record time series at a limited number of locations, or in optical fiber experiments [24,25], where time series can only be recorded at the ends of the fiber link. In contrast, many popular generic PDE identification methods such as [26] assume that the data contains derivatives w.r.t. the evolution variable, which is not possible with sparse measurement points. To the best of our knowledge, so far only a single method for the data-driven identification of Lax pairs has been proposed in the literature [27]. However, also this methods requires derivatives w.r.t. the evolution variable. The lack of a practical system identification method for Lax-integrable systems hinders the application of the strong mathematical theory behind Lax-integrable systems in practice.

In this paper, we therefore propose a novel approach to identify a Lax pair from given measurement data that does not require derivatives in the direction of propagation. Instead, we exploit that the spectral operator of any Lax pair produces infinitely many conserved quantities that should not change during propagation. The parameters of the spectral operator are chosen such that the variation in the (ideally) conserved quantities in minimal. Our method is data-driven and identifies the PDE using only measurements space series (snapshots) of solutions, taken at different time points (assuming from here on, without loss of generality, that the time t is the evolution variable of the PDE, and the location x the other variable).

In earlier work, we have already demonstrated for the NLSE [25,28] and the KdV equation [23] that parameters of specific integrable PDEs can be identified based on conserved quantities. Furthermore, conserved quantities have also successfully been applied for other applications such as training neural networks for solving Lax-integrable PDEs [29]. However, so far conserved quantities have not been exploited for the data-driven identification of Lax pairs.

In the literature, there already is a large variety of non-data-driven techniques that determine Lax pairs for a given Lax-integrable PDE (e.g., [26,30–38]). One might thus be tempted to first identify a PDE using conventional data-driven methods (e.g. [39,40]), and then find a Lax pair for it. However, this approach has several major problems. First of all, existing methods for finding Lax pairs assume that the given PDE is exactly Lax-integrable. However, in practice this will only be the case approximately. Even if the true underlying PDE is exactly Lax-integrable, the model identified by a conventional method still will not have this property due to measurement noise. Second, existing techniques to find Lax pairs from a given Lax-integrable PDE are not guaranteed to succeed, to the best of our knowledge. Third, as mentioned above, conventional PDE identification methods require measurements of the solution at closely spaced time points, which, as already pointed out earlier, is often unpractical.

In our approach, we will assume that the Lax-pair is of the AKNS-type, which is named after Ablowitz, Kaup, Newell and Segur [2]. The AKNS class captures a large class of ubiquitous Lax-integrable PDEs, such as, e.g., the (modified) Korteweg–de Vries equation (MKdV/KdV), the (de)focussing nonlinear Schrödinger equation (dNLSE/fNLSE), and the sine/sinh-Gordon equation [2], as well as transformations and higher-order variants of these equations. While the restriction to the AKNS class of course means a loss of generality,[1] we point out that on the other hand it allows us to design a more efficient method. Most importantly, we can exploit that the conservation laws of AKNS Lax pairs all have the same structure. We envision that in the future, more specialized algorithms for other classes of Lax pairs will be developed. The user can then choose the right class either using physical insight, or by simply checking which class provides the best fit. This would be similar to conventional nonlinear system identification, where specialized algorithms exist for various system classes such as Wiener–Hammerstein systems, finite Volterra series, or nonlinear ARMA models [42].

In our method, we focus on identifying the spectral L operator of the AKNS-type Lax-pair, as the Lax propagation operator A can be found by combining the spectral operator and the linearized dispersion relation [2, p. 253]. Given suitable measurement data, the dispersion relation can be found easily by comparing phase shifts in the linear Fourier domain. To identify the L operator that fits the given data best, we will exploit the explicitly known conserved global quantities associated with AKNS-type Lax pairs. Our strategy is to identify the L operator, parameterized by two potential functions, for which the global quantities vary as little as possible over time.

Our method has the advantage that it is fully automatic and data driven. Its main advantage is thus that it only requires measurements at two different points in time, which do not have to be close. (The use of more time points is possible.) We finally remark that the proposed method may of course also be used for data in the form of time series that are measured at different locations if the roles of space and time are switched in the PDE so that the location is the evolution variable. This is a common scenario in applications. See, e.g., [12,20,23].

This paper is organized as follows. Section 2 recapitulates the AKNS-type Lax pair and the associated conserved quantities. Section 3 describes the method to identify the most suitable AKNS-type PDE from measurement data by comparing global quantities. Section 4 demonstrates the method on various data sets. Finally, Section 5 concludes the paper.

[1] The "integrability ex machina" (IeM) method in [27] employs neural networks to represent the Lax pair and therefore, at first sight, appears to cover a much larger class of systems than our approach. However, this is not the case. The neural network in the IeM method implements a conventional polynomial approach [27, p. 4]. The advantage of the neural network formulation is that advanced algorithms to find the optimal weights can be employed. In [41, p. 65], it is clarified that the IeM method cannot use arbitrary neural networks (specifically, multi-layer perceptrons). The IeM method also requires physical intuition to select dictionaries [27, p. 6].

2. Theory: AKNS Lax pairs and conserved quantities

Let a nonlinear partial differential equations (PDE) be given by

$$u_t(x,t) = F(u, u_x, u_{xx}, \dots), \tag{1}$$

where u denotes the (possibly complex) signal amplitude, x the position, t the time (or more generally, the evolution variable), F the nonlinear evolution function, and subscripts denote partial derivatives. A nonlinear PDE has a Lax pair $(L(t), A(t))$, where $L(t)$ and $A(t)$ are linear operators that depend on $u(x,t)$, such that the following Lax equation is equivalent to the PDE [1, Eq. 1.4]:

$$L_t = AL - LA. \tag{2}$$

We will refer to L as the spectral operator, and to A as the propagation operator. If the Lax pair (L, A) fits the PDE of interest, then the spectrum of $L(t)$ is constant over time [1]. We say that the PDE is Lax-integrable if the Lax pair enables the solution of the PDE using the inverse scattering method [43]. A simple example of a PDE and a corresponding Lax pair is the advection equation $u_t = u_x$, where the Lax pair consists of the multiplication operator $L = u$ and the derivative operator $A = \partial_x$. For any suitable $\phi(x,t)$, we have

$$L_t\phi = (L\phi)_t - L\phi_t = (u\phi)_t - u\phi_t = u_t\phi \quad \text{and} \quad AL\phi - LA\phi = \partial_x(u\phi) - u\partial_x\phi = u_x\phi$$

$$\Rightarrow \quad L_t = AL - LA \iff u_t = u_x.$$

Many Lax-integrable PDEs exist, and the form of the associated Lax pairs can widely vary. However, the AKNS-type Lax pairs all possess the same structure, but still capture many ubiquitous PDEs. The AKNS-type Lax pairs are of the form [2]

$$L = \begin{bmatrix} i\partial_x & -iq(u) \\ ir(q) & -i\partial_x \end{bmatrix}, \qquad A = \begin{bmatrix} A_{11}(u) & A_{12}(u) \\ A_{21}(u) & -A_{11}(u) \end{bmatrix} \tag{3}$$

in which $q(u)$ and $r(q) = r(q(u))$ are potential functions depending on the x.

Not every choice of $q(u)$ and $r(q)$ is allowed, as the Lax equation in Eq. (2) imposes a compatibility condition (see [2, Eq. 2.2-2.7]). Many of the known Lax-integrable system can be obtained using only a small number of $r(q)$ relations [2]:

$$r = -1 \text{ (KdV)}, \; r = -q^* \text{ (focussing NLSE)}, \; r = q^* \text{ (defocussing NLSE)},$$

$$r = -q \text{ (sine-Gordon, focusing MKdV)}, \; r = q \text{ (sinh-Gordon, defocusing MKdV)}. \tag{4}$$

We will therefore consider only these relations for $r(q)$. Of course, more choices can be integrated into the method, although it should be checked beforehand if those relations $r(q)$ can lead to a compatible Lax pair. For most equations, such as KdV, MKdV and NLSE, $q(u)$ simply takes the form of $q = cu$ (i.e., a scaling of the signal amplitude, with c the scaling coefficient). The sine-Gordon and sinh-Gordon equations on the other hand require $q = cu_x$. Furthermore, we may also be dealing with the case that the measurement process does not directly provide u, but a transformed version of it. For example, $q(u) = u^3, r(q) = -1$ would still result in a feasible spectral operator of the KdV-type, but the associated PDE would correspond to a transformed version of the KdV (the exact PDE would depend on the linearized dispersion relation).

2.1. Conserved quantities of AKNS-type PDEs

AKNS-type PDEs possess an infinite number of conserved quantities, that are fully determined by $q(u)$ and $r(q)$. The conserved quantities can be iteratively derived using a relation from AKNS [2, Eq. 7.33]. The first five conserved quantities are

$$C_1 = \int_D qr \,\mathrm{d}x, \qquad C_2 = \int_D rq_x - r_x q \,\mathrm{d}x, \tag{5a}$$

$$C_3 = \int_D q^2r^2 + q_x r_x \,\mathrm{d}x, \qquad C_4 = \int_D -r_{xx}q_x + q_{xx}r_x + \tfrac{3}{2}\left(r^2(q^2)_x - q^2(r^2)_x\right) \,\mathrm{d}x, \tag{5b}$$

$$C_5 = \int_D 2q^3r^3 + q_{xx}r_{xx} + \left(q_x^2r^2 + r_x^2q^2\right) + 8qq_xrr_x \,\mathrm{d}x, \tag{5c}$$

where D denotes the spatial domain. Both periodic and vanishing boundary conditions are allowed.

The choices $r = -1$ and $r = \pm q$ lead to trivial conserved quantities $C_2 = 0$ and $C_4 = 0$, but the odd quantities are non-trivial for all considered choices of $r(q)$. Therefore, we only consider the odd global quantities C_1, C_3 and C_5 throughout the rest of this paper. These three conserved quantities turn out to be sufficient in our numerical experiments. Higher ones could of course be integrated, and with large libraries for the expansion of q (see Section 3), this may even become necessary for a successful identification. Unfortunately, determining the higher conserved quantities analytically is a difficult task, even with the help of symbolic computer algebra systems. Higher order conserved quantities furthermore contain derivatives of high orders, which are very sensitive to measurement noise. We therefore believe that the exploitation of higher order conserved quantities for large libraries will require a numerical method for their computation from noisy data.

We finally note that it is necessary to consider C_5, as only C_1 and C_3 are not sufficient to distinguish most PDEs. Many PDEs conserve both $\int_D u \,\mathrm{d}x$ and $\int_D u^2 \,\mathrm{d}x$, thus if $q = u$ then $r \in \{1, \pm u\}$ all lead to a conserved C_1. Furthermore, if only C_1 is used, and we found some relation $q = f(u)$ such that $C_1(t)$ is constant, then any scaling $q = cf(u)$ will also lead to a constant $C_1(t)$. To extract the scaling constant, at least C_3 is required for $r \in \{\pm q, \pm q^*\}$, as C_3 then has two terms with different scaling in q. For $r = -1$, C_3 still only has a single term ($\int_D q^2 \,\mathrm{d}x$), and therefore C_5 is required.

3. Method: Identifying the spectral operator using conserved quantities

In this section, we present our novel method to identify the spectral operator L of the AKNS Lax pair that fits a given data set best. The methods checks which $q(u)$ results in the least variation in the conserved quantities, for each of the choices of $r(u)$ in Eq. (4). The combination $(r(q), q(u))$ that results in the least variation of the conserved quantities is then considered the best candidate for the spectral operator explaining the data.

3.1. Expanding q(u) in functions from a library

Our approach starts with choosing a library of D operators of u, i.e., $g_1(u), g_2(u), \dots, g_D(u)$. We will allow $q(u)$ to be any linear combination of the operators in a library G, which we chose for this paper as the lowest-order polynomials and derivatives of u:

$$\begin{aligned} q(u) &= \sum_{d=1}^{D} c_d g_d(u), \text{ with } D = 5, \\ G &= \left\{ g_1 = u,\ g_2 = u_x,\ g_3 = u_{xx},\ g_4 = u^2,\ g_5 = u u_x \right\}. \end{aligned} \tag{6}$$

Here, c_d are the coefficients which are optimized during the identification. We chose the shown library to include at least u and u_x, such that standard versions of the KdV, MKdV and NLSE ($q = u$) and sine-Gordon and sinh-Gordon ($q = u_x$) are in the current library space. If more information of the PDE is available, the library can be expanded or shrunk accordingly. For example, if the underlying system is known to be either KdV or MKdV, the library can be shrunk to $G = \{g_1 = u\}$, and the problem then simplifies to a simple identification of the scaling constant c_1, and a choice between $r = -1$, $r = -q$ and $r = +q$.

3.2. Error function based on conserved quantities

To identify the most suitable potential function $r(q)$ as in (4) and q as in (6) for the L operator in (3), we will minimize an error based on the conserved quantities (5). We assume that our available data set consists of N independent trajectories

$$u^{(n)} = \left(u^{(n)}(\cdot, t_1), u^{(n)}(\cdot, t_2), \dots, u^{(n)}(\cdot, t_M) \right), \quad n = 1, \dots, N,$$

where the first "snapshot" $u^{(n)}(x, t_1)$ of the nth trajectory is assumed to be a measurement of the initial condition of the PDE, and the later snapshots at one or multiple later times $t_2, \dots, t_M$ are measurements of the correspondingly evolved initial conditions. Given a single trajectory $u^{(n)}$, each choice of $(q(u), r(q))$ results in three time-dependent values $C_1(t)$, $C_3(t)$ and $C_5(t)$ for the conserved quantities in (5). If the relations $(q(u), r(q))$ are chosen correctly and the data set was noiseless and obtained from an exactly Lax-integrable system, the conserved quantities will indeed be conserved, i.e., constant. We thus create an error that penalizes fluctuations in the conserved quantities over time.

We formally define the error as the standard deviation in time in the conserved quantities C_1, C_3 and C_5, relative to the average absolute conserved quantities:

$$\begin{aligned} E(c, r; u) &:= \sum_{k \in \{1,3,5\}} \left(\sum_{n=1}^{N} \frac{\sigma[C_k(t; c, r, u^{(n)})]}{\mu\left[\left|C_k(t; c, r, u^{(n)})\right|\right]} \right), \quad \text{with} \\ \mu[C_k(t)] &= \frac{1}{M} \sum_{m=1}^{M} C_k(t_m), \text{ and } \sigma[C_k(t)] = \sqrt{\frac{1}{M} \sum_{m=1}^{M} \left(C_k(t_m) - \mu[C_k(t)] \right)^2}, \end{aligned} \tag{7}$$

where $C_k(t; c, r, u^{(n)})$ denotes the corresponding conserved quantity in Eq. (5) computed at time t, which furthermore depends on the choices for the coefficient vector c, r, and the currently considered trajectory through Eqs. (4)–(6). The algorithm thus identifies r and $q = \sum_d c_d g_d$ by minimizing over the error in Eq. (7):

$$(r^{(\mathrm{ID})}, c^{(\mathrm{ID})}) = \operatorname*{argmin}_{r \in \{-1, \pm q, \pm q^*\}} \operatorname*{argmin}_{c \in C} w(c) E(c, r; u), \quad C \subseteq \mathbb{R}^d, \tag{8}$$

with $w(c)$ denoting the weight of coefficient c. For each choice of r, the error is minimized over c using a local minimization method. We used the *fminsearch* method from *Matlab* (a simplex search method), starting at multiple initial starting positions. The search space C for the coefficient vectors can in principle be chosen as the whole $\mathbb{R}^d$, but in our numerical examples we made a different choice. Similarly to conventional identification methods for PDEs (e.g. [40]), we prefer Lax pairs that are "simple" and therefore prioritize sparse coefficient vectors c. This will also prevent over-fitting, and can help to reduce the computation time. To find a sparse solution c we do not search the full space spanned by the library G at once, but rather search only its low-dimensional sub-spaces one-by-one as explained in the next section.

3.3. Sparse solutions and choice of the starting points

For large libraries G, it can be very hard to find a globally optimal $q(u)$ in the corresponding function space. However, for many of the well-known AKNS-type PDEs, the corresponding $q(u) = \sum_d c_d f_d(u)$ has only a few nonzero c_d. We exploit this by only searching the low-dimensional sub-spaces of functions in G that have at most $D^{(\mathrm{sub})}$ non-zero coefficients. In our numerical examples, we investigate all 1D, 2D and 3D sub-spaces. While the 1D and 2D sub-spaces are also contained within the 3D sub-space, it often occurs that a one- or two-dimensional solution performs nearly as good as a three dimensional solution, and should thus be preferred. Furthermore, the optimization process in lower-dimensional sub-spaces is more likely to find the global optimum in that sub-space. We thus also take the lower sub-spaces into account.

Given D basis functions, there are $\binom{D^{(\mathrm{sub})}}{D}$ sub-spaces of dimension $D^{(\mathrm{sub})}$. In each of these sub-spaces, we minimize the error in Eq. (7), and keep the optimal 1D-, 2D- and 3D-solutions. Finally, we penalize higher-dimensional solutions by multiplying each error with a heuristically chosen factor. The search space and weight for (8) are thus

$$C = \{c \in \mathbb{R}^D : \#c \leq D^{(\mathrm{sub})}, D^{(\mathrm{sub})} \leq 3\}, \quad w(c) = p^{\#c}, \tag{9}$$

with $\#c$ the number of non-zero coefficients in c. We found that $p = 1.2$ resulted in a good balance for finding a suitable but sparse coefficient vector.

We assume that the data set evokes both the linear and the nonlinear terms of the PDE, at comparable levels, so that all terms can be identified. If this is not the case, a richer data set must be considered, as many PDEs can otherwise not be distinguished. To define initial starting points, we determine a central estimate vector c^0, for which the terms $c_d^0 g_d(u)$ are all of the same order of magnitude, such that each g_d is relevant. We start with the first coefficient c_1^0, as $c_1^0 u$ is often part of the solution for q. We cannot use the conserved quantity C_1 for this purpose, as it contains only a single term; C_3 has two different terms, so we can find a c_1^0 such that $\int_D q^2 r^2 \,\mathrm{d}x$ and $\int_D q_x r_x \,\mathrm{d}x$ are of similar order. Only for $r = -1$ we have that $r_x = 0$, so in this case we consider C_5 instead. This leads to the following initial r-dependent guesses for c_1^0:

$$r = -1,\ q = c_1 u, \quad C_5 = \int_D 2c_1^3 u^3 + c_1^2 u_x^2 \,\mathrm{d}x \qquad \Rightarrow c_1^0 \approx \frac{\overline{|\int_D u_x^2(x,t)\,\mathrm{d}x|}}{\overline{|\int_D u^3(x,t)\,\mathrm{d}x|}} \tag{10a}$$

$$r = \mp q,\ q = c_1 u, \quad C_3 = \int_D c_1^4 u^4 + c_1^2 u_x^2 \,\mathrm{d}x \qquad \Rightarrow c_1^0 \approx \sqrt{\frac{\overline{|\int_D u_x^2(x,t)\,\mathrm{d}x|}}{\overline{|\int_D u^4(x,t)\,\mathrm{d}x|}}} \tag{10b}$$

$$r = \mp q^*,\ q = c_1 u, \quad C_3 = \int_D c_1^4 |u|^4 + c_1^2 |u_x|^2 \,\mathrm{d}x \qquad \Rightarrow c_1^0 \approx \sqrt{\frac{\overline{|\int_D |u_x|^2(x,t)\,\mathrm{d}x|}}{\overline{|\int_D |u|^4(x,t)\,\mathrm{d}x|}}}, \tag{10c}$$

where $\overline{(\cdot)}$ denotes the mean over all t and trajectories.

Next, we wish to ensure that all functions $c_d^0 g_d$ are of similar size for our initial guess c^0, as this will allow a change in every term to be significant. We thus choose

$$c_d^0 = \frac{c_1 \overline{|g_1|}}{\overline{|g_d|}}. \tag{11}$$

This defines our central initial guess.

Next, for each of the sub-spaces, we choose the first initial starting point as the projection of c^0 onto the corresponding subspace (i.e., we set $c_d^0 = 0$ if g_d is not in the sub-space). Next, we create a hypercube around the projected c_d^0 by multiplying or dividing each term by a scale factor $s > 1$. For example, for the subspace containing the first two basis functions, the five resulting initial starting points are the projected estimate $(c_1^0, c_2^0, 0, 0, 0)$ and the four vertices $(s^{\pm 1} c_1^0, s^{\pm 1} c_2^0, 0, 0, 0)$. Throughout this paper, we chose $s = 10$. This distribution of initial starting points sufficed for the examples considered in this paper. More initial starting points can of course be added to increase the chance of finding the global optimum when necessary, at the expense of longer computation times. We also note that all coefficients of our initial starting points were positive, as this sufficed for the considered examples. However, starting points with negative coefficients can of also be taken into account.

4. Results

In this section, we evaluate the proposed method in numerical examples for several PDEs. We first demonstrate it for noisy data obtained from the MKdV. Then, a lossy NLSE, which is only approximately integrable, is considered with noise. Next, we consider the sine-Gordon equation with noise, which requires a non-standard choice for the first potential function, namely $q = u_x$. Finally, we consider a transformed KdV, where $q(u) = u + u^2$ is a linear combination of two of the library functions. In all examples, the library of functions for $q(u)$ from Eq. (6) is used, in combination with the described low-dimensional sub-space search and the described initial starting points for c from Section 3.3.

4.1. Case 1: Noisy data obtained from the exactly integrable MKdV

In this subsection, we consider noisy data obtained from the focusing MKdV $u_t = -u_{xxx} - 6u^2u_x$, consisting of multiple periodic signals measured at two different times. This example illustrates that the algorithm can distinguish between the $r(q)$ relations for the focusing MKdV ($r = -q$), the defocusing MKdV ($r = +q$) and the (standard) KdV ($r = -1$), and also find the correct relation $q(u) = u$ from the full library function space.

The considered data set consist of $N = 21$ trajectories, where each trajectory consists of space series at only two time points, which are too far apart to determine u_t. The measurement locations and times were respectively $x = -4, -3.98, -3.96, \dots, 4$ (periodic boundary conditions), and $t \in \{0, 0.5\}$. The input signals ($t = 0$) were generated as zero-mean Gaussian distributed random numbers with variance one, ideally low-pass filtered to maximum wavenumber 1 cycle/spatial unit. Every input signal was scaled such that the mean squared amplitude of each trajectory varied between 0 and 2.5, to ensure that the dataset was rich enough. Next, all input signals were propagated for 0.1 temporal units. Finally, low-pass filtered white Gaussian noise (maximum wavenumber 1 cycle/spatial unit) with 1% of each considered signal's root-mean-squared (RMS) amplitude was added, i.e, the signal to noise ratio was $\mathrm{SNR} = \frac{\int_D |u_{(\mathrm{signal})}|^2 \,\mathrm{d}x}{\int_D |u_{(\mathrm{noise})}|^2 \,\mathrm{d}x} = 10^4$. We note here that the amount of noise was chosen such that the correct PDE could still be identified. This is often around 1% due to the use of higher derivatives and polynomials, which is typical for PDE identification methods (see e.g., [40]).

The results of the algorithm are shown in Fig. 1. One of the 21 input–output pairs is shown in the top left. For each choice of $r(q)$, we optimized c. The corresponding global quantities are shown at the bottom. The choice $r(q) = -q$ (indeed corresponding to focusing MKdV) shows the least variation between input and output, and thus the system is identified as $r(q) = -q$. Note that even for the correct choice of $r(q) = -q$, the global quantities are not entirely conserved due to the noise.

From Fig. 1(b) we find that the identified c corresponded to the correct c (up to a small error due to noise), which was identified from the 1D sub-space spanned by $\{u\}$. A near-correct c was also identified as $(1.06, 0, 0, 6 \cdot 10^{-3}, 0)$ in the $2D$ sub-spaces spanned by $\{u, u_{xx}\}$, but the corresponding error was only marginally smaller. Therefore, the 1D solution was accepted. The correct solution was not found at all in any of the 3D sub-spaces from the specified starting points, and the resulting error was therefore around 20 times larger. While it would in principle be possible to address this by just using more starting points, this example also shows that searching all lower-dimensional sub-spaces can be more effective if the reference solution can be assumed to be low-dimensional.

4.2. Case 2: Noisy data from a lossy, and thus only approximately integrable, NLSE

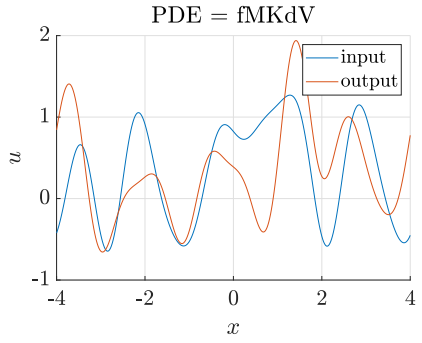
In this section, we consider complex-valued noisy data set obtained from the lossy focusing NLSE (fNLSE) $q_t = iq_{xx} + 2iq|q|^2 - 0.5\alpha q$ with $\alpha = 4$. The data is instantaneously amplified just before each measurement point to compensate the loss, and is therefore energy conserving (up to noise) between measuring points. This type of system is widely used – with reversed roles for time and space – in fiber-optical communication systems with lumped amplification [44,45]. The path-average approximation of the lossy fNLSE with amplification is obtained by averaging the decaying energy during over t, and then considering a lossless NLSE with rescaled amplitude $\overline{q}$ [45, Ch. 9]:

$$\begin{aligned} \mathcal{E}^{\mathrm{avg}} &= \int_D [q(x,t=0)]^2 \,\mathrm{d}x \int_0^T e^{-\alpha t} \,\mathrm{d}t = \int_D [q(x,t=0)]^2 \,\mathrm{d}x \frac{1-e^{-\alpha T}}{\alpha T} \\ \Rightarrow \overline{q(x,t)} &= \sqrt{\tfrac{1-e^{-\alpha T}}{\alpha T}}\, q(x,t=0), \end{aligned} \tag{12}$$

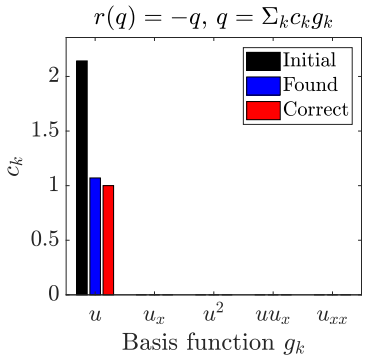
with $\mathcal{E}^{\mathrm{avg}}$ the signal energy averaged over t, $\overline{q(x,t)}$ the path-average amplitude, and T the duration between two amplifications (and thus measurement times). The measurement time points were at $t \in \{0, \frac{1}{3}, \frac{2}{3}, 1\}$, so with $T = \frac{1}{3}$ and $\alpha = 4$, we expect an amplitude normalization constant of $c_1 = \sqrt{\frac{1-e^{-\alpha T}}{\alpha T}} = 0.74$, and thus the relation $u = c_1 q$, and $r = -q^*$. Note that while the path-average approximation (12) is Lax-integrable, the true underlying lossy NLSE with amplification is not. This example thus demonstrates that the method is able to identify an integrable approximation to a non-integrable system from data.

We generated the input signals under periodic boundary conditions from complex, circularly symmetric, zero-mean Gaussian distributed random samples, and ideally low-pass filtered these to maximum wave number 0.5 cycles/spatial unit. The measurement points were $x = -4, -3.95, -3.9, \dots, 4$. Next, we propagated each signal numerically, boosting the signal just before every measurement. Finally, low-pass filtered radially symmetric white Gaussian noise (maximum wavenumber 0.5 cycles/spatial unit) with root mean squared amplitude equal to 1% of the signal amplitude was added to the input and output signals ($\mathrm{SNR} = 10^4$). We generated $N = 21$ input signals in total, with mean squared amplitudes varying between 0.2 and 0.6, to ensure that the data was rich enough. The initial starting points were chosen as described in Section 3.3.

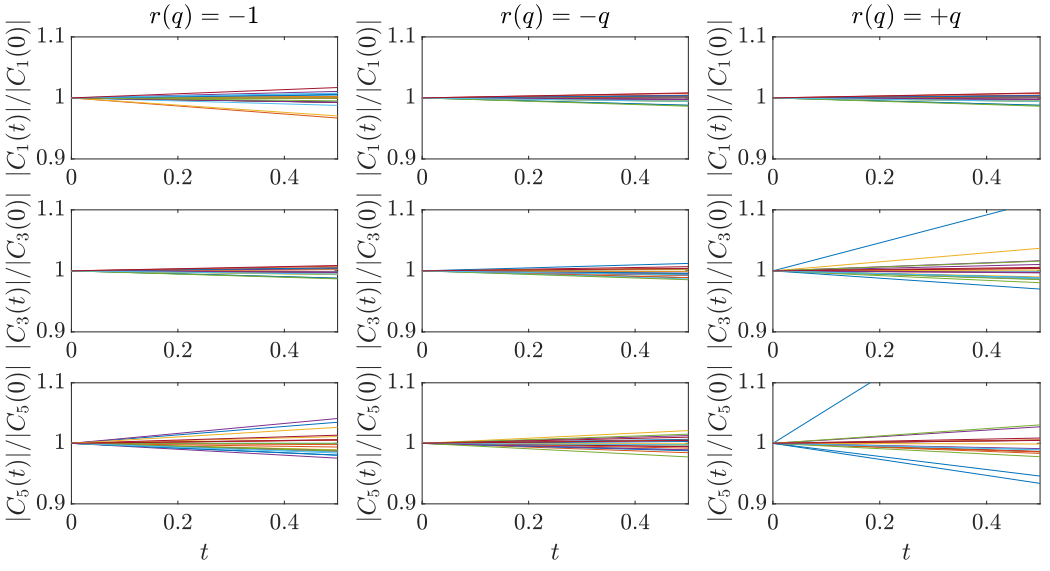
The results of the algorithm are shown in Fig. 2. From Fig. 2(c) we note that only the focusing and defocusing NLSE approximately conserve the first global quantity. We then observe that the fNLSE ($r = -q^*$) and dNLSE ($r = +q^*$) can be distinguished by the fact that C_3 and C_5 are not conserved very well for $r = +q^*$, whereas $r = -q^*$ conserves these much better. From Fig. 2(b), we observe that the guess based on Eq. (10) is indeed reasonably close to the reference value for the scaling constant c_1, despite the path-average approximation and the noise. However, we found that the identified c_1 was structurally lower than the reference value by 3%, even when the data was noiseless. This may indicate that the global quantities become slightly biased with respect to the path-averaged approximation.



(a) Noisy MKdV data, signal 1 of 21.



(b) Identified coefficients, assuming $r(q) = -q$. This solution was found in the 1D-subspace spanned by $\{u\}$.



(c) The global quantities at the identified choice for c, under the assumption of the each $r(q)$.

Fig. 1. Focusing MKdV generated dataset and the results of the identification algorithm. (a) One of the trajectories at first and last measuring time. (b) The initial, identified, and correct coefficients when assuming the relation $r = -q$, i.e., the correct relation for the focusing MKdV. (c) The conserved quantities at the identified coefficients for each of the three considered relations $r(q)$. The choice $r(q) = -q$ shows the least variation and is therefore accepted.

4.3. Case 3: Noisy data from the integrable sine-Gordon data (non-trivial q(u))

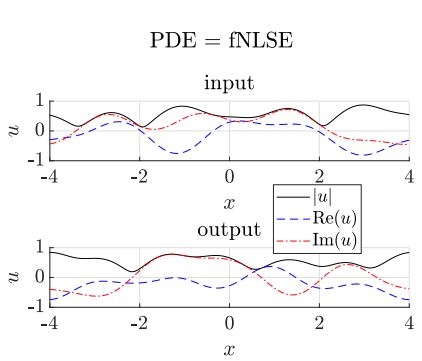
In this subsection, we consider data obtained from the sine-Gordon equation $u_{tx} = \sin(u)$ in light-cone coordinates. This equation has correct choice $q = 0.5u_x$, which is different from the more common relation $q = c_1 u$ used for the KdV and the NLSE. The data set was generated by simulating a kink-antikink collision of the sine-Gordon equation in physical coordinates (α, β), i.e. $u_{\beta\beta} = u_{\alpha\alpha} - \sin(u)$, and then transforming the data through $t = \frac{\alpha+\beta}{2}$, $x = \frac{\alpha-\beta}{2}$. The initial data was taken as a linear summation of a kink- and an antikink solution [46] that were sufficiently far apart to closely approximate an exact double soliton:

$$u(\alpha,0) = \underbrace{4\text{atan}\left(e^{\gamma_1(\alpha-\alpha_1)}\right)}_{\text{kink}} + \underbrace{-4\text{atan}\left(e^{\gamma_2(\alpha-\alpha_2)}\right)}_{\text{anti-kink}}, \tag{13a}$$

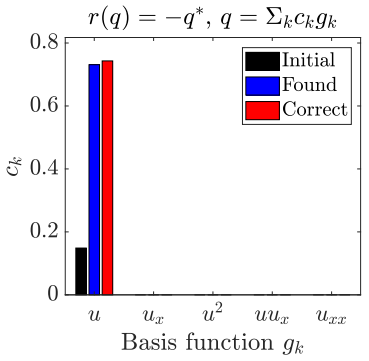
$$u_\beta(\alpha,0) = \underbrace{-2\gamma_1 v_1 \text{sech}\left(-\gamma_1(\alpha-\alpha_1)\right)}_{\text{kink}} + \underbrace{-2\gamma_2 v_2 \text{sech}\left(-\gamma_2(\alpha-\alpha_2)\right)}_{\text{anti-kink}}, \tag{13b}$$

$$\text{with } \left(v_1 = 0.2, \gamma_1 = \sqrt{\tfrac{1}{1-v_1^2}}, \alpha_1 = -2\right) \text{and } \left(v_2 = -0.1, \gamma_2 = \sqrt{\tfrac{1}{1-v_2^2}}, \alpha_2 = 3\right), \tag{13c}$$

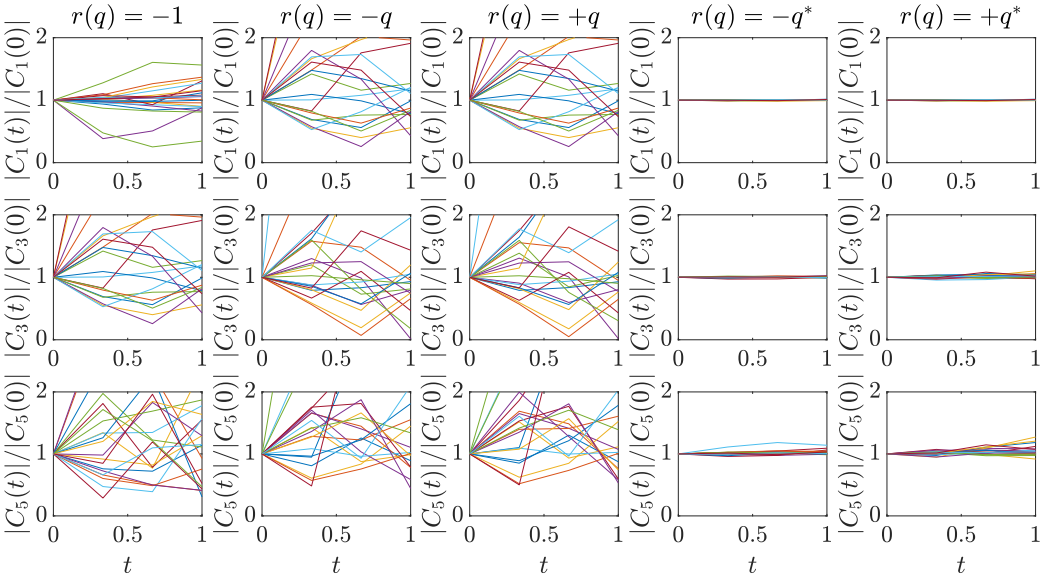
where α_1 (resp. α_2) is the initial positions of the (anti-)kink, and v_1 (resp. v_2) the velocity. The initial signal was then numerically propagated until after the collision, Measurements in physical coordinates were taken in the region $\alpha = -20, -19.96, \ldots, 20$, and



(a) Noisy fNLSE data at $t = 0$ (input) and at $t = 1$ (output), signal 1 of 21. The signals at $t = 1/3$ and at $t = 2/3$ are omitted here.



(b) Identified coefficients, assuming $r(q) = -q^*$. This solution was found in the 1D-subspace spanned by $\{u\}$.



(c) Global quantities at optimal choice for c_1 for various hierarchies $r(q)$.

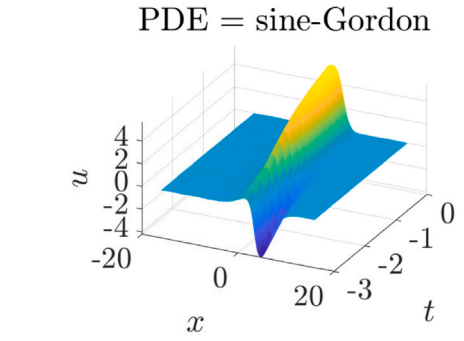
Fig. 2. Focusing NLSE generated dataset, $t \in \{0, 0.33, 0.67, 1\}$, $N = 21$, $x = -4, -3.95, -3.9, \ldots, 4$, signal bandwidth = 1 Hz. (a) One of the 21 signals from the dataset, shown at $t = 0$ and $t = 1$. (b) The initial, identified, and correct coefficients when assuming the relation $r = q^*$, i.e., the correct relation for the fNLSE. (c) The conserved quantities at the identified coefficients for each of the five considered relation $q(r)$. The choice $r(q) = -q^*$ shows the least variation and is therefore accepted.

$\beta = 0, 0.04, \ldots 20$, which were transformed to light-cone coordinates. The resulting light-cone domain becomes diamond-shaped, and we considered only a square domain from it, resulting in the final measurement locations $x = -16, -15.92, -15.84, \ldots, 16$, $t = -3, -2.8, -2.6, \ldots, 0$. Only a single ($N = 1$) trajectory was used, which measured the developing wave nearly continuously, instead of at only a few moments in time. This example thus illustrates that it is also possible to use continuous fields. Finally, we measured the smallest bandwidth containing 99% of the power of the initial condition $u(x, t = -3)$, and added white Gaussian noise, ideally low-pass filtered to this bandwidth, to all space-series measurements leading to SNR = 10^4. The full signal is shown in Fig. 3(a).

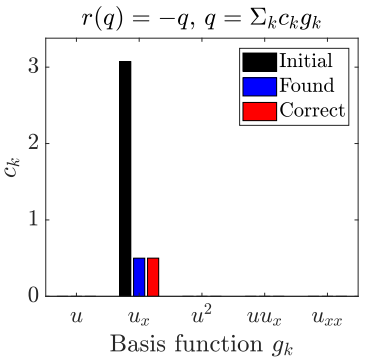
The results are shown in Fig. 3. From Fig. 3(c) we observe that the conserved quantities are best conserved by the choice $r(q) = -q$. The correct choice $q(u) = 0.5u_x$ was indeed identified up to noise as shown in Fig. 3(b).

4.4. Case 4: Noiseless and noisy transformed data from the integrable KdV (more complicated q(u))

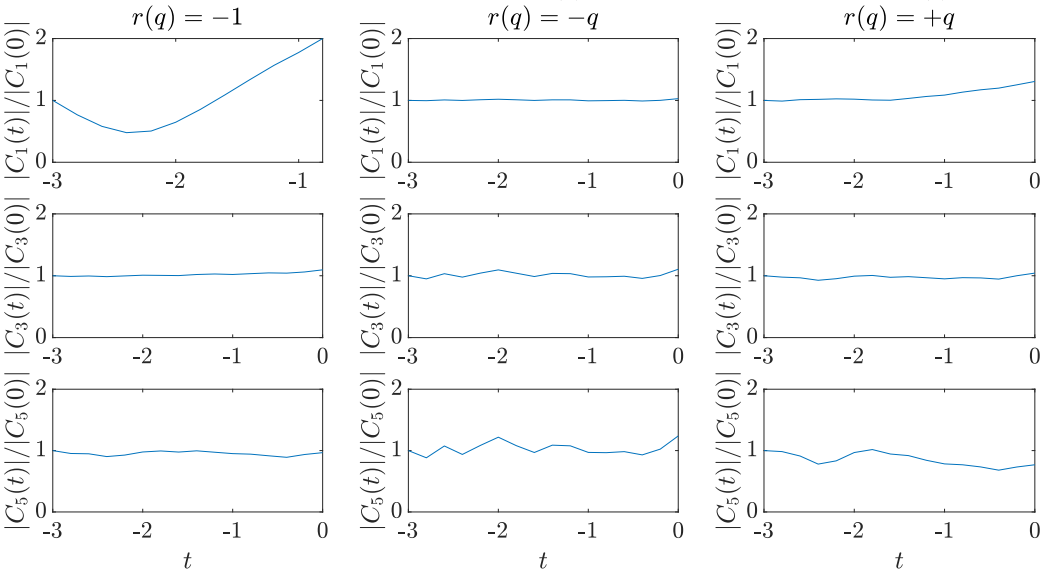
In this subsection, we again consider only two measurement time points, but this time for a transformed KdV. The data was generated by initializing a random, Gaussian distributed random signal with periodic boundary conditions, and ideally low-pass filtering it a maximum wave number of 1 cycles/spatial unit, with $x = -4, -3.95, -3.9, \ldots, 4$. The generated input signal was propagated according to the KdV $v_t = -v_{xxx} - 6vv_x$ for $T = 1$ time unit. Finally, all signals were mapped to the measurement



(a) Kink-anti-Kink collision in light-cone coordinates.



(b) Identified coefficients, assuming $r(q) = -q$



(c) The global quantities, at the optimal $q(u)$ given the shown relation for $r(q)$.

Fig. 3. Noisy Sine–Gordon data. (a) The full dataset consisting of a kink-antikink soliton in light-cone coordinates, shown between $t = -3$ and $t = 0$. (b) The initial, identified, and correct coefficients when assuming the relation $r = q^*$, i.e., the correct relation for the fNLSE. (c) The conserved quantities at the identified coefficients for each of the considered relations $q(r)$.

variable u, given by $u = -\frac{1}{2} + \sqrt{v + \frac{1}{4}} : [-\frac{1}{4}, +\infty) \to [-\frac{1}{2}, +\infty)$, such that the correct mapping back is $v = q(u) = u + u^2$. We carefully respected the domain for v to ensure that the mapping is a bijection on the given domains. We generated $N = 41$ input signals in total, with power $\overline{u^2}$ varying between 1 and 20 after the transformation, to ensure that the data was rich enough. Since we found that the method fails for the SNR used in the previous examples, we first consider the case of no noise. We then investigate other SNRs and discuss a potential explanation.

The results for the noiseless case are shown in Fig. 5. We observe from the conserved quantities in Fig. 5(c) that the conserved quantities do not vary much for all three choices of $r(q)$, but only the KdV conserved them perfectly (due to the lack of noise in the signal). From Fig. 5(b) we observe that a nearly perfect solution was identified: $q^{\mathrm{ID}}(u) = 1.0000u - 0.0009u_x + 0.9999u^2$. Although the term u_x was negligible, it decreased the error by 30% with respect to the reference solution $q^{\mathrm{ref}} = u + u^2$, and was thus accepted over the 2D solution.

Upon redoing the experiment with noise, we found that a near-correct solution was still identified at SNR $= 10^7$. However, when increasing the noise to the level of previous examples, SNR $= 10^4$, we found that the identified solution $c^{\mathrm{ID}} = (9.62, 0, -1.29, 0, 0)$ was completely different from the reference solution $(1, 0, 0, 1, 0)$, although the conserved quantities were still varying very little. Upon inspection, we found that the error at this c was about twice as low as at the reference solution.

We have the hypothesis that the significantly lower error in the identified solution than in the reference solution is due to the different noise amplification properties of the library functions. For example, consider the influence of noise in the library functions $g_1(u) = u$ and $g_3(u) = u^2$. Let $u = u_0 + \varepsilon$ denote the noiseless signal plus noise ε, with $|\varepsilon| \ll |u|$. The relative noise in the library function $u = u_0 + \varepsilon$ is then given by $\frac{u-u_0}{u_0} = \frac{\varepsilon}{u_0}$. However, in the quadratic library function $u^2 = (u_0 + \varepsilon)^2$, the relative noise is given

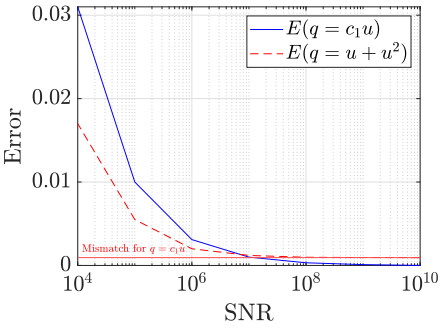


Fig. 4. The error for the reference solution $q = u + u^2$ and for $q = c_1 u$ at optimal c_1, for various levels of the SNR. At high SNR, the reference solution is near-perfect, while the error for $q = c_1 u$ consists only of a model mismatch error. However, at lower SNR, the error in the reference solution grows approximately twice as fast as for $q = c_1 u$.

by $\frac{u^2 - u_0^2}{u_0^2} = \frac{(u_0^2 + 2u_0\epsilon + \cancelto{\sim 0}{\epsilon^2}) - u_0^2}{u_0^2} \approx 2\frac{\epsilon}{u_0}$. We thus observe that the relative noise in library function $g_3(u) = u^2$ is twice as large as in library function $g_1(u) = u$. We thus expect the error to consist of a contribution due to the model mismatch, and a part due to the noise: $E = E_{\text{mismatch}} + E_{\text{noise}}$. For the reference solution, there is no contribution due to model mismatch, but the error due to noise will be approximately twice as large than for incorrect solutions of the form $q = c_1 u$. If the mismatch is small, a wrong q might thus be identified even for quite high SNRs. To test this hypothesis, we compared the errors at the reference solution $q = u + u^2$ and at $u = c_1 u$ (such that c_1 minimized the error) for various noise levels. As is shown in Fig. 4, it indeed turns out that the error contribution due to noise is approximately twice as large in $q = u + u^2$ as in $u = c_1 u$, and that the reference solution leads to a higher error already at SNR= 10^6.

As it should, the proposed method identifies the correct solution if the SNR is large enough. However, it seems likely that the range of sufficiently large SNRs could be expanded by equalizing the impact of noise for the different library functions. For example, one could try to select the library functions for u more carefully. When we removed the library functions u and u^2, and replaced those with the single library function $u + u^2$, the reference solution was identified at SNR = 10^4. We also attempted to replace the library functions u and u^2 with $u + \frac{1}{2}u^2$ and u^2, such that the library still spans the same space as before, but these two library functions are affected in similar amounts by the noise due to their quadratic term. However, the identified solution still converged to approximately $c_1 u = c_1 \left[(u + \frac{1}{2}u^2) - \frac{1}{2}(u^2)\right]$, as the higher noise in the quadratic terms of the two library functions still cancel out for this solution. Nevertheless, there might be better, systematic ways to design the library. Alternatively, one could also try to weight the contributions from the different library functions in the error to compensate for differences in their noise-sensitivity. Choosing these weights however is a non-trivial problem that is beyond the scope of this paper.

4.5. Case 5: Noiseless viscous Burgers' equation with complex viscosity coefficient

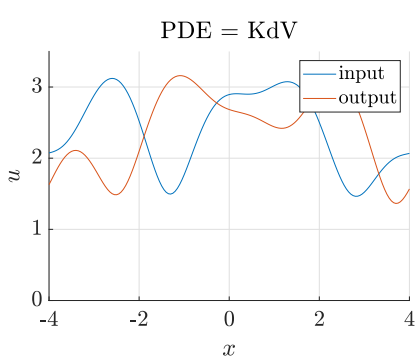
Finally, we would like to illustrate the behavior of our algorithm for an equation that does not seem to fit into our AKNS-ansatz. The equation is a viscous Burgers' equation with a complex viscosity coefficient:

$$u_t + uu_x = \nu u_{xx}, \quad \nu = -i. \tag{14}$$

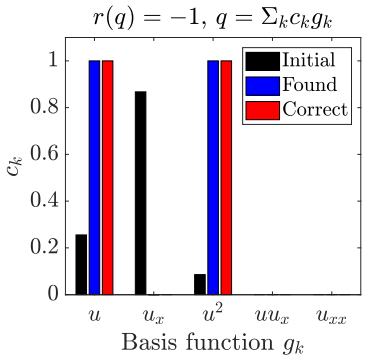
A single realization of the signal is shown in Fig. 6(a), which was propagated from $t = 0$ to $t = 1$. In total, we generated 41 sets of input–output signals. No noise was added to the signals. As we apply our algorithm on this data set, we observe in Fig. 6(b) that none of the considered choices for $r(q)$ result in well-conserved quantities. This suggests that this dataset may not fit any of the considered integrable AKNS-type PDEs.

However, the poor conservation of the quantities (5) can also arise for other reasons. It could be due to noise, or because the considered space of functions for q and r is not large enough. It could also arise when the starting points in the space for $q(u)$ were not close enough to the correct solution. A large error alone therefore does not necessarily imply that the data does not originate from an AKNS-type system (or a system close to one) in the considered search space, and further investigations are required to eliminate the alternative causes discussed above.

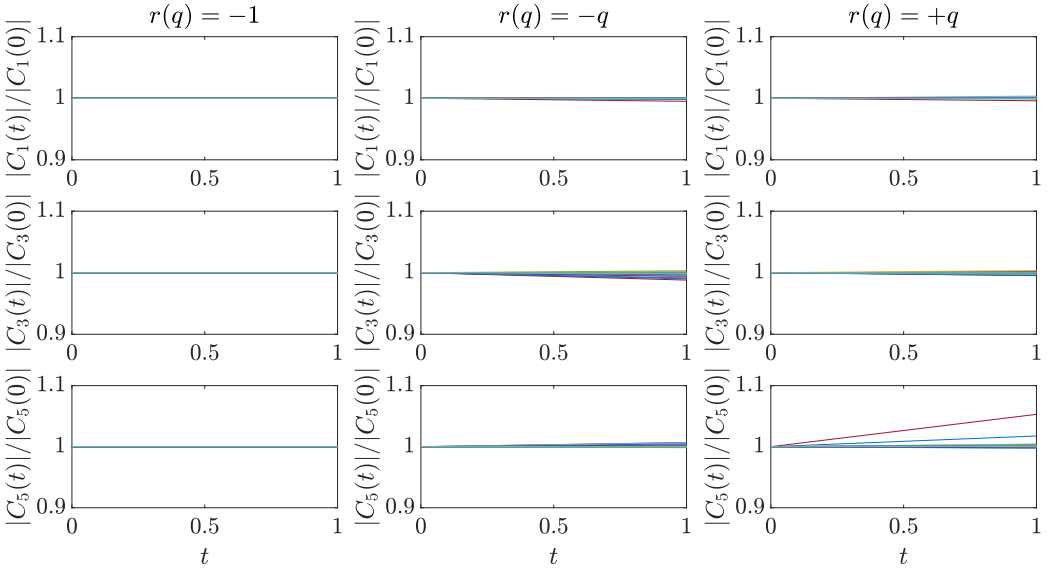
We finally remark that on the other hand, even a perfect conservation of the quantities (5) does not guarantee that the system is from the considered class of AKNS systems because only a finite number of conserved quantities is considered for the computation of the error. In such cases, the identified $r(q)$ and $q(u)$ should be further tested w.r.t. other criteria. For example, one could check numerically if the spectrum of the identified L-operator stays constant during propagation. Another possibility would be to determine the PDE that corresponds to the identified Lax pair, and use it to numerically propagate the inputs in the data set. The result should match the corresponding outputs in the data set.



(a) Noiseless transformed KdV data, signal 1 of 21.



(b) Identified coefficients, assuming $r(q) = -1$



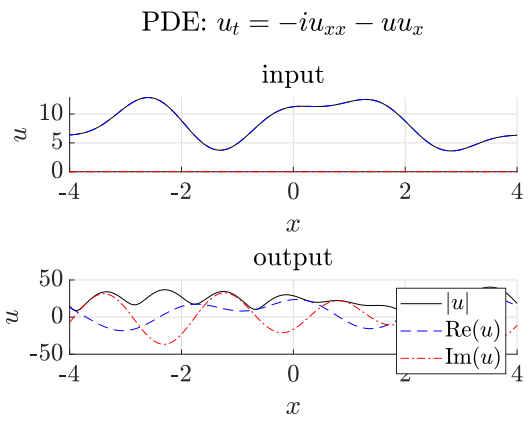
(c) Global quantities at optimal choice for c_1 for various hierarchies $r(q)$.

Fig. 5. Noiseless KdV data $v(x,t)$, transformed as $u = -\frac{1}{2} + \sqrt{v + \frac{1}{4}}$. The signal domains and parameters were $t \in \{0, 1\}$, $N = 21$, $x = -4, -3.95, -3.9, \ldots, 4$, signal bandwidth = 1 Hz. (a) One consisting of a kink-antikink soliton in light-cone coordinates, shown between $t = -3$ and $t = 0$. (b) The initial, identified, and correct coefficients when assuming the relation $r = q^*$, i.e., the correct relation for the fNLSE. (c) The conserved quantities at the identified coefficients for each of the considered relations $q(r)$.

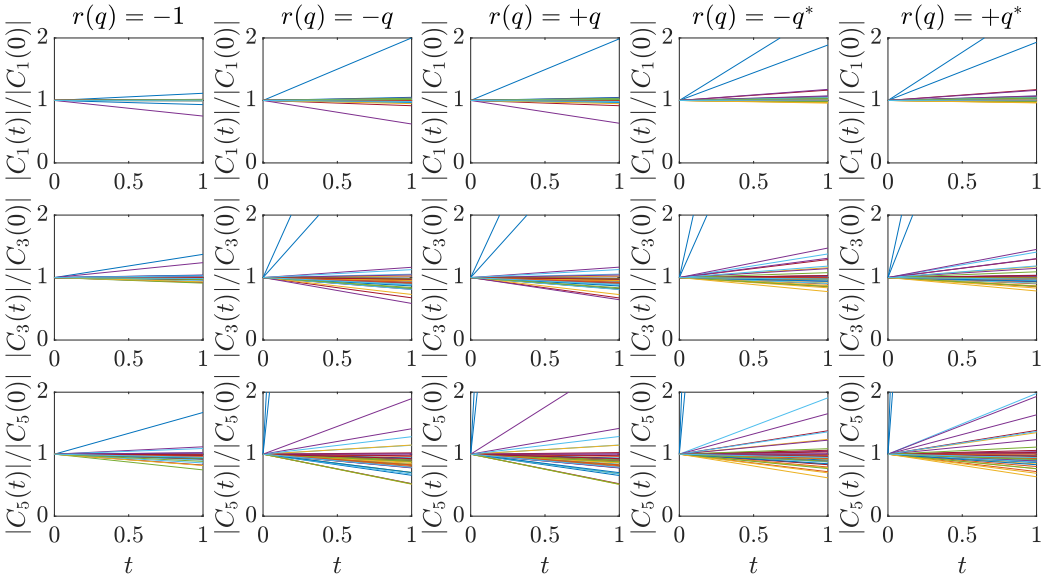
5. Conclusion

We proposed and demonstrated an automatic, data-driven identification method for the spectral operator of AKNS-type Lax pairs. Given measured data from an unknown, not necessarily Lax-integrable PDE, it finds the spectral operator L of the Lax pair that explains the given data best by matching conserved quantities. To the best of our knowledge, it is the first such method. It enables non-specialists to discover Lax-integrable systems "in the wild", and paves the way to exploiting the strong theoretical properties of Lax-integrable PDEs in new application areas. In contrast to conventional PDE identification methods, the measurements can be taken at time points that are far apart (assuming time is the evolution variable), which is an important practical advantage. Only two different time points are required. We focused on finding the L operator of the Lax pair in this paper because together with the easily measurable linear dispersion relation of the underlying PDE, it already completely specifies the Lax-integrable PDE that belongs to the Lax pair, and therefore also the propagation operator A, which is the second part of the Lax pair. The spectral operator is furthermore already sufficient to perform nonlinear Fourier analysis.

The method was demonstrated on noisy measurements from the MKdV, the NLSE and the sine-Gordon equation, as well as on noise-free and noisy data from a transformed KdV. By choosing initial guesses for the coefficient vectors that evoke different parts of the conservation laws with similar strength, and then performing searches in low-dimensional spaces of k-sparse vectors, we were able to identify the correct spectral operator for a variety of (nearly) Lax-integrable systems. We found that the algorithm generally works well if the amount of noise in the data is less than 1% of the signal amplitude, although one case required less noise. We remark that PDE identification methods in general require relatively high signal-to-noise ratios so that e.g. higher-order derivatives of the data can be computed. As a next step, the method should be investigated with real-world data. If the noises in



(a) A single realization of an input ($t = 0$) and output ($t = 1$) signal.



(b) Global quantities at the optimal $q(u)$ for various hierarchies $r(q)$.

Fig. 6. Results for the viscous Burgers equation $u_t + uu_x = \nu u_{xx}$ with complex viscosity coefficient $\nu = -i$.

such scenarios turn out to be too strong, approaches such as filtering and/or weak formulations should be incorporated to reduce the impact of the noise on the identification process. Additionally, as discussed in the last example, methods for taking the different noise amplification properties of the library functions into account should be investigated.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Pascal de Koster: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Sander Wahls:** Investigation, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] P.D. Lax, Integrals of nonlinear equations of evolution and solitary waves, Commun. Pure Appl. Math. 21 (5) (1968) 467–490.
[2] M.J. Ablowitz, D.J. Kaup, A.C. Newell, H. Segur, The inverse scattering transform–Fourier analysis for nonlinear problems, Stud. Appl. Math. 53 (4) (1974) 249–315.
[3] C.S. Gardner, J.M. Greene, M.D. Kruskal, R.M. Miura, Method for solving the Korteweg-de Vries equation, Phys. Rev. Lett. 19 (19) (1967) 1095.
[4] M. Wadati, The modified Korteweg-de Vries equation, J. Phys. Soc. Japan 34 (5) (1973) 1289–1296.
[5] V. Zakharov, A. Shabat, Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media, Sov. Phys. JETP 34 (1) (1972) 62.
[6] M.J. Ablowitz, H. Segur, Solitons and the Inverse Scattering Transform, SIAM, 1981.
[7] A.R. Osborne, Nonlinear Ocean Waves and the Inverse Scattering Transform, in: International Geophysics, vol. 97, Academic Press, 2010.
[8] S. Trillo, G. Deng, G. Biondini, M. Klein, G. Clauss, A. Chabchoub, M. Onorato, Experimental observation and theoretical description of multisoliton fission in shallow water, Phys. Rev. Lett. 117 (14) (2016) 144102.
[9] M. Brühl, H. Oumeraci, Analysis of long-period cosine-wave dispersion in very shallow water using nonlinear Fourier transform based on KdV equation, Appl. Ocean Res. 61 (2016) 81–91.
[10] S. Randoux, P. Suret, G. El, Inverse scattering transform analysis of rogue waves using local periodization procedure, Sci. Rep. 6 (1) (2016) 1–11.
[11] I. Teutsch, M. Brühl, R. Weisse, S. Wahls, Contribution of solitons to enhanced rogue wave occurrence in shallow depths: a case study in the southern North sea, Nat. Hazards Earth Syst. Sci. 23 (6) (2023) 2053–2073.
[12] M. Brühl, P.J. Prins, S. Ujvary, I. Barranco, S. Wahls, P.L.-F. Liu, Comparative analysis of bore propagation over long distances using conventional linear and KdV-based nonlinear Fourier transform, Wave Motion 111 (2022) 102905.
[13] A. Hasegawa, T. Nyu, Eigenvalue communication, J. Lightw. Technol. 11 (3) (1993) 395–399.
[14] J.E. Prilepsky, S.A. Derevyanko, K.J. Blow, I. Gabitov, S.K. Turitsyn, Nonlinear inverse synthesis and eigenvalue division multiplexing in optical fiber channels, Phys. Rev. Lett. 113 (1) (2014) 013901.
[15] M.I. Yousefi, F.R. Kschischang, Information transmission using the nonlinear Fourier transform, part I–III, IEEE Trans. Inform. Theory 60 (7) (2014) 4312–4369.
[16] Z. Dong, S. Hari, T. Gui, K. Zhong, M.I. Yousefi, C. Lu, P.-K.A. Wai, F.R. Kschischang, A.P.T. Lau, Nonlinear frequency division multiplexed transmissions based on NFT, IEEE Photonics Technol. Lett. 27 (15) (2015) 1621–1623.
[17] H. Buelow, V. Aref, W. Idler, Transmission of waveforms determined by 7 eigenvalues with PSK-modulated spectral amplitudes, in: ECOC 2016; 42nd European Conference on Optical Communication, VDE, 2016, pp. 1–3.
[18] S.T. Le, V. Aref, H. Buelow, Nonlinear signal multiplexing for communication beyond the Kerr nonlinearity limit, Nat. Photon. 11 (9) (2017) 570–576.
[19] S. Wahls, Generation of time-limited signals in the nonlinear Fourier domain via b-modulation, in: 2017 European Conference on Optical Communication, ECOC, IEEE, 2017, pp. 1–3.
[20] S.K. Turitsyn, J.E. Prilepsky, S.T. Le, S. Wahls, L.L. Frumin, M. Kamalian, S.A. Derevyanko, Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives, Optica 4 (3) (2017) 307–322.
[21] T. Gui, G. Zhou, C. Lu, A.P.T. Lau, S. Wahls, Nonlinear frequency division multiplexing with b-modulation: shifting the energy barrier, Opt. Express 26 (21) (2018) 27978–27990.
[22] X. Yangzhang, V. Aref, S.T. Le, H. Buelow, D. Lavery, P. Bayvel, Dual-polarization non-linear frequency-division multiplexed transmission with *b*-modulation, J. Lightwave Technol. 37 (6) (2019) 1570–1578.
[23] P. de Koster, M. Brühl, S. Wahls, Water-depth identification from free-surface data using the KdV-based nonlinear Fourier transform, in: International Conference on Offshore Mechanics and Arctic Engineering, Vol. 85901, American Society of Mechanical Engineers, 2022, V05BT06A056.
[24] P. De Koster, S. Wahls, Dispersion and nonlinearity identification for single-mode fibers using the nonlinear Fourier transform, J. Lightwave Technol. 38 (12) (2020) 3252–3260.
[25] P. De Koster, O. Schulz, J. Koch, S. Pachnicke, S. Wahls, Fast single-mode fiber nonlinearity monitoring: An experimental comparison between split-step and nonlinear Fourier transform-based methods, IEEE Photonics J. (2023).
[26] S.L. Brunton, J.L. Proctor, J.N. Kutz, Discovering governing equations from data by sparse identification of nonlinear dynamical systems, Proc. Natl. Acad. Sci. 113 (15) (2016) 3932–3937.
[27] S. Krippendorf, D. Lüst, M. Syvaeri, Integrability ex machina, Fortschr. Phys. 69 (7) (2021) 2100057.
[28] P. De Koster, J. Koch, O. Schulz, S. Pachnicke, S. Wahls, Experimental validation of nonlinear Fourier transform-based Kerr-nonlinearity identification over a 1600 km SSMF link, in: Optical Fiber Communication Conference, Optica Publishing Group, 2022, pp. W2A–39.
[29] S. Lin, Y. Chen, A two-stage physics-informed neural network method based on conserved quantities and applications in localized wave solutions, J. Comput. Phys. 457 (2022) 111053.
[30] H.D. Wahlquist, F.B. Estabrook, Prolongation structures of nonlinear evolution equations, J. Math. Phys. 16 (1) (1975) 1–7.
[31] N. Asano, Y. Kato, Spectrum method for a general evolution equation, Progr. Theoret. Phys. 58 (1) (1977) 161–174.
[32] H. Chen, Y. Lee, C. Liu, Integrability of nonlinear Hamiltonian systems by inverse scattering method, Phys. Scr. 20 (3–4) (1979) 490.
[33] R. Dodd, A. Fordy, The prolongation structures of quasi-polynomial flows, Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 385 (1789) (1983) 389–429.
[34] M. Ito, A REDUCE program for evaluating a Lax pair form, Comput. Phys. Commun. 34 (3) (1985) 325–331.
[35] M. Musette, R. Conte, Algorithmic method for deriving Lax pairs from the invariant Painlevé analysis of nonlinear partial differential equations, J. Math. Phys. 32 (6) (1991) 1450–1457.
[36] M. Hickman, W. Hereman, J. Larue, Ü. Göktaş, Scaling invariant Lax pairs of nonlinear evolution equations, Appl. Anal. 91 (2) (2012) 381–402.
[37] I. Habibullin, A. Khakimova, M. Poptsova, On a method for constructing the Lax pairs for nonlinear integrable equations, J. Phys. A 49 (3) (2015) 035202.
[38] I.T. Habibullin, A.R. Khakimova, A direct algorithm for constructing recursion operators and Lax pairs for integrable models, Theoret. Math. Phys. 196 (2) (2018) 1200–1216.
[39] P. Scholl, A. Bacho, H. Boche, G. Kutyniok, Well-definedness of physical law learning: The uniqueness problem, 2022, arXiv preprint arXiv:2210.08342.
[40] S.H. Rudy, S.L. Brunton, J.L. Proctor, J.N. Kutz, Data-driven discovery of partial differential equations, Sci. Adv. 3 (4) (2017) e1602614.
[41] M. Syväri, Symmetries in String Theory: Application of Machine Learning to Theoretical Physics (Doctoral thesis), Ludwig-Maximilians-Universität München, Germany, 2021, URL http://nbn-resolving.de/urn:nbn:de:bvb:19-287407.
[42] The MathWorks Inc., System identification toolbox, 2024, URL https://www.mathworks.com/products/sysid.html. Online.
[43] F. Calogero, M. Nucci, Lax pairs galore, J. Math. Phys. 32 (1) (1991) 72–74.
[44] G.P. Agrawal, Nonlinear fiber optics, Academic Press, 2013.
[45] G.P. Agrawal, Fiber-Optic Communication Systems, vol. 222, John Wiley & Sons, 2012.
[46] J. Cuevas-Maraver, P.G. Kevrekidis, F. Williams, The sine-Gordon model and its applications, Nonlinear Syst. Complexity 10 (2014).