High-fidelity numerical simulation of solitons in the nerve axon

Ken Mattsson * Jonatan Werpers

October 12, 2015

Abstract

High-order accurate finite difference schemes are derived for a nonlinear soliton model of nerve signal propagation in axons. Two types of well-posed boundary conditions are analysed. The boundary closures are based on the summation-by-parts (SBP) framework and the boundary conditions are imposed using a penalty (SAT) technique, to guarantee linear stability. The resulting SBP-SAT approximation is time-integrated with an explicit finite difference method. The accuracy and stability properties of the newly derived finite difference approximations are demonstrated for an analytic soliton solution.

Key words: finite difference methods, high-order derivative, high-order accuracy, stability, boundary treatment, nerve signals, solitons

1 Introduction

Understanding the propagation of nerve signals is of great concern in neuroscience. Specifically, the behaviour of the signal in the axon of the nerve cell has been of great interest. In 1952 Hodgkin and Huxley introduced a mathematical model for this phenomenon [21], describing the nerve axon as an electrical circuit in which the proteins are resistors and the membrane is a capacitor. Ion currents flow through the membrane and along the nerve axon leading to a propagating pulse. Since it was introduced, the model has

*Department of Information Technology, Uppsala University, P O Box 337, S-751 05 Uppsala, Sweden. telephone: +46-18-4717631, telefax: +46-18-523049, E–mail: ken.mattsson@it.uu.se

provided insight into how neurons behave. Mathematically, the Hodgkin-Huxley (HH) model consists of a scalar parabolic equation with a nonlinear source term, coupled with a set of ordinary differential equations (ODE). A high-fidelity finite difference approximation of the HH model is presented in [3].

It has been suggested (see for example [25, 4]) that the HH model does not accurately describe certain phenomena observed in experiments. While the HH model describes various aspects of the voltage pulse traveling along the nerve axon in a satisfactory manner (e.g., its velocity and the pulse amplitude), it fails to describe several other aspects of the nerve pulse that are of nonelectrical nature. It is known that nerves display thickness and length variations under the influence of the voltage pulse and the nerve pulse can be excited by a mechanical stimulus, indicating that the nerve pulse possesses a mechanical component [44]. Furthermore, heat signatures from experimental data [44] indicate that the nerve pulse is an adiabatic and reversible phenomenon such as the propagation of a mechanical wave. This later observation is in conflict with the HH model that is based on irreversible dissipative processes (currents through resistors) and should lead to dissipation of heat [18]. This, however, is not observed in nerves [18, 4].

In a series of recent publications by Heimburg *et al.* [17, 18, 26, 40, 25] an alternative thermodynamic model is proposed in which nerve pulses are described as a reversible electro-mechanical density pulse (soliton) in the axon membrane. This model will be referred to as the nerve soliton (NS) model in the present study. A thorough comparison between the NS and HH models is found in [4]. The pulse velocity in the NS model is close to the velocity of sound in the membrane, and the experimentally observed reversible heat corresponds to the reversible heating of compressible media during the propagation of adiabatic waves. The governing equation of the NS model is a nonlinear dispersive wave equation that is second order in time and fourth order in space, with an additional nonlinear second derivative term. Hence, the NS model and the HH model are governed by very different types of partial differential equations (PDE).

The focus in the present study is not to validate the correctness and significance of the NS model concerning neural activity. Nor do we claim that this novel NS model is more correct than the well-matured HH model. The focus is rather on the mathematical and numerical aspects of the NS model. The results from the present study will hopefully be useful for practitioners of the NS model and similar models.

The NS model has strong similarities to the Boussinesq equation [6, 7] that was the first theoretical description of solitary waves (in hydrodynamics) first observed by Scott Russell in 1834. In addition to their intrinsic mathematical

interest, solitons play an important role in hydrodynamics, quantum field theory [37], antiferromagnetism [8], Bose-Einstein condensates [12], nonlinear optics [11] and biological systems (DNA) [47]. The vast soliton literature also includes many variants of Boussinesq's original equation.

One of the main objectives in the present study is to derive a high-fidelity numerical finite difference method to simulate the NS model. In the literature, numerical analysis of Boussinesq-type equations in time-domain is very scarce, in particular concerning higher-order methods and the boundary treatment. Most numerical methods (see for example [10]) rewrite the model into a first order (in time) system and do not address the boundary treatment. This rewrite to first order form has the disadvantage of introducing auxiliary variables and is also less attractive from a computational point of view considering the efficiency and accuracy [22, 45].

The first main goal in the present study is to derive a well-posed initial boundary value problem (IBVP) for the NS model, for two very different types of boundary conditions (BC). The second main goal is to derive high-fidelity finite difference approximations of the well-posed IBVP. The third main goal is to numerically study soliton solutions of the NS model, and in particular investigate its stability behaviour.

It is well known that higher order methods capture wave dominated phenomena more efficiently since they allow a considerable reduction in the degrees of freedom (for a given error tolerence). (See the pioneering paper by Kreiss and Oliger [24] concerning hyperbolic problems.) The fourth derivative term in the NS model (and other Boussinesq-type equations) makes explicit time-integration computationally expensive due to a severe CFL condition, i.e., $k \simeq h^2$, where k is the time-step and h the spatial grid-size. As compared to hyperbolic models, we therefore expect the efficiency gain in using higher order methods (as opposed to lower order methods) to approximate the NS model to be even more pronounced (due to the more restrictive CFL condition). This will be carefully investigated in the present study.

The major difficulty with higher-order finite difference methods is to obtain a stable boundary treatment, that has received considerable past attention concerning hyperbolic and parabolic problems. (For examples, see [27, 42, 39, 1, 5, 19].) Roughly speaking, the numerical difficulties increase with the order of the spatial (and temporal) derivatives. The fourth derivative term in the NS model introduces additional difficulties of imposing boundary conditions, as compared to hyperbolic and parabolic problems.

The summation-by-parts–simultaneous approximation term (SBP-SAT) method is a robust and well-proven high-order finite difference methodology that ensures stability of time-dependent PDEs. The SBP-SAT method combines semi-discrete operators that satisfy a SBP formula [23], with phys-

ical BC implemented using the simultaneous approximation term (SAT) method [9]. Recent examples of the SBP-SAT approach can be found in [20, 28, 2, 38, 13].

The SBP-SAT approach has so far been developed for problems involving first and second derivatives in space. However, there are many problems where higher order derivatives are present. Some examples include the Korteweg-de Vries and the Boussinesq equations (describing nonlinear water waves), soliton models in neuroscience [40], the Euler-Lagrange equation for beams [46, 14], and the Cahn-Hilliard equation which describes the process of phase separation. Recently, high-order accurate SBP operators for third and fourth derivatives were derived [36]. Stable SBP-SAT approximations of the NS model require second- and fourth-derivative SBP operators that fulfill something we refer to as *1D compatibility* (see Definition 3.5). Second derivative SBP operators were derived in [35]. The fourth derivative SBP operators derived in [36] were however not 1D compatible with the second derivative SBP operators derived in [35]. In the present study we have therefore derived completely new 1D compatible high-order accurate SBP operators applicable to the NS model.

In Section 2 the NS model is introduced, including two different types of boundary conditions. In Section 3 the SBP-SAT method is introduced. Stability analysis of the SBP-SAT approximations is presented in Section 4. Time integration is analysed and discussed in Section 5. In Section 6 the accuracy and stability properties of the newly developed SBP-SAT approximations are verified by numerical simulations. Section 7 summarizes the work. The SBP operators are presented in the Appendix.

2 The NS model

The NS model was first presented in [17]. One of the main motivations with the present study is to establish well-posed BC for the NS model. For completeness we reproduce the derivation of the NS model in this section, which starts from the wave equation for area density changes,

$$\frac{\partial^2}{\partial \tau^2}\Delta\rho^A = \frac{\partial}{\partial z}\left[c^2\frac{\partial}{\partial z}\Delta\rho^A\right], \tag{1}$$

where τ is time, z is position, c is the speed of sound and $\Delta\rho^A$ is the density offset from the equilibrium. The sound velocity in the lipid membrane is considered a function of density, and is represented by a truncated power series,

$$c^2 = c_0^2 + p\Delta\rho^A + q\left(\Delta\rho^A\right)^2, \tag{2}$$

where p and q are parameters determined by measurements (see [17, 18, 26] for more details). To account for dispersion observed in experiments (see [4, 26]) an extra term $-\bar{h}\frac{\partial^4}{\partial z^4}\Delta\rho^A$ is added. Finally, the following mathematical model for the density of the lipid membrane was presented in [17],

$$\frac{\partial^2}{\partial\tau^2}\Delta\rho^A = \frac{\partial}{\partial z}\left[\left(c_0^2 + p\Delta\rho^A + q\left(\Delta\rho^A\right)^2\right)\frac{\partial}{\partial z}\Delta\rho^A\right] - \bar{h}\frac{\partial^4}{\partial z^4}\Delta\rho^A, \quad (3)$$

where

- $\Delta\rho^A = \rho^A - \rho_0^A$
- ρ^A - Lateral density of the membrane
- ρ_0^A - Equilibrium density in the fluid phase
- c_0 - Velocity of small amplitude sound
- p, q - Experimentally determined parameters determining the density dependence of sound velocity
- $\bar{h}$ - Determines dispersion (see [26, 4] for a theoretical argument).

In [26] the following variable changes were introduced,

$$u = \frac{1}{\rho_0^A}\Delta\rho^A, \quad x = \frac{c_0}{\sqrt{\bar{h}}}z, \quad t = \frac{c_0^2}{\sqrt{\bar{h}}}\tau, \quad \gamma_1 = \frac{\rho_0^A}{c_0^2}p, \quad \gamma_2 = \frac{(\rho_0^A)^2}{c_0^2}q, \quad (4)$$

arriving at the following dimensionless version of (3),

$$u_{tt} = \left(b(u)u_x\right)_x - u_{xxxx}, \quad (5)$$

where

$$b(u) = 1 + \gamma_1 u + \gamma_2 u^2\ . \quad (6)$$

Here we have adopted the notation u_{xxxx}, u_x and u_{tt} for the fourth, first and second partial derivatives of $u(x,t)$ in space and time respectively. The parameters $\gamma_1 = -16.6$ and $\gamma_2 = 79.5$ were determined experimentally in [17].

Remark The particular choices $\gamma_1 = -16.6$ and $\gamma_2 = 79.5$ guarantee that $b(u) > 0$ for all u.

As previously mentioned, the NS model is a dispersive nonlinear wave equation, with the principal part given by $u_{tt} = -u_{xxxx}$. The dispersion relation $\omega = \omega(\kappa)$ determines how the temporal oscillations are linked to the spatial oscillations. Plugging the plane wave solution $e^{i\kappa x}e^{i\omega(\kappa)t}$ into $u_{tt} = -u_{xxxx}$ and solving for $\omega(\kappa)$ results in the dispersion relation,

$$\omega(\kappa) = \pm\kappa^2 \,, \tag{7}$$

where $\omega(\kappa)$ is the angular frequency and κ is the wave-number. Hence, the group velocity (given by 2κ) and the wave speed (given by κ) depend on the wave-number of the wave. In addition, (7) indicates that wave propagation will occur in two opposite directions. Here we have ignored the nonlinear lower order term $(b(u)u_x)_x$, which also affects the wave speed, although a small correction. Assuming that $b(u) = b$ is constant we obtain the dispersion relation $\omega(\kappa) = \pm\kappa\sqrt{\kappa^2 + b}$.

Roughly speaking, solitons exist as a consequence of a balance between nonlinear and dispersive effects. In [4] it is shown that (5), for the Cauchy problem, supports the following soliton,

$$u(x,t,\beta) = \frac{2a_+a_-}{(a_+ + a_-) + (a_+ - a_-)\cosh((x-\beta t)\sqrt{1-\beta^2})} \,, \tag{8}$$

where

$$a_\pm = -\frac{\gamma_1}{\gamma_2}\left(1 \pm \sqrt{\frac{\beta^2 - \beta_0^2}{1-\beta_0^2}}\right) \,.$$

The parameter β, the speed of the soliton, satisfies $\beta_0 < |\beta| < 1$ where $\beta_0 = 0.649851$. Further, β also indirectly determines the amplitude and shape of the pulse. This means that this kind of soliton can only exist for certain traveling speeds, and further, that they are restricted in amplitude. In Figure 1 we present the soliton solution for $\beta = 0.65, 0.80, 0.95$, at two different times. The stability of this particular soliton solution will be verified numerically in Section 6. With initial data different from Eq. 8 the dispersive nature of the NS model is revealed, and the dispersion relation (7) is clearly visible. In Figure 7 we test the numerical method against an initial rectangular function. In Figure 8 a new type of *square* (non-analytic) soliton is found by numerical simulations.

2.1 Well-posedness

One of the main motivations with the present study is to derive two different types of BC that yield well-posed IBVP for the NS model (5). We start by

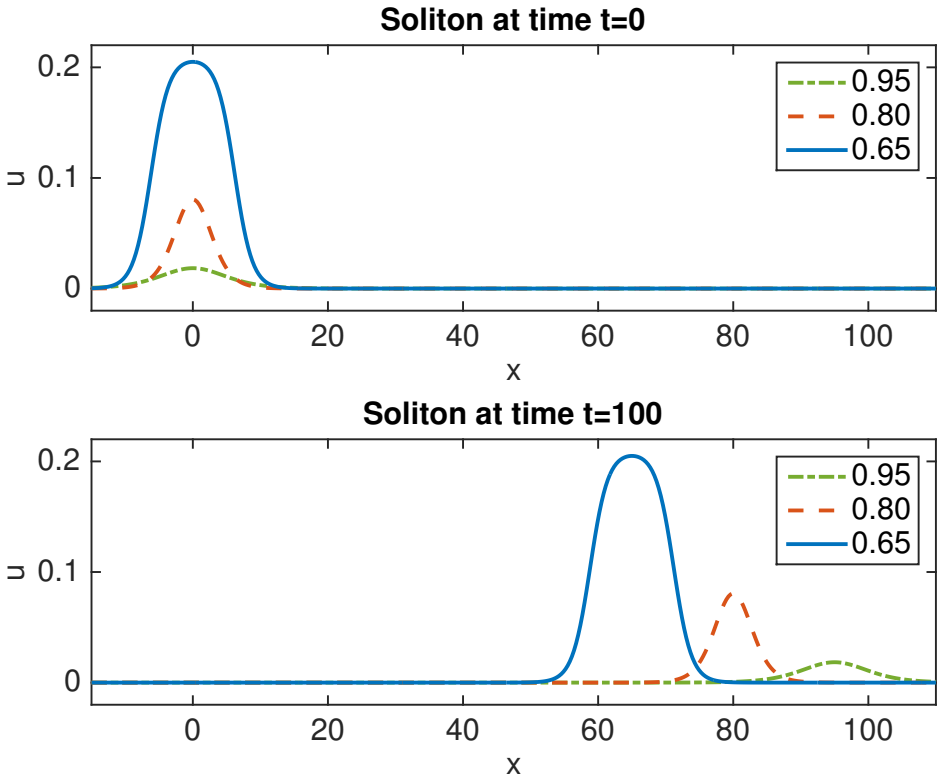

Figure 1: The soliton solutions (8) for $\beta = 0.65, 0.80, 0.95$, at two different times. At $t = 0$ the solitons are centered at $x = 0$. The solitons move to the right with speed 0.65, 0.80, 0.95 and are shown at $t = 100$.

restricting the computational domain to $0 \leq x \leq L$. The analysis of the IBVP will be performed using the energy method. Multiplying Eq. 5 by u_t and using integration by parts twice leads to,

$$\int_0^L u_t u_{tt}\, dx = BT|_0^L - \int_0^L u_{xt} b(u) u_x\, dx - \int_0^L u_{xxt} u_{xx}\, dx\ ,$$

where the boundary terms (BT) are given by

$$BT|_0^L = (u_t b(u) u_x - u_t u_{xxx} + u_{tx} u_{xx})|_0^L\ .$$

This can be written:

$$\frac{1}{2}\frac{d}{dt}\int_0^L \left(u_t^2 + b(u)u_x^2 + u_{xx}^2\right)\, dx = BT|_0^L + \frac{1}{2}\int_0^L b_t(u) u_x^2\, dx\ , \qquad (9)$$

where $\int_0^L (u_t^2 + b(u)u_x^2 + u_{xx}^2)\, dx$ represents a continuous energy, since by construction $b(u) > 0$ holds (see previous remark). To obtain a well-posed

IBVP we now have to derive well-posed BC. The nonlinear term $(b(u)u_x)_x$ in Eq. 5 introduces a potential growth term in the RHS of (9), unless $b_t(u) \leq 0$. In the following analysis we will analyse well-posedness and stability for a linearised version of (5) by freezing the nonlinear coefficient $b(u)$, and replacing it with the constant $b > 0$. The linearised problem is given by,

$$\begin{aligned} &u_{tt} = (bu_x)_x - u_{xxxx}, \quad 0 \leq x \leq L, \quad t \geq 0 \\ &u = f_1(x), \; u_t = f_2(x), \quad 0 \leq x \leq L, \quad t = 0 \end{aligned} \; . \tag{10}$$

Here $f_{1,2}(x)$ are initial data.

Remark For the linearised problem (10), $b_t = 0$ and the Cauchy problem is thus energy conservative. The linear problem is expected to accurately represent the stability behaviour of the nonlinear problem as long as u is sufficiently smooth [41]. Smoothness is a valid assumption since we are interested in the stability behaviour of soliton solutions, that are by definition smooth. However, for a non-smooth solution there is no guarantee that the nonlinear problem is energy conservative, since the additional term $\frac{1}{2}\int_0^L b_t(u)u_x^2\,dx$ in the RHS of (9) becomes large. The stability of the numerical method will be validated also for non-smooth initial data in Section 6.

Let the inner product for real-valued functions $u, v \in L^2[0, L]$ be defined by $(u, v)_a = \int_0^L ua(x)v\,dx$, (where $a(x) > 0$) and let the corresponding norm be $\|u\|_a^2 = (u, u)_a$. We will now employ the energy method to analyse well-posedness of two types of BC for the linearised problem (10). Multiplying (10) by u_t, integrating by parts twice and adding the transpose leads to,

$$\frac{d}{dt}E^{(c)} = 2(bu_tu_x - u_tu_{xxx} + u_{tx}u_{xx})|_0^L \; , \tag{11}$$

where the energy is given by,

$$E^{(c)} = \|u_t\|^2 + \|u_x\|_b^2 + \|u_{xx}\|^2 \; . \tag{12}$$

Definition 2.1 *In the present study we will use the notation $E^{(c)} = E^{(c)}(t)$ for brevity.*

The number of BC required to obtain a well-posed model is given by analysing the boundary terms (BT) in (11), rewritten as

$$\frac{d}{dt}E^{(c)} = BT|_0^L = w^TGw|_0^L \; ,$$

where

$$w = \begin{bmatrix} u_t \\ u_x \\ u_{xt} \\ u_{xx} \\ u_{xxx} \end{bmatrix}, \quad G = \begin{bmatrix} 0 & b & 0 & 0 & -1 \\ b & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

The number of BC at $x = L$ equals the number of positive eigenvalues to G, and the number of negative eigenvalues to G yields the correct number of BC at $x = 0$. We now diagonalise G, i.e., $G = S^T \Lambda S$, where $\Lambda = diag(-1, 0, 1, \sqrt{1+b^2}, -\sqrt{1+b^2})$ holds the eigenvalues to G. We have two positive and two negative eigenvalues, and hence 2 BC are to be prescribed at each boundary. To simplify notation we introduce

$$\sigma = \sqrt{1+b^2} > 0 \, . \tag{13}$$

The diagonalizer S holds the eigenvectors to G (the first column belongs to the first eigenvalue) and is given by,

$$S = \begin{bmatrix} 0 & 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sigma} & 0 & -\frac{b}{\sqrt{2}\sigma} & -\frac{b}{\sqrt{2}\sigma} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} & 0 & 0 \\ 0 & \frac{b}{\sigma} & & \frac{1}{\sqrt{2}\sigma} & \frac{1}{\sqrt{2}\sigma} \end{bmatrix} .$$

Introducing the characteristic variables $\tilde{w} = S\, w$ the boundary terms can be rewritten $w^T G w = \tilde{w}^T \Lambda \tilde{w} = \sum_{j=1}^{5} \lambda_j \tilde{w}_j^2$, where λ_j is the jth eigenvalue and $\tilde{w}_j$ the corresponding characteristic variable. By specifying the characteristic variables corresponding to the two positive eigenvalues at $x = L$, i.e., $\tilde{w}_{3,\,4}$ and the two eigenvalues corresponding to the two negative eigenvalues at $x = 0$, i.e., $\tilde{w}_{1,\,5}$ we obtain the most dissipative set of well-posed BC given by,

$$\begin{aligned} &u_{xt} - u_{xx} = g_0^{(1)}(t), && x = 0 \, , \\ &\sigma u_t - b u_x + u_{xxx} = g_0^{(2)}(t), && x = 0 \, , \\ &u_{xt} + u_{xx} = g_L^{(1)}(t), && x = L \, , \\ &\sigma u_t + b u_x - u_{xxx} = g_L^{(2)}(t), && x = L \, . \end{aligned} \tag{14}$$

Here $g_0^{(1)}(t)$, $g_0^{(2)}(t)$, $g_L^{(1)}(t)$, $g_L^{(2)}(t)$ are boundary data. In the present study (14) will be referred to as Characteristic (Char) BC.

Remark The analysis of well-posedness is done for the homogeneous case, i.e. by setting the boundary data to zero. For certain BC a stronger energy estimate can be derived (such as strong well-posedness [16]) but this will not be analysed in the present study. Well-posedness for the non-homogeneous version follows directly from the energy analysis of the homogeneous version [16]. We will similarly assume homogeneous data in the semi-discrete stability analysis.

Inserting the Char BC (14) with zero boundary data ($g_{0,\,L}^{(1,\,2)}(t) = 0$) into (11) leads to,

$$\frac{d}{dt}E^{(c)} = 2(bu_t u_x - u_t u_{xxx} + u_{tx}u_{xx})|_0^L = -2(\sigma u_t^2 + u_{xt}^2)|^L - 2(\sigma u_t^2 + u_{xt}^2)|^0 \,. \tag{15}$$

Hence, the energy $E^{(c)}$ will be damped through the boundaries. Here σ is given by (13).

A less rigorous treatment of the BC can be made by simply studying the right hand side of (11) and choosing a set of BC (2 at each side) that fulfills the condition $\frac{d}{dt}E^{(c)} \leq 0$ (here assuming homogeneous boundary data). One set of well posed BC is given by specifying u_x and u_{xxx}. The following set of BC,

$$\begin{array}{llll} u = g_0^{(1)}(t) & u_x = g_0^{(2)}(t) & x = 0 \\ u = g_L^{(1)}(t) & u_x = g_L^{(2)}(t) & x = L \end{array} \,, \tag{16}$$

will be referred to as Dirichlet-Neumann (DN) BC, and is expected to be utilised in many applications (just as Neumann and Dirichlet BC are important concerning hyperbolic wave equation problems). As shown in the coming section, DN BC is also particularly difficult to impose using the SBP-SAT method.

To see why (16) leads to wellposedness we start by taking the time derivative (and assuming zero data), resulting in

$$\begin{array}{llll} u_t = 0 & u_{xt} = 0 & x = 0 \\ u_t = 0 & u_{xt} = 0 & x = L \end{array} \,. \tag{17}$$

Inserting (17) into (11) leads to

$$\frac{d}{dt}E^{(c)} = 2(bu_t u_x - u_t u_{xxx} + u_{tx}u_{xx})|_0^L = 0 \,. \tag{18}$$

Hence, by imposing DN BC (with zero data) the energy $E^{(c)}$ is conserved.

The energy $E^{(c)}$, given by Eq. 12 is a seminorm, as the solution u itself is not included. However, as

$$\frac{d}{dt}\|u\|^2 = (u_t, u) + (u_t, u) \leq \|u\|^2 + \|u_t\|^2 \,,$$

we can add $\|u\|^2$ to the energy $E^{(c)}$ in (15) (or (18)), and get the differential inequality

$$\frac{d}{dt}\tilde{E}^{(c)} \leq \eta\, \tilde{E}^{(c)} \,, \tag{19}$$

where η is a constant and

$$\tilde{E}^{(c)} = E^{(c)} + \|u\|^2 = \|u_t\|^2 + \|u_x\|_b^2 + \|u_{xx}\|^2 + \|u\|^2 \,. \tag{20}$$

Integrating (19) in time yields

$$\tilde{E}^{(c)}(t) \leq e^{\eta\, t}\tilde{E}^{(c)}(0) \,.$$

Since $\tilde{E}^{(c)}$ includes $\|u\|^2$ we obtain a bound on $\|u\|^2$ in terms of initial data $f_{1,\,2}$ (and its derivatives). This energy estimate, although formally a valid stability proof (see [16] for more details regarding stability), is not very sharp since there is now an exponential growth in time. For simplicity we shall work with the seminorm in the following.

Remark A more detailed energy analysis (not shown in the present study) leads to a sharper energy estimate, where we can show $\|u\| \leq \|f_1\| + \xi \cdot t$, for a constant ξ.

An interesting numerical test is when a soliton interacts with a boundary. In Figure 2 we present a direct comparison between homogeneous DN and Char BC, to elucidate the behaviour of the different boundary conditions. As initial data we use the soliton (8) for $\beta = 0.80$ centered at $x = 70$. The soliton moves to the right with speed 0.80 and reaches the right boundary at approximately $t = 30$. Numerical solutions, using sixth-order accurate SBP-SAT approximations (given by (35) and (39)) are presented at $t = 30$, 40, 50 using $m = 401$ grid-points and a time-step $k = 0.03\, h^2$. (Details and analysis of the numerical method will be presented in the following sections.)

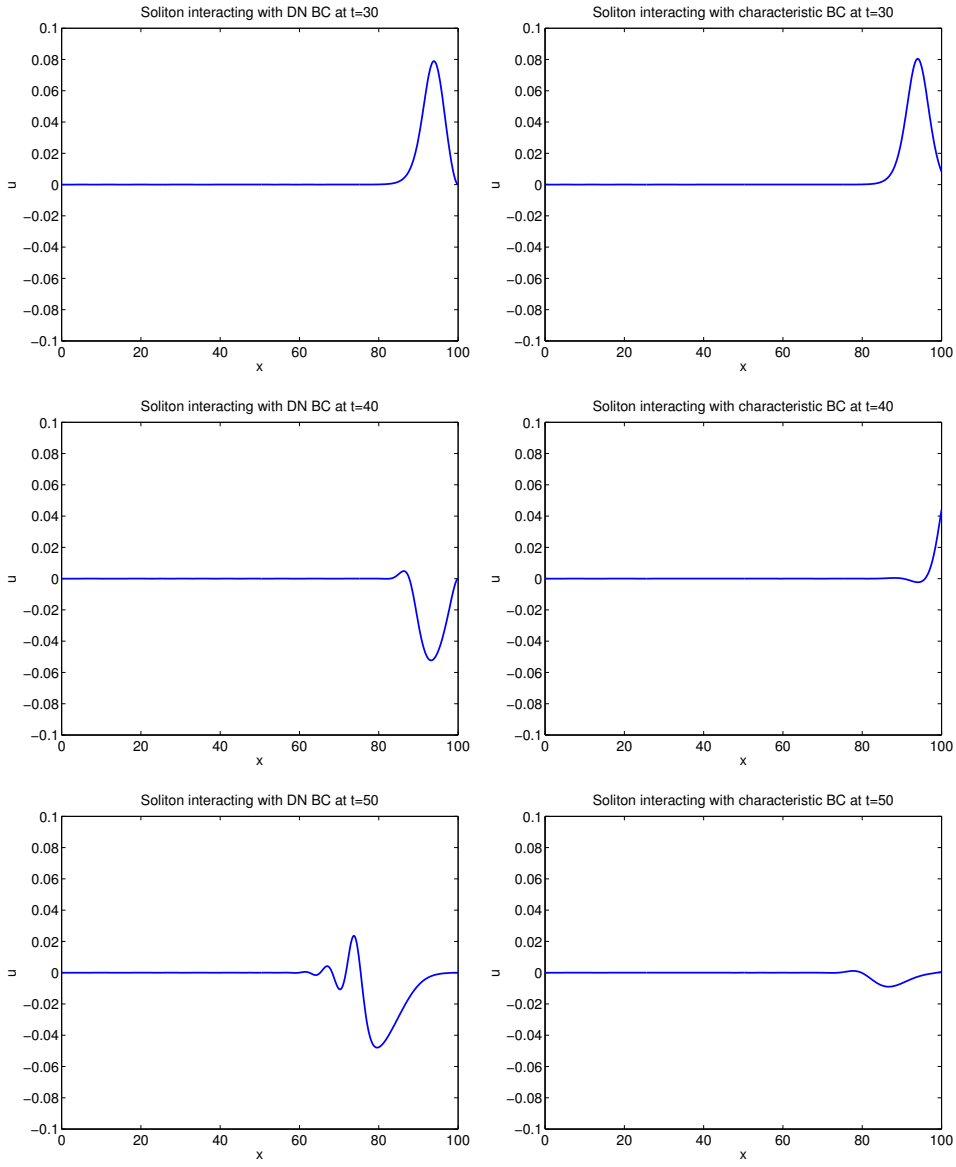

Figure 2: Numerical solution of the NS model with two different homogeneous BC, DN BC (left column) and Char BC (right column). The solutions are presented at $t = 30$, 40, 50 using $m = 401$ grid-points.

3 The finite difference method

In this section we will introduce the SBP-SAT method, which combines SBP finite difference operators and the SAT technique of weakly imposing boundary conditions. To make the paper more clear we will here introduce the

SBP-SAT technique for the wave equation on second order form. In Section 4 the SBP-SAT method will be extended to the NS model subjected to two types of BC.

We begin with some useful definitions, that are needed in the coming SBP-SAT analysis. The domain $(0 \leq x \leq L)$ is discretized using the following m equidistant grid points:

$$x_i = (i-1)\,h, \quad i = 1, 2, \ldots, m, \quad h = \tfrac{L}{m-1}\ .$$

The following vectors will be used frequently:

$$\begin{aligned} &\mathbf{x} = [x_1,\, x_2, \ldots, x_{m-1},\, x_m]^T, \quad \mathbf{x}^q = \tfrac{1}{q!}[x_1^q,\, x_2^q, \ldots, x_{m-1}^q,\, x_m^q]^T \\ &\mathbf{1} = \mathbf{x}^0 = [1,\, 1, \ldots, 1,\, 1]^T, \qquad \mathbf{0} = [0,\, 0, \ldots, 0,\, 0]^T \\ &e_1 = [1, 0, \ldots, 0]^T, \qquad\qquad e_m = [0, \ldots, 0, 1]^T\ . \end{aligned} \tag{21}$$

The following definition is important for the stability analysis,

Definition 3.1 *Let L denote the width of the domain, and $h = \frac{L}{m-1}$ the grid-spacing. A symmetric positive definite matrix H defines a discrete norm if $\mathbf{1}^T H \mathbf{1} = L$, independent of the number of grid-points m.*

The approximate solution at grid point x_i is denoted v_i, and the discrete solution vector is $v^T = [v_1, v_2, \ldots, v_m]$. Similarly, we define an inner product for discrete real-valued vector functions $u, v \in \mathbf{R^m}$ by $(u, v)_H = u^T H\, v$, where the symmetric positive definite matrix H defines a discrete norm. The corresponding norm is $\|v\|_H^2 = v^T H\, v$.

The following definition (first stated in [34, 30]) is central to the present study:

Definition 3.2 *An explicit $2p$th-order accurate finite difference scheme with minimal stencil width for the Cauchy problem is denoted a $2p$th-order accurate narrow-stencil.*

3.1 The SBP-SAT method

The spatial dependency in the NS model consists of a fourth derivative term u_{xxxx} and a nonlinear second derivative term $(b(u)u_x)_x$. To approximate the NS model using the SBP-SAT method requires accurate SBP approximations of these two terms. SBP operators are essentially central finite difference stencils closed at the boundaries with a careful choice of one-sided difference stencils, to mimic the underlying integration-by-parts formula in a discrete

norm (denoted H in the present study). In the present paper we address the SBP operators by the accuracy of the central scheme and the type of norm which they are based on. The definitions for first- and second-derivative SBP operators can be found in earlier papers (see for example [9, 32, 34, 30, 29, 33, 35]). In a recent paper [36] third- and fourth-derivative SBP operators were derived. Finite difference SBP operators may be further categorized by the structure of their norm: a) diagonal, b) diagonal interior with block boundary closures, c) fully banded. In the present study we focus on diagonal-norm SBP operators.

Consider the 1D wave equation,

$$u_{tt} = (bu_x)_x, \quad 0 \le x \le L, \quad t \ge 0 \quad , \tag{22}$$

where $b = b(x) > 0$. Multiplying Eq. 22 by u_t integrating by parts and adding the transpose leads to,

$$\frac{d}{dt}E = 2u_t b u_x|_0^L = -2u_t b u_x|^0 + 2u_t b u_x|^L \ , \tag{23}$$

where the continuous energy is defined as

$$E = \left(\|u_t\|^2 + \|u_x\|_b^2 \right) \ . \tag{24}$$

The following SBP definition (first introduced in [30]) is central to the present study:

Definition 3.3 *A difference operator* $D_2^{(b)} = H^{-1}(-M^{(b)} + b_m e_m d_{1;m} - b_1 e_1 d_{1;1})$ *approximating* $\partial/\partial\, x\, (\, b\, \partial/\partial\, x)$*, where* $b(x) > 0$*, using a* $2p$*th-order accurate narrow-stencil in the interior, is said to be a* $2p$*th-order diagonal-norm second-derivative SBP operator if the diagonal matrix* H *defines a discrete norm,* $M^{(b)} = (M^{(b)})^T \ge 0$*,* $d_{1;1}v \simeq u_x|^0$ *and* $d_{1;m}v \simeq u_x|^L$ *are finite difference approximations of the first derivative at the left and right boundary points.*

The vectors e_1 and e_m are defined in (21). Diagonal-norm second-derivative SBP operators of orders 2, 4 and 6 were derived in [35].

A semi-discretization of (22) using the above SBP definition is given by $v_{tt} = D_2^{(b)} v$. Multiplying by $v_t^T H$ from the left and adding the transpose leads to,

$$\frac{d}{dt}E_H = -2(e_1^T v_t) b_1 d_{1;1} v + 2(e_m^T v_t) b_m d_{1;m} v \ , \tag{25}$$

where the semi-discrete energy is defined as

$$E_H = \left(\|v_t\|_H^2 + v^T M^{(b)} v \right) \ . \tag{26}$$

Estimate (25) is a discrete analog of Eq. 23.

To understand how to combine SBP operators and the SAT treatment for weakly imposing the BC in the NS model, it is instructive to first study the simplified wave equation model (22) subjected to the following BC:

$$\begin{array}{ll} \gamma u_t - b u_x = g_0, & x = 0, \quad t \geq 0, \\ \gamma u_t + b u_x = g_L, & x = L, \quad t \geq 0, \end{array} \tag{27}$$

where γ is a constant. Inserting the BC (27) into (23) leads to the following energy estimate:

$$\frac{d}{dt}\left(\|u_t\|^2 + \|u_x\|_b^2\right) = -2u_t\left(\gamma u_t - g_0\right)|^0 - 2u_t\left(\gamma u_t - g_L\right)|^L . \tag{28}$$

An energy estimate is obtained if $\gamma \geq 0$. The semi-discrete approximation of (27) can be written

$$\begin{array}{ll} \gamma e_1^T v_t - b_1 d_{1;1} v = g_0, & t \geq 0 , \\ \gamma e_m^T v_t + b_m d_{1;m} v = g_L, & t \geq 0 . \end{array} \tag{29}$$

The main idea with the SBP-SAT method is to impose the semi-discrete BC weakly, by adding them as penalty terms in the RHS of the SBP finite difference approximation. The SBP-SAT approximation of (22) subjected to the well-posed BC (27) is given by,

$$\begin{array}{ll} v_{tt} = D_2^{(b)} v & +\tau H^{-1} e_1 \left((\gamma e_1^T v_t - b_1 d_{1;1} v) - g_0\right) \\ & +\tau H^{-1} e_m \left((\gamma e_m^T v_t + b_m d_{1;m} v) - g_L\right) , \end{array} \tag{30}$$

where e_1 and e_m are defined in (21), and τ is the penalty parameter chosen to obtain a semi-discrete energy estimate.

Multiplying (30) by $v_t^T H$ from the left and adding the transpose leads to,

$$\begin{array}{ll} \frac{d}{dt}\left(\|v_t\|_H^2 + v^T M^{(b)} v\right) & = -2(e_1^T v_t)^T b_1 d_{1;1} v\,(1+\tau) + 2\tau (e_1^T v_t)^T \left((\gamma e_1^T v_t - g_0\right) \\ & \quad +2(e_m^T v_t)^T b_m d_{1;m} v\,(1+\tau) + 2\tau (e_m^T v_t)^T \left((\gamma e_m^T v_t - g_L\right) . \end{array}$$

Chosing $\tau = -1$ leads to

$$\frac{d}{dt}\left(\|v_t\|_H^2 + v^T M^{(b)} v\right) = -2(e_1^T v_t)^T \left((\gamma e_1^T v_t - g_0\right) - (e_m^T v_t)^T \left((\gamma e_m^T v_t - g_L\right) . \tag{31}$$

Equation (31) exactly mimics (28)

3.2 1D compatible SBP operators

To approximate the dispersive fourth derivative term u_{xxxx} the following definition (first introduced in [36]) is central to the present study:

Definition 3.4 *A difference operator* $D_4 = H^{-1}\left(N - e_1 d_{3;1} + e_m d_{3;m} + d_{1;1}^T d_{2;1} - d_{1;m}^T d_{2;m}\right)$ *approximating* $\partial^4/\partial\, x^4$, *using a* $2p$*th-order accurate narrow-stencil in the interior, is said to be a* $2p$*th-order diagonal-norm fourth-derivative SBP operator if the diagonal matrix* H *defines a discrete norm,* $N = N^T \geq 0$, $d_{1;1}v \simeq u_x|^0$, $d_{1;m}v \simeq u_x|^L$, $d_{2;1}v \simeq u_{xx}|^0$, $d_{2;m}v \simeq u_{xx}|^L$, *and* $d_{3;1}v \simeq u_{xxx}|^0$, $d_{3;m}v \simeq u_{xxx}|^L$ *are finite difference approximations of the first, second and third derivatives at the left and right boundary points.*

For the reader interested in the construction of the SBP operators, we refer to [35, 36].

For 1D problems there are two necessary compatibility conditions (for stability reasons) when combining SBP operators of various derivative orders: 1) They should be based on the same discrete norm H, and 2) they should have the same boundary derivative approximations. In [34] we introduced a third compatibility condition between the first- and second-derivative SBP operators, necessary for multi-D problems involving mixed second derivative terms. This condition drops since we are treating a 1D problem. The following new definition is necessary in the coming stability analysis of the semi-discrete NS model.

Definition 3.5 *Let* $D_2^{(b)}$ *and* D_4 *be* $2p$*th-order diagonal-norm second- and fourth-derivative SBP operators. If* $D_2^{(b)}$ *and* D_4 *are based on the same discrete norm* H, *and boundary derivative operators* $d_{1;1}$ *and* $d_{1;m}$, *they are called 1D compatible.*

The D_4 SBP operators derived in [36] were not *1D compatible* with the $D_2^{(b)}$ SBP operators derived in [35] since they did not share the same norm. In the present study completely new fourth-derivative SBP operators are derived, *1D compatible* with the second-derivative SBP operators derived in [35]. The new SBP operators (constructed to be second-, fourth- and sixth-order accurate) are presented in the Appendix.

The following two definitions (first presented in [36]) are central to the present study:

Definition 3.6 *Let* $e_{(s)}^{(q)} = H\mathbf{x}^{q-s} - H\, D_s\mathbf{x}^q$ *be the* q*th-order error vector for the* s*th-derivative SBP operator (in the present study* $s = 1, 4$*). Here* $\mathbf{x}^{q-s}$ *is defined as* $\mathbf{0}$ *when* $q < s$, *where* $\mathbf{0}$ *and* $\mathbf{x}^q$ *are defined in (21).*

Definition 3.7 *The discrete l^2-norm of the error $e_{(s)}^{(q)}$ is defined as $\|e_{(s)}^{(q)}\|_h^2 = h\,(e_{(s)}^{(q)})^T e_{(s)}^{(q)}$. We say that the boundary closure of a diagonal-norm SBP operator D_s is rth-order accurate if $\|e_{(s)}^{(q)})\|_h^2 = 0$ for $q = 0, \ldots, s + r - 1$.*

If $r < 1$ (in Definition 3.7) the difference approximation is inconsistent at the boundaries.

It is important to understand that the formal boundary accuracy alone does not dictate the expected convergence rate of the numerical approximation of the underlying IBVP. An inconsistent boundary closure can still imply convergence. The expected convergence rate can be shown by a careful error analysis (not presented in the present study). This is summarised in the following remark:

Remark In [43] it is shown that a pointwise stable approximation of an IBVP involving derivatives up to order s yields a convergence rate of order $r+s$, where r is the order of accuracy at the boundaries. Hence, if the highest spatial derivative order is 4 in the PDE the boundary stencils may be 4 orders less accurate than the inner stencil. Thus, it is enough to employ zeroth-order accurate boundary stencils in the case of fourth-order accurate SBP operators and still obtain fourth order convergence, assuming pointwise stability. The second-order accurate fourth-derivative SBP operator is inconsistent with $r = -2$, and still this is enough to obtain second order convergence. The fourth-order accurate SBP operator has $r = 0$ (and we expect fourth order convergence). The sixth-order accurate SBP operator has $r = 1$ and we expect fifth order convergence, assuming that we employ a time discretisation that is accurate enough. In the present study we do not prove pointwise stability, but the numerical convergence study presented in Figure 3 indicates that we obtain the convergence expected when assuming pointwise stability.

To prove semi-discrete stability for the NS model (using the energy method) subjected to DN BC, the following two (new) lemmas are necessary,

Lemma 3.8 *The dissipative part N of a diagonal-norm fourth-derivative SBP operator has the following property:*

$$v^T N v = h\,\alpha_2 \left((d_{2;1} v)^2 + (d_{2;m} v)^2 \right) + v^T \tilde{N}_2 v \ , \tag{32}$$

where $\tilde{N}_2$ is symmetric and positive semi-definite, and α_2 a positive constant, independent of h.

Lemma 3.9 *The dissipative part N of a diagonal-norm fourth-derivative SBP operator has the following property:*

$$v^T N v = h^3\,\alpha_3 \left((d_{3;1} v)^2 + (d_{3;m} v)^2 \right) + v^T \tilde{N}_3 v \ , \tag{33}$$

$\alpha_2^{(2nd)}$	$\alpha_2^{(4th)}$	$\alpha_2^{(6th)}$	$\alpha_3^{(2nd)}$	$\alpha_3^{(4th)}$	$\alpha_3^{(6th)}$
1.250	0.505	0.325	0.4	0.928	0.158

Table 1: $\alpha_{2,3}$ in Lemmas 3.8 and 3.9 for the second-, fourth- and sixth-order accurate diagonal-norm fourth-derivative SBP operators.

where $\tilde{N}_3$ is symmetric and positive semi-definite, and α_3 a positive constant, independent of h.

We will refer to Lemmas 3.8 and 3.9 as the *borrowing condition*. (An analogue *borrowing condition* was introduced, for second-derivative SBP operators, in [30, 31], utilised in the stability proof for the SBP-SAT approximation of the second order wave equation subjected to Dirichlet BC.) The values of α_2 and α_3 were computed numerically, for the second-, fourth- and sixth-order accurate finite difference SBP discretizations. The results are presented in Table 1.

4 Stability analysis

In this section we present the SBP-SAT approximations of (10) combined with the two types of BC presented in Section 2.1. In the following we assume the same initial conditions $v = f_1$ and $v_t = f_2$ as in the continuous case.

4.1 Characteristic boundary conditions

The semi-discrete approximation of the Char BC (14) can be written,

$$\begin{aligned} &d_{1;1}v_t - d_{2;1}v = g_0^{(1)}(t), \quad \sigma_1 e_1^T v_t - b_1 d_{2;1}v + d_{3;1}v = g_0^{(2)}(t), \qquad t \geq 0\ , \\ &d_{1;m}v_t + d_{2;m}v = g_L^{(1)}(t), \quad \sigma_m e_m^T v_t + b_m d_{2;m}v - d_{3;m}v = g_L^{(2)}(t), \quad t \geq 0\ . \end{aligned} \tag{34}$$

Here $\sigma_{1,m}$ are defined as the local value of σ (13) at the left and right boundary, respectively.

The SBP-SAT approximation of (10) subjected to the Char BC (14) is

given by,

$$\begin{aligned} v_{tt} = D_2^{(b)} v - D_4 v &+ \tau_l^{(1)} H^{-1} d_{1;1}^T \left\{ d_{1;1} v_t - d_{2;1} v - g_0^{(1)} \right\} \\ &+ \tau_l^{(2)} H^{-1} e_1 \left\{ \sigma_1 e_1^T v_t - b_1 d_{1;1} v + d_{3;1} v - g_0^{(2)} \right\} \\ &+ \tau_r^{(1)} H^{-1} d_{1;m}^T \left\{ d_{1;m} v_t + d_{2;m} v - g_L^{(1)} \right\} \\ &+ \tau_r^{(2)} H^{-1} e_m \left\{ \sigma_m e_m^T v_t + b_m d_{1;m} v - d_{3;m} v - g_L^{(2)} \right\} \ . \end{aligned} \tag{35}$$

Here $\tau_{l,\,r}^{(1,\,2)}$ are the penalty parameters chosen to obtain a semi-discrete energy estimate. Multiplying (35) by $v_t^T H$, and adding the transpose (assuming homogeneous BC, and that $D_2^{(b)}$ and D_4 are 1D compatible) leads to,

$$\begin{aligned} \frac{d}{dt} E_H = &- b_1 (v_t)_1 d_{1;1} v + 2 (v_t)_1 d_{3;1} v - d_{1;1} v_t d_{2;1} v \\ &+ b_m (v_t)_m d_{1;m} v - 2 (v_t)_m d_{3;m} v + d_{1;m} v_t d_{2;m} v \\ &+ 2\tau_l^{(1)} (d_{1;1} v_t)^2 - 2\tau_l^{(1)} d_{1;1} v_t d_{2;1} v \\ &+ 2\tau_l^{(2)} \sigma_1 (v_t)_1^2 - 2\tau_l^{(2)} b_1 (v_t)_1 d_{1;1} v + 2\tau_l^{(2)} (v_t)_1 d_{3;1} v \\ &+ 2\tau_r^{(1)} (d_{1;m} v_t)^2 + 2\tau_r^{(1)} d_{1;m} v_t d_{2;m} v \\ &+ 2\tau_r^{(2)} \sigma_m (v_t)_m^2 + 2\tau_r^{(2)} b_m (v_t)_m d_{1;m} v - 2\tau_r^{(2)} (v_t)_m d_{3;m} v \ . \end{aligned} \tag{36}$$

E_H is given by,

$$E_H = \|v_t\|_H^2 + v^T M^{(b)} v + v^T N v \ , \tag{37}$$

and is the semi-discrete counterpart to $E^{(c)}$, given by Eq. 20.

The following lemma is one of the main results in the present study:

Lemma 4.1 *Eq. 35 is stable if $D_2^{(b)}$ and D_4 are 1D compatible second- and fourth-derivative SBP operators, and $\tau_l^{(1)} = \tau_r^{(1)} = \tau_l^{(2)} = \tau_r^{(2)} = -1$ hold.*

Proof By choosing $\tau_l^{(1)} = \tau_r^{(1)} = \tau_l^{(2)} = \tau_r^{(2)} = -1$, (36) can be formulated into,

$$\frac{d}{dt} E_H = -(d_{1;1} v_t)^2 - \sigma_1 (v_t)_1^2 - (d_{1;m} v_t)^2 - \sigma_m (v_t)_m^2 \ ,$$

which exactly mimics the corresponding continuous energy estimate (15). □

4.2 Dirichlet-Neumann boundary conditions

The semi-discrete approximation of the DN BC given by (16) can be written,

$$\begin{aligned} e_1^T v &= g_0^{(1)}(t), \quad d_{1;1} v = g_0^{(2)}(t), \quad t \geq 0 \, , \\ e_m^T v &= g_L^{(1)}(t), \quad d_{1;m} v = g_L^{(2)}(t), \quad t \geq 0 \, . \end{aligned} \tag{38}$$

The SBP-SAT approximation of (10) subjected to the DN BC (16) is given by,

$$\begin{aligned} v_{tt} = D_2^{(b)} v - D_4 v &+ H^{-1}(d_{3;1} - b_1 d_{1;1} - \tau_l^{(1)} e_1^T)^T \left\{ e_1^T v - g_0^{(1)} \right\} \\ &- H^{-1}(d_{2;1} + \tau_l^{(2)} d_{1;1})^T \left\{ d_{1;1} v - g_0^{(2)} \right\} \\ &- H^{-1}(d_{3;m} - b_m d_{1;m} + \tau_r^{(1)} e_m^T)^T \left\{ e_m^T v - g_L^{(1)} \right\} \\ &+ H^{-1}(d_{2;m} - \tau_r^{(2)} d_{1;m})^T \left\{ d_{1;m} v - g_L^{(2)} \right\} \, . \end{aligned} \tag{39}$$

Here $\tau_{l,\,r}^{(1,\,2)}$ are again the penalty parameters chosen to obtain a semi-discrete energy estimate.

Remark In (39) we have two penalty terms on each boundary. At the left boundary, for example, we can think of $(d_{3;1} - b_1 d_{1;1} - \tau_l^{(1)} e_1^T)^T$ and $(d_{2;1} + \tau_l^{(2)} d_{1;1})^T$ as penalty vectors. It is perhaps not obvious why we have made this particular Ansatz of the penalty vectors. The short answer is that this will lead to stability with a specific tuning of the penalty parameters $\tau_{l,\,r}^{(1,\,2)}$, with the aid of Lemmas 3.8 and 3.9.

The following lemma is one of the main results in the present study:

Lemma 4.2 *Eq. 39 is stable if $D_2^{(b)}$ and D_4 are 1D compatible second- and fourth-derivative SBP operators, and $\tau_{l,\,r}^{(1)} > \frac{2}{\alpha_2 h}$, $\tau_l^{(2)} > \frac{2}{\alpha_3 h^3} + \frac{\alpha_2 b_1^2}{\tau_l^{(1)} \alpha_2 - \frac{2}{h}}$, $\tau_r^{(2)} > \frac{2}{\alpha_3 h^3} + \frac{\alpha_2 b_m^2}{\tau_r^{(1)} \alpha_2 - \frac{2}{h}}$ hold.*

Proof Multiplying the homogeneous version of (39) by $v_t^T H$, and adding the transpose leads to,

$$\frac{d}{dt} \left(\tilde{E}_H + w_1^T A_1 w_1 + w_m^T A_m w_m \right) = 0 \, ,$$

where $\tilde{E}_H$ is given by,

$$\tilde{E}_H = \| v_t \|_H^2 + v^T M^{(b)} v + \frac{1}{2} v^T \tilde{N}_2 v + \frac{1}{2} v^T \tilde{N}_3 v \, ,$$

and

$$w_1 = \begin{bmatrix} v_1 \\ d_{1;1}v \\ d_{2;1}v \\ d_{3;1}v \end{bmatrix}, \qquad A_1 = \begin{bmatrix} \tau_l^{(1)} & b_1 & 0 & -1 \\ b_1 & \tau_l^{(2)} & 1 & 0 \\ 0 & 1 & \frac{h}{2}\alpha_2 & 0 \\ -1 & 0 & 0 & \frac{h^3}{2}\alpha_3 \end{bmatrix},$$

$$w_m = \begin{bmatrix} v_m \\ d_{1;m}v \\ d_{2;m}v \\ d_{3;m}v \end{bmatrix}, \quad A_m = \begin{bmatrix} \tau_r^{(1)} & -b_m & 0 & 1 \\ -b_m & \tau_r^{(2)} & -1 & 0 \\ 0 & -1 & \frac{h}{2}\alpha_2 & 0 \\ 1 & 0 & 0 & \frac{h^3}{2}\alpha_3 \end{bmatrix}.$$

Here we have assumed that $D_2^{(b)}$ and D_4 are 1D compatible and further used Lemmas 3.8 and 3.9 to obtain the following,

$$v^T N v = \frac{1}{2}v^T N v + \frac{1}{2}v^T N v = \frac{1}{2}\left(h\,\alpha_2\left((d_{2;1}v)^2 + (d_{2;m}v)^2\right) + v^T\tilde{N}_2 v\right) + \frac{1}{2}\left(h^3\,\alpha_3\left((d_{3;1}v)^2 + (d_{3;m}v)^2\right) + v^T\tilde{N}_3 v\right).$$

The values of $\alpha_{2,3}$ are presented in Table 1. By choosing $\tau_{l,r}^{(1)} > \frac{2}{\alpha_2 h}$, $\tau_l^{(2)} > \frac{2}{\alpha_3 h^3} + \frac{\alpha_2 b_1^2}{\tau_l^{(1)}\alpha_2 - \frac{2}{h}}$, $\tau_r^{(2)} > \frac{2}{\alpha_3 h^3} + \frac{\alpha_2 b_m^2}{\tau_r^{(1)}\alpha_2 - \frac{2}{h}}$, $A_{1,m}$ are positive semi-definite and $\left(\tilde{E}_H + w_1^T A_1 w_1 + w_m^T A_m w_m\right)$ defines a semi-norm. □

Remark By studying the boundary matrices $A_{1,m}$ in the above proof, the necessity of Lemmas 3.8 and 3.9 are understood. Without the coefficients involving $\alpha_{2,3}$ on the lower diagonal in $A_{1,m}$ the matrices can not be positive semi-definite, regardless of the four penalty parameters $\tau_{l,r}^{(1,2)}$.

5 Time integration

The time-discretisation of the semi-discrete SBP-SAT models presented in Section 4 will be analysed. The SBP-SAT approximations (35) and (39) can be written as the following second order ODE system (with m unknowns),

$$\begin{aligned} &v_{tt} = -H^{-1}Av - H^{-1}Cv_t + H^{-1}G(t)\,, && t \geq 0\,, \\ &v(0) = f_1\,, v_t(0) = f_2\,, && t = 0\,, \end{aligned} \tag{40}$$

where f_1 and f_2 are initial data and $G(t)$ is a known time-dependent vector function. H, A and C are $m \times m$ matrices. The following definition and lemma are introduced to simplify the coming stability analysis:

Definition 5.1 *If $A = A^T \geq 0$ and H defines a discrete norm,*

$$E_H = \|v_t\|_H^2 + v^T A v \ ,$$

is non-negative and $E_H = 0$ implies that $v_t = 0$. Hence, E_H defines a semi-norm.

Lemma 5.2 *If $A = A^T \geq 0$, $C + C^T \geq 0$ and H defines a discrete norm, the ODE system (40) is stable.*

Proof Multiplying the homogeneous version of (40) by $v_t^T H$ and adding the transpose leads to,

$$\begin{aligned} v_t^T H v_{tt} + v_{tt}^T H^T v_t &= -v_t^T A v - v^T A^T v_t - v_t^T C v_t - v_t^T C^T v_t \\ &\Downarrow \\ \frac{d}{dt} E_H &= -v_t^T (C + C^T) v_t \ . \end{aligned}$$

In the last step we use the fact that $A = A^T$ is positive semi-definite and H defines a discrete norm, so that $E_H = \|v_t\|_H^2 + v^T A v$ defines a semi-norm (see Definition 5.1). If $C + C^T \geq 0$ the time-growth of E_H is non-positive. □

The semi-discrete SBP-SAT models (35) and (39) fulfill (40) such that the conditions in Lemma 5.2 hold.

Remark The second order ODE system (40) allows for linear time-growth if A is positive semi-definite (even with homogeneous data, i.e., $G(t) = 0$). This however requires two things: 1) The zero eigenvalue is a double root, and 2) $f_2 \neq 0$. This is consistent with the characteristics of the underlying IBVP, which also allows for linear time-growth, when there are double zero roots (see [36]). If A is negative definite, there can be no time-growth. Nevertheless, the ODE system is stable if A is negative semi-definite. (See [16] for details regarding the various types of stability definitions.)

The first model (35) with Char BC has both a first and a second derivative in time while the problem (39) with DN BC has only a second derivative in time. For hyperbolic problems involving both first and second derivatives in time one often (see for example [15, 33]) rewrites the system on first order form by introducing an auxiliary variable, and time-integrates using an explicit Runge-Kutta method. In this section we start by showing why this approach for the NS model leads to a very restrictive CFL condition, when temporal first derivative (i.e., damping) terms are present. An alternative time-integration technique based on central finite differences, with a much less restrictive CFL is introduced later in this section.

5.1 Explicit Runge-Kutta methods

When temporal first derivative (i.e., damping) terms are present in the NS model, an explicit Runge-Kutta method has severe CFL restrictions (see (42)). This has been verified in numerical computations of (35) using the fourth order Runge-Kutta method, see Figure 5 and Table 3. We will now analyse this behaviour in some detail, for a simplified scalar problem.

Starting from (40) and assuming periodic BC (here referred to as the *Cauchy* problem), we can simultaneously diagonalise $H^{-1}C$ and $H^{-1}A$, and for each Fourier mode derive a scalar ODE of the form

$$\begin{aligned} &y_{tt} = -ay - cy_t + g(t) \ , \\ &y(0) = f_1 \ , y_t(0) = f_2 \ , \end{aligned} \tag{41}$$

where f_1 and f_2 are initial data, and $g(t)$ known time-dependent data. The Cauchy problem shares much of the characteristics with the corresponding IBVP, when it comes to the expected CFL condition. This motivates the study of a scalar ODE system, concerning the expected CFL condition for a similar ODE system.

The solution to the homogeneous version of (41) can be written $y(t) = p_1 e^{r_1 t} + p_2 e^{r_2 t}$, where $p_{1,2}$ are determined by the initial data and $r_{1,2}$ by solving the characteristic equation $r^2 + cr + a = 0$, i.e., $r_{1,2} = -\frac{c \pm \sqrt{c^2 - 4a}}{4}$. To time-integrate (41) using the standard fourth order Runge-Kutta method, an auxiliary variable $w = y_t$ is introduced, to obtain an ODE-system with only first derivative terms. The CFL restriction is given by $k \cdot \max |r_{1,2}| \leq 2.8$, where k denotes the time-step.

We now consider the case $c = \bar{c} \cdot h^{-3}$ and $a = \bar{a} \cdot h^{-4}$, in (41), where $\bar{a}$ and $\bar{c}$ are constants independent of h. This particular choice of h-scaling is motivated by the ODE system (35). The h-scalings (where h is the spatial grid-size) of $H^{-1}A$ and $H^{-1}C$ are given by h^{-4} and h^{-3} respectively. Hence, for sufficiently small grid-spacing h, $\max |r_{1,2}| \simeq \bar{c} \cdot h^{-3}$, and the CFL condition is given by,

$$k \leq \frac{2.8\, h^3}{\bar{c}} \ . \tag{42}$$

To summarise, we expect a very severe time-step restriction employing an explicit Runge-Kutta method to solve (35). In fact, this is exactly what we found by extensive numerical testing, see Figure 5. For this reason we will use a central finite difference method to time-integrate (35). This will be analysed in the coming section.

For the other problem (39), the first derivative component $H^{-1}Cv_t$ is not present. This leads to the CFL condition given by,

$$k \leq \frac{2.8\, h^2}{\bar{a}} \ , \tag{43}$$

which agrees with the physical requirement (see the remark below). In the present study we will therefore employ a central finite difference method that yields a CFL similar to (43) also when temporal first derivatives are present in the ODE model. This will be analysed in the coming section.

Remark The dispersion relation (7) shows that the wave speed scales as κ. Hence, the physics (in general) require $k \simeq h^2$, to capture the high-frequency part of the solution. An efficient time-integrator should therefore allow a time-step similar to what is required from the physical consideration. This is why we refer to (42) as restrictive. Ideally we want a CFL similar to (43). The soliton solution (8) on the other hand behaves like a non-dispersive solution, and for this particular solution we could expect that a high order accurate semi-implicit time-integrator could have been beneficial, choosing the time-step from an accuracy perspective, i.e., $k \simeq h$. This is something we hope to address in a coming study.

5.2 Finite difference in time

An alternative time-discretisation technique (compared to explicit Runge-Kutta methods) is to employ a central finite difference (Cdiff) method. For the case with both first and second derivative terms present in (40), higher than second-order accurate approximations are far from trivial and will not be investigated in the present study. However, we will see that the CFL condition with the central finite difference method will yield an estimate similar to (43). Hence, a second-order accurate finite difference approximation will lead to fourth order convergence due to the CFL condition (assuming that the problem is stable).

Let $t_n = n\,k$, $n = 0, 1, \ldots$ denote the discrete time-levels, where k is the time-step, and introduce the notation $v^n = v(t_n)$. We introduce the following second-order accurate finite difference approximations, $D_0 v^n = \frac{v^{n+1}-v^{n-1}}{2\,k}$ and $D_+D_- v^n = \frac{v^{n+1}-2v^n+v^{n-1}}{k^2}$, of the first and second derivatives in (40). The first derivative in the initial step can be approximated by $D_+ v^0 = \frac{v^1-v^0}{k}$. By replacing the first derivative in the initial step by $D_+ v^0$ and inserting appropriate Taylor expansions we obtain the following second-order accurate approximation of (40),

$$
\begin{aligned}
&\tfrac{v^{n+1}-2v^n+v^{n-1}}{k^2} = -H^{-1}Av^n - H^{-1}C\tfrac{v^{n+1}-H^{-1}v^{n-1}}{2\,k} + H^{-1}G^n\,, n = 1, \ldots \\
&v^0 = f_1,\ \tfrac{v^1-f_1}{k} = f_2 - \tfrac{k}{2}H^{-1}(Af_1 + Cf_2 - G^0)\ .
\end{aligned}
\tag{44}
$$

Remark Notice that the Cdiff time-integrator (44) is truly explicit only if the inverse of H is a diagonal matrix (except perhaps at a small and constant number of boundary points). This is trivially true if the discrete norm H is diagonal. This is another motivation for using diagonal-norm SBP operators.

The following definition is necessary for the stability analysis of the fully discrete finite difference approximation:

Definition 5.3 *If $A = A^T$ is positive semi definite and $P = P^T$ is positive definite,*

$$E^n = (D_+ v^n, P D_+ v^n) + \left(\frac{v^{n+1} + v^n}{2}, A \frac{v^{n+1} + v^n}{2} \right) ,$$

is non-negative for all v^n, v^{n+1} and $E^n = 0$ implies $v^{n+1} = v^n = c$, where c is a constant. Then E^n defines a semi-norm. Here we define $(D_+ v^n, P D_+ v^n) = (D_+ v^n)^T P (D_+ v^n)$.

The following Lemma is central to the coming stability analysis:

Lemma 5.4 *The matrix $(H - \frac{k^2}{4} A)$ is positive definite if the time-step is chosen as $k^2 < 4 \frac{\min(h_i)}{\max(a_i)}$, where h_i and a_i are the eigenvalues of H and A, respectively.*

Proof The eigenvalue problem $(H - \frac{k^2}{4} A)X = \lambda X$ can be written $(I - \frac{k^2}{4} H^{-1} A)X = \lambda H^{-1} X$. Since A and H are symmetric and H is positive definite, the eigenvalues of $H^{-1}A$ are given by $\frac{a_i}{h_i}$, where a_i and h_i are the eigenvalues of A and H respectively (since this is a generalized symmetric-definite eigenvalue problem). Hence, the eigenvalues to $I - \frac{k^2}{4} H^{-1} A$ are given by $1 - \frac{k^2}{4} \frac{a_i}{h_i}$. To guarantee that λ_i is positive, it is enough to require that $1 - \frac{k^2}{4} \frac{a_i}{h_i} > 0$ (since $h_i > 0$). By choosing $k^2 < 4 \frac{\min(h_i)}{\max(a_i)}$ we guarantee that $\lambda_i > 0$. $\square$

We will make use of the following two relations,

$$\begin{aligned} &\left(v^{n+1} - v^{n-1}, H(v^{n+1} + v^{n-1})\right) + \left(v^{n+1} + v^{n-1}, H(v^{n+1} - v^{n-1})\right) \\ &= 2\left(v^{n+1} - v^n, H(v^{n+1} - v^n)\right) + 4\left(v^{n+1}, H v^n\right) \\ &- 2\left(v^n - v^{n-1}, H(v^n - v^{n-1})\right) - 4\left(v^n, H v^{n-1}\right) , \end{aligned} \tag{45}$$

$$\begin{aligned} (v^{n+1}, A v^n) + \left(v^n, A v^{n+1}\right) = &\frac{1}{2}\left(v^{n+1} + v^n, A(v^{n+1} + v^n)\right) \\ &- \frac{1}{2}\left(v^{n+1} - v^n, A(v^{n+1} - v^n)\right) . \end{aligned} \tag{46}$$

The following Theorem is one of the main results of the study:

Theorem 5.5 *Eq. 44 is a stable approximation of (40) if Lemmas 5.2 and 5.4 apply.*

Proof Start by setting $G = 0$ (since G does not affect stability), and rewriting (44) to,

$$H\frac{v^{n+1}+v^{n-1}}{k^2} = -\left(A - 2\frac{H}{k^2}\right)v^n - C\frac{v^{n+1}-v^{n-1}}{2\,k}\ .$$

Multiplying by $\left(\frac{v^{n+1}-v^{n-1}}{2\,k}\right)^T$ leads to,

$$\begin{aligned}\left(\frac{v^{n+1}-v^{n-1}}{2\,k}, H\left(\frac{v^{n+1}+v^{n-1}}{k^2}\right)\right) &= -\left(\frac{v^{n+1}-v^{n-1}}{2\,k}, \left(A - 2\frac{H}{k^2}\right)v^n\right)\\ &\quad -\left(\frac{v^{n+1}-v^{n-1}}{2\,k}, C\left(\frac{v^{n+1}-v^{n-1}}{2\,k}\right)\right)\ .\end{aligned}$$

By adding the transpose, assuming $H = H^T$ and $A = A^T$, we obtain

$$\begin{aligned}&\left(\frac{v^{n+1}-v^{n-1}}{2\,k}, H\left(\frac{v^{n+1}+v^{n-1}}{k^2}\right)\right) + \left(\frac{v^{n+1}+v^{n-1}}{k^2}, H\left(\frac{v^{n+1}-v^{n-1}}{2\,k}\right)\right)\\ &= -\left(\frac{v^{n+1}-v^{n-1}}{2\,k}, \left(A - 2\frac{H}{k^2}\right)v^n\right) - \left(v^n, \left(A - 2\frac{H}{k^2}\right)\frac{v^{n+1}-v^{n-1}}{2\,k}\right)\\ &- \left(\frac{v^{n+1}-v^{n-1}}{2\,k}, (C+C^T)\left(\frac{v^{n+1}-v^{n-1}}{2\,k}\right)\right)\ .\end{aligned}$$

By using the two relations (45) and (46) we obtain,

$$\begin{aligned}&\left(D_+v^n, \left(H - \frac{k^2}{4}A\right)D_+v^n\right) + \left(\frac{v^{n+1}+v^n}{2}, A\frac{v^{n+1}+v^n}{2}\right) =\\ &\left(D_+v^{n-1}, \left(H - \frac{k^2}{4}A\right)D_+v^{n-1}\right) + \left(\frac{v^{n-1}+v^n}{2}, A\frac{v^{n-1}+v^n}{2}\right)\\ &- k\left(D_0v^n, (C+C^T)D_0v^n\right)\ .\end{aligned}$$

If Lemma 5.4 holds this can be written

$$\frac{E^{n+1}-E^n}{k} = -\left(D_0v^n, (C+C^T)D_0v^n\right)\ ,$$

where the semi-norm E^n is given by Definition 5.3, where $P = \left(H - \frac{k^2}{4}A\right)$ is positive definite. If $(C+C^T) \geq 0$ we conclude that the time-growth of the semi-norm E^n is bounded. $\square$

BC / Accuracy	Second	Fourth	Sixth
Characteristic	0.4831	0.3771	0.2755
Dirichlet-Neumann	0.4003	0.2496	0.1329

Table 2: Accurate numerical estimates of $\Gamma = \frac{k}{h^2}$ using the central finite difference time-integrator (44) for the two different types of BC and accuracy of the SBP operator.

BC / Accuracy	Second	Fourth	Sixth
Characteristic	0.3189	0.1428	0.0684
Dirichlet-Neumann	0.5665	0.3530	0.1880

Table 3: Accurate numerical estimates of $\Gamma = \frac{k}{h^2}$ for DN BC and $\Gamma = \frac{k}{h^3}$ for the Char BC for different orders of accuracy, using the fourth order Runge-Kutta method.

The fully discrete energy estimate in Theorem 5.5 is completely analogous to the semi-discrete energy estimate in Lemma 5.2, assuming that we fulfill the CFL condition, $k^2 < 4\frac{\min(h_i)}{\max(a_i)}$. For the semi-discrete SBP-SAT models presented in Section 4, h_i is proportional to h and a_i is proportional to h^{-3}, implying that the CFL conditions are given by (43) for some constant $\bar{a} > 0$. (The value of $\bar{a}$ depends on the type of BC and the specific SBP operator employed.) In Table 2 we present accurate numerical estimates of the CFL, i.e., $\Gamma = \frac{k}{h^2}$ for the different types of BC and accuracy of the SBP operator.

For completeness we also present the corresponding CFL values using the fourth order Runge-Kutta method in Table 3. Here the CFL is given by $\Gamma = \frac{k}{h^2}$ using DN BC and $\Gamma = \frac{k}{h^3}$ using Char BC, in support of the analytical estimates presented in (42) and (43).

Remark The present Cdiff method (44) has the attractive property of being time-symmetric, when only (temporal) second derivatives are present in the model. Explicit Runge-Kutta methods do not share this property. Time-symmetric integration implies discrete energy-conservation, which is an attractive property for energy-conservative models. The NS model describes an adiabatic and reversible phenomenon [44, 18] and should thus ideally be time-integrated with a time-symmetric ODE solver such as (44).

6 Computations

In this section the accuracy, stability and efficiency properties of the SBP-SAT approximations analysed in Section 4 will be verified. The SBP-SAT ap-

proximations are time-integrated with both the second-order accurate Cdiff approximation, given by (44), and the fourth order Runge-Kutta (RK4) time-integrator (to validate the claim that (44) is a more efficient time-integrator than RK4 for the semi-discrete NS model). The time-steps are adopted to the CFL limits for each method and order of accuracy, as presented in Tables 2 and 3. We also investigate the robustness of the numerical scheme and the robustness of the soliton behaviour.

6.1 Convergence and efficiency study

The convergence rate is calculated as

$$q = \log_{10}\left(\frac{\|u - v^{(m_1)}\|_h}{\|u - v^{(m_2)}\|_h}\right) / \log_{10}\left(\frac{m_1}{m_2}\right) \ , \tag{47}$$

where u is the analytic solution, and $v^{(m_1)}$ the corresponding numerical solution with m_1 unknowns. $e = \|u - v^{(m_1)}\|_h$ is the discrete l_2 norm of the error.

The accuracy properties of the SBP-SAT approximations (35) and (39) are tested against the analytic soliton solution (8) with $\beta = 0.80$ (see Figure 1). The analytic solution is used as initial and boundary data. The initial soliton is centered at $x = 0$. The spatial domain is given by $x \in [-15, 15]$. The numerical approximations are integrated to $t = 18.75$ when the soliton is centered at the right boundary at $x = 15$. We compare the results using the second order accurate Cdiff and the RK4 time-integrators.

In Figure 3 we present the convergence results using the second- fourth- and sixth-order accurate SBP-SAT approximations of (39) and (35). The Cdiff time-integrator is second-order accurate, but due to the CFL condition we expect at least fourth order convergence for the fourth- and sixth-order accurate SBP-SAT approximations. The conclusions from Figure 3 are the following: 1) the leading truncation error comes from the spatial discretisation, 2) we obtain the convergence rates expected when assuming pointwise stability (see the remark in Section 3), and 3) a completely grid-converged solution is reached for l_2-errors of approximately $e \simeq 10^{-7}$.

Remark A similar convergence study with the l_∞ (instead of the l_2) norm showed the same convergence behaviour as presented in Figure 3.

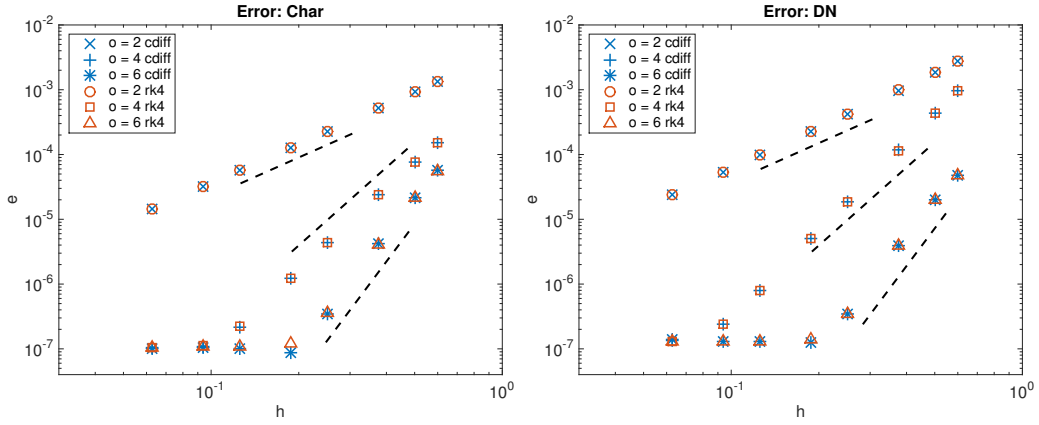

Figure 3: The l_2-error (e) as a function of h for the second-, fourth- and sixth-order accurate SBP-SAT approximations with both Char BC (left) and DN BC (right). We show the results with both the Cdiff and RK4 time-integrators. The dashed lines represent second-, fourth- and sixth-order convergence. The Cdiff and RK4 yield almost identical errors. Grid-convergence is reached at $e \simeq 10^{-7}$.

It is well known that higher order finite difference methods capture hyperbolic wave propagation more efficiently (see the pioneering paper by Kreiss and Oliger [24]). We investigate if this holds also for the NS model (assuming smooth solutions). In Figure 4 we compare the efficiency for different orders of accuracy. We plot the $l_2 - error$ as a function of runtime in seconds. The efficiency study is performed for both DN and Char BC. This can be compared to the convergence analysis found in Figure 3. The results clearly support the claim that an efficient solution of the NS model requires a stable and higher-order accurate numerical method.

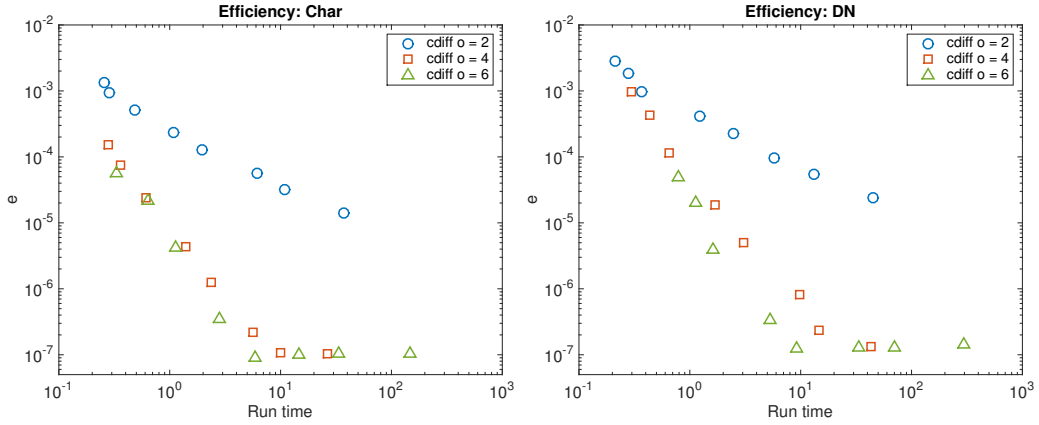

Figure 4: The l_2-error (e) as a function of runtime for the second-, fourth- and sixth-order accurate SBP-SAT approximations, using the Cdiff time-integrator. We present the results for both Char BC (left) and DN BC (right). Grid-convergence is reached at $e \simeq 10^{-7}$.

In Figure 5 we present both a convergence and an efficiency comparison between the Cdiff and RK4 time-integrators. We present the results using the fourth-order accurate SBP-SAT approximations for both DN and Char BC. The Cdiff is clearly much more efficient than RK4 for the case with Char BC, in support of the numerical CFL estimates presented in Tables 2 and 3. The convergence plot shows that the spatial discretisation dominates the l_2 errors.

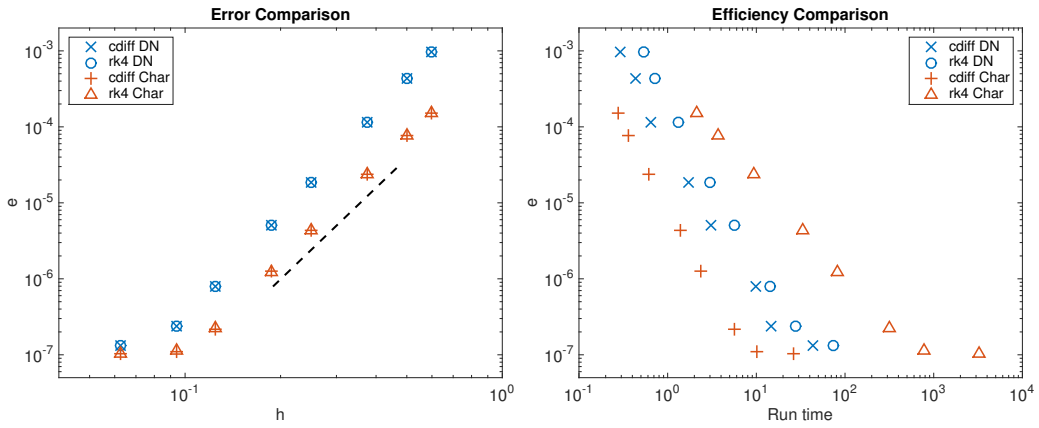

Figure 5: The l_2-error as a function of both h and runtime for the fourth-order accurate SBP-SAT approximations with both DN and Char BC. We compare the results using the RK4 and Cdiff time-integrators. The CFL conditions are presented in Tables 2 and 3. The dotted line represents fourth order convergence.

6.2 Long-time behaviour of soliton propagation

In the study of soliton problems it is very important to consider the long-time behaviour (i.e., stability and accuracy) of the numerical approximation, since the solitons are typically propagated over vast distances. This is especially true for the proposed NS model [17], for the study of nerve pulses (solitons) in the axon membrane. (The nerve axon is typically very long compared to the width of the soliton pulse.)

In Figure 6 we compare the results using the second- and sixth-order accurate SBP-SAT approximations of (39). Here the stability and accuracy properties of the numerical method for long-time integration of an analytic soliton (8) with $\beta = 0.65$ are verified. The analytic soliton, centered at $x = 40$, is used as initial data. The spatial domain is given by $x \in [0, 500]$. The numerical approximations are integrated (using Cdiff) to $t = 600$, when the analytic soliton is centered at $x = 430$. The results are compared for two different number of grid-points, $m = 250$ and $m = 500$. The dispersion error is clearly visible for the second-order approximation. The sixth-order accurate approximation have indistinguishable numerical errors in the *eyeball norm*, when $m = 500$ (and is not presented). Interestingly, the shape of the soliton is very robust, even when prone to large dispersion errors such as in the second-order case. For wave-dominated problems the gain in using a higher-order method typically grows with the propagation-time [24]. In a more realistic application the integration-time $t = 600$ is most likely to be considered small, and the necessity of using a higher order method would be even more pronounced.

Remark Since the time-integration is formally only second-order accurate, the sixth-order accurate SBP-SAT discretisation is most likely close to optimal (referring to numerical efficiency). To motivate higher than sixth-order (spatial) accuracy would require a higher order accurate time-integration (with time-step restriction proportional to h^2 or better), and that is far from trivial to achieve. Besides, second- and fourth-derivative SBP operators do not yet exist for higher than sixth-order accuracy.

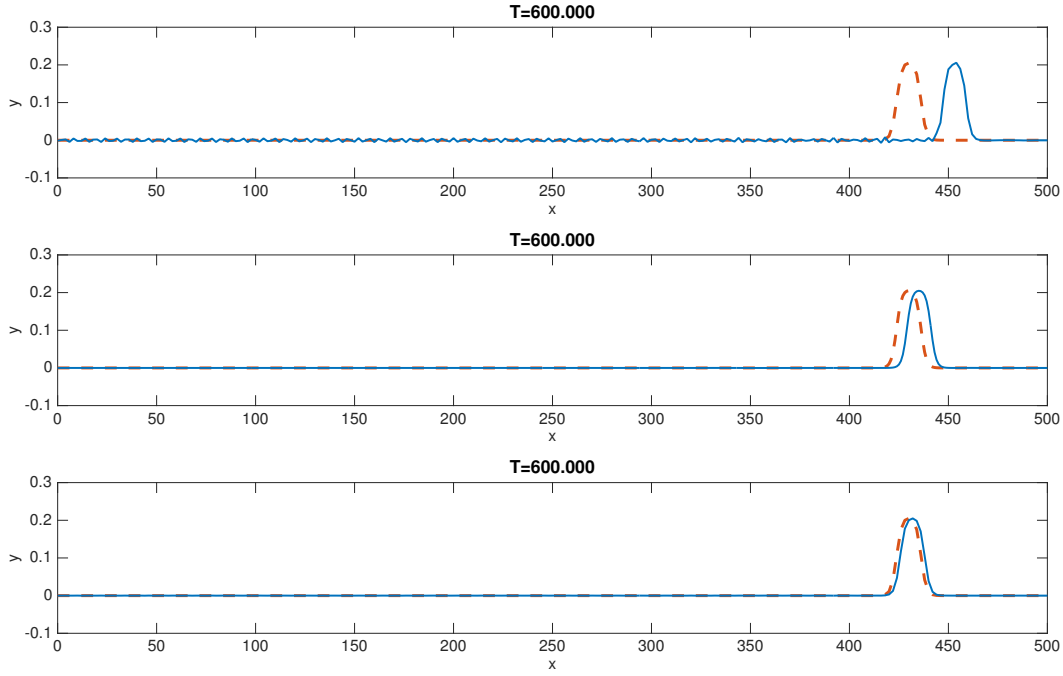

Figure 6: Comparing a second- and sixth-order accurate long-time ($t = 600$) soliton simulation. Blue (solid) line shows the numerical result and red (dotted) line the corresponding analytic solution. Top two subfigures showing the second-order results with $m = 250$ and $m = 500$, respectively. The sixth order case (bottom) with $m = 250$. Notice the huge dispersion error (top).

6.3 Stability properties with non-smooth initial data

The stability properties of the numerical scheme with non-smooth initial data is tested. Here we employ the rectangular function as initial data with its time-derivative set to zero initially, i.e., $f_2 = 0$. This function is discontinuous and contains a high-frequency spectrum and provides a challenging test case for the numerical method to investigate its robustness.

A long-time simulation with the rectangular function as initial data using 401 grid-points is presented in Figure 7. Here we employ a fourth-order accurate SBP-SAT approximation of (39). For non-smooth data with a high-frequency part in the spectrum, a time-integration to $t = 100$ (here using Cdiff) corresponds to a relatively long time-integration.

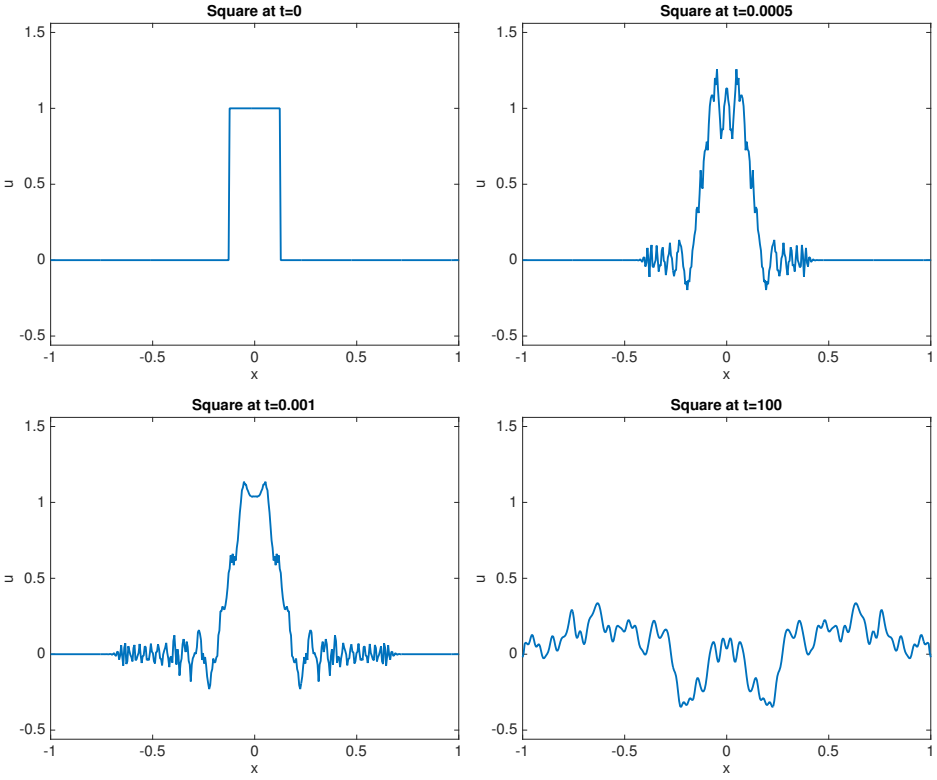

Figure 7: A fourth-order accurate simulation of (39) with $m = 401$ presented at $t = 0, 0.0005, 0.001, 100$. The initial data is the rectangular function. Notice the dispersive wave propagation with the fast high-frequency part.

6.4 Generation of non-analytic solitons

The motivation for the NS model, first presented in [17], was to explain nerve pulses as a density soliton in the axon membrane. It is therefore natural to ask if the NS model can support other types of solitons, besides the analytic soliton given by (8). To rephrase the question: how sensitive are the (soliton) nerve pulses to the initial data? We claim that a robust NS model should support solitons also when we start from smooth initial data other than the analytic soliton solution.

In Figure 8 the simulation is initiated with a Gaussian profile centered at $x = 100$ with an amplitude of 0.2, i.e., $u(x, 0) = 0.2 \exp\left(-\frac{(100-x)^2}{2\times 20^2}\right)$. The amplitud is roughly equivalent to the amplitude of an analytic soliton (8) with $\beta = 0.65$. The initial derivative is here set to $u_t(x, 0) = -u_x(x, 0)$ to trigger a right-going pulse. The spatial domain is given by $x \in [0, 400]$. The sixth-order accurate SBP-SAT approximation of (35) is time-integrated with

Cdiff. We employ $m = 400$ grid-points. The initial profile slowly develops into a new type of *square* soliton.

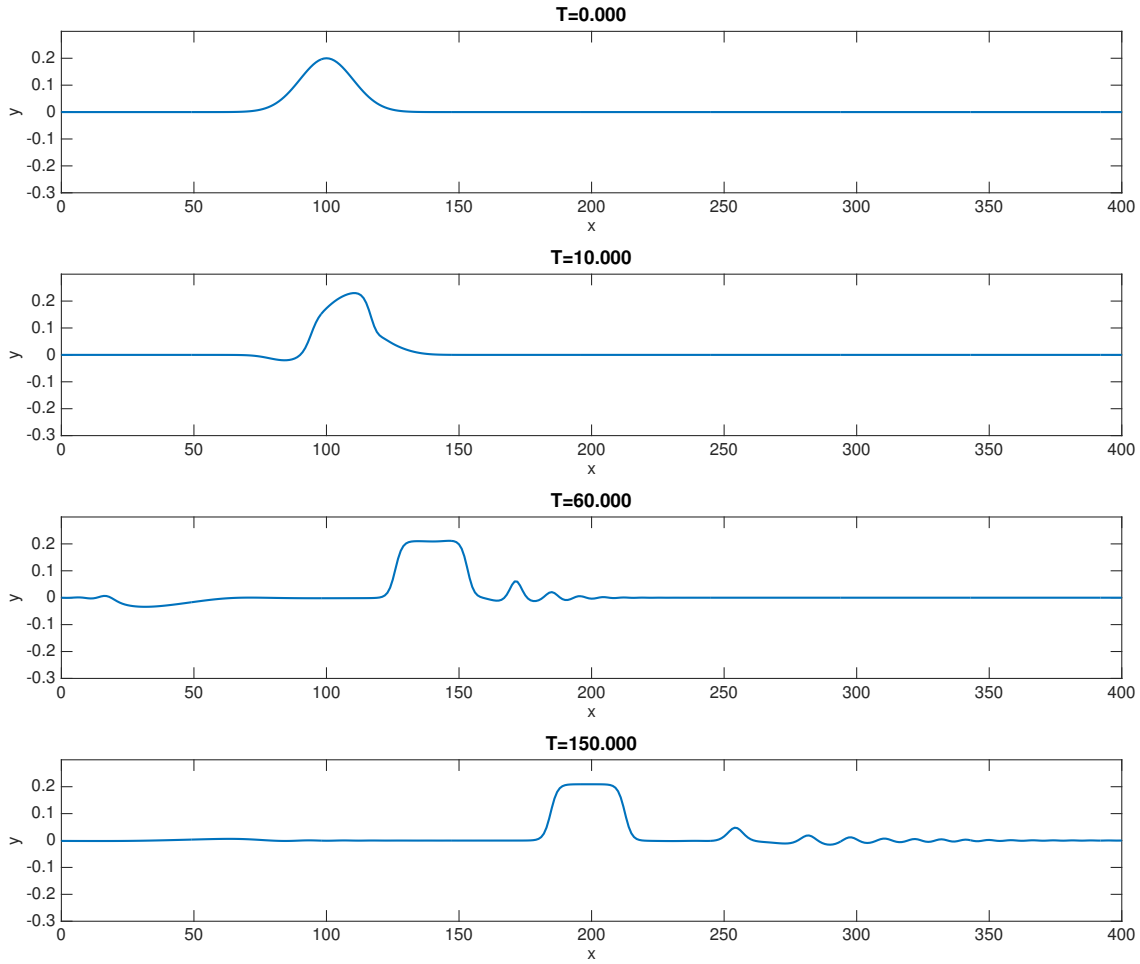

Figure 8: Initial Gaussian evolving into a *square* soliton.

By shifting the width and amplitude of the initial Gaussian profile similar (but different) solitons evolve (not reported here).

7 Conclusions and future work

The main focus has been to construct explicit, stable and high-order accurate SBP-SAT approximations of the NS model with two types of BC. To guarantee stability we employ newly constructed 1D compatible SBP operators to approximate the second- and fourth-derivative terms, combined with

the SAT technique to impose the BC weakly. One of the main results in the present study is the proof that it is necessary to utilise the borrowing condition for the fourth-derivative SBP operator in order to derive a stable boundary treatment for the case with DN BC.

Numerical soliton simulations corroborate the stability- and accuracy-properties and show that the higher-order accurate approximations are superior to the corresponding second-order accurate SBP-SAT schemes. The present work shows that the SBP-SAT finite difference method is a viable and efficient way to obtain numerical solutions of the NS model.

Another main result concerns the CFL conditions for the NS model involving both first and second derivatives in time. It is shown that explicit Runge-Kutta methods have severe CFL restrictions and a better choice is to us a centered finite difference method.

In a coming study we will derive a stable, conservative and accurate interface treatment to model axon junctions, soma- and synapse-junctions. A similar study was presented in [3] for the Hodgkin-Huxley equations. We will also investigate the efficiency of semi-implicit time-integration of the semi-discrete SBP-SAT approximations.

APPENDIX

Below we list the second-derivative SBP operators together with the corresponding 1D compatible fourth-derivative SBP operators for the second-, fourth- and sixth-order cases. The second-derivative operator

$$D_2^{(b)} = H^{-1}(-M^{(b)} + b_m e_m d_{1;m} - b_1 e_1 d_{1;1}) \ ,$$

approximates $\partial/\partial\, x\, (\, b(x)\, \partial/\partial\, x)$, where $b(x) > 0$. The fourth-derivative SBP operator is given by,

$$D_4 = H^{-1} \left(N - e_1 d_{3;1} + e_m d_{3;m} + d_{1;1}^T d_{2;1} - d_{1;m}^T d_{2;m} \right) \ .$$

The boundary closures (the coefficients) are presented for the H, $M^{(b)}$ and N matrices. Notice that the coefficients in H, $M^{(b)}$ and N should be multiplied by h, $\frac{1}{h}$ and $\frac{1}{h^3}$, respectively. Here h denotes the grid-spacing.

I Diagonal-norm SBP operators

I.1 Second order case.

The boundary derivative operators are given by

$$\begin{aligned}
d_{1;1}v &= \tfrac{-3v_1+4v_2-v_3}{2\,h}, & d_{1;m}v &= \tfrac{+3v_m-4v_{m-1}+v_{m-2}}{6\,h} \\
d_{2;1}v &= \tfrac{v_1-2v_2+1v_3}{h^2}, & d_{2;m}v &= \tfrac{v_m-2v_{m-1}+v_{m-2}}{h^2} \\
d_{3;1}v &= \tfrac{-v_1+3v_2-3v_3+v_4}{h^3}, & d_{3;m}v &= \tfrac{+v_m-3v_{m-1}+3v_{m-2}-v_{m-3}}{h^3}
\end{aligned}$$

The interior schemes of D_4 is given by

$$(D_4 v)_j = \frac{v_{j-2} - 4v_{j-1} + 6v_j - 4v_{j+1} + v_{j+2}}{h^4}$$

The norm is the traditional second order norm, i.e, the first and last coefficients are given by $H_{1,1} = H_{m,m} = \frac{1}{2}$, and the interior is a diagonal matrix with ones. Notice that the coefficients in the norm H should be multiplied by h to have the correct scaling with h.

The left boundary closure of $M^{(b)}$ (given by a 3×3 matrix) is given by

$$\begin{bmatrix}
\frac{1}{2}\, b_1 + \frac{1}{2}\, b_2 & -\frac{1}{2}\, b_1 - \frac{1}{2}\, b_2 & 0 \\
-\frac{1}{2}\, b_1 - \frac{1}{2}\, b_2 & \frac{1}{2}\, b_1 + b_2 + \frac{1}{2}\, b_3 & -\frac{1}{2}\, b_2 - \frac{1}{2}\, b_3 \\
0 & -\frac{1}{2}\, b_2 - \frac{1}{2}\, b_3 & \frac{1}{2}\, b_2 + b_3 + \frac{1}{2}\, b_4
\end{bmatrix} .$$

The corresponding right boundary closure is given by replacing $b_i \to b_{m+1-i}$ for $i = 1..4$ followed by a permutation of both rows and columns.

The interior stencil of $M^{(b)}$ at row i is given by ($i = 4 \dots m-3$):

$$\begin{aligned} m_{i,i-1} &= -\tfrac{1}{2}\, b_{i-1} - \tfrac{1}{2}\, b_i \\ m_{i,i} &= \tfrac{1}{2}\, b_{i-1} + b_i + \tfrac{1}{2}\, b_{i+1} \\ m_{i,i+1} &= -\tfrac{1}{2}\, b_i - \tfrac{1}{2}\, b_{i+1} \end{aligned}$$

The upper part of N,

$$\begin{array}{lllll} N_{1,1} = \frac{13}{10} & N_{1,3} = \frac{9}{10} & N_{2,2} = \frac{26}{5} & N_{2,4} = \frac{2}{5} & N_{3,4} = -\frac{17}{5} \\ N_{1,2} = -\frac{12}{5} & N_{1,4} = \frac{1}{5} & N_{2,3} = -\frac{16}{5} & N_{3,3} = \frac{47}{10} & N_{4,4} = -\frac{29}{5} \end{array}$$

I.2 Fourth order case.

The boundary derivative operators are given by

$$\begin{aligned} d_{1;1}v &= \tfrac{-11v_1+18v_2-9v_3+2v_4}{6\,h}, & d_{1;m}v &= \tfrac{+11v_m-18v_{m-1}+9v_{m-2}-2v_{m-3}}{6\,h} \\ d_{2;1}v &= \tfrac{2v_1-5v_2+4v_3-v_4}{h^2}, & d_{2;m}v &= \tfrac{2v_m-5v_{m-1}+4v_{m-2}-1v_{m-3}}{h^2} \\ d_{3;1}v &= \tfrac{-v_1+3v_2-3v_3+v_4}{h^3}, & d_{3;m}v &= \tfrac{+v_m-3v_{m-1}+3v_{m-2}-v_{m-3}}{h^3} \end{aligned}$$

The interior scheme of D_4 is given by

$$(D_4 v)_j = \frac{-v_{j-3} + 12v_{j-2} - 39v_{j-1} + 56v_j - 39v_{j+1} + 12v_{j+2} - v_{j+3}}{6\,h^4}$$

The upper part of the norm H,

$$H_{1,1} = \tfrac{17}{48} \quad H_{2,2} = \tfrac{59}{48} \quad H_{3,3} = \tfrac{43}{48} \quad H_{4,4} = \tfrac{49}{48}$$

The interior stencil of $-M^{(b)}$ at row i is given by ($i = 7 \dots m-6$):

$$\begin{aligned} m_{i,\,i-2} &= \tfrac{1}{6}\, b_{i-1} - \tfrac{1}{8}\, b_{i-2} - \tfrac{1}{8}\, b_i \\ m_{i,\,i-1} &= \tfrac{1}{6}\, b_{i-2} + \tfrac{1}{6}\, b_{i+1} + \tfrac{1}{2}\, b_{i-1} + \tfrac{1}{2}\, b_i \\ m_{i,\,i} &= -\tfrac{1}{24}\, b_{i-2} - \tfrac{5}{6}\, b_{i-1} - \tfrac{5}{6}\, b_{i+1} - \tfrac{1}{24}\, b_{i+2} - \tfrac{3}{4}\, b_i \\ m_{i,\,i+1} &= \tfrac{1}{6}\, b_{i-1} + \tfrac{1}{6}\, b_{i+2} + \tfrac{1}{2}\, b_i + \tfrac{1}{2}\, b_{i+1} \\ m_{i,\,i+2} &= \tfrac{1}{6}\, b_{i+1} - \tfrac{1}{8}\, b_i - \tfrac{1}{8}\, b_{i+2} \end{aligned} \quad .$$

The boundary closure of $M^{(b)}$ (given by a 6×6 matrix) is presented in [35]. (The coefficients fill almost a page.)

The upper part of N,

$$
\begin{array}{lll}
N_{1,1} = \frac{5762947}{2316384} & N_{2,3} = -\frac{2735053}{289548} & N_{3,6} = -\frac{10195}{144774} \\
N_{1,2} = -\frac{6374287}{1158192} & N_{2,4} = \frac{273109}{165456} & N_{4,4} = \frac{3259225}{579096} \\
N_{1,3} = \frac{573947}{165456} & N_{2,5} = \frac{83767}{1158192} & N_{4,5} = -\frac{324229}{72387} \\
N_{1,4} = -\frac{124637}{289548} & N_{2,6} = \frac{245549}{2316384} & N_{4,6} = \frac{1847891}{1158192} \\
N_{1,5} = \frac{67979}{2316384} & N_{3,3} = \frac{5266855}{579096} & N_{5,5} = \frac{2626501}{330912} \\
N_{1,6} = -\frac{60257}{1158192} & N_{3,4} = -\frac{1099715}{289548} & N_{5,6} = -\frac{7115491}{1158192} \\
N_{2,2} = \frac{30392389}{2316384} & N_{3,5} = \frac{869293}{1158192} & N_{6,6} = \frac{21383077}{2316384}
\end{array}
$$

I.3 Sixth order case.

The boundary derivative operators are given by

$$
\begin{array}{ll}
d_{1;1}v = \frac{-25v_1+48v_2-36v_3+16v_4-3v_5}{12\,h}, & d_{1;m}v = \frac{+25v_m-48v_{m-1}+36v_{m-2}-16v_{m-3}+3v_{m-4}}{12\,h} \\
d_{2;1}v = \frac{35v_1-104v_2+114v_3-56v_4+11v_5}{12\,h^2}, & d_{2;m}v = \frac{35v_m-104v_{m-1}+114v_{m-2}-56v_{m-3}+11v_{m-4}}{12\,h^2} \\
d_{3;1}v = \frac{-5v_1+18v_2-24v_3+14v_4-3v_5}{2\,h^3}, & d_{3;m}v = \frac{+5v_m-18v_{m-1}+24v_{m-2}-14v_{m-3}+3v_{m-4}}{2\,h^3}
\end{array}
$$

The interior scheme of D_4 is given by

$$
(D_4v)_j = \frac{7v_{j-4}-96v_{j-3}+676v_{j-2}-1952v_{j-1}+2730v_j-1952v_{j+1}+676v_{j+2}-96v_{j+3}+7v_{j+4}}{240\,h^4}
$$

The upper part of the norm H,

$$
\begin{array}{lll}
H_{1,1} = \frac{13649}{43200} & H_{3,3} = \frac{2711}{4320} & H_{5,5} = \frac{7877}{8640} \\
H_{2,2} = \frac{12013}{8640} & H_{4,4} = \frac{5359}{4320} & H_{6,6} = \frac{43801}{43200}
\end{array}
$$

The interior stencil of $M^{(b)}$ at row i is given by ($i = 10 \ldots N-9$):

$$
\begin{aligned}
m_{i,\,i-3} &= \tfrac{1}{40}\,b_{i-2} + \tfrac{1}{40}\,b_{i-1} - \tfrac{11}{360}\,b_{i-3} - \tfrac{11}{360}\,b_i \\
m_{i,\,i-2} &= \tfrac{1}{20}\,b_{i-3} - \tfrac{3}{10}\,b_{i-1} + \tfrac{1}{20}\,b_{i+1} + \tfrac{7}{40}\,b_i + \tfrac{7}{40}\,b_{i-2} \\
m_{i,\,i-1} &= -\tfrac{1}{40}\,b_{i-3} - \tfrac{3}{10}\,b_{i-2} - \tfrac{3}{10}\,b_{i+1} - \tfrac{1}{40}\,b_{i+2} - \tfrac{17}{40}\,b_i - \tfrac{17}{40}\,b_{i-1} \\
m_{i,\,i} &= \tfrac{1}{180}\,b_{i-3} + \tfrac{1}{8}\,b_{i-2} + \tfrac{19}{20}\,b_{i-1} + \tfrac{19}{20}\,b_{i+1} + \tfrac{1}{8}\,b_{i+2} + \tfrac{1}{180}\,b_{i+3} + \tfrac{101}{180}\,b_i \\
m_{i,\,i+1} &= -\tfrac{1}{40}\,b_{i-2} - \tfrac{3}{10}\,b_{i-1} - \tfrac{3}{10}\,b_{i+2} - \tfrac{1}{40}\,b_{i+3} - \tfrac{17}{40}\,b_i - \tfrac{17}{40}\,b_{i+1} \\
m_{i,\,i+2} &= \tfrac{1}{20}\,b_{i-1} - \tfrac{3}{10}\,b_{i+1} + \tfrac{1}{20}\,b_{i+3} + \tfrac{7}{40}\,b_i + \tfrac{7}{40}\,b_{i+2} \\
m_{i,\,i+3} &= \tfrac{1}{40}\,b_{i+1} + \tfrac{1}{40}\,b_{i+2} - \tfrac{11}{360}\,b_i - \tfrac{11}{360}\,b_{i+3}
\end{aligned}
\quad .
$$

The boundary closure of $M^{(b)}$ (given by a 9×9 matrix) is presented in [35]. (The coefficients fill almost a page.)

The upper part of N is given by,

$$N_{1,1} = \frac{1394226315049}{367201486080} \quad N_{2,6} = \frac{2793470836571}{12852052012800} \quad N_{4,7} = -\frac{13731270505}{64260260064}$$
$$N_{1,2} = -\frac{1137054563243}{114750464400} \quad N_{2,7} = \frac{6219558097}{428401733760} \quad N_{4,8} = \frac{2933596129}{40800165120}$$
$$N_{1,3} = \frac{16614189027367}{1836007430400} \quad N_{2,8} = -\frac{7313844559}{166909766400} \quad N_{5,5} = \frac{14871726798559}{2570410402560}$$
$$N_{1,4} = -\frac{1104821700277}{306001238400} \quad N_{3,3} = \frac{378613061504779}{12852052012800} \quad N_{5,6} = -\frac{7504337615347}{1606506501600}$$
$$N_{1,5} = \frac{1355771086763}{1836007430400} \quad N_{3,4} - \frac{9117069604217}{642602600640} \quad N_{5,7} = \frac{310830296467}{171360693504}$$
$$N_{1,6} = -\frac{27818686453}{459001857600} \quad N_{3,5} \frac{632177582849}{233673672960} \quad N_{5,8} = -\frac{55284274391}{183600743040}$$
$$N_{1,7} = -\frac{40671054239}{1836007430400} \quad N_{3,6} - \frac{1057776382577}{6426026006400} \quad N_{6,6} = \frac{106318657014853}{12852052012800}$$
$$N_{1,8} = \frac{5442887371}{306001238400} \quad N_{3,7} \frac{443019868399}{4284017337600} \quad N_{6,7} = -\frac{14432772918527}{2142008668800}$$
$$N_{2,2} = \frac{70616795535409}{2570410402560} \quad N_{3,8} - \frac{3707981}{2318191200} \quad N_{6,8} = \frac{58102695589}{22666758400}$$
$$N_{2,3} = -\frac{173266854731041}{6426026006400} \quad N_{4,4} = \frac{5029150721885}{514082080512} \quad N_{7,7} = \frac{27102479467823}{2570410402560}$$
$$N_{2,4} = \frac{28938615291031}{2570410402560} \quad N_{4,5} = -\frac{5209119714341}{1285205201280} \quad N_{7,8} = -\frac{1216032192203}{153000619200}$$
$$N_{2,5} = -\frac{146167361863}{71400288960} \quad N_{4,6} = \frac{12235427457469}{12852052012800} \quad N_{8,8} = \frac{20799922829107}{1836007430400}$$

References

[1] Saul Abarbanel and Adi Ditkowski. Asymptotically stable fourth-order accurate schemes for the diffusion equation on complex shapes. *Journal of Computational Physics*, 133(2):279 – 288, 1997.

[2] M. Almquist, K. Mattsson, and T. Edvinsson. High-fidelity numerical solution of the time-dependent Dirac equation. *J. Comput. Phys.*, 262:86–103, 2014.

[3] D. Amsallem and J. Nordström. High-order accurate difference schemes for the hodgkin-huxley equations. *J. Comput. Phys.*, 252:573–590, 2013.

[4] Revathi Appali, Ursula van Rienen, and Thomas Heimburg. A comparison of the hodgkin-huxley model and the soliton theory for the action potential in nerves. *Advances in Planar Lipid Bilayers and Liposomas*, 16:271–279, 2012.

[5] A. Bayliss, K. E. Jordan, B. J. Lemesurier, and E. Turkel. A fourth order accurate finite difference scheme for the computation of elastic waves. *Bull. Seismol. Soc. Amer.*, 76(4):1115–1132, 1986.

[6] J. Boussinesq. Théorie de l'intumescence liquide appele onde solitaire ou de translation se propageant dans un canal rectangulaire. *Comptes Rendus Acad. Sci. Paris*, 72:755–759, 1871.

[7] J. Boussinesq. Théorie des ondes et des remous qui se propagent le long d'un canal rectangulaire en communiquant au liquide contenu dans ce canal des vitesses sensiblement pareilles de la surface au fond. *J. Math. Pure Appl.*, 7:55–108, 1872.

[8] Hans-Benjamin Braun, Jiri Kulda, Bertrand Roessli, Dirk Visser, Karl W. Kramer, Hans-Ulrich Gudel, and Peter Boni. Emergence of soliton chirality in a quantum antiferromagnet. *Nat Phys*, 1(3):159–163, 12 2005.

[9] M. H. Carpenter, D. Gottlieb, and S. Abarbanel. Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes. *J. Comput. Phys.*, 111(2):220–236, 1994.

[10] Marios A. Christou. Interaction of solitons in a boussinesq equation with dissipation. *International Journal of Computer Mathematics*, 90(7):1397–1412, 2013.

[11] Claudio Conti, Giancarlo Ruocco, and Stefano Trillo. Optical spatial solitons in soft matter. *Phys. Rev. Lett.*, 95:183902, Oct 2005.

[12] Bogdan Damski and Wojciech Zurek. Soliton creation during a bose-einstein condensation. *Phys. Rev. Lett.*, 104:160404, Apr 2010.

[13] B.A. Erickson and J. Nordström. High order accurate adaptive schemes for long time, highly intermittent geophysics problems. *Journal of Computational and Applied Mathematics*, 271:328 – 338, 2014.

[14] P.J.P. Gonalves, M.J. Brennan, and S.J. Elliott. Numerical evaluation of high-order modes of vibration in uniform euler-bernoulli beams. *Journal of Sound and Vibration*, 301(3?5):1035 – 1039, 2007.

[15] M. Grote, A. Schneebeli, and D. Schötzau. Interior penalty discontinuous galerkin method for maxwell's equations: Energy norm error estimates. *Journal of Computational and Applied Mathematics*, 204:375–386, 2007.

[16] B. Gustafsson, H.O. Kreiss, and J. Oliger. *Time-Dependent Problems and Difference Methods*. Pure and Applied Mathematics: A Wiley Series of Texts, Monographs and Tracts. Wiley, 2013.

[17] T. Heimburg and A.D. Jackson. On soliton propagation in biomembranes and nerves. *Proc. Natl Acad. Sci. USA*, 102:9790–9795, 2005.

[18] T. Heimburg and A.D. Jackson. On the action potential as a propagating density pulse and the role of anesthetics. *Biophys. Rev. Lett.*, 2:57–78, 2007.

[19] Jan S. Hesthaven. A stable penalty method for the compressible Navier-Stokes equations: III. multidimensional domain decomposition schemes. *SIAM Journal on Scientific Computing*, 20:62–93, 1998.

[20] J.E. Hicken. Output error estimation for summation-by-parts finite-difference schemes. *Journal of Computational Physics*, 231(9):3828 – 3848, 2012.

[21] A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117:500–544, 1952.

[22] H.-O. Kreiss and N.A. Petersson. A second order accurate embedded boundary method for the wave equation with dirichlet data. *SIAM J. Sci. Comput.*, 27:1141–1167, 2006.

[23] H.-O. Kreiss and G. Scherer. Finite element and finite difference methods for hyperbolic partial differential equations. *Mathematical Aspects of Finite Elements in Partial Differential Equations., Academic Press, Inc.*, 1974.

[24] Heinz-Otto Kreiss and Joseph Oliger. Comparison of accurate methods for the integration of hyperbolic equations. *Tellus XXIV*, 3, 1972.

[25] B. Lautrup, R. Appali, A.D. Jackson, and T. Heimburg. The stability of solitons in biomembranes and nerves. *The European Physical Journal E*, 34(6), 2011.

[26] B. Lautrup, R. Appali, A.D. Jackson, and T. Heimburg. The stability of solitons in biomembranes and nerves. *The European Physical Journal E*, 34(6), 2011.

[27] S. K. Lele. Compact finite difference schemes with spectral-like resolution. *J. Comput. Phys.*, 103:16–42, 1992.

[28] K. Mattsson, M. Almquist, and M. H. Carpenter. Optimal diagonal-norm SBP operators. *J. Comput. Phys.*, 264:91–111, 2014.

[29] K. Mattsson and M. H. Carpenter. Stable and accurate interpolation operators for high-order multi-block finite-difference methods. *SIAM J. Sci Comput.*, 32(4):2298–2320, 2010.

[30] K. Mattsson, F. Ham, and G. Iaccarino. Stable and accurate wave propagation in discontinuous media. *J. Comput. Phys.*, 227:8753–8767, 2008.

[31] K. Mattsson, F. Ham, and G. Iaccarino. Stable boundary treatment for the wave equation on second-order form. *Journal of Scientific Computing*, 41:366–383, 2009.

[32] K. Mattsson and J. Nordström. Summation by parts operators for finite difference approximations of second derivatives. *J. Comput. Phys.*, 199(2):503–540, 2004.

[33] K. Mattsson and F. Parisi. Stable and accurate second-order formulation of the shifted wave equation. *Commun. Comput. Phys.*, 7:103–137, 2010.

[34] K. Mattsson, M. Svärd, and M. Shoeybi. Stable and accurate schemes for the compressible navier-stokes equations. *J. Comput. Phys.*, 227(4):2293–2316, 2008.

[35] Ken Mattsson. Summation by parts operators for finite difference approximations of second-derivatives with variable coefficients. *Journal of Scientific Computing*, 51:650–682, 2012.

[36] Ken Mattsson. Diagonal-norm summation by parts operators for finite difference approximations of third and fourth derivatives. *J. Comput. Phys.*, 274(0):432 – 454, 2014.

[37] R. Mishmash and L. Carr. Quantum entangled dark solitons formed by ultracold atoms in optical lattices. *Phys. Rev. Lett.*, 103:140403, Sep 2009.

[38] Anna Nissen, Katharina Kormann, Magnus Grandin, and Kristoffer Virta. Stable difference methods for block-oriented adaptive grids. *Journal of Scientific Computing*, pages 1–26, 2014.

[39] S. De Rango and D. W. Zingg. A high-order spatial discretization for turbulent aerodynamic computations. *AIAA J.*, 39(7):1296–1304, 2001.

[40] Søren S.L. Andersen, Andrew D. Jackson, and Thomas Heimburg. Towards a thermodynamic theory of nerve pulse propagation. *Progress in Neurobiology*, 88(2):104 – 113, 2009.

[41] G. Strang. Accurate partial difference methods II. non-linear problems. *Num. Math.*, 6:37–46, 1964.

[42] John C. Strikwerda. High-order-accurate schemes for incompressible viscous flow. *International Journal for Numerical Methods in Fluids*, 24:715–734, 1997.

[43] M. Svärd and J. Nordström. On the order of accuracy for difference approximations of initial-boundary value problems. *J. Comput. Physics*, 218:333–352, October 2006.

[44] I. Tasaki, K. Kusano, and M. Byrne. Rapid mechanical and thermal changes in the garfish olfactory nerve associated with a propagated impulse. *Biophys. J.*, pages 1033–1040, 1989.

[45] Kristoffer Virta and Ken Mattsson. Acoustic wave propagation in complicated geometries and heterogeneous media. *Journal of Scientific Computing*, 61(1):90–118, 2014.

[46] S.M. Wiedemann. Natural frequencies and mode shapes of arbitrary beam structures with arbitrary boundary conditions. *Journal of Sound and Vibration*, 300(1–2):280 – 291, 2007.

[47] Chun-Ting Zhang. Soliton excitations in deoxyribonucleic acid (dna) double helices. *Phys. Rev. A*, 35:886–891, Jan 1987.