# A stable and dual consistent boundary treatment using finite differences on summation-by-parts form 

Jens Berg and Jan Nordström

## Linköping University Post Print

N.B.: When citing this work, cite the original article.

Original Publication:
Jens Berg and Jan Nordström, A stable and dual consistent boundary treatment using finite differences on summation-by-parts form, 2012, European Congress on Computational Methods in Applied Sciences and Engineering.

Postprint available at: Linköping University Electronic Press
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-81890

# A STABLE AND DUAL CONSISTENT BOUNDARY TREATMENT USING FINITE DIFFERENCES ON SUMMATION-BY-PARTS FORM 

Jens Berg ${ }^{1}$ and Jan Nordström ${ }^{2}$<br>${ }^{1}$ Uppsala University, Department of Information Technology SE-751 05, Uppsala, Sweden<br>e-mail: Jens.Berg@it.uu.se<br>${ }^{2}$ Linköping University, Department of Mathematics<br>SE-581 83, Linköping, Sweden<br>e-mail: Jan.Nordstrom@liu.se

Keywords: Superconvergence, functionals, summation-by-parts, weak boundary conditions, stability, dual consistency


#### Abstract

This paper is concerned with computing very high order accurate linear functionals from a numerical solution of a time-dependent partial differential equation (PDE). Based on finite differences on summation-by-parts form, together with a weak implementation of the boundary conditions, we show how to construct suitable boundary conditions for the PDE such that the continuous problem is well-posed and the discrete problem is stable and spatially dual consistent. These two features result in a superconvergent functional, in the sense that the order of accuracy of the functional is provably higher than that of the solution.


## 1 Introduction

Functionals can represent the lift or drag on an aircraft, energy or any other scalar quantity computed from the solution to a partial differential equation (PDE). In many engineering applications, high order accurate functionals are often of greater interest than accurate solutions of the equations themselves. Whenever there is a functional involved, the concept of duality becomes important. The solution of a PDE resides in some function space, and the set of all bounded linear functionals on that space is called its dual space. Knowledge of the functional of interest can thus be obtained by studying the dual space. This is the main topic in functional analysis and references can be found in any standard textbook.

In numerical analysis, and in particular for computational fluid dynamics problems, duality have been exploited for optimal control problems [14, 8], error estimation [23, 11, 6] and convergence acceleration $[9,22,13,2]$. An extensive summary of the use of adjoint problems can be found in [10], and more recently in [5] with focus on error estimation and adaptive mesh refinement.

In [2], the theory of functional superconvergence was established for time-dependent problems using a finite difference method on summation-by-parts (SBP) form with boundary conditions imposed weakly by the simultaneous approximation term (SAT). In order to avoid additional theoretical difficulties, Dirichlet boundary conditions for both the primal and dual problem were used. The Dirichlet boundary conditions ensured that both problems were well-posed without additional efforts. In an Euler or Navier-Stokes calculation, however, Dirichlet boundary conditions are rarely used. Unless exact boundary data is known, Dirichlet boundary conditions cause reflections which pollute the solution. Other kind of boundary conditions are well-known to increase both the accuracy and stability properties of the scheme [19, 20, 21].

In this paper, we will consider time-dependent partial differential equations of the form

$$
\begin{align*}
u_{t}+L(u) & =f,  \tag{1}\\
J(u) & =(g, u)
\end{align*}
$$

where $J(u)$ is a functional output of interest, $L$ can be either linear or non-linear and $u$ can represent either a scalar or vector valued function. In [2], the concept of spatial dual consistency was introduced to avoid treating the full time-dependent dual equations when discretizing using a method of lines. The concept is motivated by the following. To find the dual problem, we follow the notation in [2, 13], and seek a function $\theta$ in some appropriate function space, such that

$$
\begin{equation*}
\int_{0}^{T} J(u) d t=\int_{0}^{T}(\theta, f) d t \tag{2}
\end{equation*}
$$

A formal computation (assume $L$ linear and $u, \theta$ to have compact support in space) gives

$$
\begin{align*}
\int_{0}^{T} J(u) d t & =\int_{0}^{T} J(u) d t-\int_{0}^{T}\left(\theta, u_{t}+L u-f\right) d t  \tag{3}\\
& =\int_{0}^{T}\left(\theta_{t}-L^{*} \theta+g, u\right) d t+\int_{\Omega}[\theta u]_{0}^{T} d \Omega+\int_{0}^{T}(\theta, f) d t
\end{align*}
$$

where $L^{*}$ is the formal adjoint, or dual, operator associated with $L$ under the inner product such that $(\theta, L u)=\left(L^{*} \theta, u\right)$. By having homogeneous initial conditions for the primal problem, we
obtain the time-dependent dual problem as

$$
\begin{equation*}
-\theta_{t}+L^{*} \theta=g \tag{4}
\end{equation*}
$$

where we have to put an initial condition for the dual problem at time $t=T$. Usually one introduces the time transformation

$$
\begin{equation*}
\tau=T-t \tag{5}
\end{equation*}
$$

which transforms (4) to

$$
\begin{equation*}
\theta_{\tau}+L^{*} \theta=g \tag{6}
\end{equation*}
$$

with an initial condition at $\tau=0$. Note that it is only the spatial part of (6) which have to be discretized differently compared to a discretization of the primal problem (1). A discretization which simultaneously approximates the spatial primal and dual operator consistently, is called spatially dual consistent and produces superconvergent time-dependent linear functionals if the scheme for the primal problem is stable [2].

A discretization of the primal problem (1) can be written as

$$
\begin{equation*}
\frac{d}{d t} u_{h}+L_{h} u_{h}=f \tag{7}
\end{equation*}
$$

where $u_{h}$ is the discrete approximation of $u$ and $L_{h}$ is a discrete approximation of $L$, including the boundary conditions. It is thus required that (7) is both stable and that the discrete dual operator, $L_{h}^{*}$, is a consistent approximation of $L^{*}$, including the dual boundary conditions.

A difference operator for the first derivative is said to be on SBP form if it can be written as

$$
\begin{equation*}
D_{1}=P^{-1} Q \tag{8}
\end{equation*}
$$

where $P$ defines a norm by

$$
\begin{equation*}
\|u\|^{2}=u^{T} P u \tag{9}
\end{equation*}
$$

and $Q$ satisfies the SBP property

$$
\begin{equation*}
Q+Q^{T}=E_{N}-E_{0}, \tag{10}
\end{equation*}
$$

where

$$
\begin{equation*}
E_{N}=\operatorname{diag}[0, \ldots, 0,1], \quad E_{0}=\operatorname{diag}[1,0, \ldots, 0] . \tag{11}
\end{equation*}
$$

The second derivative operator can be constructed either by applying the first derivative twice, i.e.

$$
\begin{equation*}
D_{2}=\left(P^{-1} Q\right)^{2} \tag{12}
\end{equation*}
$$

which results in a wide operator, or a compact operator with minimal bandwidth of the form

$$
\begin{equation*}
D_{2}=P^{-1}\left(-H+\left(E_{N}-E_{0}\right) S\right) \tag{13}
\end{equation*}
$$

as described in [3, 17, 16]. In this paper, we consider only diagonal norms [25]. A first derivative SBP operator is essentially a $2 s$-order accurate central finite difference operator which have been modified close to the boundaries such that it becomes one-sided. Together with the diagonal norm, the boundary closure is accurate of order $s$ making the SBP operator accurate of order $s+1$ in general [25]. For problems with second derivatives, the compact operator can be modified with higher order accurate boundary closures to gain one extra order of accuracy [17, 26].

The discrete inner product in an SBP setting is defined by

$$
\begin{equation*}
\left(u_{h}, v_{h}\right)_{h}=u_{h}^{T} P v_{h} \tag{14}
\end{equation*}
$$

and hence the discrete adjoint operator can be computed, according to the definition

$$
\begin{equation*}
\left(v_{h}, L_{h} u_{h}\right)_{h}=\left(L_{h}^{*} v_{h}, u_{h}\right)_{h}, \tag{15}
\end{equation*}
$$

as

$$
\begin{equation*}
L_{h}^{*}=P^{-1} L_{h}^{T} P . \tag{16}
\end{equation*}
$$

The proof that a stable and spatially dual consistent SBP scheme produces superconvergent linear functionals is presented in [2]. The proof is based on the fact that the mass matrix, $P$, in the norm is a high order accurate integration operator [12, 13].

The procedure for constructing stable schemes which superconvergent linear functionals can now be summarized as follows;

1. Determine boundary conditions such that both the primal and dual problems are wellposed
2. Discretize the primal problem and ensure stability
3. Compute $L_{h}^{*}$ and chose the remaining parameters (if any) such that the continuous adjoint $L^{*}$ is consistently approximated with the dual boundary conditions

While the procedure seems somewhat abstract, we will show using representative equations that the computations are straight forward. Note that a stable and consistent discretization of the primal problem does not imply that the dual problem is consistently approximated. In fact, that is rarely the case. Note also that spatial dual consistency, and hence superconvergence, is merely a choice of coefficients. Superconvergence is thus obtained at no extra computational cost.

In [2, 13] it was necessary to reduce the equations, which contained second derivatives, to first order form (often called the local discontinuous Galerkin, LDG, form). For systems in higher dimension, the LDG form quickly becomes tedious due to the extra variables that has to be introduced, and it is desirable that it can be avoided. The LDG form was a result of the Dirichlet boundary conditions imposed for the primal equations. Other kind of boundary conditions will improve both the stability and accuracy of the numerical scheme [20, 21], and at the same time, as we shall see, remove the need of the LDG form.

## 2 Linear incompletely parabolic system

Consider the linear incompletely parabolic system of equations on $0 \leq x \leq 1$ given by

$$
\begin{equation*}
U_{t}+A U_{x}=B U_{x x} \tag{17}
\end{equation*}
$$

where $U=[p, u]^{T}$ and

$$
A=\left[\begin{array}{cc}
\bar{u} & \bar{c}  \tag{18}\\
\bar{c} & \bar{u}
\end{array}\right], \quad B=\left[\begin{array}{ll}
0 & 0 \\
0 & \varepsilon
\end{array}\right],
$$

together with a linear functional of interest,

$$
\begin{equation*}
J(U)=\int_{0}^{1} G^{T} U d x \tag{19}
\end{equation*}
$$

Equation (17) can be viewed as the symmetrized [1], one-dimensional Navier-Stokes equations, linearized around a flow field with constant velocity $\bar{u}>0$ and speed of sound $\bar{c}>0$. In this case, we assume $\bar{u}<\bar{c}$ and hence (17) requires two boundary conditions at the inflow boundary, $x=0$, and one at the outflow boundary at $x=1$. The boundary conditions we consider are of the form

$$
\begin{equation*}
H_{L, R} U \mp B U_{x}=G_{L, R} \tag{20}
\end{equation*}
$$

where $H_{L, R}$ are to be determined for well-posedness of both the primal and dual problems.
There are many different forms of the matrices $H_{L, R}$ in (20) available which give wellposed inflow or outflow boundary conditions. The typical way to determine the structure of $H_{L, R}$ is to diagonalize the hyperbolic part of the equation and consider the ingoing or outgoing characteristics. This method will provide an energy estimate with optimal damping properties [20]. However, the dual problem associated with the linear functional (19) will most likely be ill-posed. Well-posedness of the primal problem does not imply well-posedness of the dual problem. For other than homogeneous Dirichlet boundary conditions, the choice of boundary conditions such that both the primal and dual problems are well-posed, is a non-trivial task.

Since we are only interested in the spatial dual operator, it is sufficient to consider the steady, inhomogeneous, problem

$$
\begin{equation*}
A U_{x}-B U_{x x}=F \tag{21}
\end{equation*}
$$

In this case, the differential operator $L$ is given by

$$
\begin{equation*}
L=A \frac{\partial}{\partial x}-B \frac{\partial^{2}}{\partial x^{2}} \tag{22}
\end{equation*}
$$

and we seek $\theta=[\phi, \psi]^{T}$ such that

$$
\begin{equation*}
J(U)=(\theta, F) \tag{23}
\end{equation*}
$$

Integration by parts gives

$$
\begin{align*}
J(U) & =(G, U)-(\theta, L U-F) \\
& =\left(G-L^{*} \theta, U\right)-\left[\theta^{T} A U-\theta^{T} B U_{x}+\theta_{x}^{T} B U\right]_{0}^{1}+(\theta, F), \tag{24}
\end{align*}
$$

where

$$
\begin{equation*}
L^{*} \theta=-A \theta_{x}-B \theta_{x x} \tag{25}
\end{equation*}
$$

and hence the dual operator is given by

$$
\begin{equation*}
L^{*}=-A \frac{\partial}{\partial x}-B \frac{\partial^{2}}{\partial x^{2}} . \tag{26}
\end{equation*}
$$

To determine the boundary conditions for the dual problem, we have to find a minimal set of conditions such that

$$
\begin{equation*}
\left[\theta^{T} A U-\theta^{T} B U_{x}+\theta_{x}^{T} B U\right]_{0}^{1}=0 \tag{27}
\end{equation*}
$$

after the homogeneous boundary conditions for the primal problem have been applied. This is what put restrictions on the matrices $H_{L, R}$ in (20). Not only does the boundary terms have to vanish, they will also have to satisfy the number and form for the dual equation. Wrong choice of boundary conditions for the primal problem will cause the dual problem to be ill-posed since too many, or too few, boundary conditions are placed on the dual equations and hence existence or uniqueness is lost [7].

### 2.1 Left boundary conditions

By considering only the terms at $x=0$, we can write (27) as

$$
\begin{equation*}
\theta^{T}\left(A U-B U_{x}\right)+\theta_{x}^{T} B U=U^{T}\left(\left(A-H_{L}^{T}\right) \theta+B \theta_{x}\right) \tag{28}
\end{equation*}
$$

after having applied the homogeneous boundary conditions for the primal equation. The boundary conditions for the dual equation are thus given by

$$
\begin{equation*}
\left(A-H_{L}^{T}\right) \theta+B \theta_{x}=0 . \tag{29}
\end{equation*}
$$

The form of the matrix $H_{L}$ can now be determined. Since the left boundary is an inflow boundary for the primal equation, but an outflow boundary for the dual equation, only one boundary condition is allowed for the dual equation. The boundary condition 29 hence have to be of rank one and thus it is required that

$$
A-H_{L}^{T}=\left[\begin{array}{cc}
0 & 0  \tag{30}\\
\alpha_{L} & \beta_{L}
\end{array}\right]
$$

or equivalently

$$
H_{L}=\left[\begin{array}{cc}
\bar{u} & \bar{c}-\alpha_{L}  \tag{31}\\
\bar{c} & \bar{u}-\beta_{L}
\end{array}\right] .
$$

Any other form of $H_{L}$ would impose too many boundary conditions for the dual equation and it would not be well-posed. The coefficients $\alpha_{L}, \beta_{L}$ have to be chosen such that we obtain an energy estimate for both the time-dependent primal and dual problems.

We can now turn our attention back to the primal equation to determine the coefficients $\alpha_{L}, \beta_{L}$. By applying the energy method to (17) and considering only the left boundary terms, we get

$$
\begin{equation*}
\|U\|_{t}^{2}=U^{T} A U-U^{T} B U_{x}-U_{x}^{T} B U . \tag{32}
\end{equation*}
$$

By assuming homogeneous boundary conditions,

$$
\begin{equation*}
H_{L} U-B U_{x}=0, \tag{33}
\end{equation*}
$$

we can write (32) as

$$
\begin{equation*}
\|U\|_{t}^{2}=-U^{T} M_{L} U \tag{34}
\end{equation*}
$$

where

$$
M_{L}=-A+H_{L}+H_{L}^{T}=\left[\begin{array}{cc}
\bar{u} & \bar{c}-\alpha_{L}  \tag{35}\\
\bar{c}-\alpha_{L} & \bar{u}-2 \beta_{L}
\end{array}\right]
$$

and we have to chose the coefficients $\alpha_{L}, \beta_{L}$ such that

$$
\begin{equation*}
M_{L} \geq 0 \tag{36}
\end{equation*}
$$

Since $M_{L}$ has a rather easy structure, the eigenvalues can be directly computed as

$$
\begin{equation*}
\mu_{1,2}^{(L)}=\bar{u}-\beta_{L} \pm \sqrt{\left(\bar{u}-\beta_{L}\right)^{2}-\bar{u}\left(\bar{u}-2 \beta_{L}\right)+\left(\bar{c}-\alpha_{L}\right)^{2}} \tag{37}
\end{equation*}
$$

and it is easy to see that both eigenvalues are positive if

$$
\begin{equation*}
\alpha_{L}=\bar{c}, \quad \beta_{L} \leq \frac{\bar{u}}{2} \tag{38}
\end{equation*}
$$

Since $M_{L}$ is symmetric, the choices in (38) ensures that

$$
\begin{equation*}
\|U\|_{t}^{2} \leq 0 \tag{39}
\end{equation*}
$$

and hence the boundary conditions for the primal equation are well-posed.

Remark 1. Note that the choices of $\alpha_{L}, \beta_{L}$ in (38) are not unique. Any choice such that

$$
\begin{equation*}
\left(\bar{c}-\alpha_{L}\right)^{2}-\bar{u}\left(\bar{u}-2 \beta_{L}\right) \leq 0 \tag{40}
\end{equation*}
$$

is sufficient to obtain an energy estimate and hence a well-posed problem.
It is, however, not only the primal equation which has to be well-posed. The time-dependent dual equation with its dual boundary conditions need also be well-posed with the conditions given by (38). By introducing the time transformation,

$$
\begin{equation*}
\tau=T-t \tag{41}
\end{equation*}
$$

we can write the time-dependent dual equation as

$$
\begin{equation*}
\theta_{\tau}-A \theta_{x}=B \theta_{x x} . \tag{42}
\end{equation*}
$$

The well-posedness of the dual problem is proven in
Proposition 1. The time-dependent dual problem (42) is well-posed with the dual boundary conditions (29) and the parameters given in (38).

Proof. By applying the energy method to (42), and only considering the terms at $x=0$, we obtain

$$
\begin{align*}
\|\theta\|_{\tau}^{2} & =-\theta^{T} A \theta-\theta^{T} B \theta_{x}-\theta_{x}^{T} B \theta \\
& =-\theta^{T}\left(-A+H_{L}+H_{L}^{T}\right) \theta  \tag{43}\\
& =-\theta^{T} M_{L} \theta
\end{align*}
$$

after applying the boundary conditions (29). The positiveness of $M_{L}$, with the choices (38), were already proven in the energy estimate of the primal equation.

To summarize, the left homogeneous boundary conditions for the primal problem are given by

$$
\begin{equation*}
H_{L} U-B U_{x}=0 \tag{44}
\end{equation*}
$$

and for the dual problem by

$$
\begin{equation*}
\left(A-H_{L}^{T}\right) \theta+B \theta_{x}=0 \tag{45}
\end{equation*}
$$

where $H_{L}$ are given by (31) with the coefficients defined in (38).

### 2.2 Right boundary conditions

The right boundary, $x=1$, is an outflow boundary for the primal problem and hence only one boundary condition can be used, while we have two variables in the system. This immediately puts restrictions on the boundary condition

$$
\begin{equation*}
H_{R} U+B U_{x}=G_{R} \tag{46}
\end{equation*}
$$

in such a way that $H_{R}$ is required to have the form

$$
H_{R}=\left[\begin{array}{cc}
0 & 0  \tag{47}\\
\alpha_{R} & \beta_{R}
\end{array}\right] .
$$

For any other form of $H_{R}$, too many boundary conditions are placed at the outflow boundary and the primal problem is ill-posed. The coefficients $\alpha_{R}, \beta_{R}$ has to be determined such that both the time-dependent primal and dual problems are well-posed. The energy method applied to the time-dependent primal problem results, as before, in

$$
\begin{equation*}
\|U\|_{t}^{2}=-U^{T} M_{R} U \tag{48}
\end{equation*}
$$

where

$$
M_{R}=A+H_{R}+H_{R}^{T}=\left[\begin{array}{cc}
\bar{u} & \bar{c}+\alpha  \tag{49}\\
\bar{c}+\alpha & \bar{u}+2 \beta
\end{array}\right] .
$$

We can directly compute the eigenvalues of the symmetric matrix $M_{R}$ as

$$
\begin{equation*}
\mu_{1,2}^{(R)}=\bar{u}+\beta_{R} \pm \sqrt{\left(\bar{u}+\beta_{R}\right)^{2}-\bar{u}\left(\bar{u}+2 \beta_{R}\right)+\left(\bar{c}+\alpha_{R}\right)^{2}} \tag{50}
\end{equation*}
$$

and see that they are both positive if we chose

$$
\begin{equation*}
\alpha_{R}=-\bar{c}, \quad \beta_{R} \geq-\frac{\bar{u}}{2} . \tag{51}
\end{equation*}
$$

With the choices of $\alpha_{R}, \beta_{R}$ given in (51), the boundary conditions lead to a well-posed primal problem.
Remark 2. Note that the choices of $\alpha_{R}, \beta_{R}$ in (51) are not unique. Any choice such that

$$
\begin{equation*}
\left(\bar{c}+\alpha_{R}\right)^{2}-\bar{u}\left(\bar{u}+2 \beta_{R}\right) \leq 0 \tag{52}
\end{equation*}
$$

is sufficient to obtain an energy estimate and hence a well-posed problem.
To determine the boundary conditions for the dual problem, we restrict (27) to the terms at $x=1$. After applying the homogeneous primal boundary conditions, we obtain

$$
\begin{equation*}
\theta^{T}\left(A U-B U_{x}\right)+\theta_{x}^{T} B U=U^{T}\left(\left(A+H_{R}^{T}\right) \theta+B \theta_{x}\right) \tag{53}
\end{equation*}
$$

and hence the boundary conditions for the dual problem are given by

$$
\begin{equation*}
\left(A+H_{R}^{T}\right) \theta+B \theta_{x}=0 \tag{54}
\end{equation*}
$$

Morever, it is required that the top row of $A+H_{R}^{T}$ is non-zero since two boundary conditions are required for the dual equation. This is fulfilled by (51) since

$$
A+H_{R}^{T}=\left[\begin{array}{cc}
\bar{u} & 0  \tag{55}\\
\bar{c} & \bar{u}+\beta_{R}
\end{array}\right] .
$$

The well-posedness of the dual boundary conditions for the time-dependent dual problem is given in

Proposition 2. The time-dependent dual problem (42) is well-posed with the dual boundary conditions (54) and the parameters given in (51).

Proof. As before, the energy method applied to the time-dependent dual problem gives

$$
\begin{align*}
\|\theta\|_{\tau}^{2} & =-\theta^{T}\left(A+H_{R}+H_{R}^{T}\right) \theta \\
& =-\theta^{T} M_{R} \theta, \tag{56}
\end{align*}
$$

where positiveness of $M_{R}$ is already proven from the energy estimate of the primal equation.

To summarize, the right homogeneous boundary conditions for the primal problem are given by

$$
\begin{equation*}
H_{R} U+B U_{x}=0 \tag{57}
\end{equation*}
$$

and for the dual problem by

$$
\begin{equation*}
\left(A+H_{R}^{T}\right) \theta+B \theta_{x}=0 \tag{58}
\end{equation*}
$$

where $H_{R}$ are given by (47) with the coefficients defined in (51).
Remark 3. Note how it is the problem which requires the least number of boundary conditions which sets restrictions on the form of the boundary conditions. When also considering wellposedness of the dual problem, it can be used to reduce the number of unknown parameters in the boundary conditions of the primal problem.

### 2.3 Discretization, stability and spatial dual consistency

To discretize systems of equations, it is convenient to introduce the Kronecker product which is defined for arbitrary matrices $C, D$ as

$$
C \otimes D=\left[\begin{array}{cccc}
C_{11} D & C_{1,2} D & \cdots & C_{1 n} D  \tag{5}\\
C_{21} D & C_{22} D & \cdots & C_{2 n} D \\
\vdots & \ddots & \ddots & \vdots \\
C_{n 1} D & C_{n 2} D & \cdots & C_{n n} D
\end{array}\right] .
$$

For the matrix inverse and transpose we have

$$
\begin{equation*}
(C \otimes D)^{-1, T}=C^{-1, T} \otimes D^{-1, T} \tag{60}
\end{equation*}
$$

if the usual matrix inverses are defined. Furthermore, a useful property which will be extensively used, is the mixed product property,

$$
\begin{equation*}
(C \otimes D)(\tilde{C} \otimes \tilde{D})=C \tilde{C} \otimes D \tilde{D} \tag{61}
\end{equation*}
$$

if the usual matrix products are defined.
Using the Kronecker product, equation (17) with the boundary conditions (20) can be discretized using the SBP-SAT technique as

$$
\begin{align*}
U_{t}+\left(D_{1} \otimes A\right) U & =\left(D_{2} \otimes B\right) U \\
& +\left(P^{-1} E_{0} \otimes \Sigma_{L}\right)\left(\left(I_{N+1} \otimes H_{L}\right) U-\left(D_{1} \otimes B\right) U-G_{L}\right)  \tag{62}\\
& +\left(P^{-1} E_{N} \otimes \Sigma_{R}\right)\left(\left(I_{N+1} \otimes H_{R}\right) U+\left(D_{1} \otimes B\right) U-G_{R}\right)
\end{align*}
$$

where $\Sigma_{L, R}$ are $2 \times 2$ matrices which have to be determined for stability. The second derivative is approximated using the wide operator, $D_{2}=D_{1} D_{1}=\left(P^{-1} Q\right)^{2}$. The matrices $\Sigma_{L, R}$ are given in

Proposition 3. The scheme (62) is energy stable by choosing

$$
\begin{equation*}
\Sigma_{L}=\Sigma_{R}=-I \tag{63}
\end{equation*}
$$

Proof. We let $G_{L}=G_{R}=0$ and apply the energy method to (62). By using the SBP properties of the operators, we obtain

$$
\begin{align*}
\|U\|_{t}^{2} & =U^{T}\left(E_{0} \otimes A+\Sigma_{L} H_{L}+H_{L}^{T} \Sigma_{L}^{T}\right) U-U^{T}\left(E_{N} \otimes A-\Sigma_{R} H_{R}-H_{R}^{T} \Sigma_{R}^{T}\right) U \\
& -2 U^{T}\left(E_{0} D_{1} \otimes B+\Sigma_{L} B\right) U+2 U^{T}\left(E_{N} D_{1} \otimes B+\Sigma_{R} B\right) U  \tag{64}\\
& -2\left(\left(D_{1} \otimes I_{2}\right) U\right)^{T}(P \otimes B)\left(\left(D_{1} \otimes I_{2}\right) U\right)
\end{align*}
$$

where the last term is purely dissipative. By chosing

$$
\begin{equation*}
\Sigma_{L}=\Sigma_{R}=-I \tag{65}
\end{equation*}
$$

equation (64) simplifies to

$$
\begin{equation*}
\|U\|_{t}^{2} \leq-U^{T}\left(E_{0} \otimes-A+H_{L}+H_{L}^{T}\right) U-U^{T}\left(E_{N} \otimes A+H_{R}+H_{R}^{T}\right) U \tag{66}
\end{equation*}
$$

where, by construction,

$$
\begin{equation*}
-A+H_{L}+H_{L}^{T} \geq 0, \quad A+H_{R}+H_{R}^{T} \geq 0 \tag{67}
\end{equation*}
$$

Since the Kronecker product preserves positive (semi) definiteness, we have

$$
\begin{equation*}
\|U\|_{t}^{2} \leq 0 \tag{68}
\end{equation*}
$$

and the scheme is energy stable.
The unique choice (63) of penalty matrices renders the scheme not only energy stable, but also spatially dual consistent. To prove this we must show that the discrete adjoint operator $L_{h}^{*}$ consistently approximates the discrete adjoint $L^{*}$, including the dual boundary conditions (29) and (54). This is done in
Proposition 4. The scheme (62) is spatially dual consistent with the choices of $\Sigma_{L, R}$ given in (63).

Proof. For $G_{L, R}=0$, we can write the scheme (62) as

$$
\begin{equation*}
\frac{d}{d t} U_{h}+L_{h} U_{h}=0 \tag{69}
\end{equation*}
$$

where

$$
\begin{align*}
L_{h} & =\left(D_{1} \otimes A\right)-\left(D_{2} \otimes B\right) \\
& +\left(P^{-1} E_{0} \otimes I_{2}\right)\left(\left(I_{N+1} \otimes H_{L}\right)-\left(D_{1} \otimes B\right)\right)  \tag{70}\\
& +\left(P^{-1} E_{N} \otimes I_{2}\right)\left(\left(I_{N+1} \otimes H_{R}\right)+\left(D_{1} \otimes B\right)\right) .
\end{align*}
$$

The discrete dual operator is defined by

$$
\begin{equation*}
L_{h}^{*}=\left(P \otimes I_{2}\right)^{-1} L_{h}^{T}\left(P \otimes I_{2}\right) \tag{71}
\end{equation*}
$$

A somewhat tedious calculation shows that

$$
\begin{align*}
L_{h}^{*} & =-\left(D_{1} \otimes A\right)-\left(D_{2} \otimes B\right) \\
& -\left(P^{-1} E_{0} \otimes I_{2}\right)\left(\left(I_{N+1} \otimes A-H_{L}^{T}\right)+\left(D_{1} \otimes B\right)\right)  \tag{72}\\
& +\left(P^{-1} E_{N} \otimes I_{2}\right)\left(\left(I_{N+1} \otimes A+H_{R}^{T}\right)+\left(D_{1} \otimes B\right)\right)
\end{align*}
$$

which is a consistent approximation of (26) including the dual boundary conditions (29) and (54). The scheme is hence spatially dual consistent.

### 2.4 Numerical results

A forcing function have been chosen such that an analytical solution is known, and the rates of convergence are computed with respect to the analytical solution. This is known as the method of manufactured solutions [24, 15]. The solution in this case is given by

$$
\begin{align*}
& p(x, t)=(\arctan (x)-\delta \cos (\alpha x-t)+1) e^{-x^{2}} \\
& u(x, t)=(\arctan (x)+\delta \sin (\alpha x-t)+1) e^{-x^{2}} \tag{73}
\end{align*}
$$

and the weight functions in the functionals are chosen such that

$$
\begin{align*}
& J(p)=1+\frac{\pi}{4}-\frac{\log (2)}{2}+\frac{\delta(\sin (t-\alpha)-\sin (t))}{\alpha}  \tag{74}\\
& J(u)=1+\frac{\pi}{4}-\frac{\log (2)}{2}+\frac{\delta(\cos (t)-\cos (t-\alpha))}{\alpha} . \tag{75}
\end{align*}
$$

In (62), the second derivative is approximated using the first derivative operator twice, which results in a wide operator. There are compact second derivative operators with better accuracy [17, 16]. When considering the discrete dual problem, however, spatial dual consistency can only be obtained if the second derivative operator can be factorized into the product of two consistent first derivative operators. Indeed, the scheme (62) is dual inconsistent when the second derivative is compactly discretized. In Table 1 we present the numerical results regarding the order of convergence of the solution and functionals using both a wide and compact second derivative operator. The time integration is performed until time $t=0.2$ with the classical 4th-order Runge-Kutta method using 1000 time steps. Artificial dissipation has been added as described in [18, 4].

|  | Wide operator |  |  |  |  |  |  |  |  | Compact operator |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $N$ | 64 | 96 | 128 | 160 | 64 | 96 | 128 | 160 |  |  |  |  |  |
| $s=2$ |  |  |  |  |  |  |  |  |  |  |  |  |  |
| $p$ | 3.0530 | 3.0426 | 3.0377 | 3.0345 | 3.2055 | 3.2539 | 3.1971 | 3.1560 |  |  |  |  |  |
| $u$ | 2.8870 | 2.9740 | 2.9973 | 3.0061 | 3.8368 | 4.0364 | 4.0903 | 4.1204 |  |  |  |  |  |
| $J(p)$ | 2.5584 | 3.9209 | 4.2617 | 4.4285 | 2.3289 | 3.2762 | 3.6065 | 3.7740 |  |  |  |  |  |
| $J(u)$ | 4.2536 | 4.2841 | 4.3698 | 4.4192 | 3.3403 | 3.5314 | 3.7163 | 3.8343 |  |  |  |  |  |
| $s=3$ |  |  |  |  |  |  |  |  |  |  |  |  |  |
| $p$ | 3.9528 | 4.2587 | 4.4108 | 4.4797 | 4.7882 | 4.3370 | 4.0411 | 4.0424 |  |  |  |  |  |
| $u$ | 5.1102 | 4.8158 | 4.6828 | 4.5782 | 6.0147 | 6.7091 | 6.8642 | 6.5443 |  |  |  |  |  |
| $J(p)$ | 2.8420 | 4.6607 | 5.5723 | 5.9361 | 5.8846 | 4.3321 | 2.9507 | 4.3215 |  |  |  |  |  |
| $J(u)$ | 4.4917 | 5.6177 | 5.9905 | 6.1948 | 5.9529 | 2.0560 | 3.9205 | 4.6030 |  |  |  |  |  |
| $s=4$ |  |  |  |  |  |  |  |  |  |  |  |  |  |
| $p$ | 4.8682 | 4.7599 | 4.7595 | 4.7499 | 3.6345 | 4.4454 | 4.6180 | 4.6752 |  |  |  |  |  |
| $u$ | 4.2438 | 5.1204 | 5.2586 | 5.2981 | 4.8689 | 6.4603 | 7.2947 | 7.7447 |  |  |  |  |  |
| $J(p)$ | 5.0355 | 7.7071 | 8.1376 | 8.5775 | 3.7680 | 6.0404 | 6.5043 | 6.8383 |  |  |  |  |  |
| $J(u)$ | 6.5365 | 7.3625 | 8.0564 | 8.5796 | 5.0693 | 5.8672 | 6.4344 | 6.8063 |  |  |  |  |  |

Table 1: Convergence rates for the variables and functionals using both the wide and compact second derivative operator

As can be seen from Table 1, the spatially dual consistent scheme with the wide second derivative operator results in superconvergent functionals. The same scheme with a compact
operator has no functional superconvergence due to its dual inconsistency. Note, however, that the variable $u$ has higher order of convergence than that of $p$ in the inconsistent discretization. A possible explanation for this is that the boundary closure of the compact second derivative operator has higher order of accuracy, and hence $u$ gain accuracy when compactly discretized [17].

## 3 Conclusions

We have presented a stable and spatially dual consistent boundary treatment of an incompletely parabolic system of equations in one space dimension. It was shown how to construct well-posed boundary conditions such that superconvergent functionals were obtained. Another benefit compared to having Dirichlet boundary conditions is that a reduction to a first order system can be avoided.

The choice of parameters for energy stability is unique due to the form of the boundary conditions. The unique choice of parameters also lead to a spatially dual consistent scheme which produced superconvergent linear functionals.

Spatial dual consistency is merely a choice of parameters in the scheme, and hence the superconvergence was obtained without additional computational cost compared to a standard scheme.

## Acknowledgements

This work has partly been carried out within the IDIHOM project which is supported by the European Commission under contract No. ACP0-GA-2010-265780.

## REFERENCES

[1] Saul Abarbanel and David Gottlieb. Optimal time splitting for two- and three-dimensional Navier-Stokes equations with mixed derivatives. Journal of Computational Physics, 41(1):1-33, 1981.
[2] Jens Berg and Jan Nordström. Superconvergent functional output for time-dependent problems using finite differences on summation-by-parts form. Technical Report 2012004, Department of Information Technology, Uppsala University, February 2012.
[3] Mark H. Carpenter, Jan Nordström, and David Gottlieb. A stable and conservative interface treatment of arbitrary spatial accuracy. Journal of Computational Physics, 148(2):341 - 365, 1999.
[4] Peter Diener, Ernst Nils Dorband, Erik Schnetter, and Manuel Tiglio. Optimized highorder derivative and dissipation operators satisfying summation by parts, and applications in three-dimensional multi-block evolutions. Journal of Scientific Computing, 32(1):109145, July 2007.
[5] Krzysztof J. Fidkowski and David L. Darmofal. Review of output-based error estimation and mesh adaptation in computational fluid dynamics. AIAA Journal, 49(4):673-694, 2011.
[6] Michael B. Giles, Mats G. Larson, J. Mårten Levenstam, and Endre Süli. Adaptive error control for finite element approximations of the lift and drag coefficients in viscous flow. Technical report, Report NA-97/06, Oxford University Computing Laboratory, 1997.
[7] Michael B. Giles and Niles A. Pierce. Adjoint equations in CFD: duality, boundary conditions and solution behaviour. In AIAA Paper 97-1850, 1997.
[8] Michael B. Giles and Niles A. Pierce. An introduction to the adjoint approach to design. Flow, Turbulence and Combustion, 65:393-415, 2000.
[9] Michael B. Giles and Niles A. Pierce. Superconvergent lift estimates through adjoint error analysis. Innovative Methods for Numerical Solutions of Partial Differential Equations, 2001.
[10] Michael B. Giles and Endre Süli. Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality. Acta Numerica, 11:145-236, 2002.
[11] Jason E. Hicken. Output error estimation for summation-by-parts finite-difference schemes. Journal of Computational Physics, 231(9):3828-3848, 2012.
[12] Jason E. Hicken and David W. Zingg. Summation-by-parts operators and high-order quadrature. ArXiv e-prints, March 2011.
[13] Jason E. Hicken and David W. Zingg. Superconvergent functional estimates from summation-by-parts finite-difference discretizations. SIAM Journal on Scientific Computing, 33(2):893-922, 2011.
[14] Antony Jameson. Aerodynamic design via control theory. Journal of Scientific Computing, 3:233-260, September 1988.
[15] Jens Lindström and Jan Nordström. A stable and high-order accurate conjugate heat transfer problem. Journal of Computational Physics, 229(14):5440-5456, 2010.
[16] Ken Mattsson. Summation by parts operators for finite difference approximations of second-derivatives with variable coefficients. Journal of Scientific Computing, pages 1-33, 2011.
[17] Ken Mattsson and Jan Nordström. Summation by parts operators for finite difference approximations of second derivatives. Journal of Computational Physics, 199(2):503540, 2004.
[18] Ken Mattsson, Magnus Svärd, and Jan Nordström. Stable and accurate artificial dissipation. Journal of Scientific Computing, 21(1):57-79, 2004.
[19] Jan Nordström. The influence of open boundary conditions on the convergence to steady state for the Navier-Stokes equations. Journal of Computational Physics, 85(1):210 - 244, 1989.
[20] Jan Nordström. The use of characteristic boundary conditions for the Navier-Stokes equations. Computers \& Fluids, 24(5):609-623, 1995.
[21] Jan Nordström, Sofia Eriksson, and Peter Eliasson. Weak and strong wall boundary procedures and convergence to steady-state of the Navier-Stokes equations. Journal of Computational Physics, 231(14):4867-4884, 2012.
[22] Niles A. Pierce and Michael B. Giles. Adjoint recovery of superconvergent functionals from PDE approximations. SIAM Review, 42(2):247-264, 2000.
[23] Niles A. Pierce and Michael B. Giles. Adjoint and defect error bounding and correction for functional estimates. Journal of Computational Physics, 200:769-794, November 2004.
[24] Lee Shunn, Frank Ham, and Parviz Moin. Verification of variable-density flow solvers using manufactured solutions. Journal of Computational Physics, 231(9):3801 - 3827, 2012.
[25] Bo Strand. Summation by parts for finite difference approximations for d/dx. Journal of Computational Physics, 110(1):47-67, 1994.
[26] Magnus Svärd and Jan Nordström. On the order of accuracy for difference approximations of initial-boundary value problems. Journal of Computational Physics, 218(1):333-352, 2006.

