

Traveling wave profiles for a semi-discrete Burgers equation

Uditnarayan Kouskiya^a, Robert L. Pego^{b,*}, Amit Acharya^c

^aDepartment of Civil and Environmental Engineering, Vanderbilt University, Nashville, 37212, TN, USA

^bDepartment of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, 15213, PA, USA

^cDepartment of Civil & Environmental Engineering, and Center for Nonlinear Analysis, Carnegie Mellon University, Pittsburgh, 15213, PA, USA

Abstract

We look for traveling waves of the semi-discrete conservation law $4\dot{u}_j + u_{j+1}^2 - u_{j-1}^2 = 0$, using variational principles related to concepts of “hidden convexity” appearing in recent studies of various PDE (partial differential equations). We analyze and numerically compute with two variational formulations related to dual convex optimization problems constrained by either the differential-difference equation (DDE) or nonlinear integral equation (NIE) that wave profiles should satisfy. We prove existence theorems conditional on the existence of extrema that satisfy a strict convexity criterion, and numerically exhibit a variety of localized, periodic and non-periodic wave phenomena.

Keywords: variational principles, solitary waves, Petviashvili iteration, nonlinear eigenvalue problems

2020 MSC: 49J35, 49M29, 34K31, 70G75, 35Q70

1. Introduction

A great many types of dynamic behavior are known to arise from different approximation schemes for solutions to the inviscid Burgers equation

$$\partial_t u + \partial_x \left(\frac{1}{2} u^2 \right) = 0, \quad -\infty < x < \infty, \quad t > 0. \quad (1.1)$$

In particular, conservative and dispersive approximations of this equation often exhibit *dispersive shocks*. As a dispersive shock develops, its structure is that of a modulated envelope of periodic waves, and at its leading edge one sees the emergence of several solitary waves. While a great deal is known about dispersive shocks for integrable approximations to (1.1), little appears to be understood for non-integrable approximations.

A recent study by Sprenger *et al.* [1] focuses on a non-integrable approximation generated by simple centered differences in space and a change of time scale: With $u_j(t) \approx u(jh, ht)$, one obtains the equations

$$\frac{d}{dt} u_j + \frac{1}{4} (u_{j+1}^2 - u_{j-1}^2) = 0 \quad j \in \mathbb{Z}. \quad (1.2)$$

In addition to dispersive shocks, Sprenger *et al.* illustrate many other different regimes and types of solutions in numerical simulations of these equations. They use Whitham modulation theory and weakly nonlinear asymptotics to explain some of the observed phenomena. Other interesting behaviors that were observed include highly nonlinear phenomena such as discontinuous waves connecting periodic solutions to constant states, and periodic waves emerging from both sides of a discontinuity.

In this article we focus on studying periodic and solitary traveling waves of the semi-discrete Burgers equations (1.2). For such waves, the lattice configuration duplicates itself with a spatial shift and a time shift, satisfying, for some $\tau \neq 0$,

$$u_j(t) = u_{j+1}(t + \tau) \quad \text{for all } j \in \mathbb{Z} \text{ and all } t \in \mathbb{R}.$$

*Corresponding author

Email addresses: uditnarayan.kouskiya@vanderbilt.edu (Uditnarayan Kouskiya), rpego@cmu.edu (Robert L. Pego), acharyaamit@cmu.edu (Amit Acharya)

It follows that, with $c = 1/\tau$ and $f(x) = u_0(-\tau x)$, such waves must take the form

$$u_j(t) = f(j - ct), \quad (1.3)$$

where the wave profile f satisfies the differential-difference equation (DDE)

$$-cf'(x) + \frac{1}{4}(f(x+1)^2 - f(x-1)^2) = 0. \quad (1.4)$$

Upon integration we find this is equivalent to the nonlinear integral equation (NIE)

$$-cf(x) + \frac{1}{4} \int_{x-1}^{x+1} f(z)^2 dz = C_0, \quad (1.5)$$

where C_0 is a constant.

For our study we will adapt the dual variational framework that has been developed for a variety of problems in [2, 3, 4] and computationally demonstrated in [5, 6, 7], with rigorous results presented in [7, 8]. The approach is closely related to the idea of “hidden convexity” in nonlinear PDE as developed in particular by Brenier [9, 10, 11]. Brenier shows that several types of PDE problems that nominally have no relation to convex analysis have a (“hidden”) reformulation in terms of a primal-dual pair of convex and concave optimization problems. A hidden saddle-point structure with similar flavor has recently been identified in finite-dimensional non-smooth nonlinear programming by Rockafellar [12] in connection with very general sufficient conditions for optimality. Brenier’s approach has recently been extended and employed by Vortnikov [13, 14] to establish a version of Dafermos’ maximum entropy rate principle for conservation laws, and by Mirebeau and Stampfli [15] to establish rates of convergence to smooth solutions for discretization schemes for multidimensional quadratic porous medium and (in)viscid Burgers equations.

We utilize our scheme for both numerical and theoretical reasons. Regarding the numerical computation of steady wave profiles, a method known as *Petviashvili iteration* often works rather well in practice [16, 17]. Equation (1.5) may be the simplest for which this method can be applied, in fact. The global convergence properties of Petviashvili iteration are not very well understood, however. The work of Pelinovsky and Stepanyants [17] established criteria for local convergence or divergence, but requires certain assumptions about the spectrum of the linearization at an exact wave profile, assumptions which are not known to hold in many settings, including that of (1.5).

Regarding theory, we are not aware of any proof of existence for any non-trivial traveling waves of the semi-discrete Burgers equations (1.2), except that the approach of Herrmann [18] may capture periodic solutions that are strictly of one sign after modifying the flux function $\Phi(u) = \frac{1}{2}u^2$ to be strictly monotonic and C^1 . The solitary-wave problem appears similar to that for solitary waves in Fermi-Pasta-Ulam type particle lattices involving nearest-neighbor forces, which is the subject of a recent review by Vainchtein [19]. Variational methods based on concentration compactness or mountain-pass methods have been used to prove existence theorems for solitary waves in Fermi-Pasta-Ulam lattices [20, 21, 22] and related equations of peridynamics [23, 24]. But we are not aware of an existing variational formulation for the traveling wave profile equations (1.4) or (1.5). The general variational framework of Herrmann and Matthies [25] deals with equations that resemble (1.5), but does not evidently apply due to the fact that the characteristic function of an interval has a sign-changing Fourier transform and thus is not a convolution square.

For waves that are small-amplitude long-wave perturbations of a non-zero constant state, a formal Korteweg-de Vries (KdV) approximation can be described, as one may expect, and we describe this in Appendix A. It is plausible that existence of waves in this regime could be established by adapting existing fixed-point and perturbation methods [26, 27, 28, 29, 30, 31]. On the other hand, large-amplitude waves, and any perturbations of the zero solution, are in a completely nonlinear regime inaccessible by such methods. Perhaps existence results might be had by methods based on topological degree theory, though, like those in [32, 33].

The dual variational formulation that we will study here provides a flexible method for exploring the solution space numerically. Using it we also prove conditional existence theorems for traveling waves. We prove the existence of extrema which determine an exact traveling wave solution, conditional upon a certain domain constraint being strictly satisfied. In many cases, our numerical computations strongly indicate that the domain constraint indeed strictly holds. But we do not claim any mathematically rigorous existence proof.

A useful feature of our approach is that it provides a well-set solution strategy without imposing further conditions on the fundamental problem statement for ones not naturally posed with boundary conditions (such as the traveling wave problem). We exploit this feature in our formulation and computations.

Let us briefly summarize some of the key results of our computations. We find a considerable variety of periodic and non-periodic wave profiles on long intervals. Our computations of waves with localized structure suggest that solitary wave solutions should exist having limit states

$$\bar{u} = \lim_{t \rightarrow \pm\infty} u_j(t), \quad (1.6)$$

whenever a *phase-speed non-matching condition* holds. Namely, a solitary wave of the form (1.3) that satisfies (1.6) should exist whenever the speed c does not match the phase velocity ω/ξ of any linear harmonic wave $v_j(t) = e^{i\xi j - i\omega t}$ for the equations (1.2) linearized at the constant state \bar{u} . As the dispersion relation for this linearized equation is $\omega = \bar{u} \sin \xi$, this means that the phase-speed non-matching condition reads

$$c \neq \bar{u} \operatorname{sinc} \xi \quad \text{for all } \xi \in \mathbb{R}, \quad (1.7)$$

where $\operatorname{sinc} \xi = \sin \xi / \xi$ for $\xi \neq 0$ (and $= 1$ for $\xi = 0$).

The wave profiles that we find develop different types of long-wave structure in two regimes in which phase-speed non-matching is breaking down. In one regime, $c \approx \bar{u}$. This corresponds to the KdV long-wave regime, and indeed we see wide, small-amplitude single-hump wave profiles there. In the other regime, $c \approx \bar{u} \operatorname{sinc} \xi_*$ where $\xi_* \approx 4.4934$ is the value which minimizes $\operatorname{sinc} \xi$. Here we see wave profiles oscillating with wave number near ξ_* but modulated with an envelope that slowly decays away. Note that at the critical wave number ξ_* , the phase velocity matches the group velocity $d\omega/d\xi = \bar{u} \cos \xi$ of wave packets. A recent work by Kozyreff [34] studies wave propagation in such a regime. Kozyreff's study involves the asymptotics of exponentially small terms, a topic beyond the scope of the present paper.

A further interesting result from our computations is that we find a range of cases where our optimization method finds localized waves with oscillatory decaying tails, but Petviashvili iteration, initiated with the resulting wave shape, goes unstable and fails to converge. (See Subsection 5.3.)

Our focus in the present paper is to demonstrate the feasibility and utility of the variational method as an exploratory tool for finding wave solutions. Thus we leave systematic investigation and classification of the families of nonlinear wave solutions of the semi-discrete Burgers system (1.2) for future research. For finding particular solutions or curves of solutions, other well-established discretization methods such as pseudospectral collocation combined with path-following codes (e.g., see [35]) would plausibly compute solutions more efficiently and more accurately, if one has a good initial guess to start from. One feature of the present study, however, is that even starting with base states that are poor initial guesses, our variational methods appear to converge to a solution. And the flexibility of choosing base states leads to finding a plethora of solutions of unexpected types. In any case, there clearly remain many issues to be understood more thoroughly and more rigorously.

2. Variational formulations

By a simple scaling (replacing f by $-2cf$), equation (1.4) for nontrivial traveling wave profiles can be reduced to the following DDE that corresponds to setting $c = -\frac{1}{2}$ in (1.4):

$$f'(x) + \frac{1}{2} \left((f(x+1))^2 - (f(x-1))^2 \right) = 0, \quad (2.1)$$

where $f : \mathbb{R} \rightarrow \mathbb{R}$ and $(\cdot)' \equiv \frac{d(\cdot)}{dx}$. Correspondingly, equation (1.5) is reduced to the NIE

$$f(x) + \frac{1}{2} \int_{x-1}^{x+1} f(y)^2 dy = C_1, \quad (2.2)$$

where $C_1 = \frac{1}{2}C_0/c^2$. In this section, we will describe two related dual variational formulations for this problem, one that starts directly from the DDE (2.1) and one that starts from the NIE (2.2). The formulations have slightly different numerical and analytical properties, to be compared in later sections.

Schematically, the formulations we study are dual to primal optimization problems of a classic form. Namely, as discussed in [10, 2], one seeks to minimize an integral functional $\int H(f, x) dx$, constrained by the equation in question. If $Q[f] = 0$ corresponds to the equation in question, one introduces a dual field (Lagrange multiplier) $\lambda(x)$ and writes the primal problem in the inf-sup formulation

$$\text{Find } \inf_f \sup_\lambda \int (\lambda Q[f] + H(f, x)) dx.$$

This primal problem is set only as a device, however. The point is that the dual problem, obtained by interchanging inf and sup, is one that we can exploit to obtain information about solutions of $Q[f] = 0$ both computationally and theoretically.¹

Of particular importance is the realization that $Q[f] = 0$ is a component of the first-order optimality conditions of both the $\sup_\lambda \inf_f$ and $\inf_f \sup_\lambda$ problem statements, *regardless of the choice of H* . Thus, our solution strategy exploits the use of not just a single H function, but a family of adapted, convex ones, allowing parametrization by base states which are specified functions or trajectories that can encode knowledge about approximate solutions [3, 2]. Indeed, base states are utilized in a crucial, adaptive way in our algorithm in Sec. 4.

Below, for any field (\cdot) , we will use the following notation:

$$(\cdot)(x+h) \equiv (\cdot)_{x+h}$$

In cases when $h = 0$, we will utilize the notations $(\cdot)(x) \equiv (\cdot)_x$ and $(\cdot)(x) \equiv (\cdot)$ interchangeably.

2.1. Dual DDE formulation

Considering a dual field λ corresponding to (2.1), and presuming λ and/or f decay sufficiently rapidly at infinity, we define the following *pre-dual* functional:

$$\widehat{S}[f, \lambda] = \int_{-\infty}^{\infty} \left(-\lambda' f + \frac{\lambda}{2} (f_{x+1}^2 - f_{x-1}^2) + H(f, x) \right) dx, \quad (2.3)$$

where H is a free auxiliary function of the arguments shown. Since the limits of integration involved in the last equation are from $-\infty$ to ∞ , we can write:

$$\int_{-\infty}^{\infty} \lambda f_{x+1}^2 dx = \int_{-\infty}^{\infty} \lambda_{x-1} f^2 dx \quad \text{and} \quad \int_{-\infty}^{\infty} \lambda f_{x-1}^2 dx = \int_{-\infty}^{\infty} \lambda_{x+1} f^2 dx. \quad (2.4)$$

Equation (2.3) can thus be rewritten as:

$$\widehat{S}[f, \lambda] = \int_{-\infty}^{\infty} \left(-\lambda' f + \frac{f^2}{2} (\lambda_{x-1} - \lambda_{x+1}) + H(f, x) \right) dx = \int_{-\infty}^{\infty} \mathcal{L}_H(f, \mathcal{D}, x) dx, \quad (2.5)$$

where $\mathcal{D} := (\lambda_{x-1}, \lambda', \lambda_{x+1})$ and we refer to the integrand \mathcal{L}_H as the *Lagrangian* for the current problem. The subscript H in \mathcal{L}_H denotes its dependence on the chosen auxiliary function. We now impose the following condition

$$\frac{\partial \mathcal{L}_H}{\partial f} = 0 : \quad -\lambda' + f(\lambda_{x-1} - \lambda_{x+1}) + \frac{\partial H}{\partial f} = 0, \quad (2.6)$$

to solve for f in terms of the dual objects in \mathcal{D} . The choice of the function H is made to enable this step, for a substantial class of dual fields. One such choice of H is

$$H(f, x) =: \widehat{H}(f(x), \bar{f}(x)) = \frac{\mathbf{a}}{2} (f - \bar{f})^2, \quad (2.7)$$

¹The term ‘duality’ is used in the physics literature when two superficially distinct theoretical structures can be mapped onto each other to facilitate the approximation of difficult nonlinear problems. This served as the initial motivation for the use of the term for the method adopted here, as discussed in [36, Sec. 1].

where $\bar{f} : \mathbb{R} \rightarrow \mathbb{R}$ represents a base state and $\mathbf{a} \gg 1$. The choice of \mathbf{a} is discussed at the end of this subsection and in subsection 4.2.

For such a choice of H , equation (2.6) takes the form

$$(\mathbf{a} + \lambda_{x-1} - \lambda_{x+1})f = \mathbf{a}\bar{f} + \lambda', \quad (2.8)$$

and we define the following *Dual-to-Primal (DtP)* mapping to provide a solution for f when one exists:

$$f^{(H)}(\mathcal{D}, \mathbf{a}, x) = \begin{cases} \frac{\mathbf{a}\bar{f} + \lambda'}{\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}} = \bar{f} + \frac{\lambda' - \bar{f}(\lambda_{x-1} - \lambda_{x+1})}{\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}} & \text{if } \mathbf{a} + \lambda_{x-1} - \lambda_{x+1} \neq 0 \\ 0 & \text{if } \mathbf{a} + \lambda_{x-1} - \lambda_{x+1} = 0 \text{ and } \mathbf{a}\bar{f} + \lambda' = 0 \end{cases} \quad (2.9)$$

We will often write $\hat{f} := f^{(H)}(\mathcal{D}, \mathbf{a}, x)$. The dual functional can now be defined as:

$$S[\lambda] := \widehat{S}[\hat{f}, \lambda] = \int_{-\infty}^{\infty} \left(-\lambda' \hat{f} + \frac{\hat{f}^2}{2} (\lambda_{x-1} - \lambda_{x+1}) + \frac{\mathbf{a}}{2} (\hat{f} - \bar{f})^2 \right) dx \quad (2.10)$$

which can be rewritten completely in terms of dual variables as:

$$\begin{aligned} S[\lambda] &= \int_{-\infty}^{\infty} \left(-\lambda' \hat{f} + \frac{\hat{f}^2}{2} (\lambda_{x-1} - \lambda_{x+1}) + \mathbf{a} \left(\frac{\hat{f}^2}{2} - \bar{f} \hat{f} + \frac{\bar{f}^2}{2} \right) \right) dx \\ &= \int_{-\infty}^{\infty} \left((\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}) \frac{\hat{f}^2}{2} - (\mathbf{a}\bar{f} + \lambda') \hat{f} + \mathbf{a} \frac{\bar{f}^2}{2} \right) dx \\ &= \int_{-\infty}^{\infty} \left(-(\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}) \frac{\hat{f}^2}{2} + \frac{\mathbf{a}\bar{f}^2}{2} \right) dx \\ &= \frac{1}{2} \int_{-\infty}^{\infty} \left(-\frac{(\lambda' + \mathbf{a}\bar{f})^2}{\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}} + \mathbf{a}\bar{f}^2 \right) dx \end{aligned} \quad (2.11)$$

An alternate, but equivalent, expression arises by first writing $S[\lambda]$ in terms of $(\hat{f} - \bar{f})$:

$$\begin{aligned} S[\lambda] &= \int_{-\infty}^{\infty} \left(-\lambda' \hat{f} + \frac{\hat{f}^2}{2} (\lambda_{x-1} - \lambda_{x+1}) + \frac{\mathbf{a}}{2} (\hat{f} - \bar{f})^2 \right) dx \\ &= \int_{-\infty}^{\infty} \left(-\lambda' (\hat{f} - \bar{f}) + \frac{(\hat{f} - \bar{f})^2}{2} (\lambda_{x-1} - \lambda_{x+1}) + \frac{\mathbf{a}}{2} (\hat{f} - \bar{f})^2 \right) dx \\ &\quad + \int_{-\infty}^{\infty} (\hat{f} - \bar{f}) \bar{f} (\lambda_{x-1} - \lambda_{x+1}) dx + \int_{-\infty}^{\infty} \left(-\lambda' \bar{f} + \frac{\bar{f}^2}{2} (\lambda_{x-1} - \lambda_{x+1}) \right) dx \end{aligned}$$

which reduces, after collecting and combining linear and quadratic terms in $(\hat{f} - \bar{f})$ using the DtP mapping, to

$$S[\lambda] = \int_{-\infty}^{\infty} -\frac{1}{2\Delta_\lambda} (\lambda' - \bar{f}(\lambda_{x-1} - \lambda_{x+1}))^2 dx + \int_{-\infty}^{\infty} \left(-\lambda' \bar{f} + \frac{\bar{f}^2}{2} (\lambda_{x-1} - \lambda_{x+1}) \right) dx, \quad (2.12)$$

where

$$\Delta_\lambda := \mathbf{a} + \lambda_{x-1} - \lambda_{x+1}. \quad (2.13)$$

The first variation of (2.10) in a direction $\delta\lambda$ is given by

$$\begin{aligned} \delta^{(1)} S[\lambda; \delta\lambda] &= \int_{-\infty}^{\infty} \left(\frac{\partial \mathcal{L}_H}{\partial f} \frac{\partial \hat{f}}{\partial \mathcal{D}} \delta \mathcal{D} + \frac{\partial \mathcal{L}_H}{\partial \mathcal{D}} \delta \mathcal{D} \right) dx = \int_{-\infty}^{\infty} \left(\frac{\partial \mathcal{L}_H}{\partial \mathcal{D}} \delta \mathcal{D} \right) dx \\ &= \int_{-\infty}^{\infty} \left(-\delta\lambda' \hat{f} + \frac{\hat{f}^2}{2} (\delta\lambda_{x-1} - \delta\lambda_{x+1}) \right) dx, \end{aligned} \quad (2.14a)$$

$$= \int_{-\infty}^{\infty} \delta\lambda \left(\hat{f}' + \frac{1}{2} (\hat{f}_{x+1}^2 - \hat{f}_{x-1}^2) \right) dx. \quad (2.14b)$$

where we have used

$$\delta\lambda \rightarrow 0 \quad \text{as } x \rightarrow \pm\infty. \quad (2.15)$$

On requiring $\delta^{(1)}S[\lambda; \delta\lambda] = 0$ for any $\delta\lambda$ satisfying (2.15), the Euler-Lagrange equation for $S[\lambda]$ is given by

$$\hat{f}'(x) + \frac{1}{2} \left((\hat{f}(x+1))^2 - (\hat{f}(x-1))^2 \right) = 0 \quad \forall x \in (-\infty, \infty), \quad (2.16)$$

which is the same as (2.1) now written only in terms of the dual variables.

Additionally, one can always choose $\delta\lambda(x) = 0$ for $x \notin (a, b)$ to establish (2.16) only for $x \in (a, b)$.

An important consistency check of our scheme is that for each solution, say f^* , to the primal problem, there is at least one dual functional whose critical point corresponds to that solution. That functional is constructed simply by the choice of $\bar{f} = f^*$ and the critical point is given by $\lambda = 0$, as can be directly read off from the DtP mapping and the fundamental justification of the scheme that the primal equation forms the Euler-Lagrange equation of the dual functional with DtP mapping substituted.

Concave maximization. Having demonstrated the consistency of our scheme with the problem (2.1) as a critical point problem of the dual functional (2.12), for practical purposes we consider a pure maximization problem on a bounded domain. We require the dual fields λ to vanish outside a given finite interval $\Omega = (-L, L)$, and consider a related functional given by

$$\begin{aligned} \tilde{S}[\lambda] &:= \inf_f \widehat{S}_L[f, \lambda] = \inf_f \int_{\Omega_1} \left(-\lambda' f + \frac{f^2}{2} (\lambda_{x-1} - \lambda_{x+1}) + \frac{\mathbf{a}}{2} (f - \bar{f})^2 \right) dx \\ &= \int_{\Omega_1} \inf_f \left(\frac{f^2}{2} (\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}) - f (\lambda' + \mathbf{a}\bar{f}) + \frac{\mathbf{a}}{2} \bar{f}^2 \right) dx, \end{aligned} \quad (2.17)$$

where \widehat{S}_L is exactly the functional \widehat{S} from (2.5), restricting the domain of integration to the minimal interval $\Omega_1 := (-L-1, L+1)$ that accommodates all non-vanishing values of λ_{x-1} and λ_{x+1} to be considered. Noting that the Lagrangian of \widehat{S}_L is affine in \mathcal{D} , the integrand of \tilde{S} must be concave in \mathcal{D} . Furthermore,

$$\tilde{S}[\lambda] = \begin{cases} S_L[\lambda] & \text{for } \lambda \text{ s.t. a.e. } (\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}) \geq 0 \text{ and} \\ & \text{if } (\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}) = 0 \text{ then } \lambda'(x) + \mathbf{a}\bar{f}(x) = 0, \\ -\infty & \text{otherwise,} \end{cases} \quad (2.18)$$

where S_L is the functional S from (2.10) integrated on the interval Ω_1 . This is so because at points where $\Delta_\lambda > 0$ the integrand of \tilde{S} has a unique minimizer (over f) given by the integrand of S , and at x where $\Delta_\lambda(x) = 0$, the associated condition $\lambda'(x) + \mathbf{a}\bar{f}(x) = 0$ again ensures that the integrands of \tilde{S} and S_L match, as can be seen from (2.10)-(2.9). Thus, finding a critical point of \tilde{S} corresponds to a *concave maximization* problem. Moreover, in such a maximization, λ fields which do not satisfy the *Convexity Condition*

$$\mathcal{C} : \Delta_\lambda = (\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}) \geq 0 \quad \text{a.e.} \quad (2.19)$$

cannot be competitors for being a maximizer, and it is best to not be concerned with such fields (we note that $S_L[\lambda]$ need not be concave over the set of all λ fields); alternatively, we can simply choose to seek maximizers of $S_L[\lambda]$ in the reduced set of fields which satisfy \mathcal{C} .

We will utilize the above insights in our numerical scheme by *looking for critical points of $S_L[\lambda]$ constrained by \mathcal{C}* . Some rigorous analytical properties of the concave functional \tilde{S} will be developed in Section 3 below. We note here that the condition (2.19) is the analog, in this non-local setting, of the condition that guarantees a local degenerate ellipticity of the dual critical point problem in the PDE case, as defined in [2, Sec. 3] and applied to the inviscid Burgers equation in [6, Sec. 2].

A scaling symmetry. The pre-dual and dual functionals defined above depend in a simple way upon the parameter $\mathbf{a} > 0$ that scales the amplitude of the function $H(f, x)$ in (2.7). If we make this dependence upon \mathbf{a} explicit, then it is evident from (2.5) that the pre-dual functional \hat{S} satisfies

$$\hat{S}(f, \lambda, \mathbf{a}) = \mathbf{a} \hat{S}(f, \lambda/\mathbf{a}, 1), \quad (2.20)$$

the DtP map in (2.9) satisfies

$$f^{(H)}(\mathcal{D}, \mathbf{a}, x) = f^{(H)}(\mathcal{D}/\mathbf{a}, 1, x), \quad (2.21)$$

and the functional S in (2.10) satisfies

$$S[\lambda, \mathbf{a}] = \mathbf{a} S[\lambda/\mathbf{a}, 1]. \quad (2.22)$$

Thus, increasing \mathbf{a} from 1 simply produces a proportional increase in S at a scaled down argument λ/\mathbf{a} . The first variation remains invariant with scaled arguments, satisfying

$$\delta^{(1)} S[\lambda, \mathbf{a}; \delta\lambda] = \delta^{(1)} S[\lambda/\mathbf{a}, 1; \delta\lambda/\mathbf{a}] \quad (2.23)$$

and correspondingly the second variation is inversely proportional to \mathbf{a} .

What this means is that, for our present choice of the function H in the Lagrangian, the choice of \mathbf{a} makes no difference *in theory* for the purpose of finding primal solutions in the form $\hat{f} = f^{(H)}(\mathcal{D}, \mathbf{a}, x)$ for critical points λ . With a different value of \mathbf{a} the location of the critical points simply scales proportionally while \hat{f} remains the same. In practice, however, we find it convenient to choose \mathbf{a} to be somewhat large. In particular, this makes the convexity condition (2.19) easier to satisfy with numerically chosen functions λ without having to worry about scaling down their amplitude. The choice of \mathbf{a} in principle also has some effect on numerical schemes and stopping criteria. We will discuss these issues further in Section 4.2 and Appendix B. We note that with other choices of H , parameters in its definition may not lead to this kind of scaling symmetry.

2.2. Dual NIE Formulation

We will derive an alternative dual variational formulation by starting from the nonlinear integral equation (2.2) instead of the differential-difference equation (2.1) and formulating the problem in terms of a corresponding dual field v . This leads to some differences in terms of the approximations that are natural to make and the results obtained. We will work with both approaches and compare them at the end.

In the simplest case when $C_1 = 0$, we can consider $v = -\lambda'$, and we can then write

$$\lambda_{x-1} - \lambda_{x+1} = \int_{x-1}^{x+1} v(y) dy = \int_{-1}^1 v(x+z) dz.$$

Define the right-hand side to be $Kv(x)$. Then K is a linear convolution operator satisfying

$$Kv(x) = \int_{-\infty}^{\infty} \Lambda(x-y)v(y) dy, \quad \Lambda(z) = \begin{cases} 1 & |z| \leq 1, \\ 0 & |z| > 1. \end{cases} \quad (2.24)$$

Note that if v is locally integrable on \mathbb{R} then Kv is continuous, and if v is periodic then Kv is periodic with the same period. And, if μ and v are $2L$ -periodic locally square-integrable functions, then

$$\int_{-L}^L \mu Kv dx = \int_{-L}^L v K\mu dx. \quad (2.25)$$

Indeed, since the integral of any translate of a periodic function over any full period is the same,

$$\begin{aligned} \int_{-L}^L \mu Kv dx &= \int_{-L}^L \int_{-1}^1 \mu(x)v(x+z) dz dx = \int_{-1}^1 \left(\int_{-L}^L \mu(x)v(x+z) dx \right) dz \\ &= \int_{-1}^1 \left(\int_{-L}^L \mu(x-z)v(x) dx \right) dz = \int_{-L}^L v K\mu dx. \end{aligned}$$

We will seek wave profiles f as perturbations of a constant u_∞ , taking the form

$$f(x) = u_\infty + w(x). \quad (2.26)$$

For periodic waves we require w to be $2L$ -periodic on the real line, and for solitary waves we say $L = \infty$ and require $w(x) \rightarrow 0$ as $|x| \rightarrow \infty$. In these terms, equation (2.2) takes the form

$$u_\infty + w + \frac{1}{2}K(u_\infty^2 + 2u_\infty w + w^2) = C_1.$$

We claim that it is no loss of generality to require

$$C_1 = u_\infty + u_\infty^2. \quad (2.27)$$

Indeed, if f is a solitary wave profile, evidently (2.27) must hold. If instead f is a $2L$ -periodic solution of (2.2) with $L < \infty$, then because $K(f^2)$ is $2L$ -periodic and $K(1) = 2$, we find

$$\int_{-L}^L \frac{1}{2}K(f^2) dx = \int_{-L}^L f^2 dx = \int_{-L}^L (-f + C_1) dx \leq \int_{-L}^L (f^2 + \frac{1}{4} + C_1) dx,$$

since $-f \leq -f + (f + \frac{1}{2})^2 = f^2 + \frac{1}{4}$. Thus $\frac{1}{4} + C_1 \geq 0$ and (2.27) follows with $u_\infty = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + C_1}$.

Equation (2.2) now becomes equivalent to

$$w + u_\infty K w + \frac{1}{2}K(w^2) = 0, \quad (2.28)$$

which may be written more explicitly as the equation

$$w(x) + \int_{x-1}^{x+1} \left(u_\infty w(y) + \frac{1}{2}w(y)^2 \right) dy = 0.$$

We are ready next to develop a dual variational formulation for equation (2.28). Letting v be a $2L$ -periodic dual field and noting that (2.25) should hold with $\mu = w$ and w^2 , we define a pre-dual functional by

$$\hat{\mathfrak{E}}[w, v] = \int_{-L}^L \left(w(v + u_\infty K v) + \frac{w^2}{2} K v + \mathfrak{H}(w, x) \right) dx = \int_{-L}^L \mathfrak{Q}_{\mathfrak{H}}(w, v, x) dx. \quad (2.29)$$

For convenience we take $\mathfrak{H}(w, x)$ in (2.3) in the modified form

$$\mathfrak{H}(w, x) = \frac{\mathfrak{a}}{2} \left((w - \bar{w})^2 - \bar{w}^2 \right) = \mathfrak{a} \left(\frac{w^2}{2} - w\bar{w} \right), \quad (2.30)$$

where $\bar{w} = \bar{w}(x)$ is a fixed base state, whence

$$\mathfrak{Q}_{\mathfrak{H}}(w, v, x) = \frac{w^2}{2}(\mathfrak{a} + K v) + w(v + u_\infty K v - \mathfrak{a}\bar{w}). \quad (2.31)$$

Aiming toward a well-posed dual maximization problem, notice that

$$\inf_{w \in \mathbb{R}} \mathfrak{Q}_{\mathfrak{H}}(w, v, x) = \begin{cases} -\infty & \text{if } \mathfrak{a} + K v(x) < 0, \\ -\infty & \text{if } \mathfrak{a} + K v(x) = 0 \text{ and } \mathfrak{a}\bar{w}(x) - v(x) - u_\infty K v(x) \neq 0, \\ -\frac{1}{2}(\mathfrak{a} + K v)\hat{w}^2 & \text{otherwise,} \end{cases} \quad (2.32)$$

where \hat{w} is given by a Dual-to-Primal relation in the form

$$\hat{w} = \begin{cases} \frac{\mathfrak{a}\bar{w} - v - u_\infty K v}{\mathfrak{a} + K v} & \text{if } \mathfrak{a} + K v \neq 0, \\ 0 & \text{if } \mathfrak{a} + K v = 0. \end{cases} \quad (2.33)$$

Now, for $0 < L < \infty$ and for any v that is $2L$ -periodic and locally square-integrable, we define

$$\mathfrak{S}[v] = \inf_w \hat{\mathfrak{E}}[w, v]. \quad (2.34)$$

where the inf is taken over all w that are $2L$ -periodic and locally square-integrable. If $L = \infty$ we require $v \in L^2(\mathbb{R})$ and define $\mathfrak{S}[v]$ by the same formula, taking the inf over all $w \in L^2(\mathbb{R})$.

Given any such v , define the sets

$$\begin{aligned} N_v &= \{x \in (-L, L) : \mathbf{a} + Kv < 0\}, \\ Z_v &= \{x \in (-L, L) : \mathbf{a} + Kv = 0 \text{ and } \mathbf{a}\bar{w} - v - u_\infty Kv \neq 0\}. \end{aligned} \quad (2.35)$$

Then with $|\cdot|$ denoting the Lebesgue measure of a set, from (2.32) we infer that

$$\begin{aligned} \mathfrak{S}[v] &= \int_{-L}^L \left(-\frac{1}{2}(\mathbf{a} + Kv)\hat{w}^2 \right) dx \quad \text{if } |N_v| = 0 \text{ and } |Z_v| = 0, \\ \mathfrak{S}[v] &= -\infty \quad \text{if } |N_v| > 0 \text{ or } |Z_v| > 0. \end{aligned} \quad (2.36)$$

At a state v subject to the strict convexity condition

$$\mathcal{C}_s : \quad \mathbf{a} + Kv > 0, \quad (2.37)$$

we have the formula

$$\mathfrak{S}[v] = \int_{-L}^L \left(-\frac{1}{2} \frac{(\mathbf{a}\bar{w} - v - u_\infty Kv)^2}{\mathbf{a} + Kv} \right) dx. \quad (2.38)$$

Let us compute (formally) the first variation of \mathfrak{S} in a direction δv in this case. We find

$$\begin{aligned} \delta^{(1)}\mathfrak{S}[v; \delta v] &= \int_{-L}^L \left(\hat{w}(I + u_\infty K)\delta v + \frac{\hat{w}^2}{2}(K\delta v) \right) dx \\ &= \int_{-L}^L \delta v \left((I + u_\infty K)\hat{w} + \frac{1}{2}K(\hat{w}^2) \right) dx, \end{aligned} \quad (2.39)$$

due to the self-adjointness property (2.25) of the convolution operator K . We find that for a maximizer of $\mathfrak{S}[v]$ that satisfies the strict convexity condition (2.37), *necessarily equation (2.28) holds*.

We point out that here we have a consistency property similar to that in the previous subsection. Namely, if \bar{w} happens to be a ($2L$ -periodic) solution to (2.28), then with $v = 0$ we have $\hat{w} = \bar{w}$ and the first variation vanishes in (2.39). By concavity, $v = 0$ is then a maximizer of \mathfrak{S} , regardless of any degeneracies as will be discussed in Section (3) below.

We remark that the functional \mathfrak{S} enjoys a scaling symmetry in terms of the amplitude parameter \mathbf{a} just like the one for S previously described. Thus \mathbf{a} can be chosen at will for convenience in numerical computations. Also we mention that although the base state \bar{f} from the previous subsection formally corresponds to $u_\infty + \bar{w}$ here, we may expect some differences on a bounded interval, because \bar{w} is considered here to be $2L$ -periodic outside $\Omega = (-L, L)$, while \bar{f} need not be defined there. Thus, e.g., we have no reason to expect that $\hat{f} = u_\infty + \hat{w}$ for respective maximizers of $\tilde{S}[\lambda]$ and $\mathfrak{S}[v]$ when $\bar{f} = u_\infty + \bar{w}$ on Ω .

3. Analysis of concave maximization problems

In principle, by defining the pre-dual functional and just algebraically eliminating the primal field by substituting in the DtP mapping, one obtains a dual functional whose critical points should provide solutions to the primal equation, as long as the denominator in the DtP mapping is non-vanishing. But for the purposes of analysis, it is natural to study the dual function defined by minimization of the pre-dual over primal fields, as we have described. In this section we develop several analytical facts about the resulting concave maximization problem, for both the DDE and NIE formulations.

3.1. Analysis for the dual DDE formulation

We study the DDE formulation specified on a bounded interval with homogeneous boundary condition on dual fields. Given $0 < L < \infty$, let

$$\Omega = (-L, L), \quad \Omega_1 = (-L-1, L+1). \quad (3.1)$$

Consider the functional \tilde{S} from (2.17) and (2.18) for dual fields restricted to lie the Hilbert space of functions $\lambda \in H_0^1(\Omega)$ considered as equal to zero outside Ω . We can take the norm on this space to be

$$\|\lambda\|_{H_0^1(\Omega)} = \|\lambda'\|_{L^2(\Omega)}.$$

Any such $\lambda \in C^0(\Omega)$, and satisfies the Poincare inequality. Below, we will use the notation $\|\cdot\|$ to denote either $\|\cdot\|_{L^2(\Omega)}$ or $\|\cdot\|_{L^2(\Omega_1)}$ as appropriate in context. The pre-dual functional \tilde{S}_L is affine in λ and is evidently continuous in λ for each $f \in L^2(\Omega_1)$. Then the infimum over f renders \tilde{S} an upper semicontinuous functional posed on $H_0^1(\Omega)$. We summarize the analytic properties of \tilde{S} in the following result.

Proposition 3.1. *Let $0 < L < \infty$ and let $\bar{f} \in L^2(\Omega_1)$. The functional \tilde{S} defined by (2.17) is given by (2.18) and maps $H_0^1(\Omega)$ into $[-\infty, \frac{1}{2}\mathbf{a}\|\bar{f}\|^2]$. Moreover, \tilde{S} is concave and upper semicontinuous. The interior of its domain $\text{dom}(\tilde{S}) = \{\lambda \in H_0^1(\Omega) : \tilde{S}[\lambda] > -\infty\}$ is the set*

$$\text{int dom } \tilde{S} = \{\lambda \in H_0^1(\Omega) : \Delta_\lambda = \mathbf{a} + \lambda_{x-1} - \lambda_{x+1} > 0 \text{ on } \Omega_1\}.$$

Furthermore, $-\tilde{S}$ is coercive, and \tilde{S} achieves a maximum. If some maximizer λ lies in $\text{int dom } \tilde{S}$, then the function defined on Ω_1 by

$$\hat{f}(x) = \frac{\mathbf{a}\bar{f} + \lambda'}{\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}}$$

is absolutely continuous inside Ω and is a strong solution of (2.1) there.

This result gives an existence result for some (weak) solution of the primal equation (2.1) on Ω satisfying homogeneous boundary conditions outside this interval, *conditional* on having \tilde{S} admit some maximizer at a state λ where the convexity condition (2.19) holds strictly inside Ω .

Proof. Evidently taking the infimum over f in (2.17) always yields $\tilde{S}(\lambda) \leq \hat{S}_L(0, \lambda) = \frac{1}{2}\mathbf{a}\|\bar{f}\|^2$. Clearly $\hat{S}(f, \lambda)$ depends continuously on $\lambda \in H_0^1(\Omega)$, so the concavity and upper semicontinuity follow by basic results in convex analysis, see [37, Chap. 1] or [38, Prop. 9.2].

Regarding the interior of the domain, if $\Delta_\lambda > 0$ on Ω_1 then it is bounded below there (it is continuous and approaches \mathbf{a} at the boundary) and clearly $\tilde{S}[\lambda]$ is finite in a neighborhood of λ in $H_0^1(\Omega)$. And conversely if $\Delta_\lambda = 0$ at some point in Ω_1 then a small perturbation of λ can make it negative, which makes \tilde{S} infinite.

Next we prove coercivity. Given λ at which $\tilde{S}[\lambda]$ in (2.18) is finite and using (2.9)-(2.10), we find

$$\tilde{S}[\lambda] = \frac{\mathbf{a}}{2}\|\bar{f}\|^2 - \frac{1}{2} \int_{\Omega_1 \setminus \Omega_1^*} \frac{(\lambda' + \mathbf{a}\bar{f})^2}{\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}} dx, \quad (3.2)$$

where Ω_1^* is the subset of Ω_1 where Δ_λ and $\lambda' + \mathbf{a}\bar{f}$ both vanish. Define

$$\lambda_m = \sup_{x \in \Omega} |\lambda(x)|,$$

and note that $0 \leq \mathbf{a} + \lambda_{x-1} - \lambda_{x+1} \leq \mathbf{a} + 2\lambda_m$ so that

$$\frac{-(\lambda' + \mathbf{a}\bar{f})^2}{\mathbf{a} + \lambda_{x-1} - \lambda_{x+1}} \leq \frac{-(\lambda' + \mathbf{a}\bar{f})^2}{\mathbf{a} + 2\lambda_m},$$

and that $\lambda' + \mathbf{a}\bar{f} = 0$ on Ω_1^* . Then

$$\begin{aligned} \left(\tilde{S}[\lambda] - \frac{\mathbf{a}}{2}\|\bar{f}\|^2\right) \cdot 2(\mathbf{a} + 2\lambda_m) &\leq - \int_{\Omega_1 \setminus \Omega_1^*} (\lambda' + \mathbf{a}\bar{f})^2 dx = - \int_{\Omega_1} (\lambda' + \mathbf{a}\bar{f})^2 dx \\ &\leq \int_{\Omega} \left(-(\lambda')^2 - 2\mathbf{a}\bar{f}\lambda'\right) dx \leq 2\mathbf{a}\|\bar{f}\|\|\lambda'\| - \|\lambda'\|^2. \end{aligned}$$

Since $\|\lambda'\| = \|\lambda\|_{H_0^1(\Omega)}$ it follows that $-\tilde{S}$ is coercive on $H_0^1(\Omega)$, which means that

$$\tilde{S}[\lambda] \rightarrow -\infty \quad \text{as} \quad \|\lambda\|_{H_0^1(\Omega)} \rightarrow \infty.$$

A maximizer of \tilde{S} therefore exists, because $-\tilde{S}$ is proper (i.e., finite somewhere) and a standard convex analysis result states that a proper, convex and lower semicontinuous function that is coercive has a minimizer. (E.g., see [38, Thm. 11.10].)

Since $\Delta_\lambda > 0$ for each $\lambda \in \text{int dom } \tilde{S}$, it is clear from (3.2) (with Ω_* empty) that \tilde{S} is Fréchet differentiable in a neighborhood, with (directional) derivative given by the first variation. Let us next compute this first variation at such a state. Recall that λ and its variations $\delta\lambda$ are taken to vanish outside the interval $\Omega = (-L, L)$. Then λ_{x-1} vanishes for $x < -L + 1$, so

$$\int_{\Omega_1} \hat{f}^2 \lambda_{x-1} dx = \int_{-L+1}^{L+1} \hat{f}^2 \lambda_{x-1} dx = \int_{\Omega} \hat{f}_{x+1}^2 \lambda dx,$$

and similarly for integrands with $\hat{f}^2 \lambda_{x+1}$ and with λ replaced by $\delta\lambda$. We find therefore that, like for the computation of the variation on the whole line,

$$\delta^{(1)} S_L[\lambda; \delta\lambda] = \int_{-L}^L \left(-\delta\lambda' \hat{f} + \delta\lambda \cdot \frac{1}{2} (\hat{f}_{x+1}^2 - \hat{f}_{x-1}^2) \right) dx. \quad (3.3)$$

Thus, at a maximizer of \tilde{S} where $\Delta_\lambda > 0$ in Ω_1 , we can infer that $\hat{f} \in L^2(\Omega_1)$ and is a weak solution of the primal equation (2.1) on the interval $\Omega = (-L, L)$. Since the functions $\hat{f}_{\pm 1}$ are square integrable in Ω , equation (2.1) holds strongly in $L^1(\Omega)$ and (by integration) it follows \hat{f} is absolutely continuous on Ω . \square

Remark 3.2. From (2.1), inductively we can infer higher regularity on smaller sets: \hat{f} is C^1 inside the interval $(-L + 1, L - 1)$, C^2 inside the interval $(-L + 2, L - 2)$, etc. Even if \bar{f} is smooth, however, \hat{f} may be discontinuous at $\pm L$ due to a potential discontinuity in λ' at the endpoints of Ω . In this case we could infer that \hat{f} is C^1 throughout Ω except at $\pm(L - 1)$, and further that \hat{f} is smooth in Ω outside the set of points $\pm(L - k)$ where $k \in \mathbb{N}$.

3.2. Analysis for the dual NIE formulation

In this subsection we study the NIE formulation on both finite and infinite intervals. Using this formulation we will obtain conditional existence results for C^∞ periodic solutions of the primal equation (2.28), instead of imposing homogeneous boundary conditions outside a bounded interval. The price is that the coercivity analysis turns out to be more involved, and we do not establish coercivity for all values of the parameter u_∞ .

Throughout this section, we fix $\mathbf{a} > 0$ and $\bar{u} \in L^2(\Omega)$. For any $L \in (0, \infty]$, let $\Omega = (-L, L)$. If $L = \infty$, the formula (2.24) defines K as a bounded linear operator from $L^2(\Omega)$ into $H^1(\Omega)$. The same holds if $0 < L < \infty$, by extending $v \in L^2(\Omega)$ to be $2L$ -periodic and considering Kv as the restriction of (2.24) to Ω . For later use we also define $K_0 v$ to be given by extending v to be zero outside Ω and restricting formula (2.24) to Ω .

We summarize several properties of the dual functional \mathfrak{S} in the following proposition. Note that for any v and w in $L^2(\Omega)$, the pre-dual integrand $\mathfrak{L}_{\mathfrak{S}}$ in (2.31) is integrable on Ω , which is evident since Kv is in $L^2(\Omega)$ and is continuous and bounded.

Proposition 3.3. *The functional \mathfrak{S} defined by (2.34), i.e., by*

$$\mathfrak{S}[v] = \inf_{w \in L^2(\Omega)} \int_{-L}^L \mathfrak{L}_{\mathfrak{S}}(w, v, x) dx,$$

is given by (2.36) and maps $L^2(\Omega)$ into $[-\infty, 0]$. Moreover, \mathfrak{S} is concave and upper semicontinuous. The interior of its domain $\text{dom } \mathfrak{S} = \{v \in L^2(\Omega) : \mathfrak{S}[v] > -\infty\}$ is the set where (2.37) holds, i.e.,

$$\text{int dom } \mathfrak{S} = \{v \in L^2(\Omega) : \mathbf{a} + Kv > 0 \text{ on } \Omega\}.$$

If additionally $-\mathfrak{S}$ is coercive, then a maximizer of \mathfrak{S} exists. Also, if \mathfrak{S} has some maximizer $v \in \text{int dom } \mathfrak{S}$, then the function \hat{w} given by (2.33) is a C^∞ solution of (2.28).

Proof. Since $\mathfrak{L}_\mathfrak{S}(0, \nu, x) = 0$, the infimum defining $\mathfrak{S}[\nu]$ is non-positive. Since $\hat{\mathfrak{E}}(w, \nu)$ is continuous and affine in ν , the convexity and lower semicontinuity of $-\mathfrak{S}$ follow as in the previous subsection.

From (2.32)–(2.33), evidently if (2.37) holds then $\mathfrak{S}[\nu]$ is finite and remains finite in a neighborhood of ν in $L^2(\Omega)$, so $\nu \in \text{int dom } \mathfrak{S}$. On the other hand, if $\mathfrak{S}[\nu]$ is finite but (2.37) does not hold, then $\mathfrak{a} + K\nu(x) = 0$ for some $x \in \Omega$. For any nearby $\tilde{\nu} < \nu$ locally near x , we get $\mathfrak{a} + K\tilde{\nu}(x) < 0$, whence $\mathfrak{S}[\tilde{\nu}] = -\infty$ so $\nu \notin \text{int dom } \mathfrak{S}$.

We note that $\hat{w} = \bar{w}$ when $\nu = 0$, so $\mathfrak{S}[0] = -\frac{1}{2}\mathfrak{a}\|\bar{w}\|_{L^2(\Omega)}^2$. Then $-\mathfrak{S}$ is proper, convex and lower semicontinuous, hence \mathfrak{S} has a maximizer.

Finally, if \mathfrak{S} has a maximizer ν in the interior of its domain, then \mathfrak{S} is differentiable at ν and the first variation must vanish, implying that \hat{w} solves (2.28) as shown in section 2.2 above. We claim \hat{w} is C^∞ . The function $K\nu$ is in $H^1(\Omega)$, so is continuous (and vanishes in the limit $x \rightarrow \pm\infty$ if $L = \infty$), and $\mathfrak{a} + K\nu$ has a positive minimum and bounded maximum:

$$0 < \mathfrak{a}_{\min} \leq \mathfrak{a} + K\nu \leq \mathfrak{a}_{\max} < \infty. \quad (3.4)$$

The function \hat{w} given by (2.33) is then in $L^2(\Omega)$. As $K\hat{w}$ and $K(\hat{w}^2)$ are absolutely continuous and \hat{w} satisfies (2.28), \hat{w} is absolutely continuous also, whence by bootstrapping (induction) we find \hat{w} is C^k for all k . \square

Recall that to say $-\mathfrak{S}$ is coercive means that $-\mathfrak{S}[\nu] \rightarrow \infty$ as $\|\nu\|_{L^2(\Omega)} \rightarrow \infty$. In most circumstances we can prove the function $-\mathfrak{S}$ is indeed coercive. This is easiest when the bounded operator $I + u_\infty K$ on $L^2(\Omega)$ has bounded inverse, which is evidently the case when $|u_\infty|$ is small enough, for example.

Proposition 3.4. *Suppose $I + u_\infty K$ has bounded inverse on $L^2(\Omega)$. Then $-\mathfrak{S}$ is coercive.*

Proof. Throughout the proof, $\|\cdot\|$ denotes the norm in $L^2(\Omega)$. We begin with two preliminary estimates. First, due to the invertibility hypothesis, there is a constant $\alpha > 0$ such that for all $\nu \in L^2(\Omega)$,

$$\|(I + u_\infty K)\nu\| \geq \alpha\|\nu\|,$$

whence

$$\|\mathfrak{a}\bar{w} - (I + u_\infty K)\nu\| \geq \|(I + u_\infty K)\nu\| - \|\mathfrak{a}\bar{w}\| \geq \alpha\|\nu\| - \|\mathfrak{a}\bar{w}\|.$$

Second, by the Cauchy-Schwarz inequality we find (recall we extend ν as $2L$ -periodic if needed)

$$|K\nu(x)| = \left| \int_{x-1}^{x+1} 1 \cdot \nu(y) dy \right| \leq C_K \|\nu\|,$$

where C_K is a constant independent of ν (and equal to $\sqrt{2}$ if $L > 1$). Therefore

$$\mathfrak{a}_{\max} = \max_x |\mathfrak{a} + K\nu| \leq \mathfrak{a} + C_K \|\nu\|.$$

2. Now, define the set

$$P_\nu = \{x \in \Omega : \mathfrak{a} + K\nu > 0\}. \quad (3.5)$$

If $\mathfrak{S}[\nu] > -\infty$ then by (2.36) and the definition of \hat{w} in (2.33),

$$-2\mathfrak{S}[\nu] = \int_{P_\nu} \frac{(\mathfrak{a}\bar{w} - \nu - u_\infty K\nu)^2}{\mathfrak{a} + K\nu} dx \geq \frac{1}{\mathfrak{a}_{\max}} \int_{P_\nu} |\mathfrak{a}\bar{w} - \nu - u_\infty K\nu|^2 dx.$$

However, because $|N_\nu| = |Z_\nu| = 0$, we have $\mathfrak{a}\bar{w} - \nu - u_\infty K\nu = 0$ a.e. on the complement of P_ν . Hence, the domain of integration can be extended from P_ν to all of $\Omega = (-L, L)$. Thus for all $\nu \in L^2(\Omega)$ we have

$$-2\mathfrak{S}[\nu] \geq \frac{\|\mathfrak{a}\bar{w} - (I + u_\infty K)\nu\|^2}{\mathfrak{a} + C_K \|\nu\|} \geq \frac{(\alpha\|\nu\| - \|\mathfrak{a}\bar{w}\|)^2}{\mathfrak{a} + C_K \|\nu\|}. \quad (3.6)$$

The right-hand side tends to ∞ as $\|\nu\| \rightarrow \infty$, which establishes the coercivity as claimed. \square

We can determine precisely when bounded invertibility holds by locating the spectrum of K using the Fourier transform, with the following result. Define the Fourier transform of $v \in L^2(\Omega)$ by

$$\mathcal{F}v(\xi) = \int_{-L}^L e^{-i\xi x} v(x) dx$$

on the Fourier domain given by $\xi \in \Omega^*$, where

$$\Omega^* := \begin{cases} (-\infty, \infty) & \text{if } L = \infty, \\ \{k\pi/L : k \in \mathbb{Z}\} & \text{if } 0 < L < \infty. \end{cases} \quad (3.7)$$

Then by a straightforward computation,

$$\mathcal{F}(Kv)(\xi) = (2 \operatorname{sinc} \xi) \mathcal{F}v(\xi), \quad \text{where} \quad \operatorname{sinc}(\xi) = \begin{cases} \frac{\sin \xi}{\xi} & \xi \neq 0, \\ 1 & \xi = 0. \end{cases}$$

It follows that $(I + u_\infty K)v = w$ if and only if

$$(1 + 2u_\infty \operatorname{sinc} \xi) \mathcal{F}v(\xi) = \mathcal{F}w(\xi) \quad \text{for all } \xi \in \Omega^*,$$

and $I + u_\infty K$ has bounded inverse if and only if $(1 + 2u_\infty \operatorname{sinc} \xi)^{-1}$ is uniformly bounded on Ω^* .

Notice that the range $\{2 \operatorname{sinc} \xi : \xi \in \Omega^*\}$ is a closed interval $[-\sigma_0, 2]$ with $-\sigma_0 \approx -0.434467$ for $L = \infty$, and is a discrete sequence of not-necessarily-distinct values converging to zero for $0 < L < \infty$. Thus we find the following.

Proposition 3.5. *The operator $I + u_\infty K$ has bounded inverse if and only if*

$$1 + 2u_\infty \operatorname{sinc} \xi \neq 0 \quad \text{for all } \xi \in \Omega^*.$$

In particular $I + u_\infty K$ is invertible whenever

$$-\frac{1}{2} < u_\infty < 1/\sigma_0 \approx 2.30167.$$

For $0 < L < \infty$ each number $1 + 2u_\infty \operatorname{sinc}(k\pi/L)$ ($k \in \mathbb{Z}$) is an eigenvalue of finite multiplicity.

Remark 3.6. The condition in proposition 3.5 has a physical meaning. Namely, it corresponds to the *phase-speed non-matching condition* mentioned in the introduction. If we undo the scaling $f \mapsto -2cf$, then the constant solution $f = u_\infty$ in (2.2) corresponds to the constant state $\bar{u} = -cu_\infty$ for equation (1.2). For the linearization at this state, waves $e^{i\xi j - i\omega t}$ must satisfy the dispersion relation

$$\omega = \bar{u} \sin \xi. \quad (3.8)$$

Then the phase-speed non-matching requirement that $c \neq \omega/\xi$ is equivalent to the condition that $1 + 2u_\infty \operatorname{sinc} \xi \neq 0$ for all $\xi \in \Omega^*$ as stated in the proposition.

In general, when L is finite we can also prove coercivity for any value of $u_\infty \neq -\frac{1}{2}$, with a somewhat more involved argument.

Proposition 3.7. *Suppose $\Omega = (-L, L)$ is bounded. If $u_\infty \neq -\frac{1}{2}$, then $-\mathfrak{S}$ is coercive on $L^2(\Omega)$. If $u_\infty = -\frac{1}{2}$, the functional $-\mathfrak{S}$ is coercive on the subspace of $L^2(\Omega)$ consisting of functions with mean zero.*

Proof. 1. For the value $u_\infty = -\frac{1}{2}$, the self-adjoint operator $I + u_\infty K$ has one-dimensional kernel spanned by the constant function 1. Restricting \mathfrak{S} to the orthogonal complement of this kernel, the proof of coercivity works the same as the proof of proposition 3.4.

2. Suppose $u_\infty \neq -\frac{1}{2}$ but $I + u_\infty K$ is singular. Decompose $L^2(\Omega)$ into the finite-dimensional kernel V of $I + u_\infty K$ and its orthogonal complement W on which $I + u_\infty K$ has bounded inverse. Each nonzero $v \in V$ is a trigonometric polynomial with mean zero, so it is impossible that $Kv(x) \geq 0$ for all x . Then by a compactness argument, there exists some $\sigma_1 < 0$ such that whenever $\|v\| = 1$ then $\min_x Kv(x) \leq \sigma_1$.

3. Let (v_n) be a sequence in $L^2(\Omega)$ with $\|v_n\| \rightarrow \infty$ as $n \rightarrow \infty$. For each n decompose v_n as $\mu_n + \eta_n$ with $\mu_n \in V$ and $\eta_n \in W$. There are now two cases: (i) Suppose that for some positive constant C , $\|\mu_n\| \leq C\|\eta_n\|$ for all n . Then $\|v_n\| \leq (1+C)\|\eta_n\| \rightarrow \infty$ as $n \rightarrow \infty$, and

$$\|\mathbf{a}\bar{w} - (I + u_\infty K)v_n\| = \|\mathbf{a}\bar{w} - (I + u_\infty K)\eta_n\| \geq \alpha\|\eta_n\| - \|\mathbf{a}\bar{w}\| \geq \hat{\alpha}\|v_n\| - \|\mathbf{a}\bar{w}\|,$$

where $\hat{\alpha} = \alpha/(1+C)$. One then infers that $-\mathfrak{E}[v_n] \rightarrow \infty$ by the same argument as before.

(ii) If it is false that case (i) holds for some C , then there must be a subsequence of (v_n) (denoted the same) such that $\|\mu_n\| \geq n\|\eta_n\|$ for all n , with $\|\mu_n\| \rightarrow \infty$. By step 1, we note that

$$\min_x K\mu_n(x) \leq \|\mu_n\|\sigma_1 < 0,$$

and since $Kv_n = K\mu_n + K\eta_n$,

$$\min_x (\mathbf{a} + Kv_n) \leq \mathbf{a} + \|\mu_n\|\sigma_1 + C_K\|\eta_n\| \leq \mathbf{a} + \|\mu_n\|(\sigma_1 + C_K/n).$$

This is strictly negative for sufficiently large n , and when this is the case we must have $-\mathfrak{E}[v_n] = \infty$. This finishes the proof. \square

We suspect coercivity may always hold when $L = \infty$ as well. In any case, we only get a proof that a solution to the NIE (2.28) exists *on the condition* that a maximizer of \mathfrak{E} exists that belongs to the interior of the domain of \mathfrak{E} . Presently, despite strong numerical evidence in favor as shown below, we lack any proof that such a maximizer exists, for any values of L and u_∞ .

Second variation. In general, at any point in the interior of the domain of \mathfrak{E} , its second variation can be found by substituting $v + t\delta w$ into (2.33) and (2.39) and differentiating at $t = 0$ to find that

$$\left. \frac{d}{dt} \hat{w} \right|_{t=0} = -\frac{(I + u_\infty K)\delta w}{\mathbf{a} + Kv} - \frac{\hat{w}K(\delta w)}{\mathbf{a} + Kv} = -\frac{(I + \hat{f}K)\delta w}{\mathbf{a} + Kv},$$

where

$$\hat{f} = u_\infty + \hat{w} = \frac{\mathbf{a}(u_\infty + \bar{w}) - v}{\mathbf{a} + Kv}. \quad (3.9)$$

Then differentiation of (2.39) yields

$$\delta^{(2)}\mathfrak{E}[v; \delta v, \delta w] = -\int_{-\infty}^{\infty} \delta v (I + K\hat{f})(\mathbf{a} + Kv)^{-1} (I + \hat{f}K)\delta w \, dx. \quad (3.10)$$

Here $K\hat{f}$ is regarded as composition of operators with $K\hat{f}g = K(\hat{f}g)$. Since K is self-adjoint we find

$$\delta^{(2)}\mathfrak{E}[v; \delta v, \delta v] = -\int_{-\infty}^{\infty} (\mathbf{a} + Kv)^{-1} |(I + \hat{f}K)\delta v|^2 \, dx. \quad (3.11)$$

Indicating explicitly the dependence upon the parameter \mathbf{a} , this enjoys the scaling property

$$\delta^{(2)}\mathfrak{E}_{\mathbf{a}}[v; \delta v, \delta v] = \mathbf{a}^{-1} \delta^{(2)}\mathfrak{E}_1[\mathbf{a}^{-1}v; \delta v, \delta v]. \quad (3.12)$$

Critical points and translational invariance. Now let $L \in (0, \infty]$ and suppose v is a maximizer of $\mathfrak{E}[v]$ belonging to the interior of its domain, so that the strict convexity condition (2.37) holds. Then as stated in proposition 3.3, \hat{w} as given by (2.33) is a smooth solution of (2.28) so that $\hat{f} = u_\infty + \hat{w}$ from (3.9) is a smooth solution of (2.2) with $C_1 = u_\infty + u_\infty^2$ as in (2.27).

Now, equation (2.2) is translation invariant, meaning that if f is a solution on \mathbb{R} , then the function $x \mapsto f(x+h)$ is a solution for any real h . Differentiating with respect to h at $h = 0$ we find that

$$f'(x) + \int_{x-1}^{x+1} f(y)f'(y) \, dy = 0, \quad \text{i.e.,} \quad f' + K(ff') = 0.$$

Thus the operator $I + Kf$ has f' in its kernel. Multiplying this equation by $f(x)$ we find that

$$f(x)f'(x) + f(x) \int_{x-1}^{x+1} f(y)f'(y) dy = 0 \quad \text{i.e.,} \quad (I + fK)(ff') = 0.$$

That is, the (adjoint) operator $I + fK$ has the function ff' in its kernel.

For the maximizer ν this means that the second variation vanishes in (3.11) for the variation

$$\delta\nu = \hat{f}\hat{f}' . \tag{3.13}$$

Indeed, $(I + \hat{f}K)(\hat{f}\hat{f}') = 0$. Now consider first the case when L is finite. The operator $\hat{f}K$ acting on $L^2(\Omega)$ is then always a compact operator, since the embedding of $H^1(\Omega)$ into $L^2(\Omega)$ is compact. From the Riesz-Schauder spectral theory of compact operators, the eigenvalue 0 is necessarily an isolated eigenvalue of $I + \hat{f}K$ and has a finite-dimensional (generalized) eigenspace, which we denote by Z_0 .

In the case $L = \infty$ when $\Omega = \mathbb{R}$, the operator K maps $L^2(\mathbb{R})$ into $H^1(\mathbb{R})$ but is not compact. Because \hat{w} is smooth with limit 0 at $\pm\infty$, though, the operator $\hat{w}K$ is compact on $L^2(\mathbb{R})$, due to the convenient compactness criteria of [39]. If we assume $I + u_\infty K$ has bounded inverse on $L^2(\mathbb{R})$, which is natural to ensure coercivity according to proposition 3.4, then since $\hat{f} = u_\infty + \hat{w}$, the operator $I + \hat{f}K$ will be the sum of an invertible operator and a compact one, i.e., Fredholm of index zero. Then the Riesz-Schauder theory ensures again that the eigenvalue 0 is isolated with finite-dimensional generalized eigenspace Z_0 .

Conditional strict coercivity of second variation. We expect, but are unable to prove, that Z_0 is one-dimensional and spanned by $\hat{f}\hat{f}'$. In any case, if $Y \subset L^2(\Omega)$ is any subspace complementary to Z_0 , then necessarily the operator $I + \hat{f}K$ is bounded below on Y , meaning that for some constant $\kappa_Y > 0$,

$$\|(I + \hat{f}K)u\|_{L^2} \geq \kappa_Y \|u\|_{L^2} \quad \text{for all } u \in Y.$$

This means that we have (conditional) strict coercivity in (3.11) for variations in Y , with

$$-\delta^{(2)} \mathfrak{S}[\nu; u, u] \geq \frac{\kappa_Y^2}{\mathfrak{a}_{\max}} \|u\|_{L^2}^2 \quad \text{for all } u \in Y. \tag{3.14}$$

Our numerics suggests that Z_0 is one-dimensional and the solitary wave can be chosen even, so $\hat{f}\hat{f}'$ is odd. One could take Y to consist of the even functions in L^2 , then.

4. Approximation and numerical examples for the DDE formulation

In this section and the next, we describe the numerical approximations and algorithms that we have used to obtain approximate solutions of our DDE and NIE variational formulations. Consistent with the analysis of the previous sections, we apply different boundary conditions in the two cases, imposing a Dirichlet (vanishing) condition on the dual field in the DDE formulation, and periodicity in the NIE formulation.

For the DDE case, we use finite elements and develop a modified Newton-Raphson scheme to solve the weak form of the Euler-Lagrange equation (2.16) for the dual field λ . Moreover, we implement an algorithm that adaptively adjusts the base state to facilitate convergence while respecting a discrete version of the Convexity Condition (2.19). We carry out a number of numerical experiments, exploring how different choices of the initial base state yield a considerable variety of wave profiles.

For the NIE case, we will discretize the objective functional \mathfrak{S} using a simple quadrature scheme that is spectrally accurate, and use standard optimization software to find approximate optimizers. We adjust the base state in steps along with the parameter u_∞ to perform basic path-following along a branch of solutions. This results in identifying solutions in regimes of interest different from the DDE case.

4.1. Approximation for the DDE formulation

We approximate weak solutions of the DDE (2.16), generating a weak form for solutions on a finite domain $\Omega = (-L, L)$ as follows: We generate a residual by multiplying (2.16) with a test function $\delta\lambda$ that vanishes outside Ω and integrating. After integration by parts, this yields:

$$R[\lambda; \delta\lambda] := \int_{-L}^L \left(-\delta\lambda' \hat{f} + \frac{\delta\lambda}{2} (\hat{f}_{x+1}^2 - \hat{f}_{x-1}^2) \right) dx, \quad (4.1)$$

where we have eliminated the boundary terms by imposing boundary conditions $\delta\lambda(\pm L) = 0$. Since the value of $\hat{f}(x)$ from (2.9) depends upon values of λ at $x-1$ and $x+1$, defining the terms \hat{f}_{x-1}^2 and \hat{f}_{x+1}^2 in this integrand requires that λ be defined in the extended domain $\Omega_2 = (-L-2, L+2)$. Thus we find it suffices to describe a weak form for the dual problem as follows:

Find $\lambda : (-L-2, L+2) \rightarrow \mathbb{R}$, satisfying $\lambda(x) = 0$ whenever $x \notin (-L, L)$, such that for any $\delta\lambda$ satisfying $\delta\lambda(x) = 0$ whenever $x \notin (-L, L)$,

$$\int_{-L}^L \left(-\delta\lambda' \hat{f} + \frac{\delta\lambda}{2} (\hat{f}_{x+1}^2 - \hat{f}_{x-1}^2) \right) dx = R[\lambda; \delta\lambda] = 0. \quad (4.2)$$

Here $\hat{f}(x)$ is determined for $x \in (-L-1, L+1)$ in terms of λ by the DtP map (2.9).

The weak form in (4.2) is the same problem that is satisfied by a maximizer of the functional \tilde{S} that lies in the interior of its domain, as shown in Proposition 3.1. By the same arguments as in the proof of that result, for any $\lambda \in H^1(\Omega_2)$ satisfying the weak formulation (4.2), \hat{f} is absolutely continuous inside $\Omega = (-L, L)$ and is a strong solution of (2.1) there. And \hat{f} enjoys additional regularity properties as described in the remark following the proposition.

4.1.1. A modified Newton-Raphson scheme with step-size control

The solutions to (4.2) are obtained via a Newton Raphson (N-R) scheme based algorithm: A nonlocal Galerkin Finite Element method has been implemented to approximate discrete solutions to (4.2). We start by considering the variation of $R[\lambda; \delta\lambda]$ (eq. (4.2)) in a direction $d\lambda$ given by:

$$J_{\lambda}[d\lambda; \delta\lambda] = \int_{-L}^L \left(-\delta\lambda' \frac{\partial \hat{f}}{\partial \mathcal{D}} d\mathcal{D} + \delta\lambda \left(\hat{f}_{x+1} \frac{\partial \hat{f}_{x+1}}{\partial \mathcal{D}_{x+1}} d\mathcal{D}_{x+1} - \hat{f}_{x-1} \frac{\partial \hat{f}_{x-1}}{\partial \mathcal{D}_{x-1}} d\mathcal{D}_{x-1} \right) \right) dx, \quad (4.3)$$

where

$$\mathcal{D}_{x+c} \equiv \{\lambda_{x+c-1}, \lambda'_{x+c}, \lambda_{x+c+1}\},$$

and

$$\frac{\partial \hat{f}_{x+c}}{\partial \lambda'_{x+c}} = \frac{1}{\lambda_{x+c-1} - \lambda_{x+c+1} + \mathbf{a}}; \quad \frac{\partial \hat{f}_{x+c}}{\partial \lambda_{x+c\pm 1}} = \pm \frac{\mathbf{a} \bar{f}_{x+c} + \lambda'_{x+c}}{(\lambda_{x+c-1} - \lambda_{x+c+1} + \mathbf{a})^2}. \quad (4.4)$$

In the following, we will use the summation convention on repeated indices. We discretize the extended domain and approximate various fields on it as follows:

$$\lambda(x) = \lambda^i N^i(x); \quad \delta\lambda(x) = \delta\lambda^i N^i(x); \quad d\lambda(x) = d\lambda^i N^i(x),$$

where first-order $C^0(\Omega)$ shape functions N^i are considered, and i runs over the nodes of the discretized extended domain. Let x^A denote the position of any node A on the extended domain and define a set \mathcal{S} of nodal indices as follows:

$$\mathcal{S} = \{B \mid x^B \in (-L, L)\}.$$

Our objective is to identify the coefficients λ^A for all nodes $A \in \mathcal{S}$ such that the discrete residual generated from (4.2), when equated to 0, gets satisfied. The discrete residual can be given as:

$$R^A = \int_{-L}^L \left(-(N^A)' \hat{f} + \frac{N^A}{2} (\hat{f}_{x+1}^2 - \hat{f}_{x-1}^2) \right) dx, \quad (4.5)$$

where \hat{f} is now depends on the discretized dual field of λ . Correspondingly, the Jacobian (4.3) can be discretized as:

$$J_{|\lambda}[d\lambda; \delta\lambda] = \delta\lambda^A J^{AB}(\lambda) d\lambda^B,$$

where

$$\begin{aligned} J^{AB} = \int_{-L}^L \left(- (N^A)' \left(\frac{\partial \hat{f}}{\partial \lambda'_x} (N^B)'(x) + \frac{\partial \hat{f}}{\partial \lambda_{x-1}} N^B(x-1) + \frac{\partial \hat{f}}{\partial \lambda_{x+1}} N^B(x+1) \right) \right. \\ \left. + N^A \hat{f}_{x+1} \left(\frac{\partial \hat{f}_{x+1}}{\partial \lambda'_{x+1}} (N^B)'(x+1) + \frac{\partial \hat{f}_{x+1}}{\partial \lambda_x} N^B(x) + \frac{\partial \hat{f}_{x+1}}{\partial \lambda_{x+2}} N^B(x+2) \right) \right. \\ \left. - N^A \hat{f}_{x-1} \left(\frac{\partial \hat{f}_{x-1}}{\partial \lambda'_{x-1}} (N^B)'(x-1) + \frac{\partial \hat{f}_{x-1}}{\partial \lambda_{x-2}} N^B(x-2) + \frac{\partial \hat{f}_{x-1}}{\partial \lambda_x} N^B(x) \right) \right) dx, \quad (4.6) \end{aligned}$$

and (2.9) and (4.4) can be utilized to evaluate the above expression.

To generate corrections for dual field, we implement a modification to the generic N-R scheme based on the following steps:

$$\begin{aligned} -R^A(\lambda^{(k-1)}) &= J^{AB}(\lambda^{(k-1)}) d\lambda^B; \\ \lambda^{(k)} &= \lambda^{(k-1)} + \alpha d\lambda, \end{aligned} \quad (4.7)$$

where $\lambda^{(k)}$ denotes the discretized dual field of λ at k^{th} iterate and α is a *step-size control* factor. $\alpha = 1$ in (4.7) implies a simple N-R scheme. The introduction of α has been motivated below.

In cases where the base state is set far away from any of the potential solutions of (2.1), the correction obtained via a simple N-R scheme can potentially lead to a dual field which violates the convexity condition \mathcal{C} (2.19). For such cases, we stipulate the following condition on any discretely obtained dual iterate:

$$\mathcal{C}_d : \quad \min_{\Omega} \frac{\Delta \lambda}{\mathbf{a}} > T$$

where $T \in [0, 1]$ represents a threshold value such that a large value of T indicates a large denominator. We generally opt for $T > 0.1$ with larger values implying that the dual fields take values away from the convexity boundary. However, smaller values of T can also be chosen; a value of $T = 0.01$ yields an approximation with a residual tolerance of 10^{-11} .

To satisfy \mathcal{C}_d , we control the value of α in the N-R iterates via Alg. 1. Starting from any fixed base state, we use $\lambda^{(0)} = 0$ at each node and $\alpha = 1$. Each time a correction leads to a dual field which violates \mathcal{C}_d at any point in the domain, we reduce α by a factor of 2 and re-evaluate the correction until the criteria is met. If the factor α attains too a small value judged by a threshold (set to 0.01 in the presented calculations), we stop the step-size controlled N-R and declare the primal field $f^{(H)}$ (evaluated at Gauss points) corresponding to the current dual iterate λ as the best improvement that can be obtained starting from the current base state. Using this primal field as a base state, we restart the step-size controlled N-R scheme with $\alpha = 1$. Such occasional base state resets followed by controlled N-R steps are carried out until convergence on residual (4.2) is reached while ensuring that \mathcal{C}_d remains satisfied. The algorithm has been summarized in Alg. 1.

The convergence criteria is given as:

$$\mathbf{tol} : \quad \max_A |R^A| < \mathbf{tol} \quad \forall A \in \mathcal{S}, \quad (4.8)$$

where \mathbf{tol} is a user-defined tolerance. In the following,

$$\mathbf{tol} = 10^{-12} \quad \text{for all problems solved.} \quad (4.9)$$

Algorithm for modified N-R scheme with step-size control

Choose the values for L , N and **tol**. $T = 0.95$ and $\mathbf{a} = 10^6$ and can be changed. Set $\alpha = 1$ and choose a starting base state \bar{f} . Set $c = 0$ (c : base state reset counter).

1. Global Loop:**WHILE** $c \geq 0$:

- (i) Set $k = 0$ (k : N-R counter) and $\lambda^{(0)} = 0$ at each node.
 - (ii) Find $\lambda^{(1)}$ using step no. 2 and set $k = 1$.
 - (iii) **IF** $\max_A |R^A[\lambda^{(1)}]| > \max_A |R^A[\lambda^{(0)}]|$ then **ABORT**
 - (iv) **WHILE** $k > 0$:
 - (a) **IF** $\max_A |R^A[\lambda^{(k)}]| < \mathbf{tol}$: **EXIT** and **GOTO** step no. 5.
 - (b) **IF** $\alpha < 0.01 \rightarrow$ Reset base state (**GOTO** step no. 4).
 - (c) Find $\lambda^{(k+1)}$ using using step no. 2.
 - (d) **DO** $k = k + 1$ and **GOTO** step no. 1iv.
-

2. Evaluate Increment:

- (i) Evaluate $R^A[\lambda^{(k)}]$ and $J^{AB}[\lambda^{(k)}]$.
 - (ii) Find $\lambda^{(k+1)}$ based on current α from (4.7) and check for condition (\mathcal{C}_d) (step no. 3).
IF \mathcal{C}_d is not satisfied: **DO** $\alpha = \alpha/2$ and redo current step.
ELSE return $\lambda^{(k+1)}$.
-

3. Condition \mathcal{C}_d :

- (i) Evaluate Δ_λ (using (2.13)) on discretized domain of Ω based on $\lambda^{(k+1)}$.
- (ii) Check for

$$\max_{\Omega} \frac{\Delta_\lambda}{\mathbf{a}} > T$$

4. Reset base state and restart N-R:

- (i) Evaluate $f^{(H)}$ based on λ^k (using (2.9)) on the Gauss points of discretized domain.
 - (ii) Set $\bar{f} = f^{(H)}$ at the Gauss points.
 - (iii) Do $c = c + 1$, $\alpha = 1$ and restart N-R from step no. 1.
-

5. Perform an L^2 projection to obtain $f^{(H)}$ at nodes. This establishes the solution.
-

Table 1: Algorithm to solve (2.16). N represents the number of elements used to discretize the extended domain. Step Iiii indicates that condition \mathcal{C}_d was satisfied but the Newton residual did not decrease. The algorithm currently is applicable to a wide range of base states, even those far from the solutions. It fails for base states close to being outside the function class allowed for the problem (for e.g. discontinuous base states).

4.2. Numerical examples for the DDE formulation

In each of the following examples, we discretize the extended domain where $x \in (-L-2, L+2)$, choose a base state for this extended domain, and allow the dual scheme to pick up a solution within the domain of interest, $x \in (-L, L)$. Following (4.2), we employ $\lambda_s(x) = 0$ (without loss of generality). The figures produced in the following sections are based on a standard L^2 projection performed from Gauss points to the nodal points (projection only performed in the domain of interest).

For all the following problems, $\mathbf{a} = 10^6$ and $L = 8$ unless otherwise stated. The justification for the choice of \mathbf{a} is as follows: while $\mathbf{a} \neq 0$ is a free choice in the theoretical scheme (for the critical point formulation of S), it is clear from the convexity condition (2.19) (cf. [2, Sec. 3] for conclusions on degenerate ellipticity of the dual problem in the PDE case) that in seeking solutions, a large value of $\mathbf{a} > 0$ is practically useful in allowing more freedom to sample dual states ('centered' around the state $\lambda = 0$ in the entire domain) where the problem is concave, and for obtaining solutions. In Appendix B we demonstrate this fact by a computed example.

For each of the examples presented below (except the first one since it is a trivial example), we compute the maximum of absolute difference (normalized with respect to the RMS value of the field) across the domain when the number of elements in the mesh is doubled. Based on a uniform mesh of 6400 elements, the RMS value is defined as:

$$f_{RMS} = \sqrt{\sum_A \frac{(f^A)^2}{n}},$$

where A ranges over the total number of nodes in the domain of interest (n). Accordingly, we define the difference $\mathfrak{D}(m)$ (percentage measure), where m is the number of elements in the mesh under consideration, as follows:

$$\mathfrak{D}(m) = \max_A \frac{|f_{2m}^A - f_m^A|}{f_{RMS}} \times 100, \quad (4.10)$$

where f_m^A and f_{2m}^A represent the primal fields obtained using meshes with m and $2m$ elements, respectively, at node A . $\mathfrak{D}(m)$ values for each of the examples subsequently presented can be found in Appendix E: Table E.5. The obtained primal field in each of the examples is further subjected to a Finite Difference approximation of (2.1). Details can be found in Appendix C.

In the following, we will collectively refer to the last two operations as convergence of results w.r.t mesh refinement in the discussion of computed results.

4.2.1. Approximating solutions with prior knowledge

In this section, we examine cases where the base states are chosen close to the solutions of (2.1). These base states are specifically designed so that, when used in Alg. 1, a simple N-R method can be employed. Accordingly, the Convexity Condition is met at each N-R iterate without employing the step-size control and the initial base state remains unchanged throughout the execution of the algorithm.

A fixed point iteration scheme due to Petviashvili (cf. [40, Sec. 4.3], [16]) is utilized to generate solutions to (2.1), which we will refer to as the $\mathcal{P}\mathcal{V}$ solutions. For $C_1 = 0$ in (2.2), we define:

$$g(x) := -f(x) : \quad g(x) = \frac{1}{2} \int_{x-1}^{x+1} g(s)^2 ds = \frac{1}{2} (\Lambda * g^2)(x), \quad (4.11)$$

where

$$\Lambda(x) = \begin{cases} 1 & \text{if } |x| < 1, \\ 0 & \text{otherwise,} \end{cases}$$

and employ the following iterative scheme ($n \geq 1$):

$$\begin{aligned} \text{Step 1: } \quad & \tilde{g}_{n+1}(x) = \frac{1}{2} (\Lambda * g_n^2)(x); \\ \text{Step 2: } \quad & \tilde{C}_{n+1} = \frac{\int_{-\infty}^{\infty} g_n dx}{\int_{-\infty}^{\infty} \tilde{g}_{n+1} dx}; \\ \text{Step 3: } \quad & g_{n+1}(x) = \tilde{C}_{n+1}^q \tilde{g}_{n+1}(x), \end{aligned} \quad (4.12)$$

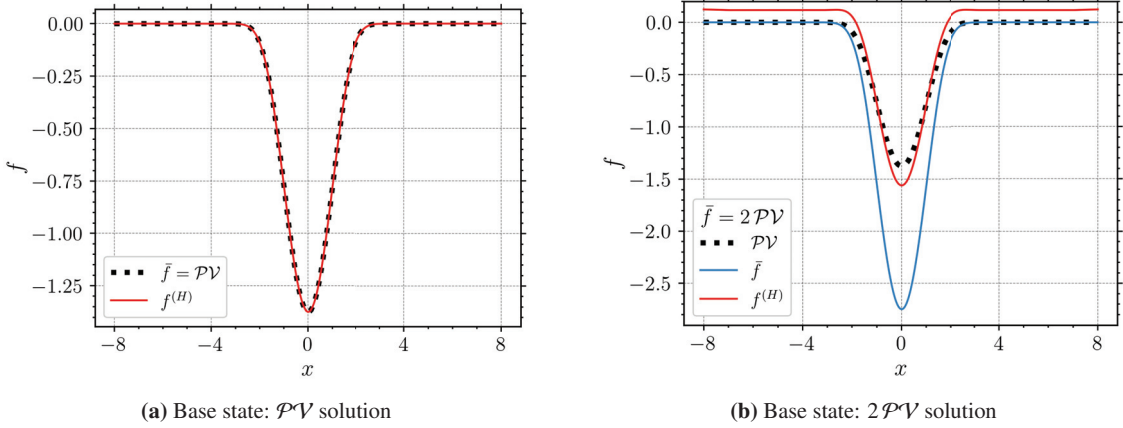


Figure 1: Fig. (a) employs the $\mathcal{P}\mathcal{V}$ solution directly as the base state, whereas Fig. (b) utilizes the $\mathcal{P}\mathcal{V}$ solution scaled by a factor of 2 before using it as the base state.

where $q = 1.4$ has been used, and the integrals in second step are for $x \in (-\infty, \infty)$. We set

$$g_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

and set the following tolerance for convergence:

$$|\tilde{C}_n - 1| < 10^{-6}.$$

The integrations in Step 2 of (4.12) are computed on a finite domain. Due to the rapid decay of the $\mathcal{P}\mathcal{V}$ solution away from $x = 0$, the domains considered are large enough to minimize the impact of the finite domain on the integrals, and it has been verified that the final $\mathcal{P}\mathcal{V}$ profiles obtained on doubling the domain are very close.

One of the $\mathcal{P}\mathcal{V}$ solutions is shown in Fig. 1a. Employing this solution as a base state for the dual scheme, Fig. 1a also demonstrates that the obtained primal field $f^{(H)}$ remains close to it.

We now examine how much we can deviate from the $\mathcal{P}\mathcal{V}$ solution used as a base state. We consider the following base state:

$$\tilde{f}(x) = \tilde{\alpha} \mathcal{P}\mathcal{V}(x), \quad (4.13)$$

where $\tilde{\alpha}$ denotes a scaling factor. For $\tilde{\alpha} = 2$, the results are shown in Fig. 1b. Comparing against Fig. 1a, it is evident that given an input condition that differs from the $\mathcal{P}\mathcal{V}$ solution, the dual scheme can pick up solutions different from the $\mathcal{P}\mathcal{V}$ solution. Convergence of the obtained primal profile w.r.t mesh refinement can be found in Appendix D: Fig. D.10a.

Results obtained for the base states set up with different scaling factors are shown in Fig. 2a, which indicate that for $\tilde{\alpha} = 0.2$, we obtain a constant-in-space type primal field. Additionally for an approximate range of $0.2 < \tilde{\alpha} < 0.8$ and $\tilde{\alpha} > 2.3$, the dual scheme fails to converge with a simple N-R. Convergence results for the obtained primal profile w.r.t mesh refinement for these examples can be found in Appendix D: Fig. D.10.

For certain examples, the obtained primal fields exhibit slight bending near the domain boundaries at $x = \pm 8$. This behavior becomes apparent when scaling the y-axis to smaller scales, as clearly illustrated in Appendix D: Fig. D.10b (primal profiles obtained for $\tilde{\alpha} = 0.2$ on different meshes: Range of plot is $\mathcal{O}(10^{-3})$). The primal field $f^{(H)}$ satisfies the governing equation in $(-L, L)$ and the solutions obtained satisfy the prescribed tolerance. Since L is an adjustable parameter, solutions on arbitrarily large domains without such bends, if deemed undesirable, can be obtained, up to computational cost.

The fact that the primal problem can be solved without any boundary condition specified on the primal field may be considered an interesting aspect of the dual scheme.

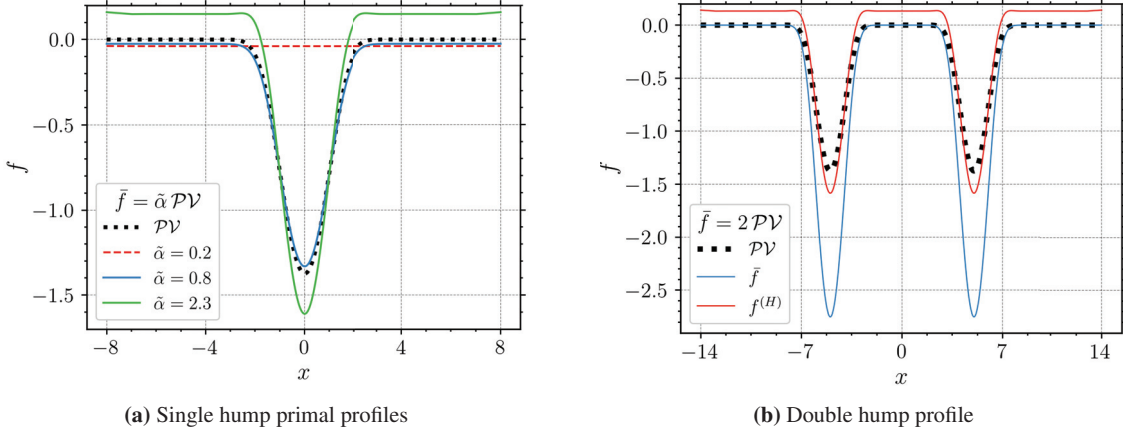


Figure 2: Profiles in Fig. (a) are produced with a single hump $\mathcal{P}\mathcal{V}$ solution scaled by $\tilde{\alpha}$ set as the base state. Fig. (b) uses a double hump $\mathcal{P}\mathcal{V}$ solution scaled by a factor of 2 as the base state.

Finally, Fig. 2b shows the result obtained when a $\mathcal{P}\mathcal{V}$ solution with two self-similar structures (humps) on the same domain and scaled by a factor $\tilde{\alpha} = 2$ is used as a base state. Corresponding profiles w.r.t mesh refinement are shown in Appendix D: Fig. D.10e.

Scaling invariance of the primal solution with \mathbf{a} for $H = \frac{1}{2}\mathbf{a}(f - \bar{f})^2$

The scaling invariance indicated and explained in Sec. 2.1, preamble of Sec. 4.2 and Appendix B has been demonstrated in Fig. 3. Numerically, a larger value of \mathbf{a} allows us to search for the dual solution in an enlarged space. This has been discussed in Appendix B.

4.2.2. Approximating solutions without any prior knowledge

Gaussian base states:

We start by considering the following standard Gaussian:

$$\mathcal{G}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad (4.14)$$

and the base states of the following type:

$$\bar{f}(x) = \gamma \mathcal{G}(x)$$

For a range of γ values, we generate primal fields from dual solutions obtained from a simple N-R scheme. The simple N-R converges only for $-5.2 < \gamma < -2.7$ and for $-0.5 < \gamma < 5$ (amongst the possibilities tried). For the latter range of γ , the dual scheme tends to pick up constant-in-space primal fields. The primal profiles obtained for a few different values of γ are shown in Fig. 4. The corresponding profiles on mesh refinement are presented in Appendix D: Fig. D.11.

Employing Alg. 1 allows us to pick up solutions to (2.16) starting from a wide range of γ in (4.14). As an example, Fig. 5 shows the results of starting the dual scheme from two nearby Gaussian base states ($\gamma = -1.7$ and $\gamma = -1.9$), each approximating to a different primal field. We note that these base states fail to converge with a simple N-R scheme on a mesh of 6400 elements. The corresponding mesh refinement profiles are presented in Appendix D: Fig. D.12.

Sinusoidal base states:

Unlike a simple N-R scheme, Alg. 1 also allows us to pick up dual solutions using a truncated, smoothed sinusoidal base state of generated from

$$\bar{f} = \begin{cases} \sin(\omega x) & \text{for } -2\pi < x < 2\pi \\ 0 & \text{otherwise} \end{cases}, \quad (4.15)$$

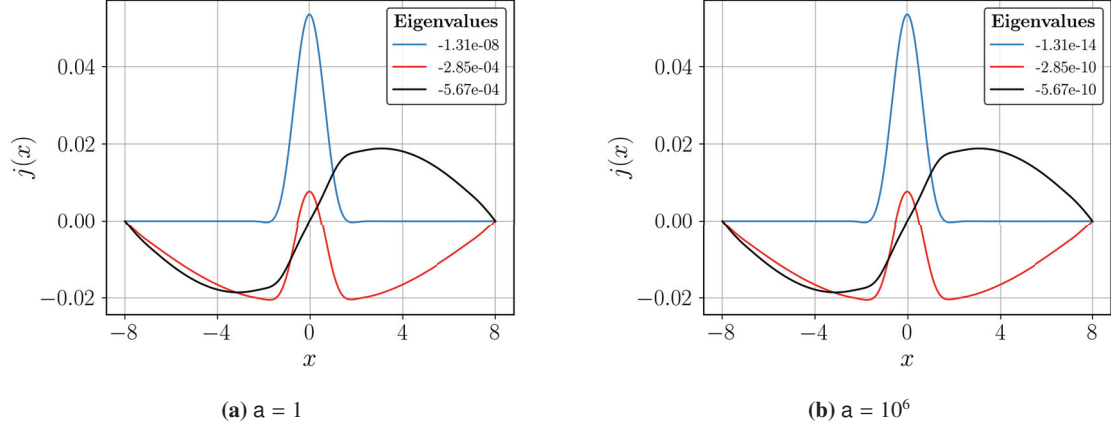


Figure 3: Invariance in the eigenfunctions upon scaling the factor a : The normalized eigenvectors j associated with the discrete Jacobian matrix (4.6) are plotted against nodal positions x for the converged dual field λ . Three smallest eigenvalues have been considered. Following parameters were set: $\bar{f} = 2\mathcal{P}\mathcal{V}$, Mesh = 6400.

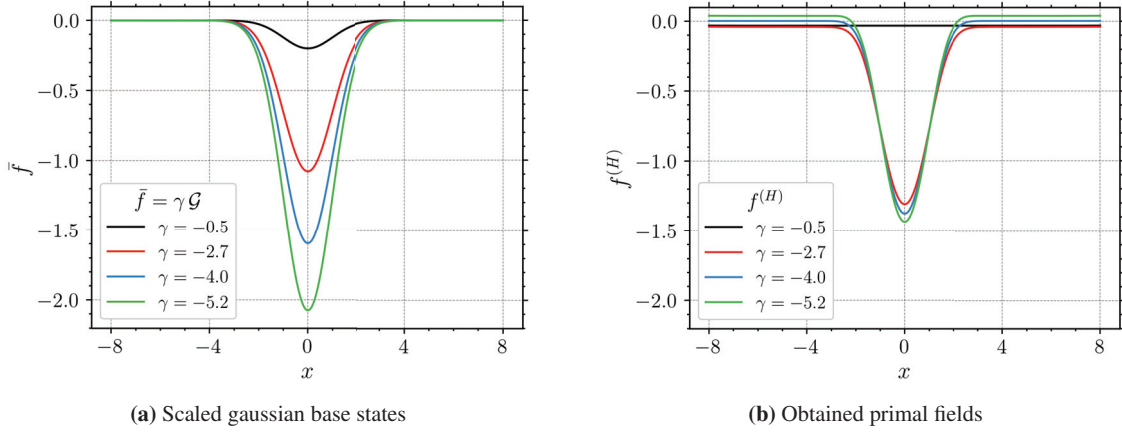


Figure 4: Primal field profiles in Fig. (b) are obtained using the corresponding scaled Gaussian profiles set as base states from Fig. (a) (drawn on same scale). Simple N-R scheme was used to produce the results. Convergence results wrt mesh refinement can be found in Appendix D: Fig. D.11.

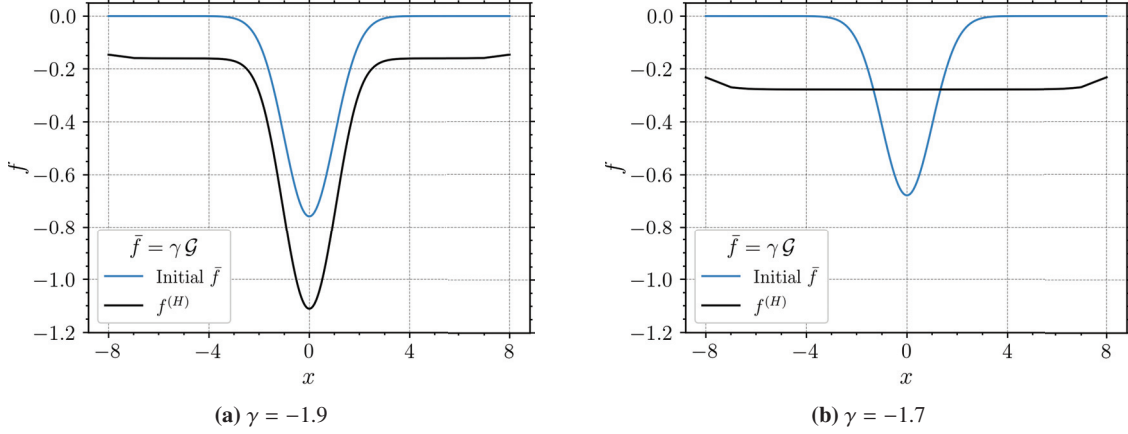


Figure 5: Fig. (a) and Fig. (b) (drawn on same scale) are produced using the step-size controlled N-R with scaled Gaussian profiles (by factor γ) set as base states. Convergence results w.r.t mesh refinement can be found in Appendix D: Fig. D.12.

where the kinks at the sharp transitions at $x = \pm 2\pi$ are smoothed out while preserving the overall shape of the profile. For $\omega = 0.5, 1$ and $\omega = 2$, the corresponding primal fields with mesh of 6400 elements are shown in Fig. 6 and the corresponding profiles on mesh refinement can be found in Appendix D: Fig. D.13. It is evident from these results that the dual scheme prefers to pick up constant primal fields for higher frequency base states.

Piecewise Linear functions as base states:

In this section, we employ piecewise linear functions as base states for the Alg. 1. We start with a negative smoothed hat function generated from

$$\bar{f} = \hat{H}(h) := \begin{cases} -(x-5)h & \text{for } 0 < x \leq 5 \\ (x+5)h & \text{for } -5 < x \leq 0; \quad h < 0 \\ 0 & \text{otherwise} \end{cases} \quad (4.16)$$

with smoothed out the kinks between its piecewise linear segments. Based on the results obtained using the dual scheme for two different heights h , as shown in Fig. 6d and Fig. 6e, it is evident that the negative peak disperses into several small humps. Mesh refinement profiles are shown in Appendix D: Fig. D.13

As a final test, we use linear profiles across the entire domain as the base state. Since no external boundary conditions (from the primal problem description) are imposed on the problem, this test aims to evaluate how the method handles the problem, when non-uniform base states are employed near the boundary.

We adopt the following base state

$$\bar{f} = -\frac{x}{4}. \quad (4.17)$$

Evident from the result for this setup (Fig. 6f, mesh refinement results shown in Appendix D: Fig. D.13f), the f profile exhibits a dip near the boundary, for the problem solved with the given boundary conditions on the dual field. However, the primal field satisfies (2.1) up to the prescribed $tol = 10^{-13}$ on each of the meshes.

4.2.3. Dispersive solitons and their disintegrations

We consider the scaled \mathcal{PV} profile with an added constant c , and define it as the base state:

$$\bar{f}(x) = \tilde{\alpha} \mathcal{PV}(x) + c.$$

Fig. 7 shows the result obtained for such a setup. As we try to move away from the \mathcal{PV} solution by increasing $\tilde{\alpha}$ and c , we start capturing solutions formed at larger u_∞ values which exhibit the emergence of modulating envelopes around the central peak, representing the characteristics of a solitary wave structure. Fig. 7a shows one such structure

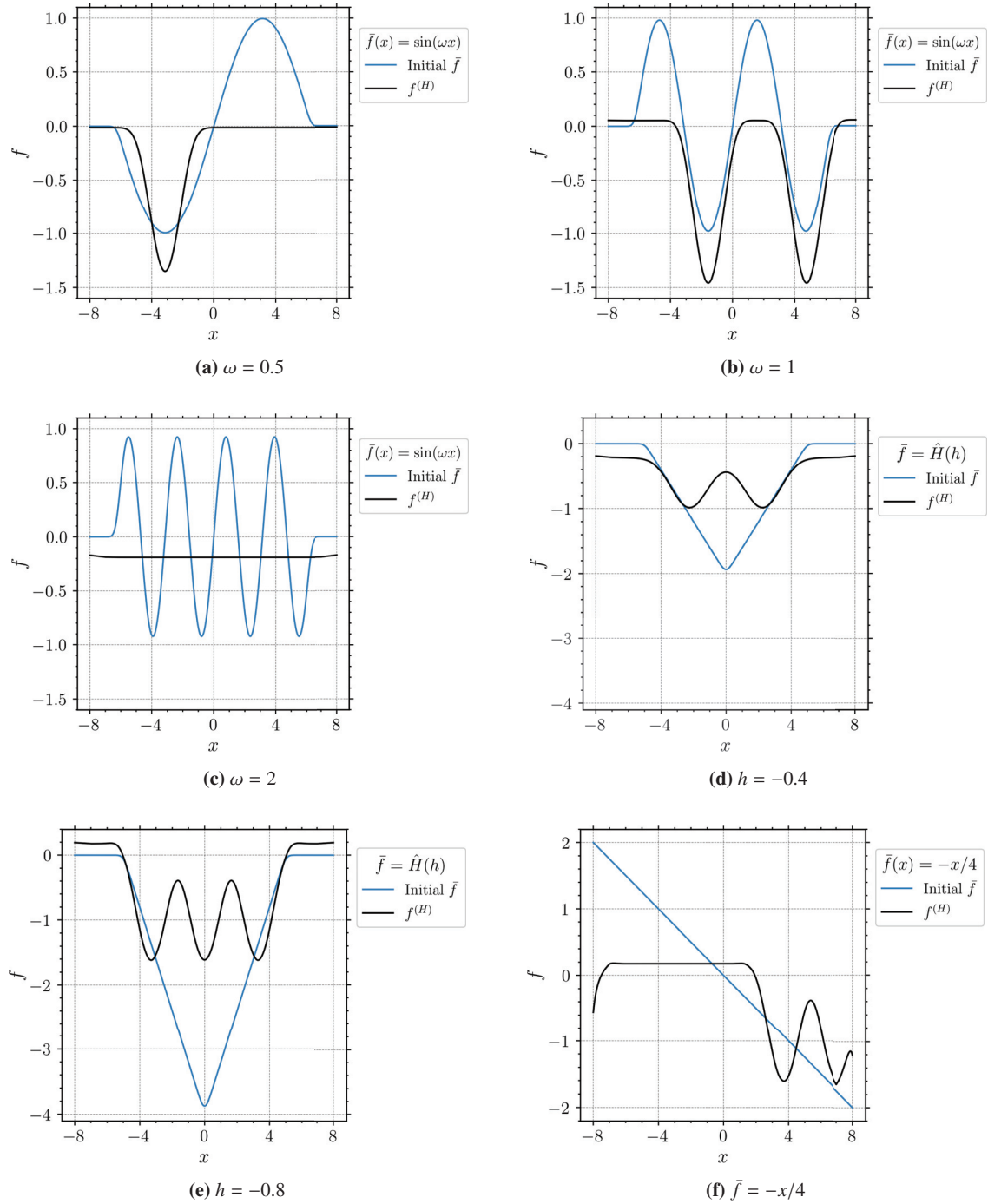


Figure 6: Fig. (a), Fig. (b) and Fig. (c) use a sinusoidal base state (4.15) (parameter: ω). Fig. (d) and Fig. (e) use a negative-smoothened hat-type initial base state (4.16) (parameter: h). Fig. (f) uses a linear profile (4.17). Results are produced using a step-size controlled N-R. Convergence results w.r.t mesh refinement for the above examples can be found in Appendix D: D.13.

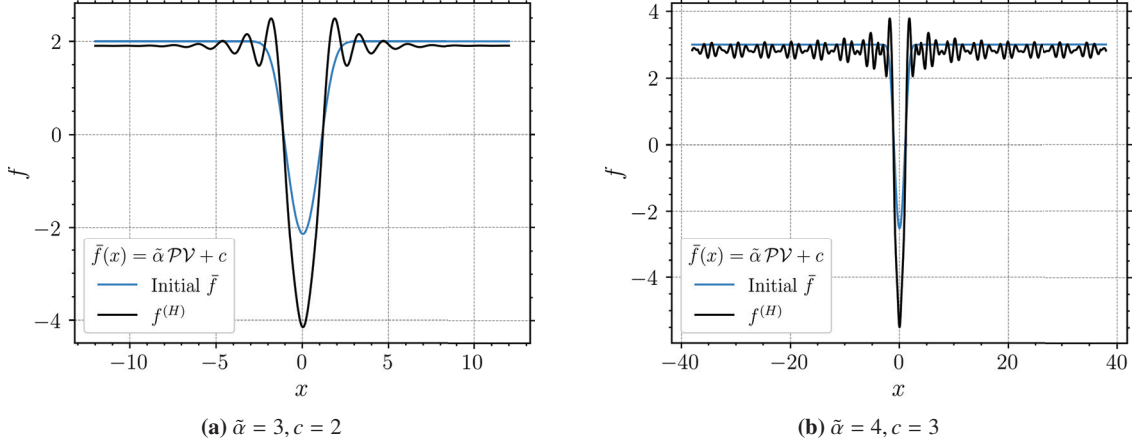


Figure 7: Both the plots are produced using scaled \mathcal{PV} profiles added to constants set as base states on $L = 38$ and a smaller domain in Fig. (a) is shown for better demonstration purposes. Fig. (a) displays a solitary wave structure and as one moves outside the shown region, $f^{(H)}$ approximates to a constant value. Fig. (b) shows the case where the dual scheme pursues a u_∞ greater than the allowed invertibility limit in proposition 3.5, leading to a disintegration of the solitary wave.

obtained using the following parameters: $L = 38$, $\tilde{\alpha} = 3$ and $c = 2$. The solution reaches an approximate value of $u_\infty = 1.87$. Primal profiles obtained w.r.t mesh refinement are presented in Fig. D.14a.

Pursuing a higher u_∞ value by adjusting $\tilde{\alpha}$ and c leads to a breakdown, such that the obtained pattern exhibits a dispersive profile throughout the domain without any ostensible compact support (here, compact support refers to the profile after subtracting u_∞) and has a smooth central dip, as shown in Fig. 7b. This result was obtained using the following parameters: $L = 38$, $\tilde{\alpha} = 4$ and $c = 3$ and we will refer to this example as a disintegrated soliton (d-soliton). This result also matches well with the proposition 3.5, where the operator $I + u_\infty K$ loses its invertibility upon pursuing $u_\infty > 2.301$ approximately and we do not obtain a solitary wave structure. Primal profiles obtained on refinement are presented in Fig. D.14b.

For quantitative comparisons of convergence w.r.t mesh refinement on par with the other computed examples, results for the dispersive soliton and the disintegrated soliton (d-soliton) are presented on a domain of $L = 8$ in Tables E.4 and E.5.

5. Approximation and numerical examples for the NIE formulation

5.1. Approximation for the NIE formulation

To approximate solutions of the NIE (2.28), we discretize the functional $\mathfrak{S}[v]$ as given in (2.38) by left-endpoint-rule quadrature. This yields a spectrally accurate approximation for smooth $2L$ -periodic functions v represented by the vector $\vec{v} = (v_j)$ of their values $v_j = v(x_j)$ on a uniform grid of N points $x_j = jh$ with grid spacing $h = 2L/N$. The integrals in the definition of K from (2.24) are approximated at x_j by the trapezoid rule, and calculated using the discrete Fourier transform. Thus the values $Kv(x_j)$ are approximated by the components $K\vec{v}_j$ of $K\vec{v}$, where $K = (K_{jk})$ is an $N \times N$ symmetric banded Toeplitz matrix whose nonzero entries are h or $h/2$.

With this notation, our approximation to $\mathfrak{S}[v]$ is given by

$$\mathfrak{S}_h(\vec{v}) = -\frac{1}{2} \sum_{j=1}^N (\mathbf{a} + K\vec{v}_j) \hat{w}_j^2 h, \quad \hat{w}_j = \frac{\mathbf{a}\bar{w}_j - v_j - u_\infty K\vec{v}_j}{\mathbf{a} + K\vec{v}_j}, \quad (5.1)$$

when the strict convexity condition $\mathbf{a} + K\vec{v}_j > \varepsilon_a$ holds, for some specified tolerance $\varepsilon_a > 0$. When this condition fails, we set $\mathfrak{S}_h(\vec{v})$ to be some large negative value. The gradient of \mathfrak{S}_h is explicitly given as

$$\frac{\partial \mathfrak{S}_h}{\partial v_j} = h\rho_j, \quad \rho_j = \hat{w}_j + \sum_{k=1}^N K_{jk}(u_\infty \hat{w}_k + \frac{1}{2} \hat{w}_k^2). \quad (5.2)$$

We minimize $-\mathfrak{E}_h$ using standard optimization software to determine an approximate minimizer $\vec{v} = (v_j)$. Provided the strict convexity condition holds, this determines an approximate solution (\hat{w}_j) to (2.28) since the gradient $\partial \mathfrak{E}_h / \partial \vec{v}$ is small.

The Hessian of \mathfrak{E}_h has the matrix entries

$$\frac{\partial^2 \mathfrak{E}_h}{\partial v_j \partial v_k}(\vec{v}) = -h \sum_{j=1}^N \frac{(\delta_j^i + K_{ij} \hat{f}_j)(\delta_k^j + \hat{f}_j K_{jk})}{\mathbf{a} + K \vec{v}_j}, \quad \hat{f}_j = u_\infty + \hat{w}_j. \quad (5.3)$$

For later reference, we note that when $\vec{v} = 0$ we have $\hat{w}_j = \bar{w}_j$, and the second variation of the functional $\mathbf{a} \mathfrak{E}[v]$ at $v = 0$ (which is independent of the parameter \mathbf{a}) is approximated by the matrix

$$\mathbf{M}(\bar{w}) = \frac{\mathbf{a}}{h} \left(\frac{\partial^2 \mathfrak{E}_h}{\partial v_j \partial v_k}(0) \right) = -(I + K \bar{\mathbf{F}})(I + \bar{\mathbf{F}}K), \quad \bar{\mathbf{F}} = \text{diag}(u_\infty + \bar{w}_j). \quad (5.4)$$

In order to carry out numerical computations for a range of values of u_∞ while maintaining the strict convexity condition, we implement a primitive kind of path-following method. We start with $u_\infty = 0$ and take the base state \bar{w} to be a numerical solution to (2.28) computed by Petviashvili iteration, as described below and in [31]. Then we change u_∞ in small increments, changing the base state \bar{w} to be the approximate solution \hat{w} obtained numerically for the previous value of u_∞ . In this way the values of v_j can be kept small and the convexity condition $\mathbf{a} + K \vec{v}_j > \varepsilon_a$ maintained.

5.2. Numerical results for the NIE formulation

The computations in this section are performed with $\mathbf{a} = 10$. We minimize $-\mathfrak{E}_h[\vec{v}]$ using the BFGS algorithm as implemented in the julia package Optim.jl [41]. We take $L = 25$ and discretize using $N = 1000$ so $h = 2L/N = 0.05$. The stopping criterion is based on the maximum norm of the gradient, with tolerance $2 \cdot 10^{-9}$.

We present two sets of solutions computed for various values of u_∞ . We start with $u_\infty = 0$ in each case, with base state $\bar{w} = \bar{f}$ as computed using 50 steps of Petviashvili iteration, similarly as in section 4.2.1. For succeeding values of u_∞ we then reset the base state \bar{w} to be the solution \hat{w} last computed, as described above in section 5.1. We note that in all cases the residual norm $\|\bar{\rho}\|_\infty < 10^{-7}$.

The first set of solutions is computed for nonpositive values of u_∞ ranging from 0 down to -0.475 , a value slightly above the lower threshold $u_\infty = -0.5$ at which the operator $I + u_\infty K$ first loses invertibility on the infinite line according to proposition 3.5. Results for these solutions appear in Fig. 8. In Fig. 8a we plot the numerically determined wave shapes $\hat{f} = u_\infty + \hat{w}$ vs. x for 8 successively decreasing values of u_∞ ranging from 0 to -0.475 , and in Fig. 8b we plot $\log_{10} |\hat{w}|$ vs. x .

We compute a second set of solutions for positive values of u_∞ ranging from 0.5 to 2.0, a value somewhat less than the upper threshold ≈ 2.30167 at which the operator $I + u_\infty K$ first loses invertibility on the infinite line according to proposition 3.5. The results for $\hat{f} = u_\infty + \hat{w}$ vs. x are shown in Fig. 9a, and we plot $\log_{10} |\hat{w}|$ vs. x in Fig. 9b.

In Table 2, for selected solutions in both sets we tabulate the lowest four non-negligible eigenvalues κ_j of the matrix $-\mathbf{a} D^2 \mathfrak{E}_h[0]$ (with base state set as $\bar{w} = \hat{w}$). This is the negative Hessian scaled by \mathbf{a} , i.e., the negative Hessian scaled to be independent of \mathbf{a} . In all cases the first eigenvalue satisfied $|\kappa_1| < 2 \cdot 10^{-15}$. We tabulate as well as the minimum of the (continuous) spectrum of the operator $(I + u_\infty K)^2$ corresponding to the ‘spectrum at infinity’ of the scaled second variation $-\mathbf{a} \delta^{(2)} \mathfrak{E}[0]$. According to the Fourier analysis in section 3.2, this value is given by

$$\min \sigma_c = \begin{cases} (1 - \sigma_0 u_\infty)^2 & \text{if } u_\infty > 0, \\ (1 + 2u_\infty)^2 & \text{if } u_\infty \leq 0, \end{cases} \quad -\sigma_0 = \min_{\xi \in \mathbb{R}} 2 \text{sinc } \xi \approx -0.434467. \quad (5.5)$$

The results in both sets of solutions appear consistent with the possibility that periodic and solitary waves exist on the line for u_∞ in the whole range from -0.5 to at least 2.2, consistent with the phase-speed non-matching condition mentioned in the introduction and with the coercivity result from propositions 3.4 and 3.5. For u_∞ between -0.5 and 0, the wave perturbation \hat{w} has a single hump shape, monotonic for $0 < x < L$. Log plots suggest that values of $|\hat{w}|$ smaller than about 10^{-6} are not computed accurately with the precision and tolerances that were used.

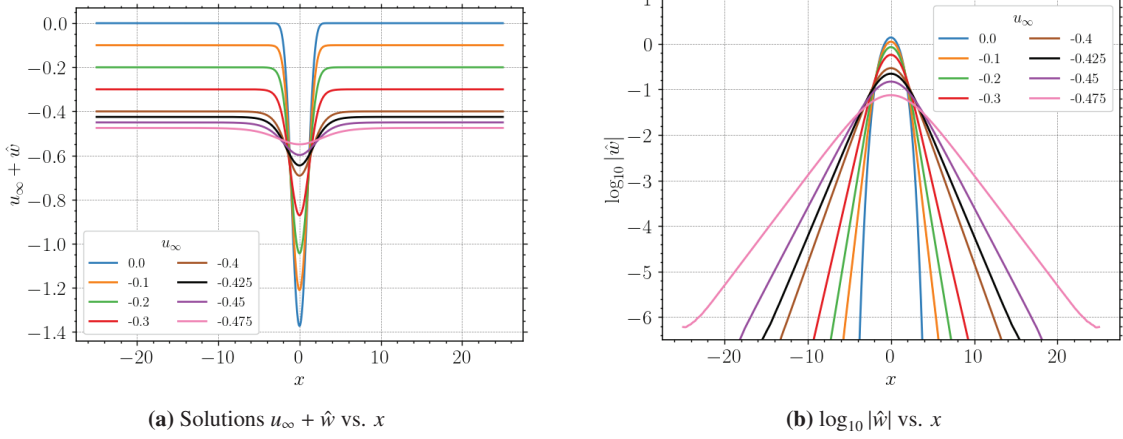


Figure 8: Solutions for $u_\infty = 0.0, -0.1, -0.2, -0.3, -0.4, -0.425, -0.45, -0.475$.

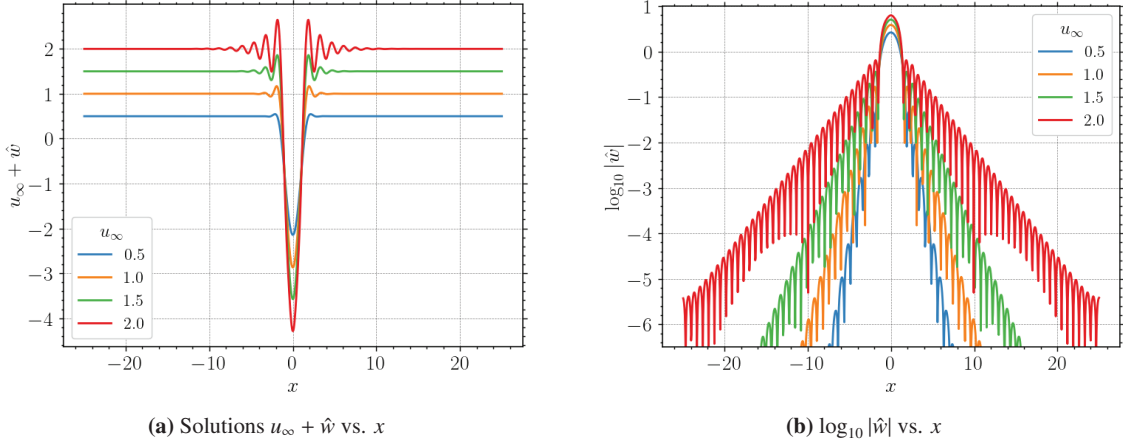


Figure 9: Solutions for $u_\infty = 0.5, 1.0, 1.5, 2.0$.

For $u_\infty \neq 0$, the values of $|\hat{w}|$ decay toward zero at an exponential rate that depends upon u_∞ and diminishes as u_∞ approaches -0.5 . This regime, where $2u_\infty \rightarrow -1$, is where we can expect the KdV approximation to be valid, in fact—the wave amplitude becomes small and wave length becomes large.

For $u_\infty = 0$, on the other hand, the wave profile $\hat{w} = \hat{f}$ may decay to zero at a rate that is faster than exponential. This is reminiscent of the solitary wave pulse for a chain of beads in Hertz contact, which was shown in [42] to decay at a rate that is faster than double exponential.

For the positive values of u_∞ between 0.2 and 2.2, on the other hand, the computed wave perturbations \hat{w} decay toward zero in an oscillatory, sign-changing way. The decay rate of the envelope diminishes as u_∞ approaches the upper threshold near 2.3. The oscillation frequency appears well approximated by the value $\xi_* \approx 4.4934$ which minimizes $\text{sinc} \xi$ and at which the phase velocity matches the group velocity of harmonic waves. This corresponds to the regime investigated in the general study by Kozyreff [34].

5.3. Usage and failure of Petviashvili iteration

As mentioned in the previous subsection, using the optimization package Optim.jl in julia we found approximate solutions (\hat{w}_j) to equation (2.28) for which the residual $\vec{\rho}$ from (5.2) has norm $\|\vec{\rho}\|_\infty < 10^{-7}$ for all the values of u_∞ listed in Table 2, ranging from -0.45 to 2.2 .

u_∞	κ_2	κ_3	κ_4	κ_5	$\min \sigma_c$
-0.45	0.00525	0.01008	0.01017	0.01155	0.01000
-0.40	0.01940	0.03960	0.04027	0.04210	0.04000
-0.30	0.06512	0.15327	0.16035	0.16209	0.16000
-0.20	0.12258	0.32787	0.36029	0.36172	0.36000
-0.10	0.18357	0.54470	0.63819	0.64042	0.64000
0.00	0.24398	0.64190	0.64190	0.78068	1.00000
0.50	0.48865	0.55764	0.55764	0.61489	0.61272
1.00	0.32294	0.32296	0.32611	0.32613	0.31983
1.50	0.12421	0.12423	0.12718	0.12721	0.12131
2.00	0.01867	0.01869	0.02025	0.02031	0.01718

Table 2: Lowest eigenvalues of $-aD^2\Xi$, and bottom of continuous spectrum. $|\kappa_1| < 2 \cdot 10^{-15}$ in all cases.

A natural idea to improve the quality of the numerical solution is by post-processing, applying a Petviashvili iteration for several steps, starting with the solution found by Optim.jl. For solving (2.28), one rewrites the equation in the equivalent form

$$w(x) = -g(x), \quad g(x) = (I + u_\infty K)^{-1} K(\frac{1}{2}g^2)(x), \quad (5.6)$$

and replaces Step 1 in the iteration scheme (4.12) by

$$\text{Step 1': } \tilde{g}_{n+1}(x) = (I + u_\infty K)^{-1} K(\frac{1}{2}g_n^2)(x). \quad (5.7)$$

The operator $(I + u_\infty K)^{-1}K$ is easily discretized and applied using the discrete Fourier transform.

With this method, we found that we could achieve residual norm $\|\tilde{p}\|_\infty < 10^{-14}$ for all values of u_∞ attempted ranging from -0.45 to 0.7 . However, the number of iterations required increased greatly for u_∞ close to 0.8 . But then for u_∞ in the range $(0.8, 2.3)$, the Petviashvili iteration became unstable and failed to improve the solution found by Optim.jl. This is a subset of the range where localized solutions with oscillatory tails are found. Presumably, the observed instability of Petviashvili iteration is related to some type of violation of criteria for local convergence as established by Pelinovsky & Stepanyants [17], but we did not investigate this in depth.

6. Discussion

In this paper, we have taken a variational approach that has been developed to solve PDEs through duality for convex optimization problems constrained by field constraints, and extended it to handle two different nonlocal equations that determine traveling waves for the semi-discrete inviscid Burgers equation: a nonlinear advance-delay differential-difference equation (DDE) and a corresponding nonlinear integral equation (NIE).

Particularly, we made extensive use of the flexibility of selecting the base state, changing it to facilitate the numerical computation of solutions in many cases when optimization algorithms produce sequences approaching the boundary of the functional domain where the objective functional is finite. For the DDE case, we implemented an algorithm that adaptively adjusts dual field increments to stay within the finiteness domain, and incorporates base state resets that have the effect of shifting or reshaping the domain to allow the search for a solution to continue and not simply stall at the boundary. For the NIE case, we used standard software to carry out each optimization for a sequence of values of the parameter u_∞ , setting as base state the solution found for the previous parameter value. This enables a new solution nearby to be found with dual field presumably of small amplitude well inside the domain.

The automatic adaptivity built into the DDE code may be responsible for the fact that it continued to work and produced “disintegrated” wave profiles that appear delocalized and may be non-periodic, in a parameter regime ($u_\infty > 2.35$, e.g.) where the periodic NIE code failed and analysis of the periodic problem indicated difficulties with coercivity.

On the other hand, the periodic NIE code performed much better to find well-localized (solitary) wave profiles with specified limiting states u_∞ . Unfortunately, we lack a convincing explanation for why this should be so.

For the nonlocal problems that we treated, truncation of the wave-profile problem on the infinite line to a bounded interval requires extended boundary conditions for base states and dual fields. We have done this in different ways for the DDE and NIE mainly as an experiment, implementing Dirichlet-type conditions for the DDE and periodic conditions for the NIE. Switching the treatments is plausibly feasible; e.g., for the DDE case one could require that the base state \tilde{f} and dual field λ be extended as periodic outside the interval $(-L, L)$. Analytically this is almost equivalent in principle to the periodic NIE formulation that we treated, with a subtle difference, in that periodic variations $\delta\lambda$ in the DDE dual field would have derivatives $\delta\nu = -\delta\lambda'$ constrained to have integral zero. This means that solutions found by the two schemes might in principle correspond to different constants C_1 in (2.2).

We re-emphasize that in this paper we have focused on the properties of the variational approach for the nonlocal wave profile problem, and not on an exhaustive exploration of the family of solutions. The use of software for continuation and path-following such as AUTO or pde2path would plausibly allow one to track branches of solutions and their possible bifurcations more systematically than we have done.

Our results nevertheless provide clues about certain parameter regimes that appear interesting to examine more closely. Because we are unable to guarantee that maximizers of the relevant concave objective functionals do not lie on the boundary of the finiteness domain, however, we have not managed to prove an unconditional existence theorem for traveling-wave profiles using either the DDE or NIE formulations.

The variational approach with base state changes does suggest a possible avenue towards a convergence proof, though. E.g., in L^2 -gradient flow for a convex functional, the norm of the gradient is non-increasing in time. In the problems we treat, the functional gradient agrees with the equation residual for the DDE or NIE. If one runs gradient flow and resets the base state with dual field reset to zero (as in our numerical algorithm which was based on Newton-type iteration rather than gradient flow, however), the equation residual would not change with the reset and would be ensured to be non-increasing in gradient flow afterwards. Perhaps one could stay away from the domain boundary and have the gradient flow equilibrate this way.

Acknowledgments

The work of UK was supported by funds from the NSF grant OIA-DMR 2021019 and the Paul P. Christiano Professorship in the Dept. of Civil & Environmental Engineering at CMU. This material is based upon work supported by the National Science Foundation under grant DMS 2106534 to RLP.

References

- [1] P. Sprenger, C. Chong, E. Okyere, M. Herrmann, P. Kevrekidis, M. A. Hoefer, Hydrodynamics of a discrete conservation law, *Studies in Applied Mathematics* (2024) e12767.
- [2] A. Acharya, A hidden convexity in continuum mechanics, with application to classical, continuous-time, rate-(in)dependent plasticity, *Mathematics and Mechanics of Solids* 30 (3) (2025) 701–719. [arXiv:https://arxiv.org/abs/2310.03201](https://arxiv.org/abs/2310.03201), [doi:10.1177/10812865241258154](https://doi.org/10.1177/10812865241258154).
- [3] A. Acharya, A dual variational principle for nonlinear dislocation dynamics, *Journal of Elasticity* 154 (1) (2023) 383–395.
- [4] A. Acharya, Variational principle for nonlinear PDE systems via duality, *Quarterly of Applied Mathematics* 81 (2023) 127–140. URL <https://arxiv.org/abs/2108.08902>
- [5] U. Kouskiya, A. Acharya, Hidden convexity in the heat, linear transport, and Euler’s rigid body equations: A computational approach, *Quarterly of Applied Mathematics* 82 (2024) 673–703.
- [6] U. Kouskiya, A. Acharya, Inviscid Burgers as a degenerate elliptic problem, *Quart. Appl. Math.* 83 (2) (2025) 315–360.
- [7] S. Singh, J. Ginster, A. Acharya, A hidden convexity of nonlinear elasticity, *Journal of Elasticity* 156 (2024) 975–1014. URL <https://link.springer.com/article/10.1007/s10659-024-10081-w>
- [8] A. Acharya, B. Strohfolini, A. Zarnescu, Variational dual solutions for incompressible fluids (2024). URL <https://arxiv.org/abs/2409.04911>
- [9] Y. Brenier, Hidden convexity in some nonlinear PDEs from geometry and physics, *J. Convex Anal.* 17 (3-4) (2010) 945–959.
- [10] Y. Brenier, The initial value problem for the Euler equations of incompressible fluids viewed as a concave maximization problem, *Communications in Mathematical Physics* 364 (2) (2018) 579–605.
- [11] Y. Brenier, Examples of hidden convexity in nonlinear PDEs (2020). URL <https://hal.science/hal-02928398/document>
- [12] R. T. Rockafellar, Augmented Lagrangians and hidden convexity in sufficient conditions for local optimality, *Math. Program.* 198 (1) (2023) 159–194. [doi:10.1007/s10107-022-01768-w](https://doi.org/10.1007/s10107-022-01768-w). URL <https://doi.org/10.1007/s10107-022-01768-w>
- [13] D. Vorotnikov, Partial differential equations with quadratic nonlinearities viewed as matrix-valued optimal ballistic transport problems, *Archive for Rational Mechanics and Analysis* 243 (3) (2022) 1653–1698.

- [14] D. Vorotnikov, Hidden convexity and dafermos' principle for some dispersive equations, arXiv preprint arXiv:2501.05389 (2025).
- [15] J.-M. Mirebeau, E. Stampfli, Discretization and convergence of the ballistic Benamou-Brenier formulation of the porous medium and Burgers' equations, working paper or preprint (Mar. 2025).
URL <https://hal.science/hal-05005367>
- [16] V. I. Petviashvili, Equation of an extraordinary soliton, *Fizika plazmy* 2 (1976) 469–472.
- [17] D. E. Pelinovsky, Y. A. Stepanyants, Convergence of Petviashvili's iteration method for numerical approximation of stationary solutions of nonlinear wave equations, *SIAM J. Numer. Anal.* 42 (3) (2004) 1110–1127. doi:10.1137/S0036142902414232.
URL <https://doi.org/10.1137/S0036142902414232>
- [18] M. Herrmann, Oscillatory waves in discrete scalar conservation laws, *Math. Models Methods Appl. Sci.* 22 (1) (2012) 1150002, 21. doi:10.1142/S021820251200585X.
URL <https://doi.org/10.1142/S021820251200585X>
- [19] A. Vainchtein, Solitary waves in FPU-type lattices, *Phys. D* 434 (2022) Paper No. 133252, 22. doi:10.1016/j.physd.2022.133252.
URL <https://doi.org/10.1016/j.physd.2022.133252>
- [20] G. Friesecke, J. A. D. Wattis, Existence theorem for solitary waves on lattices, *Comm. Math. Phys.* 161 (2) (1994) 391–418.
URL <http://projecteuclid.org/euclid.cmp/1104269908>
- [21] D. Smets, M. Willem, Solitary waves with prescribed speed on infinite lattices, *J. Funct. Anal.* 149 (1) (1997) 266–275. doi:10.1006/jfan.1996.3121.
URL <https://doi.org/10.1006/jfan.1996.3121>
- [22] M. Herrmann, Unimodal wavetrains and solitons in convex Fermi-Pasta-Ulam chains, *Proc. Roy. Soc. Edinburgh Sect. A* 140 (4) (2010) 753–785. doi:10.1017/S0308210509000146.
URL <https://doi-org.cmu.idm.oclc.org/10.1017/S0308210509000146>
- [23] R. L. Pego, T.-S. Van, Existence of solitary waves in one dimensional peridynamics, *J. Elasticity* 136 (2) (2019) 207–236. doi:10.1007/s10659-018-9701-6.
URL <https://doi.org/10.1007/s10659-018-9701-6>
- [24] M. Herrmann, K. Kleine, Korteweg–de Vries waves in peridynamical media, *Stud. Appl. Math.* 152 (1) (2024) 376–403.
- [25] M. Herrmann, K. Matthies, Nonlinear and nonlocal eigenvalue problems: variational existence, decay properties, approximation, and universal scaling limits, *Nonlinearity* 33 (8) (2020) 4046–4074. doi:10.1088/1361-6544/ab8350.
URL <https://doi.org/10.1088/1361-6544/ab8350>
- [26] G. Friesecke, R. L. Pego, Solitary waves on FPU lattices. I. Qualitative properties, renormalization and continuum limit, *Nonlinearity* 12 (6) (1999) 1601–1627. doi:10.1088/0951-7715/12/6/311.
URL <https://doi-org.cmu.idm.oclc.org/10.1088/0951-7715/12/6/311>
- [27] G. Iooss, Travelling waves in the Fermi-Pasta-Ulam lattice, *Nonlinearity* 13 (3) (2000) 849–866. doi:10.1088/0951-7715/13/3/319.
URL <https://doi.org/10.1088/0951-7715/13/3/319>
- [28] G. Iooss, G. James, Localized waves in nonlinear oscillator chains, *Chaos* 15 (1) (2005) 015113, 15. doi:10.1063/1.1836151.
URL <https://doi.org/10.1063/1.1836151>
- [29] G. James, Periodic travelling waves and compactons in granular chains, *J. Nonlinear Sci.* 22 (5) (2012) 813–848. doi:10.1007/s00332-012-9128-3.
URL <https://doi.org/10.1007/s00332-012-9128-3>
- [30] M. Herrmann, A. Mikikits-Leitner, KdV waves in atomic chains with nonlocal interactions, *Discrete Contin. Dyn. Syst.* 36 (4) (2016) 2047–2067. doi:10.3934/dcds.2016.36.2047.
URL <https://doi.org/10.3934/dcds.2016.36.2047>
- [31] B. Ingimarsen, R. L. Pego, Existence of solitary waves in particle lattices with power-law forces, *Nonlinearity* 37 (12) (2024) 125016. doi:10.1088/1361-6544/ad8c1c.
URL <https://dx.doi.org/10.1088/1361-6544/ad8c1c>
- [32] T. B. Benjamin, J. L. Bona, D. K. Bose, Solitary-wave solutions of nonlinear problems, *Philos. Trans. Roy. Soc. London Ser. A* 331 (1617) (1990) 195–244. doi:10.1098/rsta.1990.0065.
- [33] J. Bona, H. Chen, Solitary waves in nonlinear dispersive systems, *Discrete Contin. Dyn. Syst. Ser. B* 2 (3) (2002) 313–378. doi:10.3934/dcdsb.2002.2.313.
- [34] G. Kozyreff, Speed of wave packets and the nonlinear Schrödinger equation, *Phys. Rev. E* 107 (1) (2023) Paper No. 014219, 15. doi:10.1103/physreve.107.014219.
URL <https://doi.org/10.1103/physreve.107.014219>
- [35] J. C. Eilbeck, R. Flesch, Calculation of families of solitary waves on discrete lattices, *Phys. Lett. A* 149 (4) (1990) 200–202. doi:10.1016/0375-9601(90)90326-J.
URL [https://doi.org/10.1016/0375-9601\(90\)90326-J](https://doi.org/10.1016/0375-9601(90)90326-J)
- [36] A. Acharya, An action for nonlinear dislocation dynamics, *Journal of the Mechanics and Physics of Solids* 161 (2022) 104811.
- [37] H. Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.
- [38] H. H. Bauschke, P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, 2nd Edition, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, Cham, 2017, with a foreword by Hédya Attouch. doi:10.1007/978-3-319-48311-5.
URL <https://doi.org/10.1007/978-3-319-48311-5>
- [39] R. L. Pego, Compactness in L^2 and the Fourier transform, *Proc. Amer. Math. Soc.* 95 (2) (1985) 252–254. doi:10.2307/2044522.
URL <https://doi.org/10.2307/2044522>
- [40] B. Ingimarsen, R. Pego, On long waves and solitons in particle lattices with forces of infinite range, *SIAM Journal on Applied Mathematics* 84 (3) (2024) 808–830.
- [41] P. K. Mogensen, A. N. Riseth, Optim: A mathematical optimization package for Julia, *Journal of Open Source Software* 3 (24) (2018) 615.

doi:10.21105/joss.00615.

[42] J. M. English, R. L. Pego, On the solitary wave pulse in a chain of beads, Proc. Amer. Math. Soc. 133 (6) (2005) 1763–1768. doi: 10.1090/S0002-9939-05-07851-2.
URL <https://doi.org/10.1090/S0002-9939-05-07851-2>

Appendix A. Formal KdV asymptotics

Here we describe solutions of the semi-discrete Burgers equation (1.2) that are long-wave perturbations of a constant state $u_* \neq 0$, by a well-known formal asymptotic approximation. The approximation takes the form

$$u_j(t) = u_* + \varepsilon^2 v(x, \tau), \quad \text{with } x = \varepsilon(j - c_* t), \quad \tau = \varepsilon^3 t, \quad (\text{A.1})$$

where $c_* = u_*$ is the limiting speed of long waves in the linearization of (1.2). Then straightforward use of the chain rule and Taylor expansion yields

$$\frac{d}{dt} u_j = \varepsilon^5 \partial_\tau v - \varepsilon^3 c_* \partial_x v, \quad (\text{A.2})$$

$$u_{j\pm 1} = u_* + \varepsilon^2 \left(v \pm \varepsilon \partial_x v + \frac{1}{2} \varepsilon^2 \partial_x^2 v \pm \frac{1}{6} \varepsilon^3 \partial_x^3 v + \frac{1}{24} \varepsilon^4 \partial_x^4 v \right) + O(\varepsilon^7). \quad (\text{A.3})$$

By straightforward substitution this results in the residual, or equation error,

$$\frac{d}{dt} u_j + \frac{1}{4} (u_{j+1}^2 - u_{j-1}^2) = \varepsilon^5 \left(\partial_\tau v + v \partial_x v + \frac{u_*}{6} \partial_x^3 v \right) + O(\varepsilon^7). \quad (\text{A.4})$$

Thus (1.2) is formally approximated by the KdV equation

$$\partial_\tau v + v \partial_x v + \frac{u_*}{6} \partial_x^3 v = 0. \quad (\text{A.5})$$

This KdV equation has solitary wave solutions with any wave speed \hat{c} having the same sign as u_* , given by

$$v(x, \tau) = 3\hat{c} \operatorname{sech}^2 \left(\frac{x - \hat{c}\tau}{2} \sqrt{\frac{6\hat{c}}{u_*}} \right). \quad (\text{A.6})$$

This provides an approximate traveling wave for the semi-discrete Burgers equation (1.2) in the form

$$\hat{u}_j(t) = u_* + 3\gamma \operatorname{sech}^2 \left(\frac{j - ct}{2} \sqrt{\frac{6\gamma}{u_*}} \right), \quad c = u_* + \gamma, \quad \gamma = \varepsilon^2 \hat{c}. \quad (\text{A.7})$$

Appendix B. Effect of the choice of \mathbf{a}

This appendix motivates the usage of a larger value of \mathbf{a} in the auxiliary potential function H (2.7). We consider the problem for a scaled \mathcal{PV} set as base state (see Sec. 4.2.2 and (4.13)) with $\tilde{\alpha}$ set as 2. We employ a simple N-R scheme with tolerance set as per (4.9) for this problem with two different types of initial guesses on λ given by:

$$\lambda^{(0)}(x) = 0 \quad \text{and} \quad \lambda^{(0)}(x) = e^{-x^2}.$$

The convergence results and the corresponding residual values are shown in the Table B.3.

The difference in the (non)convergence to a solution of the N-R iterations for the two cases can be understood as follows. Direct inspection shows that our algorithm produces invariant results under the following scaling transformation: if a residual $R^{(k,1)}$ is the result produced at iteration k for a choice of $\mathbf{a} = 1$ and initial guess $\lambda^{(0)}$ then, for a choice of $\mathbf{a} = \mathbf{a}^*$, if the initial guess is scaled to $\mathbf{a}^* \lambda^{(0)}$, the residual remains invariant, i. e.,

$$R^{(k,\mathbf{a}^*)} = R^{(k,1)}.$$

Moreover, $\lambda^{(k,\mathbf{a}^*)} = \mathbf{a}^* \lambda^{(k,1)}$.

$\lambda^{(0)}(x) = 0$	
a	N-R Result
10^{-6}	Converged
1	Converged
10^6	Converged

$\lambda^{(0)}(x) = e^{-x^2}$	
a	N-R Result
10^{-6}	No Convergence
1	No Convergence
10^6	Converged

Table B.3: Convergence results for different initial conditions: “Converged” and “No Convergence” indicate whether the residual in the N-R iterations converged below the set tolerance (4.9), $tol = 10^{-12}$, or not.

Indeed, from (2.9) we observe that

$$f^{(H)}(\mathcal{D}, 1, x) = f^{(H)}(\mathbf{a}\mathcal{D}, \mathbf{a}, x),$$

where $\mathbf{a}\mathcal{D}$ denotes scaling each element of \mathcal{D} by the scalar \mathbf{a} . Based on (4.4), it can also be verified that

$$\frac{\partial f^{(H)}}{\partial \mathcal{D}}(\mathbf{a}\mathcal{D}, \mathbf{a}, x) = \frac{1}{\mathbf{a}} \frac{\partial f^{(H)}}{\partial \mathcal{D}}(\mathcal{D}, 1, x).$$

Thus, in the N-R iterations

$$\lambda^{(k+1)} - \lambda^{(k)} = -\mathbf{J}^{(k-1)}\mathbf{R}^{(k)},$$

if $\lambda^{(k)} \rightarrow \mathbf{a}\lambda^{(k)}$, $\mathbf{J}^{(k-1)} \rightarrow \mathbf{a}\mathbf{J}^{(k-1)}$, $\mathbf{R}^{(k)} \rightarrow \mathbf{R}^{(k)}$ when \mathbf{a} changes from $1 \rightarrow \mathbf{a}$ then this implies that at each step $\lambda^{(k+1)} \rightarrow \mathbf{a}\lambda^{(k+1)}$.

For $\lambda^{(0)} = 0$, the scaling hypothesis on the initial guess is satisfied and there is no change in the N-R iteration convergence profile as \mathbf{a} is varied. When $\lambda^{(0)}$ is not scaled with \mathbf{a} , the hypothesis is not satisfied and there is a significant difference in the ability of the algorithm to obtain solutions with varying \mathbf{a} .

Appendix C. Testing primal fields using a Finite Difference approximation

Based on the dual solution obtained at nodes of any FE mesh, we use a finite difference to approximate the terms in (2.1) and check how well the equation gets satisfied at these nodes. For a primal field f and corresponding to any node A , this agreement is evaluated based on the following expression:

$$Err^A = \frac{f^{A+1} - f^{A-1}}{2dx} + \frac{1}{2} \left(\left(f^{A+\frac{1}{dx}} \right)^2 - \left(f^{A-\frac{1}{dx}} \right)^2 \right), \quad (\text{C.1})$$

where dx represents the element length and the meshes are chosen in such a way that for a node at x , there always exists nodes at $x + 1$ and $x - 1$.

Focusing on the internal part of the domain, the maximum absolute value of Err^A within the region $x \in (-L + 2.5, L - 2.5)$ for examples in this work is summarized in Table E.4.

Appendix D. Results obtained on mesh refinement

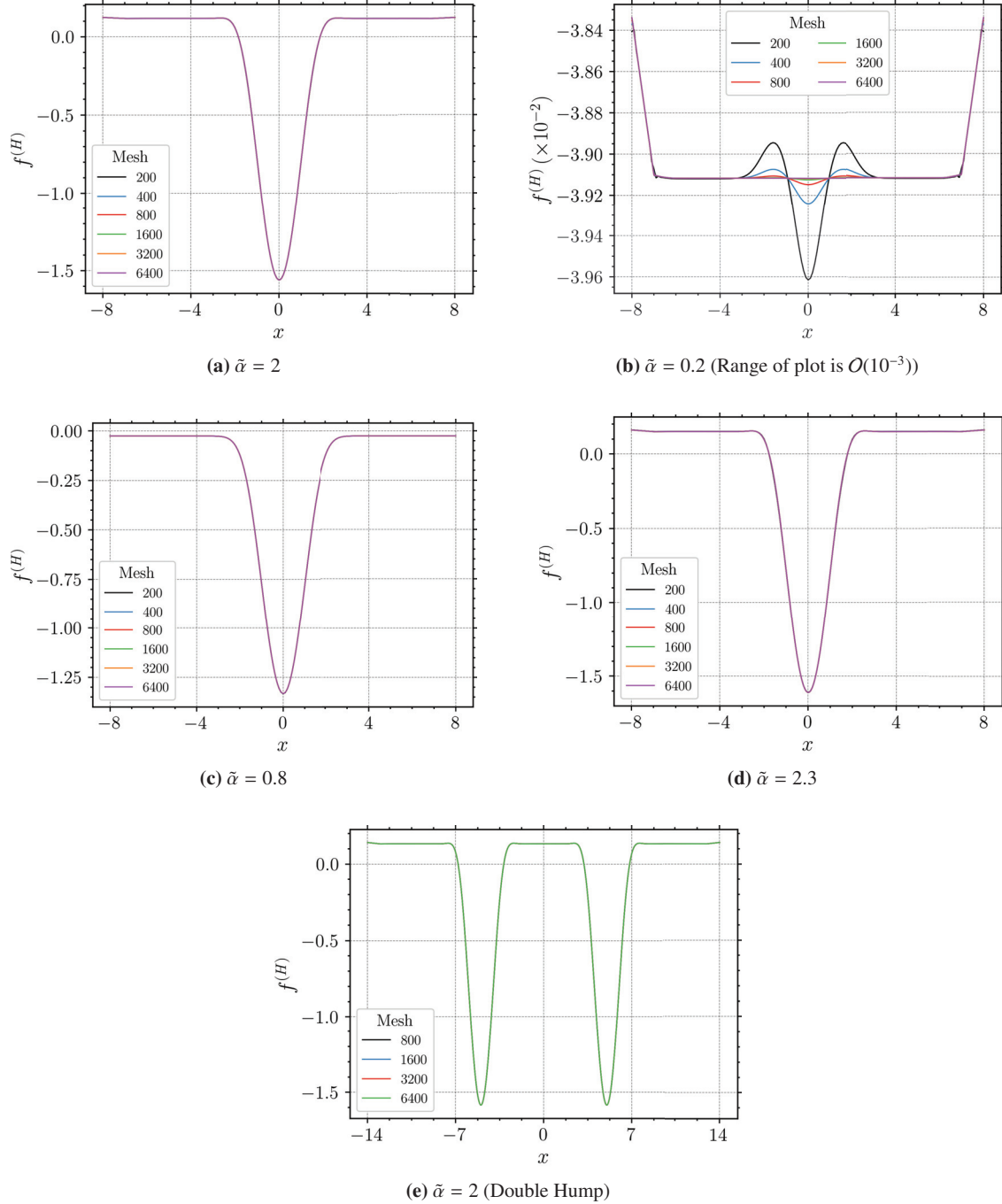
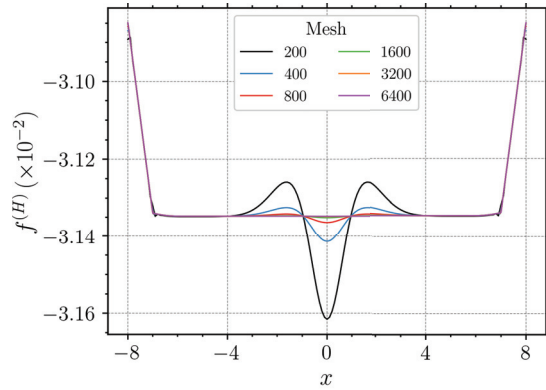
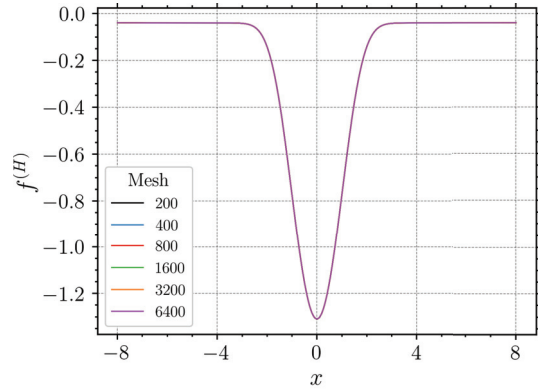


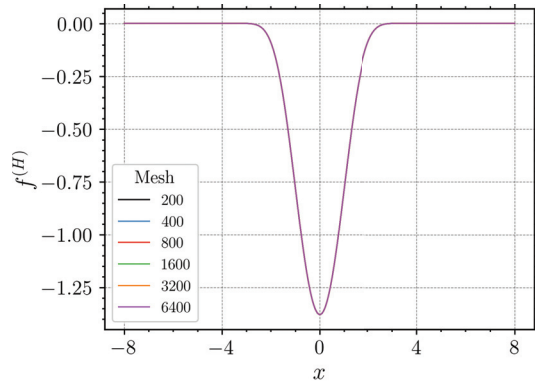
Figure D.10: Convergence of obtained primal profiles w.r.t mesh refinement with $\mathcal{P}^{\mathcal{V}}$ profiles scaled by a factor of $\tilde{\alpha}$ set as the base states: Fig. (a)-(d) were produced with a single hump in the base state, whereas Fig. (e) was produced using a double hump in the base state.



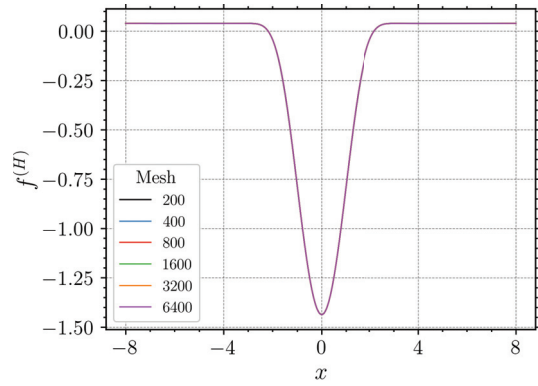
(a) $\gamma = -0.5$ (Range of plot is $O(10^{-4})$)



(b) $\gamma = -2.7$

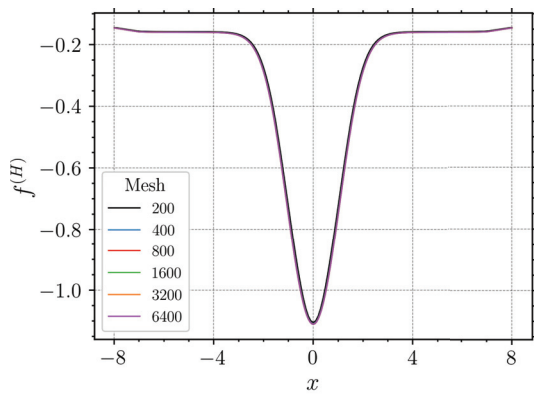


(c) $\gamma = -4.0$

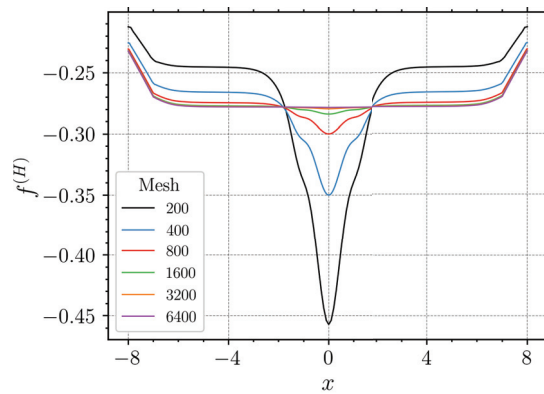


(d) $\gamma = -5.2$

Figure D.11: Convergence of obtained primal profiles w.r.t mesh refinement for scaled Gaussian profiles (by a factor of γ) set as base states. Results were produced using a simple N-R scheme.



(a) $\gamma = -1.9$



(b) $\gamma = -1.7$ (Range of plot is $O(10^{-1})$)

Figure D.12: Convergence of obtained primal profiles w.r.t mesh refinement for two closely related scaled Gaussian profiles, each scaled by a factor of γ , set as the base state. Alg. 1 was utilized.

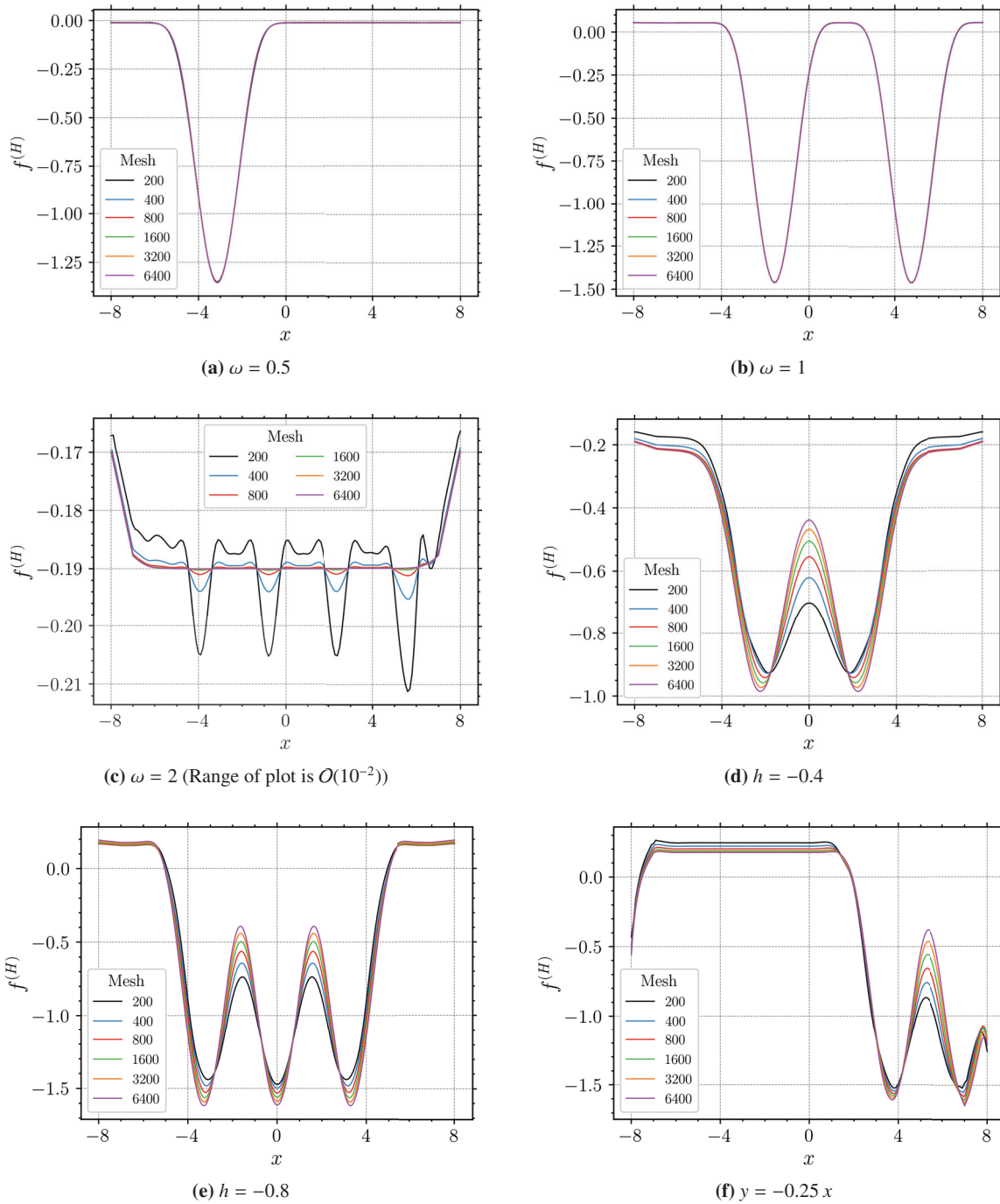


Figure D.13: Convergence of obtained primal profiles w.r.t mesh refinement for the following functions set as base states: sinusoidal functions (4.15), negative hat functions (4.16) and a linear function (4.17) are shown in Fig. (a)-(c), Fig. (d)-(e) and Fig. (f) respectively. Alg. 1 was utilized.

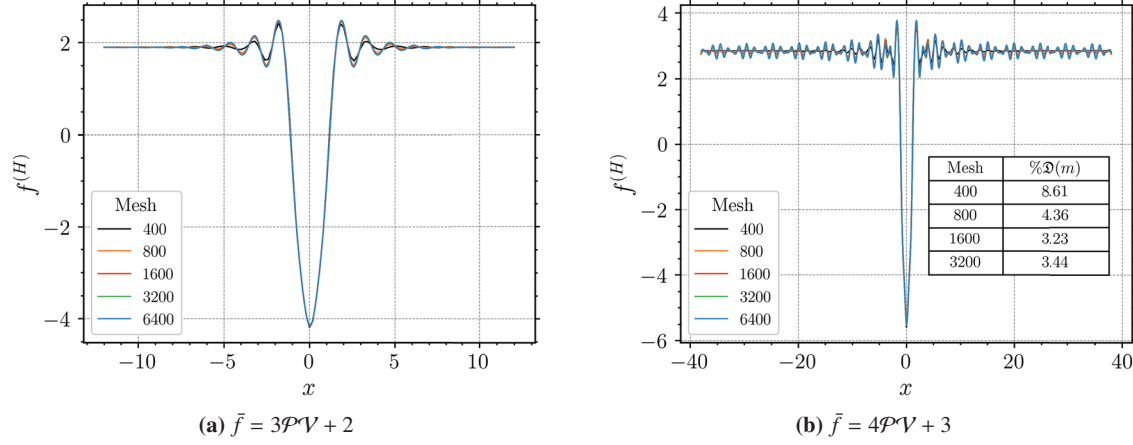


Figure D.14: Convergence of obtained primal profiles w.r.t mesh refinement for base states set as a scaled \mathcal{P}^V solution shifted by a constant (Fig. 7). Profiles were drawn based on tests performed on $L = 38$.

Appendix E. Supporting tables

In Table E.4, we tabulate values of the numerically estimated residual Err^A based on (C.1). In Table E.5, we tabulate values of $\mathcal{D}(m)$ from (4.10), testing convergence of the finite element scheme under mesh refinement. Values for $\mathcal{D}(m)$, (%) are provided for all examples. In this table, the second column, titled ‘Fig.’, corresponds to the figures illustrating the convergence of obtained primal profiles w.r.t mesh refinement.

Example			Mesh						Multiplier
Base State	Fig.	Parameter	200	400	800	1600	3200	6400	
Scaled \mathcal{P}^V	D.10a	$\tilde{\alpha} = 2$	198	52	13	3	1	0.8	$\times 10^{-4}$
	D.10b	$\tilde{\alpha} = 0.2$	6.8	1.7	0.4	0.1	0.04	0.02	
	D.10c	$\tilde{\alpha} = 0.8$	61	16	4	1	0.4	0.4	
	D.10d	$\tilde{\alpha} = 2.3$	514	140	35	9	3	1	
Double Hump	D.10e	$\tilde{\alpha} = 2$	-	-	39	9.8	2.7	0.8	$\times 10^{-4}$
Gaussian	D.11a	$\gamma = -0.5$	3.6	0.9	0.22	0.05	0.01	0.003	$\times 10^{-4}$
	D.11b	$\gamma = -2.7$	37.9	9.48	2.37	0.59	0.14	0.03	
	D.11c	$\gamma = -4.0$	66	17	4	1	0.2	0.06	
	D.11d	$\gamma = -5.2$	108	28	7	1	0.4	0.1	
	D.12a	$\gamma = -1.7$	153	60	18	5	1.6	0.3	$\times 10^{-3}$
	D.12b	$\gamma = -1.9$	53	14.9	3.9	1	0.2	0.06	
Sine wave	D.13a	$\omega = 0.5$	217	55	14	3.5	0.9	0.2	$\times 10^{-4}$
	D.13b	$\omega = 1$	166	42.8	10.8	2.7	0.7	0.2	
	D.13c	$\omega = 2$	396	99	24	6.2	1.5	0.3	
Hat functions	D.13d	$h = -0.4$	13.3	7.17	3.86	2.35	2.02	1.81	$\times 10^{-2}$
	D.13e	$h = -0.8$	25.9	14.2	12.8	11.4	9.75	8.2	
Linear	D.13f	$y = -0.25x$	35.2	22.6	18	14.1	10.3	7.29	$\times 10^{-2}$
Soliton	-	$\tilde{f} = 3\mathcal{P}^V + 2$	-	19.6	4.97	2.29	1.13	0.56	$\times 10^{-3}$
d-soliton	-	$\tilde{f} = 4\mathcal{P}^V + 3$	-	11.4	9.3	5.7	3.56	2.14	$\times 10^{-1}$

Table E.4: Accuracy test: Numerically estimated residual Err^A based on (C.1)

Example			$\mathfrak{D}(m)$, (%) on Mesh				
Base State	Fig.	Parameter	200	400	800	1600	3200
Scaled $\mathcal{P}\mathcal{V}$	D.10a	$\tilde{\alpha} = 2$	1.06	0.27	0.06	0.02	0.004
	D.10b	$\tilde{\alpha} = 0.2$	1.07	0.27	0.07	0.02	0.006
	D.10c	$\tilde{\alpha} = 0.8$	0.67	0.17	0.04	0.01	0.002
	D.10d	$\tilde{\alpha} = 2.3$	1.78	0.45	0.11	0.03	0.007
Double Hump	D.10e	$\tilde{\alpha} = 2$	-	-	0.16	0.04	0.01
Gaussian	D.11a	$\gamma = -0.5$	0.71	0.18	0.05	0.01	0.005
	D.11b	$\gamma = -2.7$	0.40	0.10	0.03	0.006	0.002
	D.11c	$\gamma = -4.0$	0.43	0.11	0.03	0.007	0.002
	D.11d	$\gamma = -5.2$	0.89	0.22	0.06	0.01	0.003
	D.12a	$\gamma = -1.7$	43.62	20.40	6.53	1.76	0.44
	D.12b	$\gamma = -1.9$	5.55	1.60	0.41	0.10	0.03
Sine wave	D.13a	$\omega = 0.5$	2.96	0.72	0.18	0.04	0.01
	D.13b	$\omega = 1$	1.59	0.34	0.19	0.17	0.14
	D.13c	$\omega = 2$	9.44	2.38	0.58	0.14	0.03
Hat functions	D.13d	$h = -0.4$	16.09	13.25	9.88	7.52	5.86
	D.13e	$h = -0.8$	12.35	10.68	8.89	7.52	6.52
Linear	D.13f	$y = -0.25x$	17.5	17	16.5	15.61	13.64
Soliton	-	$\tilde{f} = 3\mathcal{P}\mathcal{V} + 2$	-	0.29	0.14	0.071	0.036
d-soliton	-	$\tilde{f} = 4\mathcal{P}\mathcal{V} + 3$	-	8.56	6.82	5.2	3.64

Table E.5: Convergence test: Values of $\mathfrak{D}(m)$, (%) from (4.10)

In both tables, the label ‘Soliton’ refers to the dispersive solitary wave profile (Fig. 7a), and ‘d-soliton’ denotes the disintegrated profile (Fig. 7b), both up to adjusting for value of u_∞ as explained in subsection 4.2.3. All the tests (except the double-hump example) are performed for $L = 8$.