Communications in Nonlinear Science and Numerical Simulation 125 (2023) 107311

Contents lists available at ScienceDirect

Communications in Nonlinear Science and Numerical Simulation

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

Research paper

Phase computation for the finite-genus solutions to the focusing nonlinear Schrödinger equation using convolutional neural networks

Stepan Bogdanov [a,*], Dmitry Shepelsky [b], Anastasiia Vasylchenkova [c], Egor Sedov [a], Pedro J. Freire [a], Sergei K. Turitsyn [a], Jaroslaw E. Prilepsky [a,*]

[a] Aston Institute of Photonic Technologies, Aston University, Birmingham B4 7ET, UK
[b] B. Verkin Institute for Low Temperature Physics and Engineering, 61103 Kharkiv, Ukraine
[c] Optical Networks Group, University College London, London WC1E 7JE, UK

A R T I C L E I N F O

Article history:
Received 8 January 2023
Received in revised form 21 April 2023
Accepted 16 May 2023
Available online 22 May 2023

Keywords:
Nonlinear Schrődinger equation
Finite-genus solutions
Nonlinear Fourier transform
Riemann–Hilbert problem
Convolutional neural networks

A B S T R A C T

We develop a method for retrieving a set of parameters of a quasi-periodic finite-genus (finite-gap) solution to the focusing nonlinear Schrödinger (NLS) equation, given the solution evaluated on a finite spatial interval for a fixed time. These parameters (named "phases") enter the jump matrices in the Riemann-Hilbert (RH) problem representation of finite-genus solutions. First, we detail the existing theory for retrieving the phases for periodic finite-genus solutions. Then, we introduce our method applicable to the quasi-periodic solutions. The method is based on utilizing convolutional neural networks optimized by means of the Bayesian optimization technique to identify the best set of network hyperparameters. To train the neural network, we use the discrete datasets obtained in an inverse manner: for a fixed main spectrum (the endpoints of arcs constituting the contour for the associated RH problem) and a random set of modal phases, we generate the corresponding discretized profile in space via the solution of the RH problem, and these resulting pairs – the phase set and the corresponding discretized solution in a finite interval of space domain – are then employed in training. The method's functionality is then verified on an independent dataset, demonstrating our method's satisfactory performance and generalization ability.

1. Introduction

The Inverse Scattering Transform (IST) method, introduced about half a century ago, allows the solution of certain classes of nonlinear partial differential equations (PDEs) (the so-called *integrable* equations), by performing the analysis of the corresponding linear problems [1–6]. The IST method consists of three steps: (i) solving a linear auxiliary set of equations and defining a spectral (direct) problem that maps the given initial data (say, the initial conditions) onto a set of special quantities (spectral data); (ii) establishing the evolution of these spectral characteristics; and (iii) solving an inverse problem, which allows retrieving the solution of the PDE in question at a particular value of the evolution (time or propagation distance) variable. Here, the direct and inverse problems refer to a linear operator from the so-called Lax pair [3] of linear, ordinary differential equations, whose compatibility condition is equivalent to the PDE in question. In the

* Corresponding authors.
E-mail addresses: 190244042@aston.ac.uk (S. Bogdanov), y.prylepskiy1@aston.ac.uk (J.E. Prilepsky).

https://doi.org/10.1016/j.cnsns.2023.107311

cases of probably the most famous integrable PDEs, the Korteweg–de Vries equation [1], and the focusing/defocusing NLS equation [2], the associated linear equations are the Sturm–Liouville equation (or one-dimensional linear Schrödinger equation) and the Dirac/Zakharov–Shabat systems of two coupled equations [1,2]. Note that the direct and inverse problems entering the IST are often called nonlinear Fourier transform (NFT) in the optical communications literature [7,8]. For the initial value problems, where the data and the solution are assumed to be vanishing sufficiently fast as the spatial (non-evolution) variable approach infinities (for which the IST method was originally developed), the direct spectral problems take the form of scattering problems and the associated spectral characteristics establish the relation of the different solutions to the linear equations from the Lax pair [9].

As for the inverse problems, the IST method in the original formulation uses the Gelfand–Levitan–Marchenko integral equations [9]. An alternative approach to the inverse problem is to consider a factorization problem of the Riemann-Hilbert (RH) type, formulated in the complex plane of the spectral parameter [10]. The spatial and temporal variables enter the RH problem as parameters in an extremely simple way. This allows (loosely) viewing the (matrix-valued) RH problems as an analog to integral representations for special functions of mathematical physics viewed as the solutions of respective linear differential equations (e.g. Airy equation, Bessel equation, Hypergeometric equations, etc.). Consequently, as it has been realized since the 90th of the last century, the RH problem representations can be efficiently used for studying various properties of solutions of the respective PDEs, particularly various asymptotic regimes [11].

On the other hand, for problems on the circle (corresponding to problems on a finite spatial interval with periodic boundary conditions), it is the introduction of the so-called finite-genus (or finite-gap) integration methodology in the 1970s that had groundbreaking implications. It is particularly possible to generate large classes of exact solutions, the so-called finite-gap solutions, obtained from solving a Jacobi inversion problem on a finite-genus Riemann surface [12,13]. These solutions can be given "explicitly", in terms of associated theta functions. Here we put "explicitly" in quotes since, from a computational point of view, these functions are not easily accessible, even though some fast algorithms for their computation have been proposed [14,15].

Notice that not all periodic solutions are of a finite-gap type: the corresponding spectral representation may consist of an infinite number of spectral bands. In turn, not all finite-gap solutions are periodic: their construction may involve a combination of periodic functions with incommensurable frequencies. On the other hand, recent progress in the adaptation of the RH approach to initial boundary value problems for integrable PDEs (on a half-line or an interval), associated with the development of the Unified Transform Method (a.k.a. the Fokas Method [16]), suggests a similar approach for linking the IST methods to the periodic (and/or finite-gap) problems. In the recent papers [17,18], it was shown (in the case of NLS equation) that the solution of the initial boundary value problem on a finite interval with x-periodic boundary conditions can be formulated in terms of the solution of an associated RH problem, where the relevant jump matrices can be expressed explicitly in terms of the spectral (actually, scattering) data, which are computed via the initial data.

When discussing the applications of the IST methods to real-world problems, a crucially important issue is how efficiently the proposed method can be implemented numerically. The practicality of approaches based on the RH problem formalism is supported by recent progress in developing an efficient numerical framework for approximating the solutions to matrix-valued RH problems [10,19–21].

The theory of finite-genus solutions to the NLS equation has been intensively used in the water (ocean) waves analysis [22,23]. At the same time, over recent years, there has been a growing interest in the development of optical communication methods based on the utilization of non-decaying (periodic) solutions to NLS equation [24–30] as a more efficient alternative to the soliton-based communications [31]. The IST operations associated with periodic finite-genus NLS solutions were named periodic nonlinear Fourier transform (PNFT) in these works. In Refs. [28,29], an optical signal modulation and digital signal processing (DSP) method has been proposed for a PNFT-based transmission, where the inverse problem step (constructing a signal in the physical domain, at the transmitter side, given spectral parameters on which the transmitted information is encoded) harnesses the numerical solution of an RH problem. The data for this RH problem are 2×2 jump matrices, which are off-diagonal matrices satisfying a certain symmetry condition, with constant entries on each separated part of the jump contour consisting of a finite number of arcs (see Appendix). The solution of such an RH problem gives rise to a finite-gap (finite-genus) solution to the NLS equation [13,32]: the endpoints of the arcs fix the associated Riemann surface, while the constants in the jump matrices specify an individual solution. Here, the NLS evolution is translated, using spectral terms, into a linear evolution of the constants (naturally understood as real-valued *phases* varying from 0 to 2π) in the jumps matrices, with the individual frequencies uniquely determined by the *main spectrum* (the ends of the spectral arcs). Finally, the decoding procedure at the receiver side consists in solving the direct problem, mapping the received signal as (periodic) function to the (finite-dimensional) space of *phases*.

In [28,29], restricting to the case of solutions which are both periodic and of a finite-genus type, calculations for the direct problem were presented based on the appropriate RH problem deformation found by the authors. However, the proposed approach faced challenges decreasing the efficiency of the method: (i) the *phases* were determined modulo π (au lieu of theoretically possible 2π); (ii) the procedure required not only the calculation of the *auxiliary spectrum* (the eigenvalues of the associated Dirichlet boundary value problem for the Zakharov–Shabat operator), but also a correct specification of those eigenvalues lying on an appropriate sheet of the (two-sheet) Riemann surface. One more drawback of the (current state of) implementation of the direct part of the RH formalism is that it can be applied to the periodic problem only, thus requiring additional efforts (particularly for numerical implementations) to ensure the commensurability of frequencies. However, from the optical communications viewpoint, the latter rather looks like an artificial complication of the signal encoding stage, resulting from a technical but not fundamental limitation. Therefore, the new approaches where the periodicity requirement can be circumvented are particularly interesting from the optical communications perspective.

In the end, we notice that artificial neural networks (NNs) have been widely implemented to solve various tasks related to inverse scattering [33,34]. Turning closer to the subject of the current paper, we first mention Refs. [35,36], where the NN-based methods were used for the demodulation (identification of nonlinear parameters) in optical (NLS) solitonic systems. Ref. [37] proposes a receiver based on regression NNs that can demodulate information for both single- and dual-polarization NLS systems. In Ref. [38], the authors propose a demodulation method for eigenvalue-modulated signals using an eigenvalue-domain NN, and demonstrate its effectiveness through simulation and experimental results. Recent Ref. [39] presents the development of two convolutional neural network (CNN)-based approaches with different complexity to compute the continuous nonlinear spectrum associated with decaying signals. Based on Wavenet architecture [40], the special CNN structure called NFT-Net was developed in [41]: it allows determining the continuous nonlinear spectrum associated with a decaying (actually finite-extent) profile. We will use this NFT-Net convolutional NN structure from [41] as a starting point for our study below.

In the current paper, we describe a method for using a NN to extract the "phases" associated with the arcs of an RH problem from the finite-genus NLS solution waveform. To do this, we designed a special NN architecture and trained it using a large dataset of waveforms and corresponding phase information. Determining the optimal hyperparameters for a NN can be a complex and time-consuming task, as it involves multidimensional optimization. The NN architecture presented in this article has a few dozen hyperparameters, which makes grid-based optimization methods computationally expensive and inefficient. As an alternative, we used the Bayesian optimization method, which is effective in tasks with a high number of parameters [42]. To find the best architecture for the NN, we used Bayesian optimization starting with NFT-Net [41] as an initial value for NN's hyperparameters.

2. Direct spectral problem for finite-genus NLS solutions parametrized via RH problem

As mentioned in the introduction, a wide variety of solutions of an integrable nonlinear evolution equation can be constructed in terms of a solution to a family of RH problems (parametrized by x and t) whose data depend on x and t in a way specific to an integrable nonlinear equation in question. Specifically, the so-called finite-genus solutions of the focusing NLS equation,

$$iq_t + q_{xx} + 2|q|^2 q = 0, \tag{1}$$

(the NLS equation is written for a dependent field variable q and time/evolution (t) and space (x) independent variables), can be characterized as follows [13,32].

1. They are specified by two sets of numbers: (i) complex-valued $\{\lambda_j\}_0^N$ for some integer N, with $\operatorname{Im}\lambda_j > 0$; and (ii) $\{\phi_j\}_0^N$ with real $\phi_j \in [0, 2\pi)$. Note that ϕ_j, the phases, are exactly the parameters, the computation of which is addressed in our study here.
2. They can be constructed in terms of the solution of a family of the RH problems, formulated in the complex λ-plane and parametrized by x and t, the data for which are given in terms of sets $\{\lambda_j\}_0^N$ and $\{\phi_j\}_0^N$.

The associated RH problem and the relations relating the RH problem solution to finding the finite-genus solutions to NLS equation are detailed in Appendix for the completeness of our exposition.

2.1. Direct problem in the periodic case

The direct problem associated with the inverse problem presented in Appendix consists of the following: given a N-genus solution $q(x, t)$ of the NLS equation, associated with a prescribed Riemann surface (prescribed branch points λ_j, $j = 0, 1, \ldots, N$) and evaluated as a function of x at some fixed $t = t^*$, determine the underlying phase parameters ϕ_j. In our setting, we start from the inverse problem (given specific spectral characteristics, find the underlying solution of the NLS equation), and so the direct problem reproducing these spectral characteristics given an appropriate solution of the NLS equation at a fixed evolutionally variable t can be viewed as "an inverse problem to the inverse problem".

In the case where C_j^f together with f_0 are commensurable (see Eqs. (A.7)–(A.16) for the definitions of respective parameters), and thus, the underlying solution of the NLS equation is periodic in x, a possible way to solve the direct problem is based on the idea of finding a RH representation for the solution of the initial boundary value problem (IBVP) for the NLS equation, where the initial data given for x varying on an interval (of the periodicity length L), i.e., $q(x, 0) = q_0(x)$ for $x \in (0, L)$, are supplemented by the periodicity conditions:

$$q(0, t) = q(L, t), \quad q_x(0, t) = q_x(L, t) \quad \text{for all } t \geq 0. \tag{2}$$

If the RH problem in this representation has the same structure as the original RH problem, i.e., it is described by RH problem conditions 2 and 3 from Appendix, then the constants ϕ_j appearing in the jump construction will give the sought solution of our direct problem.

To get the appropriate representation, one can proceed in two steps: (i) first, provide *some* RH representation (with *some* contour and jumps), where the data for the RH problem can be constructed from the data of the IBVP, i.e., the initial data $q_0(x)$ for $x \in (0, L)$; (ii) second, using the flexibility of the RH representation for the solution of nonlinear equations, transform this (original) RH problem to that having the above-mentioned *desired form*.

The first step has been recently addressed in [17,18], where it was shown that in the case of focusing NLS equation, the solution of the periodic IBVP can be given in terms of the solution of a RH problem, where (i) the contour is the union of $\Gamma = \cup_{j=0}^{N} \Gamma_j$ and the real and imaginary axes, and (ii) the jump matrices can be constructed in terms of the entries $a(\lambda)$ and $b(\lambda)$ of the scattering matrix

$$s(\lambda) = \begin{pmatrix} a^*(\lambda^*) & b(\lambda) \\ -b^*(\lambda^*) & a(\lambda) \end{pmatrix}$$

associated with the Zakharov–Shabat spectral problem (the x-equation of the Lax pair for the NLS equation):

$$\Phi_x = U\Phi, \tag{3}$$

with

$$U = -i\lambda\sigma_3 + \begin{pmatrix} 0 & q \\ -q^* & 0 \end{pmatrix}$$

considered on the whole line, with the potential $q = q(x, 0)$ involved in U being continued on the whole line (outside $(0, L)$) by setting it to 0. In turn, this step can be performed in two sub-steps. In sub-step 1, a RH problem is constructed using the spectral functions $a(\lambda)$ and $b(\lambda)$ supplemented by the spectral functions $A(\lambda)$ and $B(\lambda)$ that enter the scattering matrix:

$$S(\lambda) = \begin{pmatrix} A^*(\lambda^*) & B(\lambda) \\ -B^*(\lambda^*) & A(\lambda) \end{pmatrix},$$

associated with the t-equation from the Lax pair,

$$\Phi_t = V\Phi, \tag{4}$$

with $V(t) = -2i\lambda^2\sigma_3 + 2\lambda V_1(t) + V_2(t)$, where

$$V_1(t) = \begin{pmatrix} 0 & q(0,t) \\ -q^*(0,t) & 0 \end{pmatrix}, \quad V_2(t) = -i\left(V_1^2(t) + \begin{pmatrix} 0 & q_x(0,t) \\ -q_x^*(0,t) & 0 \end{pmatrix}\right)\sigma_3.$$

Namely, assuming for a moment that $q(0, t)$ and $q_x(0, t)$ are given for $t \in (0, T)$ with some $T > 0$, Eq. (4) can be considered, similarly to (3), as a spectral problem for a matrix equation with coefficients determined in terms of $q(0, t)$ and $q_x(0, t)$. Since V in Eq. (4) is a polynomial of the second order w.r.t. λ, it follows that the contour where the scattering relation is established consists of two lines, the real and imaginary axes (where $\operatorname{Im}\lambda^2 = 0$).

An obvious drawback of this construction (of the RH problem) is that neither $q(0, t)$ nor $q_x(0, t)$ are given as the data for the IBVP. Then sub-step 2 addresses the problem of replacing the RH problem constructed in terms of $a(\lambda)$, $b(\lambda)$, $A(\lambda)$, and $B(\lambda)$ by an equivalent one (in the sense that $q(x, t)$ obtained following (A.4) from the both problems are the same), whose formulation involves $a(\lambda)$ and $b(\lambda)$ only. A key for performing this sub-step is the so-called "global relation" [16,17,43], which is a relation amongst $a(\lambda)$, $b(\lambda)$, $A(\lambda)$, and $B(\lambda)$ reflecting the fact that the IBVP with periodic boundary conditions is well-posed (particularly, has a unique solution) without prescribing the boundary values $q(0, t)$ and $q_x(0, t)$.

In the current setting (i.e., for the periodic problem in x), the global relation takes the form of the equation:

$$e^{2i\lambda L}\left(A(\lambda)a^*(\lambda) + B(\lambda)b^*(\lambda)\right)B(\lambda) + (A(\lambda)b(\lambda) - a(\lambda)B(\lambda))\,A(\lambda) = \tag{5}$$

$$= e^{4i\lambda^2 T} O\left(\frac{1 + e^{2i\lambda L}}{\lambda}\right), \tag{6}$$

where the r.h.s. is not given precisely but only asymptotically, as $\lambda \to \infty$. Noticing that the r.h.s. in (5) approaches 0 as $\lambda \to \infty$ staying in the first quadrant of the complex λ-plane suggests replacing the r.h.s. by zero, which leads to a quadratic equation for the ratio $B(\lambda)/A(\lambda)$, with the coefficients given in terms of $a(\lambda)$ and $b(\lambda)$. Define $R(\lambda)$ as the solution of this equation:

$$R(\lambda) = \frac{e^{-i\lambda L}a(\lambda) - e^{i\lambda L}a^*(\lambda^*) + \sqrt{(e^{-i\lambda L}a(\lambda) - e^{i\lambda L}a^*(\lambda^*))^2 - 4b^*(\lambda^*)b(\lambda)}}{2e^{i\lambda L}b^*(\lambda^*)}, \tag{7}$$

where the branch of the square root is chosen such that the branch cuts are the arcs connecting the pairs of complex conjugate points (actually, they are λ_j and λ_j^*) and that $R(\lambda) \to 0$ as $\lambda \to \infty$. Then, one can show that the RH problem sought at sub-step 2 is that obtained from the original RH problem, where $B(\lambda)/A(\lambda)$ is replaced by $R(\lambda)$. Due to the jumps of $R(\lambda)$ across the arcs connecting λ_j and λ_j^*, additional jump conditions on these arcs arise and, thus, the jump contour takes the form: $\cup_{j=0}^{N} \Gamma_j \cup \mathbb{R} \cup i\mathbb{R}$, whereas the jump matrix on all parts of the contour can be algebraically given in terms of $a(\lambda)$, $b(\lambda)$, and $R(\lambda)$. To complete the formulation of the RH problem from step 1, the jump conditions have to be complemented by the residue conditions at zeros of $a(\lambda)$ and singularities of $R(\lambda)$ in the respective quadrants, if any (these are also given in terms of spectral quantities determined by the initial data only). For the exact formulation of the RH problem of step 1, see [18], Theorem 4.6.[1]

We note that the scattering matrix $s(\lambda)$ in our setting is closely related to the monodromy matrix $M(\lambda)$, a key object of the spectral theory associated with the Zakharov–Shabat equation with periodic conditions. The monodromy matrix is defined as $M(\lambda) = \Phi(L, 0, \lambda)$, where $\Phi(x, 0, \lambda)$ is the solution of Eq. (3) satisfying the condition $\Phi(0, 0, \lambda) = I$:

$$M(\lambda) = e^{-i\lambda L\sigma_3} s^{-1}(\lambda) = \begin{pmatrix} e^{-i\lambda L}a(\lambda) & -e^{-i\lambda L}b(\lambda) \\ e^{i\lambda L}b^*(\lambda^*) & e^{i\lambda L}a^*(\lambda^*) \end{pmatrix}, \tag{8}$$

where we have taken into account that $\det M(\lambda) = \det s(\lambda) \equiv 1$. Accordingly, $R(\lambda)$ can be expressed in terms of the entries of the monodromy matrix:

$$R(\lambda) = \frac{M_{11}(\lambda) - M_{22}(\lambda) + \sqrt{\Delta^2(\lambda) - 4}}{2M_{21}(\lambda)}, \tag{9}$$

where $\Delta(\lambda) := M_{11}(\lambda) + M_{22}(\lambda)$. In this context, $\{\lambda_j, \lambda_j^*\}_0^N$ are called the *main spectrum*; they are simple zeros of $\Delta^2(\lambda) - 4$ and thus the branch points of (9). On the other hand, the simple zeros of $M_{12}(\lambda)$ which are not double zeros of $\Delta^2(\lambda) - 4$ (as well as the multiple zeros of $M_{12}(\lambda)$) constitute the *auxiliary spectrum* $\{\mu_j\}_1^N$.

The second step consists of transforming the RH problem described above (with jumps across $\cup_{j=0}^N \Gamma_j \cup \mathbb{R} \cup i\mathbb{R}$ and residue conditions) to a RH problem of the form (A.2)+(A.3)+(A.1) with some constants ϕ_j. This step can also be divided into two sub-steps: (i) transforming the RH problem to that with jumps across $\Gamma = \cup_{j=0}^N \Gamma_j$ only (thus getting rid of jumps across $\mathbb{R}$ and $i\mathbb{R}$ and singularity conditions); (ii) making the jumps on each Γ_j to have the structure as in (A.1).

In the case $N = 0$, this step has been done in [17,18]; in this case, the contour for the RH problem consists of a single arc, and there are no singularity conditions. The associated (0-genus) solution of the NLS equation is a simple exponential function: $q(x, t) = \alpha e^{-2i\beta x + 2i\omega t + i\phi_0}$, where $\alpha = \operatorname{Im}\lambda_0$, $\beta = \operatorname{Re}\lambda_0$, and $\omega = \alpha^2 - 2\beta^2$.

The cases with $N \geq 1$ turn out to be much more involved. Particularly in the realization of the first sub-step, one faces the following problems:

(i) We need to get rid of singularity conditions at the singularity points of $R(\lambda)$, which can be characterized as zeros of the denominator in (9) that are not zeros of the numerator in (9). In terms of the spectral theory of the Zakharov–Shabat equation with periodic boundary conditions, the (possibly empty) set of such singularity points $\{\mu_j\}_1^{N_1}$, $N_1 \leq N$ consists of those conjugated *auxiliary spectrum* points for this problem which are located on the sheet (of the two-sheeted Riemann surface of R) characterized by the condition $R(\lambda) \to 0$ as $\lambda \to \infty$.

(ii) The resulting (λ-dependent) jump matrix is as follows [29]:

$$\check{J}(x, t, \lambda) = \begin{pmatrix} 0 & iJ_0(\lambda)e^{-2i\lambda(x-L)-4i\lambda^2 t} \\ iJ_0^{-1}(\lambda)e^{2i\lambda(x-L)+4i\lambda^2 t} & 0 \end{pmatrix}, \qquad \lambda \in \Gamma \tag{10}$$

(in the notations of [29], $P(\lambda) = \log J_0(\lambda)$), where it is the square of $J_0(\lambda)$ that has a simple expression in terms of $R_\pm(\lambda)$ or, alternatively, of the entries of the monodromy matrix:

$$J_0^2(\lambda) = \frac{R_+(\lambda)}{R_+^*(\lambda)} Q^2(\lambda) = \frac{R_-(\lambda)}{R_-^*(\lambda)} Q^2(\lambda) = -\frac{M_{12}(\lambda)}{M_{21}(\lambda)} Q^2(\lambda), \tag{11}$$

where $Q(\lambda) = \prod_{j=1}^{N_1} \frac{\lambda - \mu_j}{\lambda - \mu_j^*}$.

Having $J_0(\lambda)$ obtained, the second sub-step (reducing the jump to that as in (A.1)) is well-defined; it can be done using the solution of the scalar RH problem

$$d_+(\lambda)d_-(\lambda) = J_0(\lambda)e^{i\phi_j}, \quad \lambda \in \Gamma_j, \quad j = 0, \dots, N, \tag{12}$$

$$d(\lambda) \to 1, \quad \lambda \to \infty. \tag{13}$$

Here the constants ϕ_j are not prescribed but determined uniquely by the conditions in (12), see (13)–(16) in [29]; they turn out to be the phases sought in the direct problem.

On the other hand, because of items (i) and (ii) given above, the problem of computing $J_0(\lambda)$ numerically turns out to be challenging at least in two aspects: (i) choosing the relevant auxiliary spectrum points and (ii) choosing the appropriate branch of the square root $\sqrt{-M_{12}(\lambda)/M_{21}(\lambda)}$. For instance, ignoring (ii) reduces the range of uniquely determined ϕ_j to $[0, \pi)$ [29].

The above-mentioned problems in the realization of the "direct problem step" already in a situation limited by the periodicity requirement incites us to use a methodologically different approach for this step: soft-computing. In the next Section, we present the method based on convolutional neural networks, which, though being an approximation, can be used for non-periodic finite-genus solutions with incommensurable frequencies C_j^f.

[1] In [18], the notation $\tilde{\Gamma}$ is adopted for R.

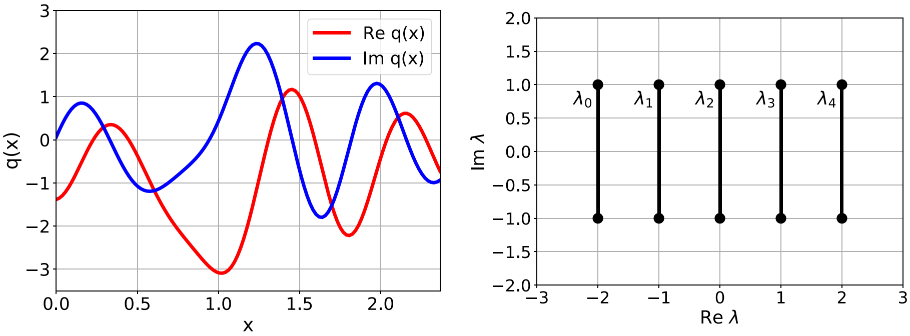

Fig. 1. The profile of 4-genus solution in physical domain (left) corresponding to phases ϕ values $\{\pi, \frac{\pi}{6}, \frac{3\pi}{2}, \frac{2\pi}{3}, \frac{7\pi}{6}\}$ and its main spectrum (right). The arcs (vertical cuts) in the λ-plane connecting the complex conjugated pairs of the main spectrum points are also depicted.

3. Neural networks for phases computation

Soft computing can be defined as a set of techniques to address complex, real-world problems that cannot be solved using traditional, deterministic models. It includes approaches such as artificial NNs, genetic algorithms, and fuzzy logic, which are based on learning from data rather than precise modeling. An NN typically consists of an input layer that receives a set of known parameters and one or more hidden layers that can be fully-connected, convolutional, recurrent, etc. The hidden layers contain a large number of adjustable parameters (weights), which are optimized through a training process using the available data. The output layer, which determines the number of classification classes or values to be predicted in the regression case, is determined by the specific task.

Convolutional NNs (CNNs) are inspired by the principles of animal vision and are designed to extract features from input signals [44]. CNNs are a type of NNs that are particularly well-suited for processing and analyzing data that has a spatial or temporal structure. One reason why CNNs are efficient for nonlinear transformations is that they can learn and adapt to complex patterns and relationships in data. They do this through convolutional layers, which apply filters to the input data to extract features and create a representation of the data more suitable for learning. These filters are able to recognize patterns in the data and extract the relevant features, even if the relationships between the features are nonlinear. Another reason why CNNs are effective for nonlinear transformations is their ability to handle a large number of input variables.

The benefit of using NNs for our problem is that they can handle high complexity and nonlinearity. The IST is an operation that decomposes a solution into its nonlinear spectrum components, which can be quite challenging for complex profiles (signals). Another advantage of using NNs is their ability to generalize to new data. Once an NN has been trained on a data set, it can make predictions on new, unseen data. This can be particularly useful in the case of the IST (nonlinear Fourier transform) for optical signals, as it allows the neural network to make predictions based on patterns it has learned from a large dataset rather than relying on a specific set of predetermined rules.

3.1. Solutions set preparation and data collection

A finite-genus solution can be described in terms of the parameters, the spectral data, defining the jump matrices of the RH problem, see Appendix. These parameters include the endpoints of the jump arcs, λ_j, the main spectrum points, and the constants ϕ_j, referred to as *phases*, that are included in the jump matrices (A.1). A nonlinear harmonic is a solution component defined by a single arc connecting a main spectrum point to its complex conjugate, along with a phase represented by a jump matrix entrance on the arc. An example of the main spectrum of a 4-genus solution and its profile in the physical space x (taken at a fixed value of evolution variable t) is depicted in Fig. 1.

In this work, three different configurations of the main spectrum were used to test the performance of the proposed algorithm for solving the direct problem. Two of these configurations correspond to 4-genus solutions with different values of the imaginary part of the main spectrum ($\operatorname{Im}\lambda_j = 1$ and $\operatorname{Im}\lambda_j = 5$ for all values of j), corresponding to different power levels of the solution, while the third configuration is an 8-genus solution with $\operatorname{Im}\lambda_j = 1$. The algorithm's flexibility is demonstrated by its ability to handle these different configurations; the latter are listed in Table 1 and characterized by their main spectrum points in the upper complex half-plane. These variations in the scattering data structure of the signals allow us to evaluate the algorithm's performance under various conditions.

The data for the NN have been collected by solving the inverse problem for the given main spectrum configurations with uniformly randomly distributed phases ϕ_j in the interval from 0 to 2π. An implemented algorithm is based on RH problem solver [10,19,45]. The number of solutions generated for each main spectrum configuration was 4×10^5. Each dataset was split into three unequal parts: the train, validation, and test data, corresponding to 80%, 17.5%, and 2.5% of the entire dataset, respectively. The training set was used to train the NN model. The validation set was used to evaluate

Table 1
The main spectra for 4-genus solutions with $\operatorname{Im}\lambda = 1$ and $\operatorname{Im}\lambda = 5$, and for 8-genus solution with $\operatorname{Im}\lambda = 1$.

Genus	Main spectrum
4-genus	-2+i, -1+i, i, 1+i, 2+i
4-genus	-2+5i, -1+5i, 5i, 1+5i, 2+5i
8-genus	-4+i, -3+i, -2+i, -1+i, i, 1+i, 2+i, 3+i, 4+i

the model's performance during the training process, and the test set was used to evaluate the final performance of the model with unseen data. The test set should represent the types of data the model is expected to encounter in real-world applications.

Generally, if the frequencies of a finite-genus solution (constants $C_1^f, C_2^f, \ldots$) are not commensurable, the resulting solution is not periodic. To represent the solution on a final duration (via a finite number of discrete points), a maximum period L_0, corresponding to the minimal frequency, was chosen. This ensures that each nonlinear harmonic has at least one period within L_0. The number of x-discretization points of $q(x)$ in L_0 is kept constant for all cases and equals 128. To take advantage of the cyclical nature of phases in the RH problem approach, their values used during the training of the NN were chosen as complex points on the unit circle, even though a single real number determines every phase of the solution. In this research, two real numbers were used to represent each phase: the real part, $\operatorname{Re} e^{i\phi_j}$, and the imaginary part, $\operatorname{Im} e^{i\phi_j}$. Such a representation allowed us to eliminate the problems arising at the ends of $[0, 2\pi)$ interval of phase variation.

3.2. Neural network architecture

The architecture of the NN used in our work is based on NFT-Net NN [41]. The convolutional structure of NFT-Net provided high effectiveness for implementing conventional nonlinear Fourier transform (for rapidly decaying solutions). Since the basic principles of the scattering transforms are the same for the case of vanishing and periodic boundary conditions, the NFT-Net seems to be a good candidate for the initial iteration of the structure optimization procedure in our case.

The nonlinear transformation of an initial signal to the phases provided by the NFT-Net-type NN can be described as follows: the discretized solution is fed into the CNN, which applies filters to extract the desirable features. The filters are typically small, with a fixed size and shape, and are designed to recognize specific patterns in the data. As the filters move across the solution, they produce feature maps representing the solution at different scales and orientations. These filters can recognize patterns in the data and extract the relevant features, even if the relationships between the features are nonlinear. The size of these feature maps is determined by the original solution's size, the filters' size, and the number of filters used. The feature map provided by the previous convolutional layer serves as an input for the next layer. Multiple consecutive convolutional layers extract a set of primitive features from the input data in the first layer and increasingly complex and abstract features in subsequent layers. The filters used in these layers are trainable, meaning that they can be adjusted and adapted to a specific transformation based on the input data under consideration. This flexibility allows the NN to learn and extract relevant features from the data. The output of the convolutional layers is then passed through one or more fully connected layers, which perform a linear operation on the data and apply a nonlinear activation function. These layers are used to learn complex relationships between the features extracted by the convolutional layers and the output of the CNN. The output of the NN is a set of real values representing the real and imaginary parts of $e^{i\phi_j}$, from which we extract the desired phases ϕ_j.

To optimize the NN performance, we used Bayesian optimization, employing the NFT-Net architecture [41] as an initial structure guess. One advantage of Bayesian optimization is that it is typically able to search the hyperparameter space more efficiently than other methods, such as grid search or random search [42]. This is because it takes into account the previous results and uses them to guide the search toward the most "promising" areas of the hyperparameter space. The optimization process begins by defining the range of possible values for each hyperparameter and the objective function that will be used to evaluate the model's performance. A probabilistic model is then fit to the data using the initial set of hyperparameter configurations and their corresponding performance scores. This model is used to predict the performance for different hyperparameter configurations. The next step is to select the hyperparameters that are expected to maximize the model's performance, according to the probabilistic model. This configuration is then evaluated using the objective function, and the performance score is recorded. The probabilistic model is then updated with the new data point (the selected hyperparameters and its corresponding performance score). This updated model is used to guide the search for the following hyperparameter configuration. The process is then repeated, with the updated probabilistic model being used to select the next hyperparameter configuration, and the process continues until the desired number of iterations is reached.

The hyperparameters with which the Bayesian optimizer manipulated included the number and type of layers (convolutional and fully-connected), the size and stride of the filters in the convolutional layers, the number of neurons in the fully-connected layers, and the activation function types used. The Bayesian hyperparameters optimization was

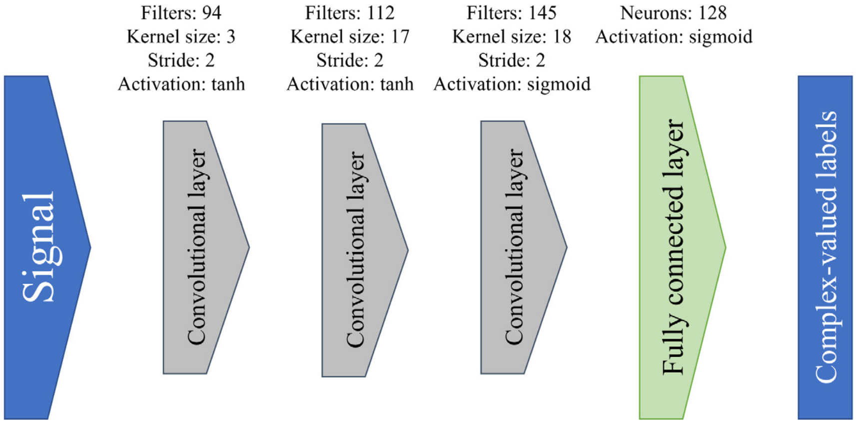

Fig. 2. An optimized neural network architecture used for retrieving the phases for all test cases. Each shape on the figure corresponds to a particular NN layer, where the values of parameters are given above the shapes.

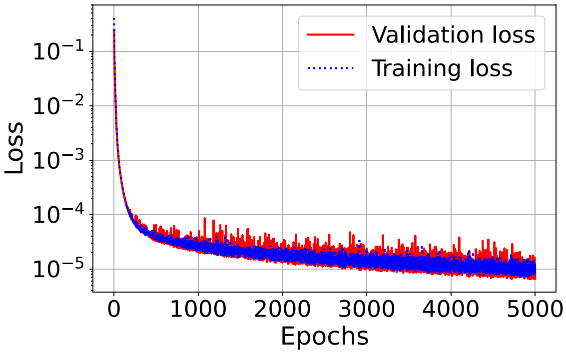

Fig. 3. Behavior of the loss function vs. epoch number for the training and validation runs. The figure corresponds to the case where we were retrieving the phases for the 4-genus solution with $\operatorname{Im}\lambda = 1$.

applied to the data corresponding to the 4-genus solution with $\operatorname{Im}\lambda_j = 1$, and the particular architecture obtained is given in Fig. 2. We used this architecture as an initial input for the Bayesian optimization for other solutions (4-genus with $\operatorname{Im}\lambda_j = 5$ and 8-genus with $\operatorname{Im}\lambda_j = 1$) but observed no further prediction improvement within some time. Therefore, the resulting optimized NN architecture from Fig. 2 was used in all the test cases.

Let us introduce the value that defines the error between the predicted and true phase values:

$$\Delta_j = e^{i\phi_j^{\text{pred}}} - e^{i\phi_j^{\text{true}}},$$

where index j denotes the corresponding nonlinear harmonic. $\operatorname{Re}\Delta_j$ and $\operatorname{Im}\Delta_j$ are real and imaginary parts of the error, correspondingly. To measure the error in the NN's predictions, we used the mean squared error (MSE) loss function:

$$\text{Loss} = \sum_j \left[\left(\operatorname{Re}\Delta_j\right)^2 + \left(\operatorname{Im}\Delta_j\right)^2 \right]. \tag{14}$$

To optimize the weights of the NN, we used the Adam algorithm with the 10^{-4} learning rate. After 5000 epochs of training, the NNs did not show any improvement in their performance, so the respective weights' values were chosen as optimal and used in the inference. An example of the behavior of the loss function for the training and validation of our NN (for the case of 4-genus solutions with $\operatorname{Im}\lambda_j = 1$) is depicted in Fig. 3. The NN was developed and trained using the Keras framework.

4. Results

We define the error in the predicted phase as follows:

$$\phi_{err} = |\phi_{true} - \phi_{pred}|.$$

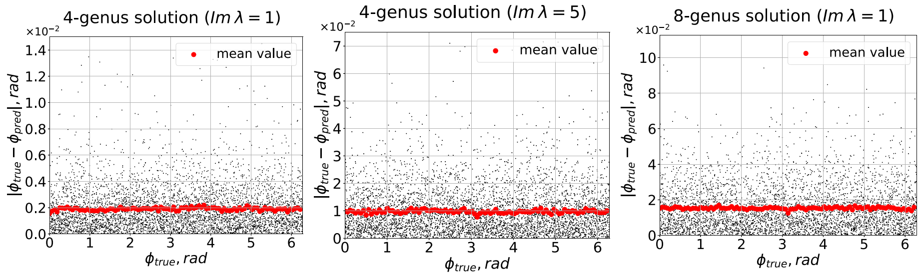

Fig. 4. The distributions of an error of predicted phases and their mean values. The black points are predictions of individual phases, while the mean value is average over all cases in a subinterval and over all nonlinear harmonics.

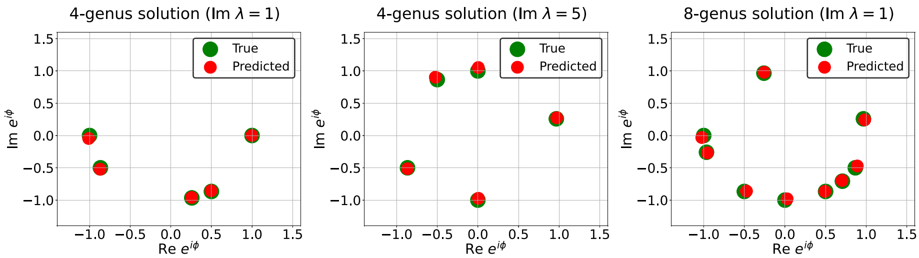

Fig. 5. The true and predicted labels for both 4-genus solutions (Im $\lambda = 1$ and Im $\lambda = 5$) and 8-genus solution with phase ϕ sets $\{0, \frac{5\pi}{3}, \frac{19\pi}{12}, \frac{7\pi}{6}, \pi\}$, $\{\frac{7\pi}{6}, \frac{2\pi}{3}, \frac{\pi}{12}, \frac{3\pi}{2}, \frac{\pi}{2}\}$, $\{\frac{7\pi}{12}, \frac{7\pi}{4}, \frac{3\pi}{2}, \pi, \frac{4\pi}{3}, \frac{11\pi}{6}, \frac{\pi}{12}, \frac{5\pi}{3}, \frac{13\pi}{12}\}$ correspondingly. The green circles corresponding to true values are sized up to be seen better.

The distributions of the errors of individual predictions are given in Fig. 4. We divide the interval $[0, 2\pi)$ into 100 subintervals and calculate a mean error value over all phases in the subinterval for all nonlinear harmonics. The mean values of the error are distributed homogeneously over the phase: $1.9 \times 10^{-3}, 9.7 \times 10^{-3}$ and 1.5×10^{-2} for 4-genus solutions (Im $\lambda_j = 1$ and Im $\lambda_j = 5$) and 8-genus solution with Im $\lambda_j = 1$, correspondingly. These results demonstrate that we achieved sufficient accuracy using a NN. However, later on, we will discuss the factors that influence the accuracy of the prediction.

We also present a result of individual predictions for three solution configurations, Fig. 5. We construct the points with Cartesian coordinates Re $e^{i\phi_j}$ and Im $e^{i\phi_j}$ that correspond to the output of our NN (red circles) and compare them with true values (green circles).

One interesting effect we observed is that the phases of the different nonlinear harmonics determined via the NN vary in accuracy. To demonstrate this more clearly, we drew the distribution of prediction errors for different nonlinear harmonics. As a measure of accuracy, we calculate the mean phase error over the interval $[0, 2\pi)$ for each nonlinear harmonic separately, see Fig. 6.

In Fig. 6, the distributions of errors in phase prediction are given for all considered main spectrum configurations. For simplicity, we located every error value above a cut connecting the corresponding main spectrum point λ_j with its complex conjugate. The phases of central nonlinear harmonics have a lower accuracy than those closer to the ends of the Reλ band interval. So far, we have not found a plausible explanation for this effect. The difference in the errors for two types of 4-genus solution (blue and red dots in the left panel of Fig. 6) arises because the 4-genus solution with Im $\lambda_j = 1$ (blue dots in the left panel of Fig. 6) has a lower imaginary part of the main spectrum for all points, and so the lower nonlinearity in comparison with 4-genus with Im $\lambda_j = 5$ (red dots in the left panel of Fig. 6). Taking into account that the solution of the RH problem involves an integration over the cuts of the main spectrum, and these cuts are five times longer in the case of genus 4 with Im $\lambda_j = 5$, the solver provides lower accuracy for 4-genus solution (Im $\lambda_j = 5$) as long as we keep the same discretization rate for both cases.

5. Discussion and conclusions

In the current paper, we proposed solving the direct problem for the finite-genus solutions to the NLS equation (i.e., given the x-profile of a finite-genus solution, retrieve the parameters in its RH representation) using convolutional neural networks. Our approach eliminates a set of restrictions for the direct problem.

The previous method allowed us to deal only with exactly-periodic solutions and determine phases modulo π, the respective procedure required the involved computation and analysis of the auxiliary spectrum. The proposed method is devoid of the listed shortcomings as we can obtain the training data for the NN using the existing RH problem solver:

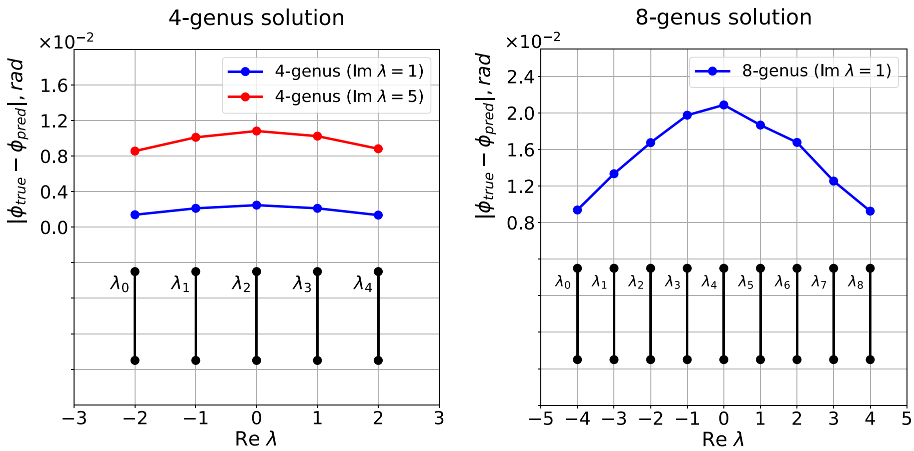

Fig. 6. Error distribution between different nonlinear harmonics. The error value for each nonlinear harmonic is located above the corresponding points of the main spectrum.

the resulting NN structure demonstrated the capability to deal with arbitrary phase distributions and generalize to other cases (larger power and higher genus).

An essential novelty and advantage of our approach are expanding a class of solutions for which the direct problem can be solved: the exact periodicity (modal frequency commensurability) ceases to be a mandatory condition within the soft-computing method presented in our work. Insofar as the proposed method works for any finite-genus solutions, its potential applications are larger. We notice that the theory allowing us to calculate the direct problem for N-genus solution to the NLS equation with arbitrary non-commensurable frequencies does not exist, and our soft-computing approach is only a way up to date to deal with this difficulty.

A choice of neural network architecture was a crucial part of our study. Though CNNs have proved viable in previous similar research, some adaptations and optimizations have been required. In our case, as far as we do not utilize any specially designed solver for the direct problem, it is the solver for the *inverse* RH problem that provided us with the training and testing datasets.

Many factors determine the accuracy of the NN-based phase retrieving. The most critical issues are the accuracy of solving the inverse problem in the data generation stage and the amount of training data. Moreover, the neural network architecture and hyperparameters values significantly impact on the performance.

A possible reason for the non-uniform distribution of errors for different nonlinear harmonics is that the central harmonics are more affected by the accuracy of calculations for neighboring harmonics than the side harmonics, and we had lower computational errors when dealing with the data generated for the latter. This is indirectly confirmed by the fact that the distribution in the case of 5 harmonics for different imaginary parts of λ_j, demonstrates approximately the same shape (left pane in Fig. 6). However, this issue remains an open question and requires further study.

In the end, we note that the generalization of our approach to predict not only the phases (when having the main spectrum fixed) but also the main spectrum together with phases requires more effort. We also note that here we did not attempt to compress the NN model to reduce its computational complexity, while the usage of model compression techniques [46] is definitely important if we aim at developing some efficient signal processing solutions.

Finally, we note that in application to optical fiber communications, we need to consider a dual-polarization system [47], which is conventionally modeled by a coupled system of the NLS equations, also known as the Manakov system [48–50]. Since the Manakov system is integrable (possesses the Lax pair representation), it is natural to generalize our approach to this system. However, such a generalization faces at least two problems: (i) the underlying Riemann–Hilbert formalism involves 3 × 3 matrices, which makes its numerical implementation more complicated (compared to the 2 × 2 case); (ii) currently, there is no available theoretical description of the connection between algebro-geometric solutions to the Manakov system and solutions to the respected Riemann–Hilbert problem with piece-wise constant jump matrices across the arcs of a certain contour. Thus, generalizing our approaches to the Manakov equations case is still an open question requiring more studies.

CRediT authorship contribution statement

Stepan Bogdanov: Conceived the study, Proposed the neural network model, Collected the data, Numerical simulations, Designed the figures, Wrote the manuscript, Reviewed the manuscript. **Dmitry Shepelsky:** Conceived the study, Wrote the manuscript, Reviewed the manuscript. **Anastasiia Vasylchenkova:** Reviewed the manuscript. **Egor Sedov:** Proposed the neural network model, Performed the architecture optimization, Reviewed the manuscript. **Pedro J. Freire:** Reviewed the manuscript. **Sergei K. Turitsyn:** Reviewed the manuscript, Supervision. **Jaroslaw E. Prilepsky:** Conceived the study, Proposed the neural network model, Wrote the manuscript, Reviewed the manuscript, Supervision.

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.

Acknowledgments

SB, JEP, and SKT acknowledge the support from Leverhulme Trust, Grant No. RP-2018-063. PJF is supported by the EU Horizon 2020 program under the Marie Sklodowska-Curie grant agreement 813144 (REALNET). ES and SKT acknowledge the support of the EPSRC project TRANSNET. JEP and DS thank the Erasmus+ ERC mobility programme between Aston University and B. Verkin Institute for Low Temperature Physics and Engineering. AV is grateful to the Leverhulme Trust Early Career Fellowship ECF-2020-150.

Appendix. Riemann-Hilbert formulation of the inverse problem associated with finite-genus solutions to focusing NLS equation

Consider the following RH problem.
Given

- the oriented contour $\Gamma = \cup_{j=0}^{N} \Gamma_j$, where $\Gamma_j = (\lambda_j, \lambda_j^*)$ (an arc connecting λ_j with λ_j^*), and
- the 2×2-valued function

$$J(x,t,\lambda) = \begin{pmatrix} 0 & ie^{-i\phi_j - 2i\lambda x - 4i\lambda^2 t} \\ ie^{i\phi_j + 2i\lambda x + 4i\lambda^2 t} & 0 \end{pmatrix}, \quad \lambda \in \Gamma_j, \tag{A.1}$$

where $\phi_j \in [0, 2\pi)$, $j = 0, \dots, N$, find a 2×2-valued function $\Psi(x,t,\lambda)$ such that:

1. For all $x \in \mathbb{R}$ and $t \in \mathbb{R}$, $\Psi(x,t,\lambda)$ is analytic w.r.t. λ for $\lambda \in \mathbb{C} \setminus \bar{\Gamma}$ and continuous up to Γ from the both sides of Γ;
2. The limiting values $\Psi^+(x,t,\lambda)$ and $\Psi^-(x,t,\lambda)$, $\lambda \in \Gamma$ of $\Psi(x,t,\lambda)$, as λ approaches Γ from the $+$ and - side respectively, are related by $J(x,t,\lambda)$:
$$\Psi^+(x,t,\lambda) = \Psi^-(x,t,\lambda) J(x,t,\lambda), \qquad \lambda \in \Gamma; \tag{A.2}$$
3. At λ_j and λ_j^*, $\Psi(x,t,\lambda)$ has singularities of the order of no bigger than the inverse fourth root;
4. As $\lambda \to \infty$,
$$\Psi(x,t,\lambda) = I + O(1/\lambda), \tag{A.3}$$
where $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Having the RH problem (A.1)–(A.3) solved for all x and t, determine $\Psi_1(x,t)$ from the asymptotic behavior of $\Psi(x,t,\lambda)$ as $\lambda \to \infty$: $\Psi(x,t,\lambda) = I + \frac{\Psi_1(x,t)}{\lambda} + \dots$. Then $q(x,t)$ is given by

$$q(x,t) := 2i[\Psi_1]_{12}(x,t) \tag{A.4}$$

(where $[\cdot]_{12}$ stands for the 12 entry of a matrix), is a solution of the NLS equation of finite-genus type: it can be expressed in terms of Riemann theta functions associated with the Riemann surface of genus N, with the branch points at λ_j and λ_j^*, $j = 0, \dots, N$.

The statement that $q(x,t)$ determined as above can be explicitly expressed in terms of certain Riemann theta functions (with parameters depending on x and t) is based on the possibility to express $q(x,t)$ in terms of the solution of another RH problem (see Proposition 1 below), which can be considered as a transformation of the original RH problem evoking the so-called "g-function mechanism" [32].

In order to formulate the transformed RH problem, we need a set of parameters uniquely defined by the set of the branch points λ_j and λ_j^*, $j = 0, \dots, N$. First, define $w(\lambda)$ by

$$w(\lambda) = \prod_{j=0}^{N} \sqrt{(\lambda - \lambda_j)(\lambda - \lambda_j^*)} \tag{A.5}$$

as a function analytic in $\mathbb{C}\setminus\Gamma$ whose branch is fixed by the asymptotic condition $w(\lambda)\simeq\lambda^{N+1}$ as $\lambda\to\infty$. Let each arc Γ_j be oriented upward and let $w^+(\lambda)$ be the values of w at the "+" side of the corresponding Γ_j. Further, define the $N\times N$ matrix $\mathbf{K}$ by

$$\mathbf{K}_{mj} := \int_{\Gamma_j}\frac{\xi^{m-1}d\xi}{w^+(\xi)},\qquad m,j=1,\dots,N \tag{A.6}$$

and determine the vectors $\mathbf{C}^f := (C_1^f,\dots,C_N^f)^T$ and $\mathbf{C}^g := (C_1^g,\dots,C_N^g)^T$ as the solutions of the following linear equations:

$$\mathbf{K}\cdot\mathbf{C}^f = [0,\dots,0,-2\pi i]^T,\qquad \mathbf{K}\cdot\mathbf{C}^g = -4\pi i\left[0,\dots,0,2,\sum_{j=0}^{N}(\lambda_j+\lambda_j^*)\right]^T. \tag{A.7}$$

Finally, determine the constants f_0 and g_0 from the large-λ developments of two scalar functions analytic in $\mathbb{C}\setminus\Gamma$:

$$f(\lambda) := \frac{w(\lambda)}{2\pi i}\sum_{j=1}^{N}\int_{\Gamma_j}\frac{C_j^f\,d\xi}{w^+(\xi)(\xi-\lambda)} = \lambda+f_0+O(1/\lambda), \tag{A.8}$$

$$g(\lambda) := \frac{w(\lambda)}{2\pi i}\sum_{j=1}^{N}\int_{\Gamma_j}\frac{C_j^g\,d\xi}{w^+(\xi)(\xi-\lambda)} = 2\lambda^2+g_0+O(1/\lambda) \tag{A.9}$$

Proposition 1 ([32]). *Given $\{\lambda_j\}_0^N$ with $\operatorname{Im}\lambda_j>0$ and $\{\phi_j\}_0^N$ with $\phi_j\in[0,2\pi)$, the N-genus solution $q(x,t)$ of the NLS equation obtained from the solution of the RH problem (A.1)–(A.3) can also be expressed by*

$$q(x,t) = 2i[\hat{\Phi}_1]_{12}(x,t)e^{2if_0x+2ig_0t}, \tag{A.10}$$

where $\hat{\Phi}_1$ enters the large-λ development

$$\hat{\Phi}(x,t,\lambda) = I+\frac{\hat{\Phi}_1(x,t)}{\lambda}+\dots \tag{A.11}$$

of the solution $\hat{\Phi}(x,t,\lambda)$ of the following RH problem: find $\hat{\Phi}(x,t,\lambda)$ analytic in $\mathbb{C}\setminus\Gamma$ and satisfying the jump conditions:

$$\hat{\Phi}^+(x,t,\lambda) = \hat{\Phi}^-(x,t,\lambda)\hat{J}_j(x,t),\qquad \lambda\in\Gamma_j,\ j=0,\dots,N, \tag{A.12}$$

with

$$\hat{J}_j(x,t) = \begin{pmatrix} 0 & ie^{-i(\phi_j+C_j^f x+C_j^g t)} \\ ie^{i(\phi_j+C_j^f x+C_j^g t)} & 0\end{pmatrix} \tag{A.13}$$

and the normalization condition

$$\hat{\Phi}(x,t,\lambda) = I+O(1/\lambda),\qquad \lambda\to\infty. \tag{A.14}$$

Here $C_0^f=C_0^g=0$, whereas the constants f_0 and g_0 in (A.10) and C_j^f, C_j^g, $j=1,\dots,N$ in (A.13) are determined by $\{\lambda_j\}_0^N$ and $\{\phi_j\}_0^N$ via (A.6)–(A.9).

Remark 2. $\Psi(x,t,\lambda)$ and $\hat{\Phi}(x,t,\lambda)$ are related as follows:

$$\Psi(x,t,\lambda) = e^{(if_0x+ig_0t)\sigma_3}\hat{\Phi}(x,t,\lambda)e^{(i(\lambda-f(\lambda))x+i(2\lambda^2-g(\lambda))t)\sigma_3}, \tag{A.15}$$

where $\sigma_3=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$. Here, the exponential in (A.15) presents an appropriate realization of the "g-function mechanism", which allows reducing the original RH problem to that characterized by jump data independent of λ (on each part Γ_j of the contour) combined with the standard normalization at infinity (which is independent of λ as well).

Remark 3. It is the RH problem (A.12)–(A.14) that can be solved explicitly in terms of the theta functions of the N-genus Riemann surface associated with $w(\lambda)$ (A.5) and characterized by the branch points λ_j and λ_j^*, $j=0,\dots,N$.

Remark 4. The parameters x and t enter the jump matrix (A.13) and thus the solution $\hat{\Phi}$ through the linear expressions

$$\tilde{\phi}_j(x,t) := \phi_j+C_j^f x+C_j^g t. \tag{A.16}$$

Consequently (and taking into account (A.10)), the mapping $\phi_j|_{t=0}\mapsto\phi_j|_{t=T}=\phi_j|_{t=0}+C_j^gT-2g_0T$ represents the evolution, in spectral terms, of a particular N-genus solution, which provides simple means to retrieve the original parameters ϕ_j (associated with the transmitter, at $t = 0$) given these parameters at the receiver (at $t = T$). Particularly, if all C_j^f together with f_0 turn to be commensurable, then the underlying solution of the NLS equation is periodic in x. Similarly, in t.

References

[1] Gardner CS, Greene JM, Kruskal MD, Miura RM. Method for solving the Korteweg-deVries equation. Phys Rev Lett 1967;19:1095–7.
[2] Zakharov VE, Shabat AB. Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media. J Exp Theor Phys 1972;34:62–9.
[3] Lax PD. Integrals of nonlinear equations of evolution and solitary waves. Comm Pure Appl Math 1968;21(5):467–90.
[4] Wadati M. The exact solution of the modified Korteweg-de Vries equation. J Phys Soc Japan 1972;32:1681.
[5] Ablowitz MJ, Kaup DJ, Newell AC, Segur H. The inverse scattering transform-Fourier analysis for nonlinear problems. Stud Appl Math 1974;53(4):249–315.
[6] Ablowitz MJ, Clarkson PA. Solitons, nonlinear evolution equations and inverse scattering. Vol. 149. Cambridge University Press; 1991.
[7] Yousefi MI, Kschischang FR. Information transmission using the nonlinear Fourier transform, Part I: Mathematical tools. IEEE Trans Inform Theory 2014;60(7):4312–28.
[8] Turitsyn SK, Prilepsky JE, Le ST, Wahls S, Frumin LL, Kamalian M, et al. Nonlinear Fourier transform for optical data processing and transmission: advances and perspectives. Optica 2017;4(3):307–22.
[9] Novikov S, Manakov SV, Pitaevskii LP, Zakharov VE. Theory of solitons: the inverse scattering method. Springer Science & Business Media; 1984.
[10] Trogdon T, Olver S. Riemann–Hilbert problems, their numerical solution, and the computation of nonlinear special functions. Philadelphia, PA: Society for Industrial and Applied Mathematics; 2015, http://dx.doi.org/10.1137/1.9781611974201.
[11] Deift P, Zhou X. A steepest descent method for oscillatory Riemann-Hilbert problems. Asymptotics for the MKdV equation. Ann Math 1993;137:295–368.
[12] Matveev VB. 30 years of finite-gap integration theory. Phil Trans R Soc A 2008;366(1867):837–75.
[13] Belokolos ED, Bobenko AI, Enolskii VZ, Its AR, Matveev VB. Algebro-geometric approach to nonlinear integrable equations. Vol. 550. Springer; 1994.
[14] Osborne AR. Rogue waves: Classification, measurement and data analysis, and hyperfast numerical modeling. Eur Phys J Spec Top 2010;185(1):225–45.
[15] Chimmalgi S, Wahls S. On computing high-dimensional Riemann theta functions. Commun Nonlinear Sci Numer Simul 2023;107266. http://dx.doi.org/10.1016/j.cnsns.2023.107266.
[16] Fokas AS. A unified approach to boundary value problems. Society for Industrial and Applied Mathematics; 2008, http://dx.doi.org/10.1137/1.9780898717068.
[17] Deconinck B, Fokas A, Lenells J. The implementation of the unified transform to the nonlinear Schrödinger equation with periodic initial conditions. Lett Math Phys 2021;111(1):1–18.
[18] Fokas A, Lenells J. A new approach to integrable evolution equations on the circle. Proc R Soc Lond Ser A Math Phys Eng Sci 2021;477(2245):20200605.
[19] Olver S. A general framework for solving Riemann–Hilbert problems numerically. Numer Math 2012;122(2):305–40.
[20] Trogdon T, Olver S. Numerical inverse scattering for the focusing and defocusing nonlinear Schrödinger equations. Proc R Soc A Math Phys Eng Sci 2013;469(2149):20120330.
[21] Olver S, Trogdon T. Nonlinear steepest descent and numerical solution of Riemann-Hilbert problems. Comm Pure Appl Math 2014;67(8):1353–89.
[22] Osborne A. Nonlinear ocean waves and the inverse scattering transform. Academic Press; 2010.
[23] Osborne AR. Nonlinear fourier analysis: Rogue waves in numerical modeling and data analysis. J Mar Sci Eng 2020;8(12):1005.
[24] Goossens J-W, Hafermann H, Jaouën Y. Data transmission based on exact inverse periodic nonlinear Fourier transform, Part I: Theory. J Lightwave Technol 2020;38(23):6499–519.
[25] Kamalian M, Prilepsky JE, Le ST, Turitsyn SK. Periodic nonlinear Fourier transform for fiber-optic communications, Part I: theory and numerical methods. Opt Express 2016;24(16):18353–69.
[26] Kamalian M, Prilepsky JE, Le ST, Turitsyn SK. Periodic nonlinear Fourier transform for fiber-optic communications, Part II: eigenvalue communication. Opt Express 2016;24(16):18370–81.
[27] Kamalian-Kopae M, Vasylchenkova A, Kotlyar O, Pankratova M, Prilepsky J, Turitsyn S. Artificial neural network-based equaliser in the nonlinear Fourier domain for fibre-optic communication applications. In: 2019 Conference on lasers and electro-optics Europe & European quantum electronics conference. 2019, http://dx.doi.org/10.1109/CLEOE-EQEC.2019.8871832, ci_1_4.
[28] Kamalian M, Vasylchenkova A, Shepelsky D, Prilepsky JE, Turitsyn SK. Signal modulation and processing in nonlinear fibre channels by employing the Riemann–Hilbert problem. J Lightwave Technol 2018;36(24):5714–27.
[29] Kamalian-Kopae M, Vasylchenkova A, Shepelsky D, Prilepsky JE, Turitsyn SK. Full-spectrum periodic nonlinear Fourier transform optical communication through solving the Riemann-Hilbert problem. J Lightwave Technol 2020;38(14):3602–15.
[30] Goossens J-W, Jaouën Y, Hafermann H. Experimental demonstration of data transmission based on the exact inverse periodic nonlinear Fourier transform. In: Optical fiber communication conference. Optica Publishing Group; 2019, p. M1I–6.
[31] Mollenauer LF, Gordon JP. Solitons in optical fibers: Fundamentals and applications. Elsevier; 2006.
[32] Kotlyarov V, Shepelsky D. Planar unimodular Baker-Akhiezer function for the nonlinear Schrödinger equation. Ann Math Sci Appl 2017;2(2):343–84.
[33] Chen X, Wei Z, Li M, Rocca P. A review of deep learning approaches for inverse scattering problems (invited review). Prog Electromagn Res 2020;167:67–81.
[34] Arridge S, Maass P, Öktem O, Schönlieb C-B. Solving inverse problems using data-driven models. Acta Numer 2019;28:1–174.
[35] Jones RT, Gaiarin S, Yankov MP, Zibar D. Time-domain neural network receiver for nonlinear frequency division multiplexed systems. IEEE Photonics Technol Lett 2018;30(12):1079–82.
[36] Yamamoto S, Mishina K, Maruta A. Demodulation of optical eigenvalue modulated signal using neural network. IEICE Commun Express 2019;8(12):507–12.
[37] Wu Y, Xi L, Zhang X, Zheng Z, Wei J, Du S, et al. Robust neural network receiver for multiple-eigenvalue modulated nonlinear frequency division multiplexing system. Opt Express 2020;28(12):18304–16.
[38] Mishina K, Sato S, Yoshida Y, Hisano D, Maruta A. Eigenvalue-domain neural network demodulator for eigenvalue-modulated signal. J Lightwave Technol 2021;39(13):4307–17.
[39] Zhang WQ, Chan TH, Vahid SA. Serial and parallel convolutional neural network schemes for NFDM signals. Sci Rep 2022;12(1):7962.

[40] van den Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, et al. WaveNet: A generative model for raw audio. In: 9th ISCA speech synthesis workshop. 2016, p. 125.
[41] Sedov EV, Freire PJ, Seredin VV, Kolbasin VA, Kamalian-Kopae M, Chekhovskoy IS, et al. Neural networks for computing and denoising the continuous nonlinear Fourier spectrum in focusing nonlinear Schrödinger equation. Sci Rep 2021;11(1):22857.
[42] Pelikan M, Goldberg DE, Cantú-Paz E, et al. BOA: The Bayesian optimization algorithm. In: Proceedings of the genetic and evolutionary computation conference GECCO-99. Vol. 1. Citeseer; 1999, p. 525–32.
[43] Fokas AS, Its AR. The nonlinear Schrödinger equation on the interval. J Phys A: Math Gen 2004;37:6091–114.
[44] Dhillon A, Verma GK. Convolutional neural network: a review of models, methodologies and applications to object detection. Progr Artif Intell 2020;9(2):85–112.
[45] Olver S. A Julia package for solving Riemann–Hilbert problems. 2019, https://github.com/juliaholomorphic/riemannhilbert.jl.
[46] Freire PJ, Napoli A, Ron DA, Spinnler B, Anderson M, Schairer W, et al. Reducing computational complexity of neural networks in optical channel equalization: From concepts to implementation. J Lightwave Technol 2023;1–26. http://dx.doi.org/10.1109/JLT.2023.3234327.
[47] Marcuse D, Manyuk C, Wai PKA. Application of the Manakov-PMD equation to studies of signal propagation in optical fibers with randomly varying birefringence. J Lightwave Technol 1997;15(9):1735–46.
[48] Manakov SV. On the theory of two dimensional stationary self-focusing of electromagnetic waves. Sov Phys—JETP 1974;38:248–53.
[49] Christiansen PL, Eilbeck JC, Enolskii VZ, Kozlov NA. Quasi-periodic and periodic solutions for coupled nonliner Schr''odinger equation of Manakov type. Proc R Soc Lond Ser A Math Phys Eng Sci 2000;456:2263–81.
[50] Wu L, Geng Z, He G. Algebro-geometric solutions to the Manakov hierarchy. Appl Anal 2016;95:769–800.