Water-depth identification from free-surface data using the KdV-based nonlinear Fourier transform

de Koster, P.B.J.; Brühl, M.; Wahls, S.

DOI
10.1115/OMAE2022-78882

Publication date
2022

Document Version
Final published version

Published in
Proceedings of the ASME 2022 41th International Conference on Ocean, Offshore and Arctic Engineering (OMAE 2022)

Proceedings of the ASME 2022 41th International Conference on Ocean, Offshore and Arctic Engineering
OMAE2022
June 5-10, 2022, Hamburg, Germany

OMAE2022-78882

WATER-DEPTH IDENTIFICATION FROM FREE-SURFACE DATA USING THE KDV-BASED NONLINEAR FOURIER TRANSFORM

Pascal de Koster
Delft Center of Systems and Control
Delft University of Technology
Delft, 2628 CD
Netherlands
Email: p.b.j.dekoster@tudelft.nl

Markus Brühl
Delft Center of Systems and Control
Delft University of Technology
Delft, 2628 CD
Netherlands

Sander Wahls
Delft Center of Systems and Control
Delft University of Technology
Delft, 2628 CD
Netherlands

ABSTRACT

We propose a novel method to determine the average water depth from shallow, weakly nonlinear water waves that are approximated by the Korteweg-de Vries equation. Our identification method only requires free-surface measurements from two wave gauges aligned in the direction of wave propagation. The method we propose is based on comparing solitonic components in wave packets, which are computed using the nonlinear Fourier transform (NFT) (typical time-series data often contains at least some solitonic components, even when these components are not directly visible). When the correct water depth is used for the normalisation of the wave, the solitonic components found by the NFT remain constant as the wave packet propagates, whereas any other water depth will result in solitonic components that do not remain constant. The basic idea is thus to iteratively determine the water depth that leads to a best fit between the solitonic components of time series measurements at two different gauge positions. We present a proof-of-concept on experimental bore data generated in a wave flume, where the identified water depth is within 5% of the measured value.

1 INTRODUCTION

The Korteweg-de Vries (KdV) equation closely models the propagation of progressive free-surface waves in shallow water with depth-to-wavelength ratio $h/L < 0.22$ [1,2], and finds many applications in coastal engineering [3]. The water depth is the essential governing parameter in the KdV equation. However, the water depth may not always be known, as it may be hard to measure, or is slowly changing over time. This is for example the case in wave flumes with moving-bed experiments, where the average water depth may change due to sediment transportation [4, p.108-111], or water losses due to overtopping [5, p.4]. Another example occurs for arrays of buoys in front of the coast, where the average water depth can slowly vary due to drifting of the buoys over an uneven bottom, sediment transportation, or the changing tidal elevation [6, 7].

Existing methods for water depth measurement at the coast, in waterways, and in flumes usually rely on direct measurement of water depths, using sound or light [8]. Alternatively, statistical properties such as wavelength and period may be derived from aerial photographs, which may be related to water depth as well [9]. However, the mentioned examples use very specialised devices to determine the depth, whereas these are not always present or accessible. In contrast, buoys at the coast and in waterways measuring the free surface amplitude are often already installed for more general purposes, and wave flumes are often already equipped with wave gauges, providing plentiful free surface measurements at certain fixed locations in both cases. We investigate the possibility to identify the average water depth from this type of data. It has already been shown that the bathymetry can be roughly estimated from space-time series free-surface data, given that the average water depth is approximately known [10]. However, this method used space-time data

(two-dimensional data), as opposed to just time series at two locations, and would therefore not be applicable on buoy or gauge data. A final approach for identifying the water depth is through the constants of motion of the KdV equation, analogous to a method discussed in [11]. The water depth can directly be extracted as the ratio of certain global quantities of the wave at two different positions, as shown in Appendix D. However, this method strongly depends on the pre-processing method, and is sensitive to noise and distortion.

In this paper, we propose a simple method to use the free-surface elevation data at just two locations aligned in the direction of wave propagation to determine the water depth, under the assumption that the wave propagates in constant depth and is approximately governed by the Korteweg-de Vries equation. Neither the exact distance between the gauges (e.g., in the case of buoy data) nor precisely time-synchronised measurements are required in our proposed method, as long as the measurements are of the same progressive free-surface waves. The method we propose is based on the nonlinear Fourier transform (NFT) for the KdV equation (KdV-NFT) [12]. In this paper, we will demonstrate a proof-of-concept by focussing on wave packets, so that the NFT with vanishing boundary conditions can be applied.

The NFT decomposes a wave packet into a discrete spectrum and a continuous spectrum, representing two types of waves. The discrete spectrum represents translatory stable waves (solitons), while the continuous spectrum represents dispersive oscillatory wave components (radiation). If a free-surface wave in constant depth evolves perfectly according to the KdV equation, both the solitonic components and the amplitudes of the continuous spectrum remain constant during propagation. As the water depth governs the propagation of a wave packet, the water depth is also the governing parameter in the NFT and its resulting spectra. Therefore, the basic idea is to iteratively determine the water depth that leads to a best fit between the nonlinear Fourier spectra of time series measurements at two different gauge positions. For our measurements, most of the energy was contained in the solitonic components, and we will therefore focus on matching the discrete spectra in this paper. The value at which the NFT spectra fit best is then identified as the optimal water depth, under the assumption that the water depth between the gauges is constant.

Although NFT-based identification has not been applied for water depth identification as of yet to the best of our knowledge, we already showed in earlier work that NFT-based parameter identification is possible for optical fibre systems [11, 13]. Light propagation through optical fibres is governed by the nonlinear Schrödinger equation (NLSE), which also allows for an (NLSE-based) nonlinear Fourier transform. Similar to the KdV-NFT, the NLSE-NFT decomposes a signal into solitonic components and dispersive components, allowing us to use similar approaches from [11] in this research.

The structure of this paper is as follows. First, we present the KdV model and its nonlinear Fourier transform. Second, we propose our novel water-depth identification algorithm, based on comparing solitonic components. Third, we validate the identification method on simulated and experimental data. Finally, we conclude the paper.

2 THE KDV EQUATION AND THE NONLINEAR FOURIER TRANSFORM

The development of long unidirectional progressive free-surface waves in shallow water is modelled by the time-like Korteweg-de Vries equation [14],

$$\eta_l + c_0'\eta_\tau + \alpha'\eta\eta_\tau + \beta'\eta_{\tau\tau\tau} = 0, \tag{1}$$

where $\eta(t,s)$ [m] denotes the free-surface elevation compared to the still-water level, l [m] the position, τ [s] the time, c_0' [s/m] the inverse wave celerity, α' [s/m^2] the nonlinearity coefficient and β' [s^3/m] the dispersion coefficient. Subscripts denote partial derivatives. For progressive free-surface waves, the governing coefficients depend only on the water depth [15]:

$$\begin{aligned} c_0 &= \sqrt{gh}, & \alpha' &= -\frac{3}{2hc_0} = -\frac{3}{2h\sqrt{gh}}, \\ c_0' &= \frac{1}{c_0} = \frac{1}{\sqrt{gh}}, & \beta' &= -\frac{h^2}{6c_0^3} = -\frac{h^2}{6(\sqrt{gh})^3}, \end{aligned} \tag{2}$$

where c_0 [m/s] is the shallow water wave celerity, c_0' [s/m] the inverse wave celerity, g [m/s^2] the gravitational acceleration, and h [m] the still-water depth.

The idea of this paper is to identify the still-water depth at which the KdV equation best relates free surface measurements at consecutive wave gauges. One option would be to numerically propagate the signal at the first gauge to the second gauge using Eqn. (1) for different water depths, and keep the water depth at which the simulated and measured signal at the second gauge fit best. However, we instead consider the nonlinear spectrum of the signals, which has the advantage that it does not depend on the distance travelled by the wave or the time synchronisation of the signals. This property is especially useful when the distance between measurement devices is not known. Finally, this method has the advantage that the equation does not have to be solved with time-stepping methods, which increase in computation time as the distance increases.

2.1 Normalisation

The nonlinear Fourier transform is often determined from the normalised and dimensionless KdV equation:

$$q_x + 6qq_t + q_{ttt} = 0, \tag{3}$$

in which $q(t,x)$ is the normalised free-surface elevation, t the normalised time and x the normalised location. To obtain the normalised KdV in Eqn. (3) from the dimensional one in Eqn. (1), we apply the following change of variables:

$$t = \frac{1}{T_0}(\tau - c_0' l), \quad x = \frac{1}{T_0^3} c_x l, \quad q(t,x) = T_0^2 c_q \eta(\tau, l), \tag{4a}$$

$$\text{with} \quad c_x = \beta' = -\frac{\sqrt{h}}{6g^{3/2}}, \qquad c_q = \frac{\alpha'}{6\beta'} = \frac{3g}{2h^2}. \tag{4b}$$

The amplitude-normalisation coefficient c_q and space-normalisation coefficient c_x both only depend on the water depth. Note that the amplitude normalisation coefficient c_q represents the ratio of the nonlinearity α' versus the dispersion β'. The time normalisation T_0 [s] is entirely free to choose, and will only linearly scale the signal, its linear Fourier spectrum and its NFT spectrum. For simplicity, we will use $T_0 = 1$. We also note here that the value of c_0' is not strictly necessary for the normalisation: using a wrong c_0' only translates the signal without influencing its shape, and therefore does not influence the solitonic components or amplitudes of the continuous spectrum of the NFT (see Property 3 in Appendix B). Throughout this paper, we will therefore use $c_0' = 0$ for the normalisation step. In this paper, we will first identify the amplitude normalisation coefficient c_q (i.e., the ratio between nonlinear and linear effects), and then extract h from c_q:

$$h = \sqrt{\frac{3g}{2c_q}}. \tag{5}$$

2.2 Solitons and the nonlinear Fourier transform

It is well known that the KdV equation supports both dispersive waves and waves of translation, so called solitons. The energy of dispersive wave components will spread out more and more over time, but each soliton remains localised, and will be visible indefinitely. More specifically, any normalised wave packet $q(t)$ will evolve into a train of $N \geq 0$ solitons after sufficiently long time [16, p. 83]:

$$q(t,x) \approx \sum_{n=1}^{N} 2k_n^2 \operatorname{sech}^2\left(k_n(t - 4k_n^2 x - t_n^0)\right), \tag{6}$$

in which t_n^0 is a time shift depending on the initial conditions. Most importantly, the height $(2k_n^2)$, the width $(1/k_n)$ and the celerity $(4k_n^2)$ of a soliton are all determined by a generalised wave number k_n.

Although the detection of solitons and their k_n is straightforward when all solitons have separated, it is not clear which solitons will come out of some arbitrary wave packet. However, the NFT is able to precisely determine which solitons are present in any wave packet. If we normalised the KdV equation correctly, the same solitonic components will remain present in a wave packet during propagation. All solitons present in a signal are represented by the discrete spectrum of the NFT, which is a set of pairs of purely imaginary eigenvalues $\lambda_n = ik_n$ (directly related to the generalised wave number of each soliton in Eqn. 6) and residues r_n (indirectly related to the location of each soliton):

$$\Lambda^{\text{ds}} = \{(\lambda_n, r_n), n = 1, \ldots, N\}. \tag{7}$$

Assuming propagation according to Eqn. (3), the eigenvalues remain constant, and the residues grow exponentially with the travelled distance

$$\lambda_n(x = X) = \lambda_n(x = 0), \quad r_n(x = X) = r_n(x = 0)e^{8i\lambda_n^3 X}. \tag{8}$$

We used the *FNFT*-software library [17] to determine the eigenvalues and residues from the normalised signal. Further details on the exact definition and calculations of the NFT may be found in Appendix A. Some relevant properties of the NFT are discussed in Appendix B.

The NFT can also extract the oscillatory wave components from a wave packet, represented by a continuous spectrum. However, we will only focus on the discrete spectrum, as we find that the experimental signals considered in this paper are strongly soliton-dominated.

3 SPECTRAL MATCHING ALGORITHM

As stated in the previous section, the eigenvalues in the nonlinear Fourier spectrum of a KdV-governed signal remain constant, given that the correct amplitude normalisation constant c_q was applied. As c_q relates directly to the water depth h, our strategy is to consider a wave packet at two consecutive wave gauges (an 'input' and an 'output' gauge), normalise both signals with a certain $c_q(h)$ and compare the eigenvalues of their NFT spectrum. The value of c_q is iteratively adapted until one is found at which the NFT eigenvalues at the two wave gauges match optimally. This identified normalisation constant c_q^{ID} is then converted to the identified water depth h^{ID} using Eqn. (4b).

To quantify the error between the NFT eigenvalues of the two signals, we first sort the eigenvalues in each spectrum from largest to smallest imaginary part, $k_1 > k_2 > \cdots > k_N > 0$. We then match the highest input eigenvalue λ_1^{in} to the highest output eigenvalue λ_1^{out}, the second highest to the second highest, and so on. If either spectrum contains more eigenvalues than the other ($N^{\text{in}} \neq N^{\text{out}}$), the remaining (lowest) eigenvalues are matched to artificial '0-eigenvalues' at the origin. We finally considered two

possibilities for the error-norm: $p = 1$, corresponding to the absolute difference between input and output eigenvalues; or $p = 3$, proportional to the absolute difference between the *energy* $\mathscr{E}_n$ of input and output eigenvalues ($\mathscr{E}_n \propto k_n^3$, see Eqn. (23b) in Appendix B). Precisely, we define the error as

$$E^p = \frac{\sum_n \left| \left(k_n^{\text{in}}\right)^p - \left(k_n^{\text{out}}\right)^p \right|}{\sum_n \left(k_n^{\text{in}}\right)^p + \sum_m \left(k_m^{\text{out}}\right)^p}, \quad p \in \{1,3\}. \tag{9}$$

The error is normalised, such that the maximum possible error is 1. Due to the observed continuity of the eigenvalues as c_q is varied, and the fact that eigenvalues can only (dis)appear at the origin (see Appendix B), the error is also continuous in c_q.

We finally identify the optimal c_q^{ID} and corresponding water depth h^{ID} by applying local minimisation over c_q using Eqn. (9) as cost function. To obtain an initial starting position, we perform a rough grid search over realistic values of h, and take the one with lowest error.

This concludes the spectral matching algorithm that we will apply in this paper. We note that the spectral matching algorithm is somewhat similar to the identification method from global conserved quantities in Appendix D, as the eigenvalues can also be considered as conserved quantities. As shown in Appendix D, the method of conserved quantities can give good results, while being easy to implement. However, it should be kept in mind that the estimate of the method is not complemented by an error, so it cannot be deduced how reliable the estimate is. The method may also be sensitive to noise and pre-processing, and can give biased results in some cases [11]. The NFT-based algorithm is therefore often the more reliable choice.

Finally, we mention here that other parts of the NFT spectrum can also be taken into account for identifying the water depth, namely the continuous spectrum and the residues of the discrete spectrum. The moduli of the continuous spectrum of the NFT (representing oscillatory wave components) also remain constant during propagation, and could therefore be compared as well. As mentioned before however, the continuous spectrum of our experimental data contained too little energy to test continuous-spectrum-based identification. On the other hand, the residues of the discrete spectrum can be of use, as discussed below.

3.1 Extension using residues

Although the proposed spectral-matching algorithm only uses the eigenvalues, each eigenvalue is complemented by a residue containing information relating to its position. As each soliton moves with its own speeds, so too do the residues grow with different speeds as the wave packet propagates, as shown in Eqn. (8). We may thus check how well the solitons match by checking if their residues indeed grew proportionally to the distance between the wave gauges (if known). Alternatively, after the water depth was identified using only the eigenvalues, we may estimate the distance between the input and output gauges using the residues, and compare how well this matches the measured distance.

Instead of considering the residues directly, we will convert each residue to a 'soliton (time-)location', t_n, which we define as the position of the soliton peak in the case the wave packet was a pure single-soliton (the presence of multiple solitons and continuous spectrum cause shifts, as well as that no clear peak may be distinguishable when multiple solitons are close [18, 19]). The propagation speed of this soliton location is equal to the soliton celerity $c_n = 4k_n^2$:

$$t_n(x) = \frac{1}{2k_n} \ln\left(\frac{r_n(x)}{2ik_n}\right), \tag{10}$$

$$t_n(x+X) = t_n(x) + \underbrace{4k_n^2}_{=c_n} X. \quad \text{(propagation relation)} \tag{11}$$

Recall that the normalised position is given by $x = \frac{1}{T_0^3} c_x l$, thus the normalised travelled distance is given by $X = \frac{1}{T_0^3} c_x L$, with L the physical distance between the wave gauges.

Finally, the normalised frames of the wave packets at input and output will experience some (unknown) time shift t_0 due to the wave celerity c_0' and due to a possible synchronisation mismatch. The influence of this t_0 is the same for all soliton locations, and may thus be taken into account with one t_0 for all pairs (k_n, t_n):

$$t_n^{\text{out}} - t_n^{\text{in}} = 4(k_n^{\text{in}})^2 \frac{c_x}{T_0^3} L + t_0. \tag{12}$$

The propagated distance L may thus be identified by fitting both L and t_0 simultaneously in a least-squares approach to the shifts in soliton locations of all solitons, when there are at least two solitons present.

4 RESULTS

Within this section, we demonstrate the proposed water-depth identification algorithm using free-surface measurements from a wave flume at two different wave gauges. Furthermore, we also test the identification algorithm on simulation data, where the wave packet at one wave gauge was taken as input, and then numerically propagated with the KdV equation to the position of a consecutive wave gauge.

4.1 Experimental setup

The experimental data used in this study was measured at the Hydraulic Engineering Laboratory at the National University of

Singapore (NUS). The data was originally generated for a different study [19], and shared with us for the analysis in this paper. The experimental setup is shown in Fig. 1. The flume is 0.9 m wide, 0.9 m deep, and 36 m long. At the left-end side is a movable piston-type wave maker. After 28.7 m, the flume bottom slopes upward. Four wave gauges marked CG1 to CG4 (Capacity Gauge) are present, with respective positions $x_1 = 5.193$ m, $x_2 = 9.886$ m, $x_3 = 14.882$ m, and $x_4 = 18.869$ m. The wave gauges consist of two thin metal wires, which measure the water-height dependent voltage difference. The voltage is then linearly converted to the wave height. Measurements were taken every 0.05 s, for a duration of 90 s. Initially, the water was at rest with measured depth $h = 9.80$ cm. Then, the wave maker was moved from $x = 0$ m to $x = 4.0$ m in approximately 3 s to generate a trapezoidal-shaped wave with a relative height of 2.0 cm. Note that the wave maker remained at $x = 4.0$ m, so the filled part of the flume became shorter, and the still water level would increase after all waves have died out. This takes a very long time, and for the course of the experiment, we took the initial still water level $h^{\text{ref}} = 9.80$ cm as reference level.

The measurements of the free-surface elevation are shown in Fig. 1. Initially, we observe the initial trapezoidal-shaped wave at CG1, right behind the wave maker. As the wave travels, its fastest solitons start to separate from the bore, and reach the next gauges first. Travelling at approximately $c_0 = 1$ m/s, the front of the wave reaches the right end of the flume around $t = 28$ s, reflects imperfectly from the sloped end, and propagates to the left, reaching the gauges in opposite order. Around $t = 51$ s, the (now left-going) wave reflects nearly perfectly at the piston at the left side of the flume, and starts moving to the right again. The second reflection and first reflection interfere at most of the wave gauges, contradicting the assumption of unidirectional waves in the KdV approximation.

We will compare only the incoming waves at the wave gauges to identify the water depth, as these measurements are not affected by interference. We compare several combinations to investigate the influence of the travelled distance on the identified water depth: if two gauges are very close, the waves are very similar, and thus the influence of the water depth will be small, making it harder to identify; for longer distances, the accumulated effect of the water depth will be greater, and thus the water depth should be easier to identify.

4.2 Pre-processing

For the pre-processing of the data, we extracted a time window as large as possible without being distorted by the reflected wave. We truncated the signals measured at CG1 to CG4 to the time interval $t \in [0, 34]$. We note here that the NFT assumes zero-boundary conditions at the left and right of the signals. Although the signal smoothly goes to zero at $t = 0$, this is not the case at $t = 34$, where the new steady water level was measured

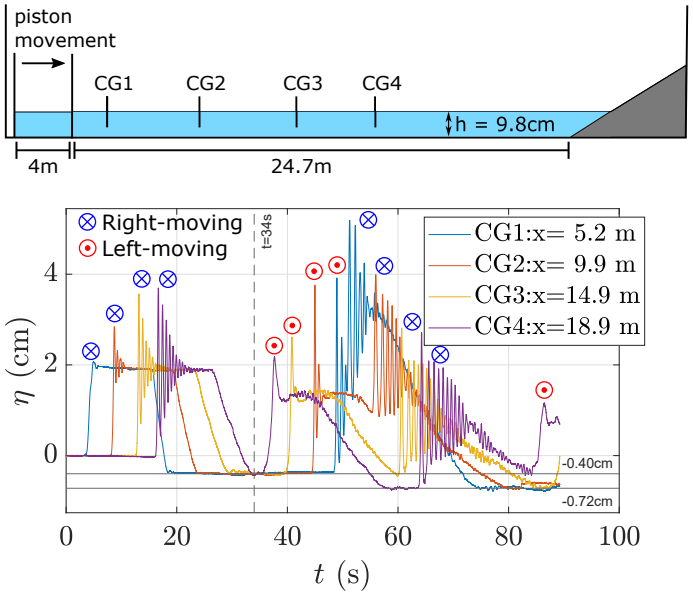

FIGURE 1. TOP: A SIDE VIEW OF THE EXPERIMENTAL SETUP, WITH THE POSITIONS OF THE FOUR WAVE GAUGES. BOTTOM: THE MEASURED WAVE HEIGHT WITH RESPECT TO INITIAL STILL-WATER LEVEL AT THE FOUR WAVE GAUGES. THE VERTICAL DASHED LINE MARKS THE TIME UP TO WHICH THE WAVE GAUGES ARE NOT AFFECTED BY THE REFLECTED WAVE. RE-MADE AFTER [19].

to be -0.40 cm. The block signal is followed by a trough, which does not seem to go back to the initial still-water level within the time frame of the experiment (in fact the water level seems to decrease further down to -0.72 cm after the reflected wave has passed over). Fortunately, it was observed in [19] that the presence or absence of a trough mainly influences the continuous spectrum, while leaving the discrete spectrum nearly unchanged. We confirmed that this is indeed the case, by comparing the discrete spectra of the signal at CG1 with and without trough (see Appendix C).

As the trough influences the discrete spectrum only marginally, we cut off the right part of the signal at the time that the tail of the wave packet reaches zero for the first time, and replace the trough with zeros, following the method of [19]. This ensures that the vanishing boundary conditions are satisfied without a jump in water level at the right hand side.

Finally, we perform a small rescaling of the data, similar to the pre-processing in [19]. We rescale the data by comparing the first moment $\int_{-\infty}^{\infty} \eta \, dt$ and the second moment $\int_{-\infty}^{\infty} \eta^2 dt$ of the wave packets. These two integrals are conserved quantities of the KdV equation [20], and thus should be equal at all four wave gauges. Cutting off the trough is important for this recalibration, as both moments would otherwise have depended on the considered length of the trough. We rescaled the data such

that the second moment (proportional to the wave energy) of all signals are equal to the one at the first wave gauge, resulting in the following rescaling factors for the free-surface elevation for CG1 to CG4: $[1, 0.990, 0.989, 1.013]$. Although these changes are small, we found that they significantly improve the matching of the discrete spectra and the estimation of h.

After the pre-processing of the data, we calculated the energy in the discrete and the continuous NFT spectrum for the signal at CG1, to check our assumption that the signal is soliton dominated. Using $h^{\text{ref}} = 9.80\,\text{cm}$ for the normalisation in the NFT, we found from Eqn. (23b) in Appendix B that the discrete spectrum contained 99.9% of the signal energy, satisfying our assumption.

4.3 Numerical validation of the algorithm

Before considering the experimental data, we first validate the spectral matching algorithm on simulated data, generated using the KdV equation. We took the pre-processed data at wave gauges CG1, CG2 and CG3 as input data, and numerically propagated each input signal to the position CG4, using the dimensional KdV equation in Eqn. (1) with h^{ref} for all coefficients. For this numerical case, the eigenvalues should remain constant when normalising the signal with h^{ref}, and the spectral matching error should be zero up to numerical errors. Fig. 2 shows from top to bottom the input-output data (the blue and black lines), the corresponding eigenvalues as function of h, and the spectral matching error as function of h. From the normalised eigenvalues we observe that as h decreases (i.e., higher nonlinearity), the number of detected solitons increases, while the eigenvalues all drift upwards. We observe consistently that eigenvalues of the input signal (blue) and the eigenvalues of the simulated output (black) exactly coincide at $h = h^{\text{ref}}$, while drifting apart for other values of h. As a result, the spectral matching errors E^1 and E^3 show a clear minimum at h^{ref} for the simulated data set, while the errors increase as h is further from the reference value. This validates that the spectral-matching algorithm performs well when the considered data is exactly governed by the KdV equation.

However, we observe from the pre-processed signals in Fig. 2 that the numerical output signals (black) differ somewhat from the experimental signals (red). The KdV only models the wave propagation approximately, and the model mismatch becomes apparent here. In particular, the KdV seems to develops the shape of the input wave too fast, which can also be visually deduced from the fact that the highest numerical soliton arrives earlier than the experimental soliton, and the distance between the solitons is larger than in the measured wave. Also the right tail of the numerical wave packet flattens faster than the tail the measured output.

Although the KdV does not perfectly describe the wave propagation of our experimental data, it was shown in [19] that the solitonic components predicted by the KdV-NFT were indeed present. Therefore, our proposed method of comparing solitonic components at different gauges may still perform well despite the observed model mismatch.

4.4 Experimental results of NFT-based water-depth identification

We now consider the spectral matching algorithm for the pre-processed experimental data. First, we determined the E^1-error and E^3-error over a grid, as shown in Fig. 2. We observe that the E^3-error seems to be much smoother than the E^1-error. This is due to the appearance of new eigenvalues, that can quickly grow in amplitude and contribute significantly to the error. The energy-based error E^3 overcomes this drawback, as the energy of these new eigenvalues is only very small. Due to the smoothness of E^3, we will use this error norm for determining our final estimation of the water depth for each of the data sets. After the grid search, we apply local minimisation of E^3 to identify the water depth h^{ID} at which E^3 is minimised.

For the experimental data, the optimal matching with the input eigenvalue is close to the reference value for all considered cases. In particular, the minimum error for the CG2-CG4 case was extremely low, and corresponds very well to the reference value. The largest error observed (for the CG1-CG4 case, $h^{\text{ID}} = 9.34\,\text{cm}$) was less than 5% off the reference value. When using the reference water depth, we found that for all signals the highest eigenvalues matched well, except for CG1-CG4. Our hypothesis is that the wave at CG1 is still influenced by the deceleration of the piston, as it is only 119 cm behind the piston. This deceleration may have caused a larger mismatch with the KdV equation during the first few meters of propagation, causing the spectra at the first gauge to be different from the other gauges.

Next, we also observe that the identified water depth for the CG3-CG4 case ($h^{\text{ID}} = 10.10\,\text{cm}$) is also relatively far off (3%). This is probably due to the fact that these gauges are very close, so the total influence of the water depth on the propagation becomes harder to measure. This may also be observed from the fact that the error well for the CG3-CG4 case is less steep than in the other cases.

The CG2-CG4 case provides the best setup: it features a sufficiently long distance, while not being influenced (too much) by the start-up interference due to the piston. As a result, we observe that the minimum observed E^3 is very close to 0, indicating an excellent match between input and output eigenvalues. The identified water depth ($h^{\text{ID}} = 9.70\,\text{cm}$) is only 1% off the reference depth. This validates that the water depth can be identified very accurately when using suitable data. However, even when our data contained more distortion or was measured at close gauges, we still managed to get estimates with less than 5% error.

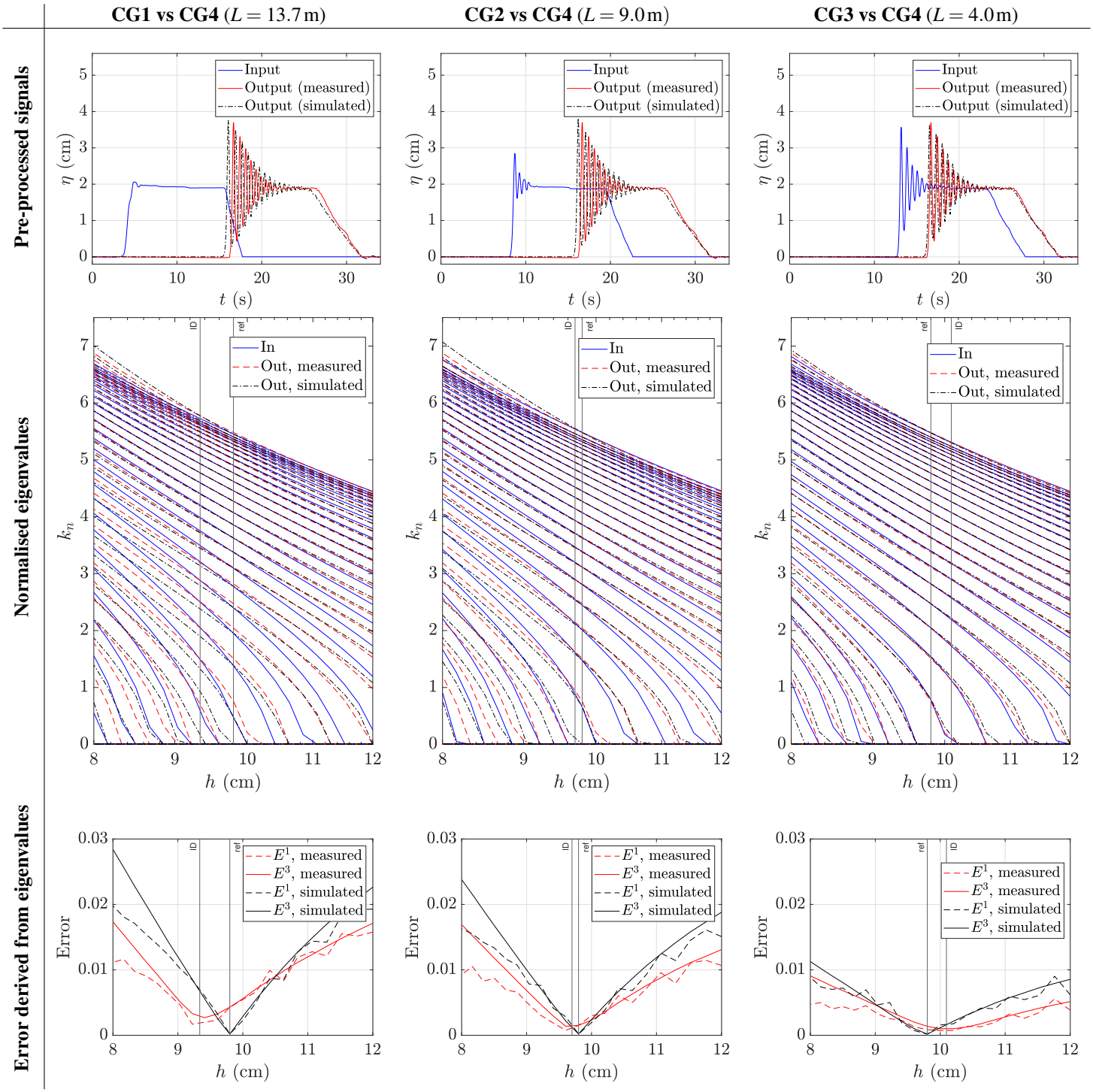

FIGURE 2. THE USED DATA FOR THE IDENTIFICATION ALGORITHM. TOP: THE PRE-PROCESSED DATA AT VARIOUS WAVE GAUGES, AND THE THEORETICAL OUTPUT WHEN THE INPUT SIGNAL WAS PROPAGATED ACCORDING TO THE KDV AT THE REFERENCE WATER DEPTH. MIDDLE: THE NORMALISED EIGENVALUE ENERGIES ($\sim k_n^3$) FOR VARIOUS WATER DEPTHS. BOTTOM: THE ERROR BETWEEN THE INPUT AND OUTPUT EIGENVALUES, AND THE IDENTIFIED WATER DEPTH USING LOCAL MINIMISATION OF E^3.

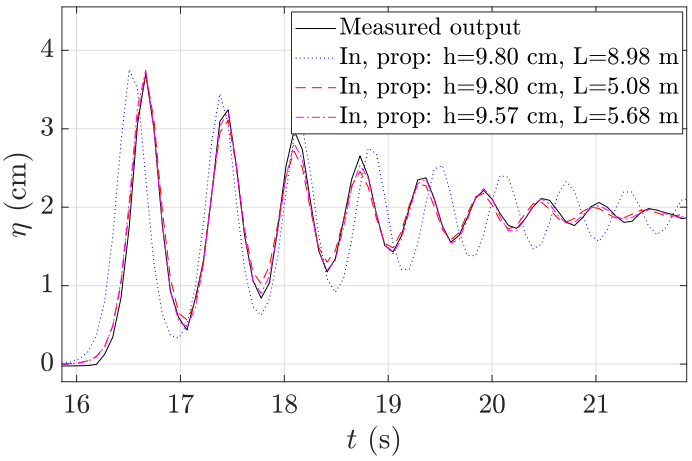

FIGURE 3. THE BORE FRONT OF THE MEASURED OUTPUT AT CG4, COMPARED TO THE KDV-PROPAGATED SIGNAL FROM CG2 FOR VARIOUS COMBINATIONS OF DISTANCE L AND WATER DEPTH h, WHILE TIME-SYNCHRONISING AS WELL AS POSSIBLE. THE SECOND CURVE USES THE REFERENCE VALUES FOR h AND L, THE THIRD CURVE IS THE RESULT OF LOCAL OPTIMISATION FOR THE TRAVELLED DISTANCE, AND THE FOURTH CURVE FOR SIMULTANEOUSLY OPTIMISING FOR BOTH THE DISTANCE AND THE WATER DEPTH.

4.5 Validation of NFT-based identification using numerical propagation

To further validate the NFT-based water depth identification method, we compare it to an approach based on numerically forward propagating the input wave to the position of the output gauge, using the dimensional tKdV from Eqn. (1) for various values of h. Although this method lacks the benefits from NFT-based identification (solution depends on numerical space-step size, propagation distance must be known or has to be identified as well, time synchronisation is often required), it is one of the most straightforward approaches to find the governing water depth in the KdV equation.

This method propagates the input signal for a distance L (space step size $\Delta l = 0.02\,\mathrm{m}$) to the position of the output wave gauge using the dimensional KdV in Eqn. (1), for different values of h. We then identify the water depth as the h at which the forward-propagated input signal $\eta^{\text{in, prop}}(t;h,L)$ matches $\eta^{\text{out}}(t)$ as well as possible. As it may often occur that the time measurements at wave gauges are not properly synchronised, we will only consider the shapes of the wave packets, and allow for a horizontal time-translation t_0. Our error-norm for this propagation-based matching is

$$E^{\text{prop}}(h,L) = \min_{t_0} \frac{\int_{-\infty}^{\infty} \left|\eta^{\text{in, prop}}(t;h,L) - \eta^{\text{out}}(t-t_0)\right| \mathrm{d}t}{\int_{-\infty}^{\infty} |\eta^{\text{out}}(t)|\,\mathrm{d}t}. \quad (13)$$

From Fig. 2, we have already observed that the simulated output using h^{ref} is already quite different form the measured output. We observed that this was caused by a model mismatch, where the KdV equation would develop the shape of the wave faster than the measured wave. Therefore, we will optimise over both the water depth and the travelled distance L. We consider the case with CG2-CG4, as this gave the best fit for the eigenvalue-matching, and we thus expect that this dataset fits KdV-propagation best, although probably for a different propagation distance.

The result is shown in Fig. 3. As observed before, the measured signal and the signal propagated with $h_{\text{ref}} = 9.8\,\mathrm{cm}$ and $L = 9\,\mathrm{m}$ do not fit well, as the solitons have already separated too much, indicating that the simulated wave has indeed developed too much. The third line was optimised only for the travelled distance, and shows a good fit, but an effective propagation distance of only $L = 5\,\mathrm{m}$. Optimising over both the depth and the propagation distance results in the fourth line, and we identify $h = 9.57\,\mathrm{cm}$ and $L = 5.68\,\mathrm{m}$. This identified water depth is only 2% off the measured value, and in close correspondence with the identified value of NFT-based matching ($h = 9.70\,\mathrm{cm}$). For CG1-CG4 we identified $h = 9.33\,\mathrm{cm}$, which is also in close correspondence with the value from NFT-based identification ($h = 9.34\,\mathrm{cm}$), which shows that NFT-based matching indeed gives similar results as the propagation-based method. Only for the CG3-CG4 case, the identified values were further apart, $h = 9.20\,\mathrm{cm}$ for the propagation-based method, and $h = 10.10\,\mathrm{cm}$ for the NFT-based method. This difference may be explained due to the short distance between the wave gauges, so the effects of the underlying mechanics are hard to measure.

All together, KdV-propagation-based water depth identification indeed gives similar results as NFT-based water-depth identification, when also allowing the distance and the time-offset as parameters in the propagation-based matching. However, we did observe that the NFT-based estimates were all slightly closer to the measured water depth, as well that the propagation-based method required careful considerations regarding the numerical instabilities and the space-step size. Although we did not pay special attention to computation time, we do mention here that the NFT-based method was somewhat faster than the propagation-based method using the KdV-solver from the software *Chebfun* [21].

4.6 Validation of NFT-based identification using discrete spectrum residues

The NFT-based method managed to identify the water depth with at most 5% error by only using the eigenvalues. However, it did not take the soliton location into account. In this section, we show that the soliton locations can be used for validation purposes as well, for example to identify the travelled distance with the propagation relation in Eqn. (12). During the propagation-

based identification, we observed that the identified propagation distance was about 30% lower than expected when only considering the wave shape. We show here that the soliton locations indicate a similar result.

We consider here only the case CG2-CG4, as this data showed the closest agreement with the KdV model. Fig. 4 shows the calculated soliton locations at the input (CG2) and at the output (CG4). As expected, the solitons arrived later at the output gauge than at the input gauge, and thus their t_n are higher. The right figure shows the difference in soliton locations between the input and output gauge. We observe that the solitons with higher k_n moved faster in general, as their soliton locations required less time to cover the distance between CG2 and CG4 (the fastest required only 7.8 s, the slowest 9.1 s). Against expectation however, we observe that the theoretically fastest (with the largest k_n) has not moved faster than some other solitons. This could be due to random effects or a slight model mismatch. However, the general trend looks in correspondence with Eqn. (12): each soliton location should have moved ahead of the wave packet quadratically in its generalised wave number k_n.

We fitted the difference in soliton location for L and t_0 according to Eqn. (12), using all soliton locations except for those of the lowest three solitons, as these contain little energy and are prone to noise. We identified a propagation distance of $L = 6.4\,\mathrm{m}$, which is significantly lower than the actual wave gauge distance of 9.0 m for the CG2-CG4 case. We emphasise that this distance takes only takes into account how much the shape of the packet has developed, which is unrelated to the wave speed c_0 or inverse wave speed c_0'. We note that the identified value of $L = 6.4\,\mathrm{m}$ is close to value identified by the propagation based method ($L = 5.68\,\mathrm{m}$). Also, upon performing a similar analysis for CG1-CG4, and for CG3-CG4, we find that the identified propagation distance is systematically about 30% lower than the measured gauge distance. Apparently, the KdV equation does provide the correct solitons, but overestimates the wave development speed.

5 CONCLUSION

We proposed a method to identify the water depth from wave packets approximately governed by the KdV equation, based on the nonlinear Fourier transform of free-surface measurements. By comparing the solitonic wave components in a wave packet in a wave flume at two consecutive locations, we were able to determine the water depth with at most 5% error, and only a 1% error under the most suitable circumstances. Furthermore, when using simulated data, the identified water depth matched the reference value almost exactly for all considered cases, and outperformed a method based on numerical propagation of the KdV equation. We finally demonstrated that the residues of the NFT spectrum could also be used to extend the identification method. Further research could focus on validation and application of the method outside a lab environment, such as water-depth identification from coastal buoy data.

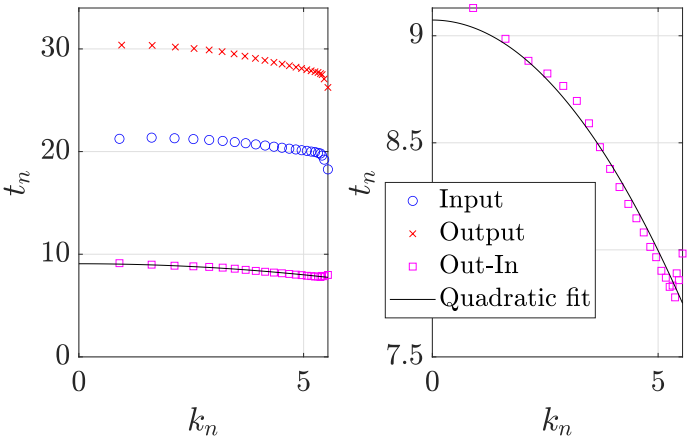

FIGURE 4. THE SOLITON LOCATIONS ACCORDING TO EQS. 10, AND THE QUADRATIC FIT BETWEEN THE DIFFERENCE (ZOOMED IN AT THE RIGHT).

ACKNOWLEDGMENT

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 716669).

The experimental data have been provided by P. L.-F. Liu and I. Barranco from National University of Singapore. The experiments have been conducted with support from the National Research Foundation of Singapore through a grant to the National University of Singapore (Grant No. NRF2018NRF-NSFC003ES-002).

REFERENCES

[1] Dingemans, M. W., 1997. *Water wave propagation over uneven bottoms: Linear wave propagation*, Vol. 13. World Scientific.

[2] Hammack, J. L., and Segur, H., 1974. "The Korteweg-de Vries equation and water waves. Part 2. Comparison with experiments". *Journal of Fluid mechanics,* ***65***(2), pp. 289–314.

[3] Crighton, D., 1995. "Applications of KdV". In *KdV'95*. Springer, pp. 39–67.

[4] Frostick, L. E., McLelland, S. J., and Mercer, T. G., 2019. *Users guide to physical modelling and experimentation: Experience of the HYDRALAB network*. CRC Press.

[5] Williams, H. E., Briganti, R., Romano, A., and Dodd, N., 2019. "Experimental analysis of wave overtopping: A new small scale laboratory dataset for the assessment of uncer-

tainty for smooth sloped and vertical coastal structures". *Journal of Marine Science and Engineering,* ***7***(7), p. 217.
[6] Jacob, B., and Stanev, E. V., 2021. "Understanding the impact of bathymetric changes in the German Bight on coastal hydrodynamics: One step toward realistic morphodynamic modeling". *Frontiers in Marine Science,* ***8***, p. 576.
[7] Grabemann, I., Gaslikova, L., Brodhagen, T., and Rudolph, E., 2020. "Extreme storm tides in the German Bight (North Sea) and their potential for amplification". *Natural Hazards and Earth System Sciences,* ***20***(7), pp. 1985–2000.
[8] Mitchell, S., Thayer, J. P., and Hayman, M., 2010. "Polarization lidar for shallow water depth measurement". *Applied optics,* ***49***(36), pp. 6995–7000.
[9] Polcyn, F., Brown, W. L., and Sattinger, I., 1970. The measurement of water depth by remote sensing techniques. Tech. rep., Michigan University Ann Arbor intitute of Science and Technology.
[10] Vasan, V., Manisha, and Auroux, D., 2021. "Ocean-depth measurement using shallow-water wave models". *Studies in Applied Mathematics,* ***147***(4), pp. 1481–1518.
[11] de Koster, P., and Wahls, S., 2020. "Dispersion and nonlinearity identification for single-mode fibers using the nonlinear Fourier transform". *Journal of Lightwave Technology,* ***38***(12), pp. 3252–3260.
[12] Osborne, A. R., 2002. "Nonlinear ocean wave and the inverse scattering transform". In *Scattering*. Elsevier, pp. 637–666.
[13] de Koster, P., and Wahls, S., 2022. "Experimental validation of nonlinear Fourier transform-based Kerr-nonlinearity identification over a 1600 km SSMF link". In Optical Fiber Communications Conference (OFC), to appear.
[14] Korteweg, D. J., and De Vries, G., 1895. "On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves". *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* ***39***(240), pp. 422–443.
[15] Osborne, A., 1991. "Nonlinear Fourier analysis for the infinite-interval Korteweg-de Vries equation I: An algorithm for the direct scattering transform". *Journal of Computational Physics,* ***94***(2), pp. 284–313.
[16] Schuur, P. C., 2006. *Asymptotic analysis of soliton problems: an inverse scattering approach*, Vol. 1232. Springer.
[17] Wahls, S., Chimmalgi, S., and Prins, P. J., 2018. "FNFT: a software library for computing nonlinear Fourier transforms". *Journal of Open Source Software,* ***3***(23), p. 597.
[18] Prins, P. J., and Wahls, S., 2019. "Soliton phase shift calculation for the Korteweg-de Vries equation". *IEEE Access,* ***7***, pp. 122914–122930.
[19] Bruehl, M., Prins, P., Ujvary, S., Barranco, I., Wahls, S., and Liu, P., 202. "Comparative analysis of bore propagation over long distances using conventional linear and KdV-based nonlinear Fourier transform". *UNDER REVIEW,* ***1***(1), pp. 1–50.
[20] Miura, R. M., Gardner, C. S., and Kruskal, M. D., 1968. "Korteweg-de Vries equation and generalizations. II. existence of conservation laws and constants of motion". *Journal of Mathematical physics,* ***9***(8), pp. 1204–1209.
[21] Driscoll, T. A., Hale, N., and Trefethen, L. N., 2014. Chebfun guide.
[22] Gardner, C. S., Greene, J. M., Kruskal, M. D., and Miura, R. M., 1967. "Method for solving the Korteweg-deVries equation". *Physical review letters,* ***19***(19), p. 1095.
[23] Lax, P. D., 1968. "Integrals of nonlinear equations of evolution and solitary waves". *Communications on pure and applied mathematics,* ***21***(5), pp. 467–490.
[24] Ablowitz, M. J., Kaup, D. J., Newell, A. C., and Segur, H., 1974. "The inverse scattering transform-Fourier analysis for nonlinear problems". *Studies in Applied Mathematics,* ***53***(4), pp. 249–315.
[25] Zakharov, V. E., and Faddeev, L. D., 1971. "Korteweg-de Vries equation: A completely integrable Hamiltonian system". *Funktsional'nyi Analiz i ego Prilozheniya,* ***5***(4), pp. 18–27.
[26] Ablowitz, M. J., and Segur, H., 1981. *Solitons and the inverse scattering transform*. SIAM.
[27] Lax, P. D., 1975. "Periodic solutions of the KdV equation". *Communications on pure and applied mathematics,* ***28***(1), pp. 141–188.

Appendix A: Definition and calculation of the NFT spectrum

Given a system governed by the normalised KdV equation from Eqn. (3), we may determine the NFT while assuming vanishing boundary conditions for the wave packet, $q(t) \to 0$ as $t \to \pm\infty$ fast enough. The NFT is defined through the Schrödinger eigenvalue problem, which uses the wave packet as potential [22–24]:

$$\phi_{tt} + q(t)\phi = (i\lambda)^2\phi, \tag{14}$$

where ϕ is the eigenfunction and $\lambda \in \mathbb{C}$ the spectral parameter. For a real signal $q(t)$, the Schrödinger eigenfunction problem only allows two types of solutions such that the eigenfunctions ϕ do not blow up. The first type of solutions ϕ correspond to $\lambda \in \mathbb{R} \setminus \{0\}$, which result in oscillatory eigenfunctions of infinite energy, but finite power. These solutions relate to the so-called continuous spectrum, and represent dispersive wave components in the signal $q(t)$. The second type of solutions consist of a finite number of discrete, purely-imaginary eigenvalues in the upper-half plane $\lambda_n = ik_n$, $0 < k \in \mathbb{R}$, $n = 1,\ldots,N$. The corresponding eigenfunctions ϕ_n decay exponentially in both tails and are finite-energy. It is well known that these discrete eigenvalues

correspond to the solitonic wave components present in the signal $q(t)$, which remain stable during propagation of $q(t)$ [22,24], and eventually separate into the train of solitons from Eqn. (6).

One of the most convenient ways of solving the Schrödinger eigenvalue problem is by rewriting Eqn. (14) as a system of two first order differential equations. We do so by switching to the basis from [18], in which the Schrödinger scattering problem is given as follows:

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \frac{1}{2i\lambda} \begin{bmatrix} i\lambda - \frac{\partial}{\partial x} \\ i\lambda + \frac{\partial}{\partial x} \end{bmatrix} \phi, \qquad \text{(change of basis)} \tag{15a}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} \begin{bmatrix} v_1(t,\lambda) \\ v_2(t,\lambda) \end{bmatrix} = \begin{bmatrix} -i\lambda + \frac{q(t)}{2i\lambda} & \frac{q(t)}{2i\lambda} \\ -\frac{q(t)}{2i\lambda} & i\lambda - \frac{q(t)}{2i\lambda} \end{bmatrix} \begin{bmatrix} v_1(t,\lambda) \\ v_2(t,\lambda) \end{bmatrix}, \tag{15b}$$

$$\begin{bmatrix} e^{-i\lambda t} \\ 0 \end{bmatrix} \xleftarrow[q\to 0]{t\to-\infty} \begin{bmatrix} v_1(t,\lambda) \\ v_2(t,\lambda) \end{bmatrix} \xrightarrow[q\to 0]{t\to+\infty} \begin{bmatrix} a(\lambda)e^{-i\lambda t} \\ b(\lambda)e^{+i\lambda t} \end{bmatrix}. \quad \text{(BCs)} \tag{15c}$$

Here, Eqn. (15b) corresponds with the Schrödinger eigenvalue problem, and Eqn. (15c) are the boundary conditions (BCs) that we impose to obtain the so-called scattering coefficients $a(\lambda)$ and $b(\lambda)$.

The full nonlinear Fourier transform spectrum consists of a discrete spectrum and a continuous spectrum. The discrete spectrum (ds) Λ^{ds}, contains all purely imaginary eigenvalues $\lambda_n = ik_n$ in the upper-half plane such that the eigenfunction $[v_1, v_2]^T$ is finite energy. This can only occur if $a(\lambda_n) = 0$. To complete the discrete spectrum, the eigenvalues are supplemented by their residues r_n, which relate to the locations of the solitons:

$$\Lambda^{\mathrm{ds}} := \left\{ \left(\lambda_n = ik_n, r_n = \frac{b}{a_\lambda}(\lambda_n) \right) : a(ik_n) = 0, \text{ and } 0 < k_n \in \mathbb{R} \right\}, \text{ with } a_\lambda(\lambda_n) = \left. \frac{\partial a(\lambda)}{\partial \lambda} \right|_{\lambda=\lambda_n}. \tag{16}$$

The continuous spectrum (cs), Λ^{cs}, consists of the so-called reflection coefficient b/a on the real line:

$$\Lambda^{\mathrm{cs}} := \left\{ \frac{b}{a}(\xi), \xi \in \mathbb{R} \setminus 0 \right\}. \tag{17}$$

As the wave packet evolves over x according to the normalised KdV equation, the eigenvalues remain constant, and the scattering coefficients a and b evolve trivially [15, 24]:

$$\lambda_n(x) = \lambda_n(0), \tag{18a}$$

$$a(\lambda, x) = a(\lambda, 0), \tag{18b}$$

$$b(\lambda, x) = b(\lambda, 0)e^{8i\lambda^3 x}. \tag{18c}$$

As the λ_n remain constant, so do the k_n.

Appendix B: Properties of the NFT

Property 1 (Scattering coefficients codomain). *The scattering coefficients $a(\lambda)$ and $b(\lambda)$ satisfy [24]:*

$$|a(\lambda)| \in [1, \infty), \quad |b(\lambda)| \in [0, \infty), \quad \left|\tfrac{b}{a}(\lambda)\right| \in [0, 1), \tag{19a}$$

$$|a(\lambda)|^2 - |b(\lambda)|^2 = 1. \tag{19b}$$

Property 2 (Linear Fourier transform as $q \to 0$). *In the small amplitude limit, $q(t) \to 0$, the Fourier transform degenerates to the linear Fourier transform. No eigenvalues will be present, and the continuous spectrum and linear Fourier spectrum relate as follows [15, Eqn. (4.7)]:*

$$\lim_{q(t)\to 0, |b/a|\to 0} 2i\xi \frac{b}{a}(\xi) = -\mathscr{F}\{q\}(\omega = 2\xi), \tag{20}$$

$$\text{with } \mathscr{F}\{q\}(\omega) = \int_{-\infty}^{\infty} q(t)e^{-i\omega t}\,\mathrm{d}t \quad \text{the Fourier transform},$$

in which ω denotes the linear angular frequency.

Property 3 (Time translation). *Let $\{a(\lambda), b(\lambda)\} = NFT\{q(t)\}$ be the scattering coefficients of the NFT corresponding to the signal $q(t)$. Then a time translation $t \to t - t_0$ in the signal results only in an exponential term in the b-coefficient [24]:*

$$\left\{ a(\lambda), b(\lambda)e^{-2i\lambda t_0} \right\} = NFT\{q(t - t_0)\}. \tag{21}$$

Note that, the eigenvalues λ_n (i.e., the zeros of $a(\lambda)$) remain unchanged, as $a(\lambda)$ remains unchanged.

Property 4 (Energy in the KdV-NFT). *Let the energy of a normalised KdV-governed signal be defined as*

$$\mathscr{E}^q = \int_{-\infty}^{\infty} q(t)^2\,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\mathscr{F}\{q\}(\omega)|^2\,\mathrm{d}\omega, \tag{22}$$

then the energy of the discrete spectrum $\mathscr{E}^{ds}$ and of the continuous spectrum $\mathscr{E}^{cs}$ are given as follows [25, p.286], [26, 1.6.21b]:

$$\mathscr{E}^q := \mathscr{E}^{cs} + \mathscr{E}^{ds}, \tag{23a}$$

$$\mathscr{E}^{ds} = \sum_{n=1}^{N} \underbrace{\frac{16}{3}k_n^3}_{\mathscr{E}_n}, \tag{23b}$$

$$\mathscr{E}^{cs} = \frac{1}{\pi}\int_{-\infty}^{\infty} -4\xi^2 \ln\left(1 - \left|\tfrac{b}{a}(\xi)\right|^2\right)\mathrm{d}\xi, \tag{23c}$$

where $\mathscr{E}_n$ is the energy of solitonic component $\lambda_n = ik_n$. We note here that for the low signal-amplitude case, the continuous-spectrum energy reduces to the energy in the linear Fourier spectrum: $\mathscr{E}^{cs} \xrightarrow{|b/a|\to 0} \frac{1}{\pi}\int_{-\infty}^{\infty} 4\xi^2 \left|\frac{b}{a}(\xi)\right|^2 \mathrm{d}\xi = \frac{1}{2\pi}\int_{-\infty}^{\infty} |\mathscr{F}\{q\}(\omega = 2\xi)|^2 \,\mathrm{d}\omega$, aligning with Property 2

Conjecture 5 (Continuity of NFT in c_q (unproven)). *Let $q(t;c_q) := c_q\eta(t)$, and $\{\lambda_n(c_q), n = 1\ldots N(c_q)\} = NFT\{c_q\eta(t)\}$ denote the eigenvalues of a signal $\eta(t)$ as a function of the normalisation coefficient c_q. Then the position of each eigenvalue $\lambda_n(c_q)$ is a continuous function of c_q over the imaginary axis. Moreover, eigenvalues can only (dis)appear at the real axis. Increasing c_q increases the signal energy, and most often causes existing eigenvalues to drift upwards, while new eigenvalues appear from the real axis.*

Appendix C: The influence of a trough on the discrete spectrum

We validate that the presence or absence of a trough after a signal has little influence on the discrete spectrum, which was also mentioned (although not shown) in [19]. We show here that indeed the trough has little influence on the eigenvalues for the incoming signal at wave gauge 1. We consider three different cut-off points: 1) at 18 s, when the bore reaches the still-water level for the first time again; 2) at 34 s, the time frame considered in all experiments; 3) at 48 s, the largest time possible before the reflected wave reaches wave gauge 1 again. The three signals with different cut-off point are shown in Fig. 5. The eigenvalues were determined for the reference water depth $h^{\mathrm{ref}} = 9.80\,\mathrm{cm}$. All eigenvalues are visually identical, except for the lowest eigenvalue (which has very little signal energy). This confirms that the trough indeed has very little influence on which solitons are present in the discrete spectrum.

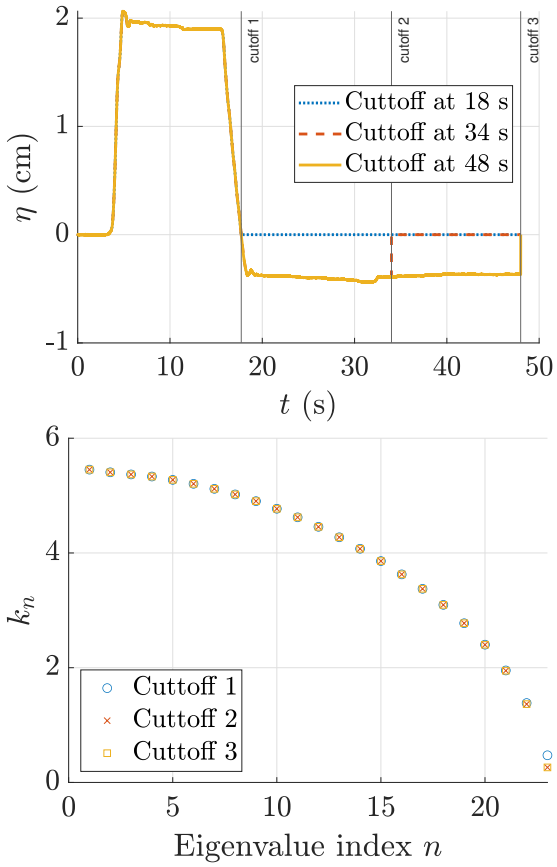

FIGURE 5. SIGNALS WITH DIFFERENT CUT-OFF POINTS FOR THEIR TROUGHS, AND THEIR CORRESPONDING EIGENVALUES FOR THE REFERENCE WATER DEPTH. THE EFFECT OF THE TROUGH IS ONLY VISIBLE FOR THE LAST (LOWEST-ENERGY) SOLITON.

Appendix D: Identification using conserved quantities

It is well known that the normalised KdV equation in Eqn. (3) has an infinite number of conserved quantities or constants of motion. These quantities are conserved during propagation for tKdV-governed experimental data, given that the correct amplitude normalisation coefficient c_q was used. It is thus possible to identify c_q as the value for which these quantities are indeed equal at input signal and output. A similar method has also been demonstrated for NLSE-governed signals in [11].

The first two constants of motion of the KdV are related to conservation of mass and conservation of energy. The consecutive constants are less intuitive, and will be referred to as the third moment, fourth moment, etc. The first conservation laws for KdV in Eq. 3 are given as follows [20, 27]:

$$\begin{aligned} C_1 &= \int q\,\mathrm{d}t, & C_3 &= \int \left(-2q^3 + q_t^2\right)\mathrm{d}t, \\ C_2 &= \int q^2\,\mathrm{d}t, & C_4 &= \int \left(9q^4 - 18qq_t^2 + \tfrac{9}{5}q_{tt}^2\right)\mathrm{d}t. \end{aligned} \tag{24}$$

When we equate $C_n^{\mathrm{in}} = C_n^{\mathrm{out}}$, and substitute $q^{\mathrm{in}}(t) = T_0^2 c_q \eta^{\mathrm{in}}((\tau - c_0 l)/T_0)$, and $q^{\mathrm{out}}(t) = T_0^2 c_q \eta^{\mathrm{out}}((\tau - c_0 l)/T_0)$, we will find that T_0 always drops out. However, c_q also drops out if only a single power of q is present, as is the case for C^1 and C^2. Therefore, C^3 is the first quantity for which c_q does not drop out upon equating

$C_3^{\text{in}} = C_3^{\text{out}}$ (we set $T_0 = 1$ to simplify the demonstration):

$$\begin{aligned}
&\int_{-\infty}^{\infty} -2\left(c_q \eta^{\text{in}}(\tau)\right)^3 + \left(c_q \eta_\tau^{\text{in}}(\tau)\right)^2 \mathrm{d}\tau \\
&\qquad = \int_{-\infty}^{\infty} -2\left(c_q \eta^{\text{out}}(\tau)\right)^3 + \left(c_q \eta_\tau^{\text{out}}(\tau)\right)^2 \mathrm{d}\tau, \\
\Rightarrow & c_q^2 \int_{-\infty}^{\infty} \left(\eta_\tau^{\text{in}}(\tau)\right)^2 - \left(\eta_\tau^{\text{out}}(\tau)\right)^2 \mathrm{d}\tau \\
&\qquad = 2c_q^3 \int_{-\infty}^{\infty} \left(\eta^{\text{in}}(\tau)\right)^3 - \left(\eta^{\text{out}}(\tau)\right)^3 \mathrm{d}\tau, \\
\Rightarrow & c_q = \frac{\int_{-\infty}^{\infty} \left(\eta_\tau^{\text{in}}(\tau)\right)^2 - \left(\eta_\tau^{\text{out}}(\tau)\right)^2 \mathrm{d}\tau}{2\int_{-\infty}^{\infty} \left(\eta^{\text{in}}(\tau)\right)^3 - \left(\eta^{\text{out}}(\tau)\right)^3 \mathrm{d}\tau}.
\end{aligned} \tag{25}$$

This relation provides a fast and easy method to obtain estimates for c_q. Although all conserved quantities except for C_1 and C_2 can be used to find c_q in a similar fashion, the higher conserved quantities contain higher derivatives and powers, which are increasingly sensitive to noise. Furthermore, we can see that in C_4 three different orders of c_q will pop up (order 4, 3 and 2), which will lead to an underdetermined system if we only have measurements at two wave gauges. Even when three gauges are considered, a system of equations will have to be solved, which is probably more prone to noise. We will therefore only use C_3.

Using this method on the simulated data from CG2-CG4, and converting c_q to h, we successfully recovered the correct water of $h = 9.8\,\text{cm}$, validating the method for ideal KdV-governed data. Also when applying this method to the pre-processed experimental data, we identified depths close to the measured water depth, as shown in Table. 5. Although this method is very fast and can yield good results, it can be sensitive to noise and offsets due to its dependence on derivatives and high powers in q. Also due to the fact that only a single number is given as output, it is unknown how reliable this number is. By itself, this method is therefore often not suitable for application in practice, but may in many cases provide initial estimates or additional validation.

TABLE 1. THE IDENTIFIED WATER DEPTHS FROM THE PRE-PROCESSED WAVE GAUGE DATA.

Data set	Measured	CG1-CG4	CG2-CG4	CG3-CG4
h	9.80 cm	9.40 cm	9.59 cm	9.90 cm