INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2003; **41**:1295–1318 (DOI: 10.1002/fld.435)

A high-order mass-lumping procedure for B-spline collocation method with application to incompressible flow simulations

O. Botella[1,2,*,†]

[1]*Center for Turbulence Research, Stanford University, Stanford, CA 94305, U.S.A.*
[2]*LEMTA, 2 avenue de la Forêt de Haye, 54504 Vandœuvre-lès-Nancy, France*

SUMMARY

This paper presents new developments of the staggered spline collocation method for cost-effective solution to the incompressible Navier–Stokes equations. Maximal decoupling of the velocity and the pressure is obtained by using the fractional step method of Gresho and Chan, allowing the solution to sparse elliptic problems only. In order to preserve the high-accuracy of the B-spline method, this fractional step scheme is used in association with a sparse approximation to the inverse of the consistent mass matrix. Such an approximation is constructed from local spline interpolation method, and represents a high-order generalization of the mass-lumping technique of the finite-element method. A numerical investigation of the accuracy and the computational efficiency of the resulting semi-consistent spline collocation schemes is presented. These schemes generate a stable and accurate unsteady Navier–Stokes solver, as assessed by benchmark computations. Copyright © 2003 John Wiley & Sons, Ltd.

KEY WORDS: B-splines; collocation method; mass matrix; Navier–Stokes equations; fractional step methods

1. INTRODUCTION

The development of numerical methods based on B-spline methodology is motivated by the substantial computational cost of large-eddy simulations (LES) of complex turbulent flows. Indeed, the large number of grid points needed in turbulent boundary layers remains one of the principle obstacles to a wider application of LES to flows of engineering interest. An active part of research in LES is devoted to reducing these resolution requirements, by the formulation of approximate wall conditions (see e.g. Reference [1]), and by the development of highly accurate numerical methods for the precise representation of near-wall structures.

Several works [2–4] have been devoted to the development of B-spline methods on semi-structured embedded meshes. This technique allows a substantial reduction in the

*Correspondence to: O. Botella, LEMTA, CNRS UMR 7563, 2 avenue de la Forêt de Haye, Vandœuvre-lès-Nancy, 54504, France.
†E-mail: obotella@ensem.inpl-nancy.fr

Received 27 July 2000
Accepted 30 August 2002

computational cost of a simulation by using fine grids in physically significant flow regions only. The use of B-splines is motivated by the development of robust and non-dissipative LES schemes on arbitrary meshes. The conservation of physical invariants such as kinetic energy is highly desirable for the simulation of turbulent flows [5], and these requirements are reproduced with difficulty by finite-difference schemes on non-uniform meshes [6]. Moreover, the resolution power of B-splines of maximum continuity allows the representation of a broad range of scales of a turbulent flow [3].

The work of Kravchenko *et al.* [3] and Kravchenko and Moin [2] has shown the high suitability of B-spline methods for the computation of complex turbulent flows. However the Galerkin approximation that is employed is too CPU intensive. The method is burdened by the cost of evaluating non-linear terms where, as observed in Reference [3], 50% of the computational time is spent on their evaluation.

This paper represents a follow-up to the work initiated in References [7, 8] for developing a cost-effective B-spline Navier–Stokes solver. The equations are discretized with the collocation method, which allows a drastic reduction of the cost of evaluating non-linearities. A stable approximation to the pressure is obtained by constructing staggered bases for the velocity and pressure which are, in a sense, the B-spline equivalent to the popular staggered finite-difference discretization [9]. The time-discretization employs a fractional step scheme [10, 11].

In association with 'local' (or 'explicit') discretizations such as finite-difference or finite-volume approaches, fractional step techniques are widely considered as the most cost-effective method for solving the Navier–Stokes equations. Indeed, they provide a maximum decoupling of the velocity and the pressure, so that only sparse elliptic problems need to be solved at each time-cycle. However, for 'global' discretizations such as B-spline methods that yield a non-diagonal mass matrix, a straightforward application of these methods retains some coupling between the velocity and pressure: the pressure operator associated with the projection step, that involves the dense inverse of the mass matrix, is dense and can only be constructed for modestly sized problems. As suggested in References [7, 8], the pressure equation can nonetheless be solved by means of an Uzawa algorithm, but the CPU cost of this iterative solution, even accelerated by modern Krylov subspace methods [12, 13], is prohibitively high for large scale problems.

In order to make this B-spline method attractive with respect to CPU cost, we have made an effort to modify the fractional step scheme in order to obtain a simpler linear system for the pressure that would be sparse, and eliminate the need for Uzawa iterations. A modification of the mass matrix to get a sparse approximation to its inverse is a key element in this endeavor.

The modification of the 'consistent approximation', which generates the non-diagonal mass matrix, has always been a critical issue for finite-element type methods. A common *ad hoc* simplification consists in approximating the mass matrix with a diagonal matrix, usually by summing its rows and putting the result on the diagonal (the 'lumped mass' approximation, see e.g. References [14, 15]). For compressible flow simulations, this simplification is motivated by the use of explicit time-stepping such that, when the time-derivative term is lumped, the inversion of the consistent mass matrix is no more needed at each time cycle. This mass lumping technique, however, diminishes the accuracy of the resulting scheme, most notably for unsteady flows dominated by convection effects [16], since this approximation is, in general, a first-order in space approximation to the consistent mass matrix.

So far, the most satisfying application of the lumping technique to incompressible flow computations is represented by the 'projection 2' scheme of Gresho [17] and Gresho and Chan [18]. It uses a semi-consistent mass matrix approximation (SCM), i.e. the mass matrix is lumped in front of the pressure gradient in the momentum equation only, while the continuity equation is unaltered. As a consequence, the pressure operator is sparse and can efficiently be inverted by standard elliptic solvers. Early applications of the SCM technique to the B-spline collocation method were reported in References [7, 8]. For low-order spline approximations, this scheme performed accurate Navier–Stokes benchmark computations with only a fraction of the CPU time needed by the original consistent scheme. However, due to the crude approximation represented by the lumping of the mass matrix, the SCM scheme led to a loss of the accuracy that would be expected for high-order B-splines.

In order to preserve, as far as possible, the high accuracy of the B-spline method, it is thus necessary to build more accurate approximations of the consistent mass matrix than the lumped approximation. These considerations led us in this work to the development of approximate inverses of the mass matrix, i.e. highly-accurate sparse approximations to the inverse of the consistent mass matrix. This concept of approximate inverse is somewhat similar to the one developed for the iterative solution of linear systems, where the approximate inverse is an explicit preconditioner whose application in an iterative procedure requires a sparse matrix-vector multiplication only (see Reference [19] and references therein). The main difference is that we are able to replace the solution of a mass matrix problem by a single sparse matrix-vector multiplication, while keeping the order of accuracy of the B-splines.

For the B-spline collocation method, such sparse approximations are obtained by application of local interpolation schemes. These schemes of quasi-interpolation were developed in e.g. References [20–22] to build spline representation of a function from data values. The B-spline coefficients are not determined as the solution to a collocation system, as the consistent approximation would require, but rather as the linear combination of the function values at a small number of data points. When these data points are chosen among the collocation points, this linear combination defines the entries of the approximate inverse of the mass matrix. The number of data points affects the order of accuracy of the approximate inverse. The case with a single data point corresponds to the low-order lumped approximation. The increase in the number of data points raises the order, and a sufficient number of points yields the order of accuracy of the consistent approximation. These approximate inverses represent thus a high-order generalization of the mass lumping technique.

The concepts of approximate inverse and local interpolation may have important application to numerical algorithms where a fast transformation from the physical (collocation) space to the B-spline coefficients space is required. Among others, we cite the development of restriction operators for spline multigrid methods [23], and the evaluation of non-linearities in collocation space, as in the pseudospectral method. In the present work, the use of an approximate inverse allows us to solve the pressure equation of the Navier–Stokes scheme with a fraction of the CPU time required by the consistent approximation, with the same order of accuracy. The slight loss of the resolving power of the semi-consistent schemes, caused by the replacement of the consistent mass matrix, is anyway greatly counterbalanced by their computational efficiency. These issues are carefully addressed here by presenting numerical tests and Fourier analysis. The combination of these approximate inverses with the SCM fractional-step technique allows the construction of a highly-accurate cost-effective Navier–Stokes solver, as proved by the benchmark tests reported in this paper.

2. BACKGROUND ON B-SPLINE NUMERICAL SCHEMES

2.1. Construction of B-spline bases

A spline function is a piecewise polynomial of order k (i.e. its polynomial degree is $k-1$ at most) defined on the interval $\Lambda =]a,b[$, whose derivatives at some order possess jump-discontinuities at breakpoints $\xi = \{\xi_i,\ i = 1,\ldots,l+1\}$ defined by

$$a = \xi_1 < \xi_2 < \cdots < \xi_i < \cdots < \xi_l < \xi_{l+1} = b \tag{1}$$

In the following, we focus on the characterization of the so-called smoothest splines, which have jump discontinuities in their $k-1$ derivative, since previous studies in References [7, 8] have assessed their superior resolving power in the collocation approach.

For the approximation to a given function $f(x)$, the spline function $\tilde{f}(x)$ is commonly described in its B-representation

$$\tilde{f}(x) = \sum_{i=1}^{N} \alpha_i B_i^k(x) \tag{2}$$

where $B_i^k(x)$ is a special spline function of order k called a B-spline which has, in particular, the property of having compact support (see e.g. Reference [20]), and the number of the B-splines is

$$N = l + k - 1 \tag{3}$$

The B-splines of order 1 are step functions defined by

$$B_i^1(x) = \begin{cases} 1 & \text{if } x \in [t_i, t_{i+1}] \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

and an efficient construction of the B-splines of order $k \geqslant 2$ is given by the recurrence relation of Curry and Schoenberg (see e.g. Reference [20]):

$$B_i^k(x) = \frac{x - t_i}{t_{i+k-1} - t_i} B_i^{k-1}(x) + \frac{t_{i+k} - x}{t_{i+k} - t_{i+1}} B_{i+1}^{k-1}(x) \tag{5}$$

Formulae (4) and (5) introduce the set of knots

$$X = \{t_i,\ i = 1,\ldots,N+k\} \tag{6}$$

which enforce the regularity of the B-spline basis by requiring

$$t_{k+i-1} = \xi_i \quad \text{for } i = 2,\ldots,l \tag{7}$$

i.e. the knots coincide with the breakpoints in the interior of the domain.

The construction of the basis given by Equations (4)–(7) leaves freedom in the first k and last k of the knots. A convenient choice for the approximation to boundary value problems is to set these end-knots as

$$t_1 = \cdots = t_k = a, \quad t_{N+1} = \cdots = t_{N+k} = b \tag{8}$$

In that case, by using basic properties [20] such as that of compact support,

$$B_i^k(x) = 0 \quad \text{for } x \notin [t_i, t_{i+k}] \tag{9}$$

and partition of unity,

$$\sum_{i=1}^{N} B_i^k(x) = 1 \quad \text{for } x \in [a, b] \tag{10}$$

the spline function (2) satisfies

$$\tilde{f}(a) = \alpha_1 \quad \text{and} \quad \tilde{f}(b) = \alpha_N \tag{11}$$

so that Dirichlet boundary conditions are imposed strongly.

Periodic boundary conditions are imposed by setting

$$t_k = a \tag{12a}$$

and performing a periodic extension of the interior knots (7) as

$$t_i = t_{i+N-k+1} - (b-a) \quad \text{for } i = 1, \ldots, k-1 \tag{12b}$$

$$t_i = t_{i-N+k-1} + (b-a) \quad \text{for } i = N+1, \ldots, N+k \tag{12c}$$

The periodicity of the spline function is then enforced by requiring periodicity on its coefficients, i.e. for the last $k-1$ coefficients with indices $i = N-k+2, \ldots, N$,

$$\alpha_i = \alpha_{i-N+k-1} \tag{13}$$

A useful property is that a B-spline basis of order k can represent elements of the space $\mathbb{P}_k(\Lambda)$, i.e. polynomials of degree $k-1$ at most. More precisely, Lyche and Schumaker [22] established the identity

$$\sum_{i=1}^{N} \gamma_{im} B_i^k(x) = x^{m-1}, \quad m = 1, 2, \ldots, k, \tag{14a}$$

where

$$\gamma_{im} = (-1)^{m-1} \frac{(m-1)!}{(k-1)!} \psi_i^{(k-m)}(0), \quad \text{with } \psi_i(x) = \prod_{p=1}^{k-1} (x - t_{i+p}) \tag{14b}$$

In the following, the superscript referring to the order of the B-splines will be dropped for the sake of brevity.

2.2. Semi-consistent fractional step scheme

Numerical approximation to the unsteady Navier–Stokes equations for a viscous incompressible fluid is performed in the domain $\Omega =]0,1[^2$. For ease of discussion, homogeneous Dirichlet boundary conditions are imposed upon the velocity.

In order to obtain a B-spline approximation that is not plagued by spurious pressure modes (see e.g. Reference [14]), the staggered B-spline collocation discretization introduced in References [7, 8] is used. In this method, the velocity and the pressure are represented with distinct bases as

$$\boldsymbol{v} = \sum_{i,j=1}^{N} \boldsymbol{v}_{i,j} B_i(x) B_j(y), \quad p = \sum_{i,j=1}^{N-2} p_{i,j} \tilde{B}_i(x) \tilde{B}_j(y) \tag{15}$$

where $\{B_i(x),\ i=1,\ldots,N\}$ is the velocity basis of order k with knots X_v, while the pressure basis $\{\tilde{B}_i(x),\ i=1,\ldots,N-2\}$ is of order $k-1$ with knots X_p staggered with respect to X_v.

The equations are discretized on the collocation grid $\{(x_i, y_j);\ i,j=1,\ldots,N\}$ that will be defined later.

The time-integration is based on the following prototype fractional step scheme, where the non-linear terms are discarded,

$$\mathscr{M} \frac{\bar{U} - U^n}{\Delta t} - \mathscr{K} \bar{U} + \mathscr{M} \mathscr{M}_A^{-1} \tilde{\mathscr{D}} P^n = F^{n+1} \tag{16a}$$

$$\bar{U}_{|\partial\Omega} = 0 \tag{16b}$$

and

$$\mathscr{M} \frac{U^{n+1} - \bar{U}}{\Delta t} + \mathscr{M} \mathscr{M}_A^{-1} \tilde{\mathscr{D}} (P^{n+1} - P^n) = 0 \tag{17a}$$

$$\mathscr{D} U^{n+1} = 0 \tag{17b}$$

$$U^{n+1}_{|\partial\Omega} = 0 \tag{17c}$$

In these equations, Δt is the time step, U and P are vectors representing the unknown spline coefficients of the velocity and the pressure, respectively, F is a source term, $\mathscr{M}$ is the (non-diagonal) mass matrix, $\mathscr{K}$ is the viscous diffusion matrix, $\mathscr{D}$ and $\tilde{\mathscr{D}}$ represent first derivative operators of velocity and pressure, respectively. The prediction step (16) amounts to solving a discrete Helmholtz equation for the provisional velocity $\bar{U}$. To obviate the need of artificial boundary conditions on the pressure, the projection step (17) is left written as a Div–Grad problem (see e.g. Reference [24]) instead of casting it as a Poisson equation for the pressure with Neumann conditions at the boundary.

In contrast to the standard B-spline discretization considered in References [7, 8], which will be referred to as the consistent method (CM), this scheme considers a modification of the pressure gradient in Equations (16a) and (17a) by introducing the matrix $\mathscr{M}_A^{-1}$ which is the approximate inverse of the mass matrix $\mathscr{M}$, in a sense to be defined later. The divergence equation (17b) is identical for both methods, and expresses the fact that the continuity condition be satisfied at the inner collocation points. Note that when $\mathscr{M}_A^{-1} = \mathscr{M}^{-1}$, the semi-consistent scheme (SCM) (16)–(17) reduces to the original CM scheme.

The main interest of the SCM scheme is that the projection step (17) yields the pressure equation

$$\mathscr{A}_A (P^{n+1} - P^n) = \mathscr{D} \bar{U} / \Delta t \tag{18}$$

where the pressure operator

$$\mathscr{A}_A = \mathscr{D}\mathscr{M}_A^{-1}\tilde{\mathscr{D}} \tag{19}$$

is sparse when $\mathscr{M}_A^{-1}$ is sparse, resulting in a pressure equation that can be efficiently solved by standard iterative methods for elliptic problems.

Scheme (16)–(17) was introduced under the name 'projection 2' by Gresho and Chan [18] for the finite-element method with $\mathscr{M}_A^{-1} = \mathscr{M}_L^{-1}$, i.e. an approximate inverse generated by the lumped approximation which is, in general, a first-order approximation to the mass matrix. We refer to References [18] and [14] for further analysis of this scheme.

The use of a highly accurate approximate inverse is motivated by investigating the truncation error of the SCM scheme. When combining Equations (16a) and (17a) to eliminate the provisional velocity $\bar{U}$, we get

$$\mathscr{M}\frac{U^{n+1}-U^n}{\Delta t} - \mathscr{K}U^{n+1} + \tilde{\mathscr{D}}P^{n+1} + \mathscr{E}_S + \mathscr{E}_A = F^{n+1} \tag{20}$$

where, in addition to the $O(\Delta t^2)$ splitting error

$$\mathscr{E}_S = -\Delta t\,\mathscr{K}\mathscr{M}_A^{-1}\tilde{\mathscr{D}}(P^{n+1}-P^n) \tag{21}$$

common to fractional-step schemes, the approximation error

$$\mathscr{E}_A = (\mathscr{M}-\mathscr{M}_A)\mathscr{M}_A^{-1}\tilde{\mathscr{D}}P^{n+1} \tag{22}$$

is a spatial error expressing the degree of accuracy to which $\mathscr{M}_A$ approximates the consistent mass matrix $\mathscr{M}$. The use of an approximate inverse whose accuracy is consistent with the B-spline discretization is thus mandatory for preserving the accuracy of the SCM scheme.

3. CONSTRUCTION OF APPROXIMATE INVERSE OF THE MASS MATRIX USING LOCAL SPLINE APPROXIMATION

3.1. Consistent interpolation vs local interpolation

The consistent interpolation of a function $f(x)$ consists in finding a spline function

$$\tilde{f}(x) = \sum_{i=1}^{N} \alpha_i(f)\,B_i(x) \tag{23}$$

that takes on the values of $f(x)$ at a given set of collocation points $\{x_j,\ j=1,\ldots,N\}$, i.e.

$$\sum_{i=1}^{N} \alpha_i(f)\,B_i(x_j) = f(x_j), \quad \text{for } j=1,\ldots,N \tag{24}$$

This linear system takes the matrix form

$$\bar{\mathscr{M}}\alpha = F \tag{25}$$

where $\alpha = (\alpha_1(f),\ldots,\alpha_N(f))$, $F = (f(x_1),\ldots,f(x_N))$ and $\bar{\mathscr{M}} = (B_i(x_j))_{i,j=1,\ldots,N}$ is the consistent, non-diagonal mass matrix of bandwidth k (for clarity, the matrix operators corresponding

to one-dimensional spline discretization are overlined). The solution of a linear system of equations is thus needed for determining the spline coefficients.

In contrast, local interpolation methods were developed in e.g. References [20–22] such that the determination of the coefficients does not require solving a collocation system. These methods are local in the sense that the evaluation of the coefficients depends on the value of the function and/or its derivatives at a small number of data points. In the following, we focus on local schemes involving function values, i.e. schemes such that the ith spline coefficient is determined as

$$\alpha_i(f) = \sum_{j=1}^{k_1} \beta_{ij} f(\tau_{ij}) \tag{26}$$

where, for $i = 1, \ldots, N$, $\{\tau_{ij},\ j = 1, \ldots, k_1\}$ is a given set of distinct data locations in Λ, $k_1 \leqslant k$ is the number of data points used for the evaluation of each spline coefficient, and will be referred to as the order of the local scheme, and $\{\beta_{ij},\ j = 1, \ldots, k_1\}$ are coefficients to be determined.

In the case where the data points are chosen from the set of collocation points, Scheme (26) can be written in matrix form as

$$\alpha = \bar{\mathscr{M}}_A^{-1} F \tag{27}$$

where the coefficients $\{\beta_{ij}\}$ in (26) define the entries of the square matrix $\bar{\mathscr{M}}_A^{-1}$ that is precisely the approximate inverse of the consistent mass matrix $\bar{\mathscr{M}}$ we are seeking. The matrix $\bar{\mathscr{M}}_A^{-1}$ is sparse, each of its row possessing k_1 non-zero entries at most. Thus, the linear system solution required by the consistent approximation is now replaced by a sparse matrix-vector multiplication involving the same right-hand-side.

3.2. Derivation of the local interpolant

The local interpolation scheme we use for the determination of the entries of $\bar{\mathscr{M}}_A^{-1}$ is the scheme based on point evaluations considered in Example 3.4 of Lyche and Schumaker's paper [22]. This construction is valid for B-spline bases of any order k, and an arbitrary distribution of knots.

Given $k_1 \leqslant k$ and some data points $\{\tau_{ij}; i = 1, \ldots N,\ j = 1, \ldots, k_1\}$ such that $\{\tau_{i1}, \ldots, \tau_{ik_1}\}$ are distinct for all i, the coefficients $\{\beta_{ij}; i = 1, \ldots, N,\ j = 1, \ldots, k_1\}$ in (26) are determined so that the local scheme reproduces polynomial of order k_1, i.e.

$$\tilde{f} = f, \quad \forall f \in \mathbb{P}_{k_1}(\Lambda), \quad k_1 \leqslant k \tag{28}$$

For this purpose, it is convenient to write Eq. (26) as

$$\alpha_i(f) = \sum_{j=1}^{k_1} \mu_{ij} [\tau_{i1}, \tau_{i2}, \ldots, \tau_{ij}] f \tag{29}$$

where $[\cdot, \ldots, \cdot] f$ represents the divided difference of $f(x)$, defined by the recursion formulae (see e.g. Reference [20])

$$[\tau_{i1}, \tau_{i2}, \ldots, \tau_{ij}] f = \frac{[\tau_{i1}, \tau_{i2}, \ldots, \tau_{ij-1}] f - [\tau_{i2}, \tau_{i3}, \ldots, \tau_{ij}] f}{\tau_{i1} - \tau_{ij}}, \quad [\tau_{i1}] f = f(\tau_{i1})$$

Condition (28) amounts to representing each monomial x^{m-1}, $m=1,\ldots,k_1$, as

$$\sum_{i=1}^{N}\left(\sum_{j=1}^{k_1}\mu_{ij}\,[\tau_{i1},\tau_{i2},\ldots,\tau_{ij}]\,x^{m-1}\right)B_i(x)=x^{m-1} \tag{30}$$

By using (14), the coefficients μ_{ij} are obtained as the solution of the lower triangular linear system, for each $i=1,\ldots,N$

$$\sum_{j=1}^{k_1}\mu_{ij}[\tau_{i1},\tau_{i2},\ldots,\tau_{ij}]x^{m-1}=\gamma_{im},\quad \text{for } m=1,\ldots,k_1 \tag{31}$$

from which the μ_{ij} can be obtained by back-solution. The values of μ_{ij} for $j=1,\ldots,4$ are given in Reference [22], and are listed here for completeness:

$$\mu_{i1}=1 \tag{32a}$$

$$\mu_{i2}=\gamma_{i2}-\tau_{i1} \tag{32b}$$

$$\mu_{i3}=\gamma_{i3}-(\tau_{i1}+\tau_{i2})\mu_{i2}-\tau_{i1}^2 \tag{32c}$$

$$\mu_{i4}=\gamma_{i4}-(\tau_{i1}+\tau_{i2}+\tau_{i3})\,\mu_{i3}-(\tau_{i1}^2+\tau_{i1}\tau_{i2}+\tau_{i2}^2)\mu_{i2}-\tau_{i1}^3 \tag{32d}$$

Due to the structure of system (31), the values of the $\{\mu_{ij}\}$ do not depend on k_1. Thus, the coefficients of a scheme of order $k_1\leqslant 4$ are given by the first k_1 lines in Equation (32). The extra effort to determine the coefficients of a scheme of order 5 would only be to calculate μ_{i5} from (31). Once the μ_{ij} are determined, the β_{ij} can be obtained by equating (29) to (26).

Lyche and Schumaker [22] have shown that this local scheme yields an accuracy of order k_1 in the maximum norm. In particular, for the case $k_1=k$, it is thus possible to obtain a local interpolant that preserves the order of accuracy of the consistent approximation.

Another important case occurs for $k_1=2$: if, to obtain the ith B-spline coefficient, the first data point τ_{i1} is chosen to be equal to $x_i^\star$ defined as

$$x_i^\star=\gamma_{i2},\quad \text{where from (14}b\text{)}\quad \gamma_{i2}=\sum_{p=1}^{k-1}t_{i+p}/(k-1) \tag{33}$$

then $\mu_{i2}=0$ in Equation (32b) for any choice of the second data point τ_{i2}. Thus, the spline function

$$\tilde{f}(x)=\sum_{i=1}^{N}f(x_i^\star)\,B_i(x) \tag{34}$$

is a second-order approximation to $f(x)$. This scheme is precisely the variation-diminishing approximation of Marsden and Schoenberg (see e.g. Reference [20]), and the collocation points (33) will subsequently be referred to as the Marsden–Schoenberg points.

A last important remark concerns the accuracy of the imposition of Dirichlet boundary conditions with the local scheme, in the case where the end-knots are set using Equations (8). From Equations (11) and (29), the value of the spline at the end-point $x=a$ is

$$\tilde{f}(a)=\alpha_1(f),\quad \text{with } \alpha_1(f)=\sum_{j=1}^{k_1}\mu_{1j}[\tau_{11},\ldots,\tau_{1j}]f \tag{35}$$

where the coefficients $\{\mu_{1j},\ldots,\mu_{1k_l}\}$ are solution to

$$\sum_{j=1}^{k_l}\mu_{1j}[\tau_{11},\ldots,\tau_{1j}]x^{m-1}=a^{m-1},\quad m=1,\ldots,k_l \tag{36}$$

If τ_{11} is chosen to be equal to a, the unique solution is then $\mu_{11}=1,\ \mu_{12}=\cdots=\mu_{1k_l}=0$. From Equation (35) we get $\tilde{f}(a)=f(a)$, and correspondingly $\tilde{f}(b)=f(b)$ at the other end-point, showing that Dirichlet boundary conditions are satisfied exactly just as in the consistent approximation (Equation (11)).

3.3. *Approximate inverse of the mass matrix*

The local interpolant allows us to build an approximate inverse $\bar{\mathscr{M}}_A^{-1}$ of the mass matrix when the data points $\{\tau_{i,j}\}$ are chosen from the set of collocation points $\{x_i,\ i=1,\ldots,N\}$. For the imposition of Dirichlet conditions with the end-knots (8), an additional constraint would be to set $\tau_{11}=a$ and $\tau_{N1}=b$, the choice of the remaining data points $\{\tau_{1j},\ j=2,\ldots,k_l\}$ and $\{\tau_{Nj},\ j=2,\ldots,k_l\}$ having no consequence.

A case of special interest arises when the approximate inverse reduces to $\bar{\mathscr{M}}_A^{-1}=\mathscr{I}$. This approximate inverse corresponds to the mass matrix lumped with the row-sum technique widely used in the finite-element community (see e.g. References [14, 15]), obtained by summing the rows of the consistent mass matrix $\bar{\mathscr{M}}$, putting the result on the diagonal and using the property of partition of unity (10). For the spline-collocation method, the local interpolant of order $k_l=1$ with $\tau_{i1}=x_i$ generates such a matrix, and thus yields first-order accuracy in general. However, as shown by the variation-diminishing scheme (34), the lumped mass matrix is second-order accurate on nonuniform grids (i.e. for nonuniform distributions of knots) when the Marsden–Schoenberg collocation points (33) are used.

The latter case identifies the Marsden–Schoenberg points as an alternative to the usual choice of the collocation points, i.e. the location of the maximum of the B-splines. These two definitions are equivalent in particular cases only, such as a periodic domain with uniform knots. They nonetheless yield the same characterization of the first and last collocation points when the end-knots (8) are used, namely $x_1=a$ and $x_N=b$. In the following, the Marsden–Schoenberg points will be used as much as possible, even though no advantages have been observed yet when local interpolants of order higher than 2 are employed.

As sketched in Figure 1, the approximate inverse generated by a local scheme of order $k_l\geqslant 3$ can be viewed as a high-order generalization of the mass lumping technique. Since $\bar{\mathscr{M}}^{-1}$ is dense, the evaluation of the spline coefficient $\alpha_i(f)$ by the consistent approximation involves values of $f(x)$ at all collocation points. In contrast, the mass lumping technique consists in identifying $\alpha_i(f)$ to the value of $f(x)$ at the collocation point associated with the ith B-spline. More generally, the local approximation to $\alpha_i(f)$ involves values of $f(x)$ at several collocation points, which are located in Figure 1 in the support of $B_i(x)$. This approximation has the effect of increasing the bandwidth of $\bar{\mathscr{M}}_A^{-1}$ while raising the accuracy of the evaluation of $\alpha_i(f)$.

Several issues have to be addressed for generating approximate inverse of practical interest. A loss in the spatial resolution power of the spline-collocation method would predictabily result from the replacement of the consistent mass by the local mass approximation. Furthermore, while $O(N^{-k_l})$ asymptotical accuracy is assured when k_l data points per coefficient are used, the definition of the local interpolant leaves freedom in their positioning. The influence

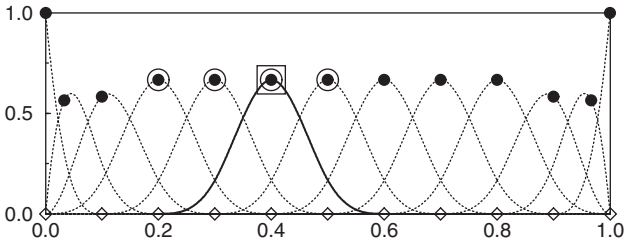

Figure 1. Sketch to illustrate the approximation of a function $f(x)$ with a B-spline basis of order $k=4$, on $l+1=10$ equidistant breakpoints ($\diamond$) with end-knots (8). The coefficient associated to the 6th B-spline (——) is evaluated with the values of $f(x)$ at: all Marsden–Schoenberg collocation points ($\bullet$) with the consistent approximation, a unique collocation point (framed) with the mass lumping approximation, and 4 collocation points (circled) when a local scheme of order $k_l=4$ is used.

of the location of the data points on the accuracy of the resulting local scheme is investigated in the next section.

4. NUMERICAL RESULTS

This section is devoted to the assessment of the accuracy of the sparse approximate inverses and the application to the Navier–Stokes equations. All computations are carried out on the collocation grid defined by the Marsden–Schoenberg points (33), thus yielding second-order accuracy to the lumped approximation of the mass matrix.

4.1. Local approximation in a one-dimensional periodic domain

It is convenient to analyze the resolution properties of the local approximation for B-spline bases on a uniform distribution of breakpoints with periodic boundary conditions, where the end-knots are set as (12). The investigation of the influence of the position of the data points used for generating $\bar{\mathscr{M}}_A^{-1}$ is then greatly simplified since, in this configuration, the bases are generated by translation of the same cardinal B-spline [20]. As a result, the ith Marsden–Schoenberg point is characterized as the maximum of the spline $B_i(x)$, i.e. for k even

$$x_i^\star = t_{i+k/2}, \quad i=1,\ldots,N \tag{37}$$

The matrix $\bar{\mathscr{M}}_A^{-1}$ is then a banded circulant matrix whose entries have the form

$$(\bar{\mathscr{M}}_A^{-1})_{ij} = (m_A^{-1})_{j-i \bmod N} \tag{38}$$

This configuration will allow us to perform a modified wavenumber analysis of the semi-consistent schemes.

The influence of the choice of the data points on the accuracy of the resulting schemes is performed for local interpolant of order $k_l=4$. For completeness, results are also provide for the popular variation-diminishing scheme (i.e. mass lumping approximation, $k_l=2$). Table I displays the various sets of data points that we consider. Those sets of points are located as close as possible to the support $[t_i, t_{i+k}]$ of $B_i(x)$, in order to minimize the bandwidth of $\bar{\mathscr{M}}_A^{-1}$.

Table I. Description of the sets of data points used for local interpolation. The index i refers to the collocation point associated with the ith B-spline.

Set of data points	Index of data points used for local interpolation
Set I	$\{i\}$
Set II	$\{i-2, i-1, i, i+1\}$
Set II′	$\{i-1, i, i+1, i+2\}$
Set III	$\{i-3, i-2, i-1, i\}$
Set III′	$\{i, i+1, i+2, i+3\}$
Set IV	$\{i-2, i-1, i+1, i+2\}$

Table II. Entries $(\bar{\mathscr{M}}_A^{-1})_{ij} = m^{-1}_{j-i \bmod N}$ of the approximate inverse for the various sets of data points.

	Order $k=4$							Order $k=6$						
Set of data points	m_{-3}^{-1}	m_{-2}^{-1}	m_{-1}^{-1}	m_0^{-1}	m_1^{-1}	m_2^{-1}	m_3^{-1}	m_{-3}^{-1}	m_{-2}^{-1}	m_{-1}^{-1}	m_0^{-1}	m_1^{-1}	m_2^{-1}	m_3^{-1}
Set I				1							1			
Set II		0	−1/6	4/3	−1/6				0	−1/4	3/2	−1/4		
Set II′			−1/6	4/3	−1/6	0				−1/4	3/2	−1/4	0	
Set III	1/6	−2/3	5/6	2/3				1/4	−1	5/4	1/2			
Set III′				2/3	5/6	−2/3	1/6				1/2	5/4	−1	1/4
Set IV		−2/9	13/18		13/18	−2/9			3/4	−1/4		−1/4	3/4	

The entries of $\bar{\mathscr{M}}_A^{-1}$ with respect to the sets of data points are given in Table II for splines of order $k=4$ and 6.

Set I corresponds to the variation-diminishing scheme, while the five other sets use the local interpolant of order 4.

Sets II and II′ represent the two possible choices of data points that yield approximate inverses with the shortest bandwidth. For the particular case of equidistant knots with periodicity conditions considered in this section, these two sets happen to generate the same symmetric tridiagonal matrix $\bar{\mathscr{M}}_A^{-1}$ (see Table II). In general, this property is lost when the knots are not distinct and equally spaced. As an example, for a B-spline basis on a uniform distribution of breakpoints with Dirichlet boundary conditions, the tridiagonality of $\bar{\mathscr{M}}_A^{-1}$ is lost at its first and last $k-1$ lines, due to the fact that the end-knots are identical (see Equation (8)).

Sets III and III′ correspond, respectively, to a left and right biasing of the data points with respect to $x_i^\star$. These two sets are the only ones that introduce a non-symmetric approximate inverse (see Table II). As it will be seen shortly, this feature has important effects on the nature of the differencing errors of the semi-consistent schemes generated by these two sets.

Finally, set IV is designed to generate an approximate inverse with a sparsity pattern that is symmetric on arbitrary grids, by not considering the datum at point $x_i^\star$.

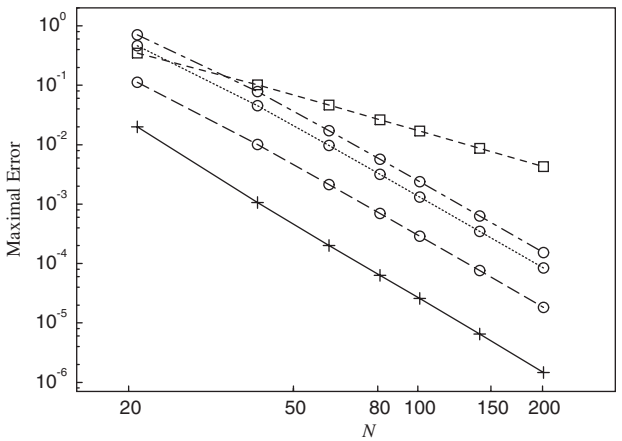

Figure 2. Maximal error vs N, number of B-splines of order $k=4$, for the various approximation methods. Local scheme with $k_l=2$ (□): set I —-; local schemes with $k_l=4$ (○) : set II – – –, set III ....., set IV —·—· ; consistent approximation ——.

The first numerical test concerns the interpolation of the periodic function

$$f(x)=\sin(10\pi x)+\cos(2\pi x+2), \quad \text{in } \Lambda=]0,1[\tag{39}$$

We recall that the consistent approximation requires a linear system solution, while the semi-consistent schemes require a single sparse matrix-vector multiplication. The maximal value of the error, sampled on a fine grid of 1001 equidistant points, is displayed in Figure 2 for splines of order $k=4$. The poor accuracy of the lumping approximation is obvious, yielding second-order accuracy. The order of accuracy of the consistent approximation is recovered for all the local schemes of order 4. These results show the importance of data point positions. Not surprisingly, the lowest error of the local schemes is obtained with set II. Note also that for a moderate spatial resolution ($N\leqslant 25$), sets III and IV yield results inferior to those obtained with the lumping approximation. For this interpolation test, sets III and III′ give identical results.

For completeness, Figure 3 displays analogous results obtained with splines of order 6. As in the previous case, fourth-order accuracy is obtained with local schemes of order 4 and, again, set II displays the lowest magnitude error. The results obtained with these local schemes are, of course, far from the 6th-order accuracy of the consistent approximation. This rate of convergence would nonetheless be obtained with local schemes of order 6.

A classical evaluation of the resolving abilities of a numerical scheme is given by the Fourier analysis of differentiating errors (e.g. References [25, 26]). A complete analysis of the resolving power of (consistent) B-spline methods is performed by Kwok *et al.* [27]. In the following, we focus on investigating the resolving power of the semi-consistent approximation to the first derivative, since the gradient of the pressure only is altered by the introduction of the approximate inverse in the SCM fractional scheme (16)–(17).

Following Kwok *et al.* [27], we consider the eigenvalue problem

$$u'=\lambda u \quad \text{in } \Lambda=]-\pi,\pi[\tag{40}$$

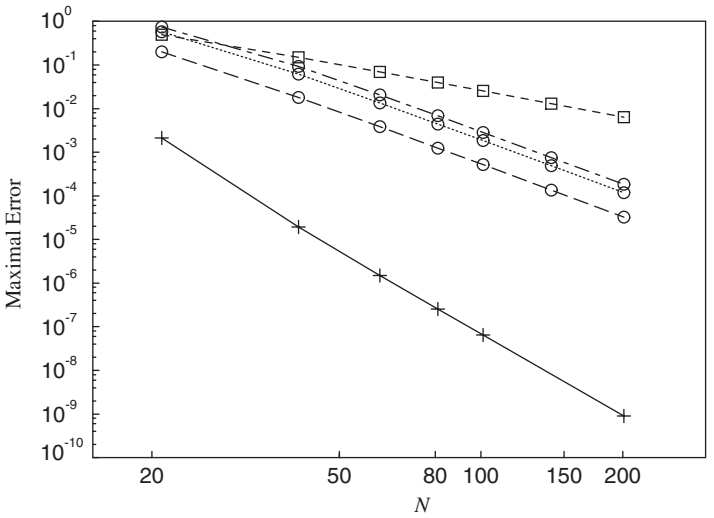

Figure 3. Maximal error vs the number of points N with splines of order $k=6$. See the caption of Figure 2 for the labelling.

with periodic boundary conditions, to be solved on a uniform mesh with grid spacing $h=2\pi/N$. The discretization of this problem by the consistent approximation leads to the generalized eigenvalue problem

$$\bar{\mathscr{D}}\alpha=\lambda\bar{\mathscr{M}}\alpha \tag{41a}$$

where $\bar{\mathscr{D}}=(B_i'(x_j^{\star}))$ is the first derivative collocation operator. Correspondingly, the semi-consistent discretization leads to

$$\bar{\mathscr{M}}_A^{-1}\bar{\mathscr{D}}\alpha=\lambda\alpha \tag{41b}$$

The eigenvalue spectrum of (41), that reads

$$\lambda=\mathrm{i}\omega'(\omega),\quad \mathrm{i}^2=-1$$

with $\omega=jh$ for $j=0,\ldots,N-1$, corresponds to the B-spline differencing error, which has to be compared with exact differentiation, $\lambda=\mathrm{i}\omega$.

We recall that the real and complex part of the modified spectrum ω' are, respectively, associated with errors of dispersive and dissipative nature. Consistent B-spline approximations yield centered differencing schemes (i.e. $\bar{\mathscr{M}}$ and $\bar{\mathscr{D}}$ are symmetric) with purely real ω', yielding thus errors of dispersive nature only. Similarly, semi-consistent approximations with sets I, II and IV are purely dispersive since $\bar{\mathscr{M}}_A^{-1}\bar{\mathscr{D}}$ is symmetric. On the contrary, the biasing of the data points in sets III and III$'$ generates forward and backward differencing schemes, respectively, yielding thus modified wavenumber with non-zero imaginary part.

For the various discretizations with splines of order $k=4$ and 6, respectively, Figures 4 and 5 sketches the real part of the modified wavenumber spectrum $\omega'(\omega)$ versus the wavenumber ω. As expected, the mass lumping approximation (set I) degrades the resolving power

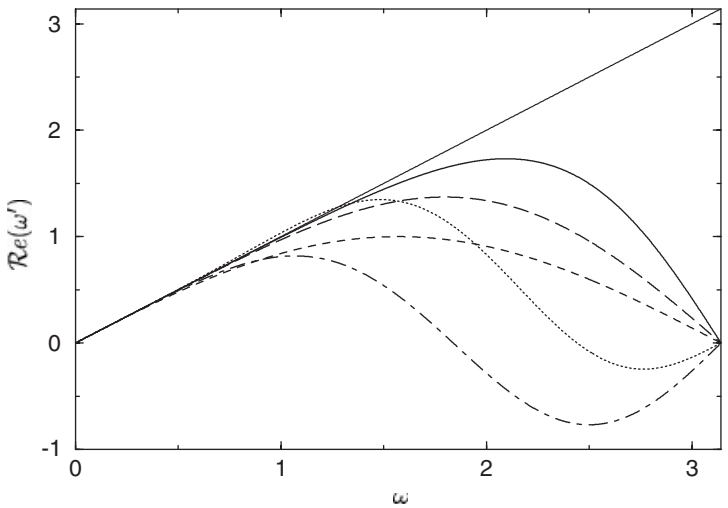

Figure 4. Real part of modified wavenumber of the first derivative yielded by the different methods for splines of order $k=4$. Local scheme with $k_l=2$: set I- - - -; local schemes with $k_l=4$: set II– – – –, set III ·····, set IV —·—; consistent scheme ——; exact ——.

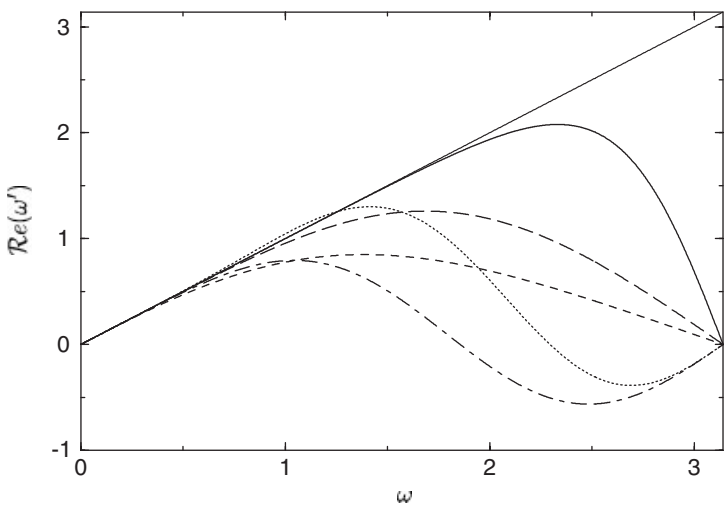

Figure 5. Real part of modified wavenumber of the first derivative for splines of order $k=6$. See the caption of Figure 4 for the labelling.

compared to the consistent approximation: as an example, set I with $k=4$ yields only the resolving power of the common second-order centered finite differencing. For the local schemes of order 4, the location of the data points has a great influence on their resolving ability, yielding very disparate wavenumber plots. It is particularly striking that set IV gives lower resolving power than even the lumped mass approximation.

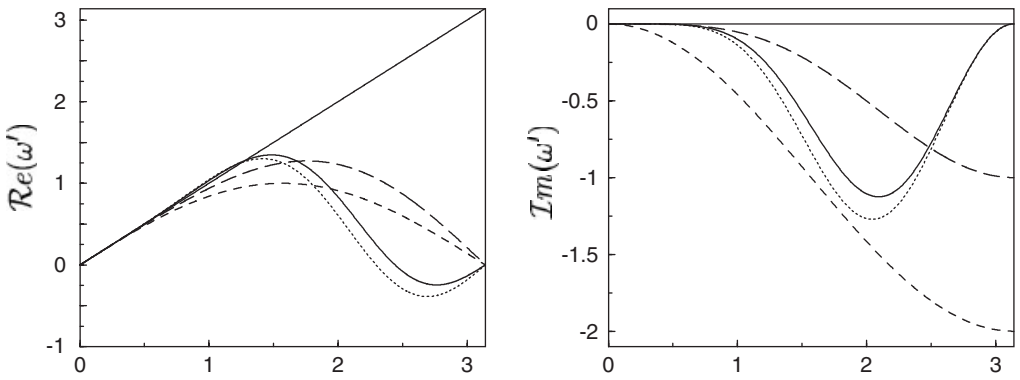

Figure 6. Real part (left) and imaginary part (right) of the modified wavenumber for: local scheme with set III, $k=6$—; local scheme with set III; $k=4\cdots\cdots$; upwind scheme: – – – –; Quick scheme — — —; exact ——.

Table III. Resolving efficiency r_ε of the first derivative for the various approximations.

	Order $k=4$			Order $k=6$		
Schemes	$\varepsilon=0.1$	$\varepsilon=0.01$	$\varepsilon=0.001$	$\varepsilon=0.1$	$\varepsilon=0.01$	$\varepsilon=0.001$
Consistent	0.59	0.36	0.20	0.73	0.55	0.39
Set I	0.25	0.08	0.02	0.21	0.06	0.02
Set II	0.44	0.24	0.13	0.40	0.21	0.12
Set III	0.31	0.17	0.09	0.29	0.15	0.09
Set IV	0.27	0.14	0.08	0.26	0.14	0.08

The imaginary part of $\omega'(\omega)$ is displayed in Figure 6 for the schemes obtained with set III and, for comparison purpose, for the first-order upwind biased finite differencing and the second-order Quick scheme [28]. For completeness, we have also plotted the corresponding real part. This figure shows that numerical dissipation is introduced in the semi-consistent schemes with sets III and III′, with a significant portion of high wavenumbers being artificially damped, as observed with the upwind finite difference schemes.

A more quantitative measurement of the resolving ability of a scheme is the resolving efficiency introduced by Lele [26]. The resolving efficiency r_ε is defined as the fraction of wavenumbers ω verifying

$$\frac{|\omega'(\omega)-\omega|}{\omega}\leqslant\varepsilon \tag{42}$$

for a given value of ε, i.e. the fraction of the entire range of wavenumbers that are accurately represented within a relative error tolerance of ε. The resolving efficiency r_ε of the various schemes is given in Table III for representative values of ε. We observe that this measurement gives, once more, set II as the most accurate of the local schemes. In particular, for splines of order $k=4$, set II recovers up to 76, 67 and 65% of the resolved fraction r_ε of

the consistent approximation for $\varepsilon = 10^{-1}$, 10^{-2} and 10^{-3}, respectively. In comparison, the lumping approximation recovers only 42, 22 and 10% of this resolved fraction.

In summary, accurate approximate inverses that preserve the order of accuracy of the consistent approximation can be generated when a sufficient number of data points is used. The location of these data points is nonetheless of crucial importance. These numerical tests have found the best position of these data points on a periodic uniform grid to be set II. The resolution properties of the consistent approximation are, nonetheless, not fully recovered by the local schemes. The marginal loss in accuracy is greatly counterbalanced by their much lower computational cost, as will be illustrated in the next section.

In the case of Dirichlet boundary conditions, set II is used as a template for building approximate inverses of order $k_l = 4$ on a uniform distribution of breakpoints. The local approximation for coefficients with indices $i = k, \ldots, N-k+1$ is performed with this set, generating a centered tridiagonal approximation as in the periodic case. A modification has to be performed for the first and last $k-1$ deficient B-splines, i.e. those having multiple knots in their support. The first and last coefficients use the datum at the end-point a and b, respectively, leading to the exact imposition of the Dirichlet conditions. For the remaining B-spline coefficients with indices $i = 2, \ldots, k-1$ and $i = N-k, \ldots, N-1$, we use the collocation points with indices $\{i-1, i, i+1, i+2\}$ and $\{i-2, i-1, i, i+1\}$, respectively. This distribution of data points will be used for the remaining part of the paper.

4.2. Semi-consistent approximation of the Div-Grad problem

We are now ready to describe the semi-consistent approximation (SCM) of the projection step (17), by considering as model equations the Div-Grad problem:

$$\sigma \boldsymbol{v} + \nabla p = \mathbf{f} \quad \text{in } \Omega =]0,1[^2 \tag{43a}$$

$$\nabla \cdot \boldsymbol{v} = 0 \quad \text{in } \Omega =]0,1[^2 \tag{43b}$$

$$\boldsymbol{v} = \mathbf{g} \quad \text{on } \partial\Omega \tag{43c}$$

This new discretization follows essentially along the lines of the consistent method (CM) introduced in References [8, 9]. Equations (43a) and (43b) are evaluated on the $(N-2) \times (N-2)$ inner collocation points, while the remaining boundary points are used for the determination of the boundary conditions (43c). The discretization of the divergence equation (43b) is identical for both methods and reads in matrix form:

$$\mathscr{D} U = G \tag{44}$$

where the velocity coefficients determined from the boundary conditions are put in the right-hand side (RHS).

The SCM approximation to Equation (43a) is now described with some details related to the imposition of non-homogeneous boundary conditions. For this purpose, we denote by $\mathscr{I}_\mathrm{I}$ the set of indices of collocation points in the interior of domain, and correspondingly $\mathscr{I}_\mathrm{B}$ refers to the indices of the boundary nodes. The discretization of Equation (43a) at the interior collocation point (x_i, y_k) is

$$\sigma \sum_{(j,l) \in \mathscr{I}_\mathrm{I}} \mathscr{M}^A_{i,k,j,l} U_{j,l} + \sum_{(j,l) \in \mathscr{I}_\mathrm{I}} \tilde{d}_{i,k,j,l} P_{j,l} = F_{i,k} \tag{45a}$$

where $\{\tilde{d}_{i,k,j,l}\}$ are the gradient coefficients of the pressure spline, and

$$F_{i,k} = \mathbf{f}(x_i, y_k) - \sigma \sum_{(j,l)\in\mathscr{I}_B} \mathscr{M}^A_{i,k,j,l} U_{j,l} \tag{45b}$$

corresponds to a RHS augmented with boundary velocity coefficients. The matrix form of (45a) reads

$$\sigma U + \mathscr{M}_A^{-1}\tilde{\mathscr{D}}P = \mathscr{M}_A^{-1}F \tag{46}$$

where $\mathscr{M}_A^{-1}$ is constructed from the tensor product of one-dimensional matrices $\bar{\mathscr{M}}_A^{-1}$ in each spatial direction, which use the distribution of data points for Dirichlet conditions described in the previous section.

On the other hand, the entries $\{\mathscr{M}^A_{i,k,j,l}\}$ in Equation (45b) need to be determined for the imposition of non-homogeneous boundary conditions. For this purpose, the sum in (45b) is expanded as

$$\sum_{(j,l)\in\mathscr{I}_B} \mathscr{M}^A_{i,k,j,l} U_{j,l} = \sum_{(j,l)\in\mathscr{I}_B} \bar{\mathscr{M}}^A_{i,j}\bar{\mathscr{M}}^A_{k,l} U_{j,l} \tag{47}$$

by using tensor product properties. For $j = 1,\ldots,N$, the coefficients $\{\bar{\mathscr{M}}^A_{i,j};\ i = 1,\ldots,N\}$ in the x-direction are then obtained by solving the one-dimensional problems

$$\bar{\mathscr{M}}_A^{-1}\boldsymbol{x}^j = \delta^j, \quad \text{with } \delta^j = (\delta_{j,m};\ m = 1,\ldots,N) \tag{48}$$

giving as solution $\boldsymbol{x}^j = (\bar{\mathscr{M}}^A_{i,j};\ i = 1,\ldots,N)$. An analogous procedure is performed for determining the coefficients in the y-direction.

Equations (44) and (46) yield the pressure equation

$$\frac{1}{\sigma}\mathscr{A}_A P = \frac{1}{\sigma}\mathscr{D}\mathscr{M}_A^{-1}F - G \tag{49a}$$

where the sparse pressure operator is

$$\mathscr{A}_A = \mathscr{D}\mathscr{M}_A^{-1}\tilde{\mathscr{D}} \tag{49b}$$

As in the CM discretization [7, 8], this operator displays only one zero eigenvalue, showing that no spurious pressure modes occur in the SCM discretization. Furthermore, all other eigenvalues are distinct with negative real part, and the condition number scales like N^2 on uniform grids.

Raising the order of the approximate inverse increases the number of non-zero entries of $\mathscr{A}_A$. In two dimensions, when a natural ordering of the unknowns is used, a modification of the block structure of $\mathscr{A}_A$ is observed. As an illustration, Figure 7 compares the structure of pressure operators generated by the lumping approximation and the approximate inverse of order 4. These operator have similar block-structured pattern, where the 11×11 blocks are more or less filled according to the order of the approximate inverse. For splines of order 4, we observe that an approximate inverse of order 4 doubles the number of non-zero entries of the pressure operator compared to the lumped mass approximation. Correspondingly, an increase of 70% of the entries is observed for splines of order 6.

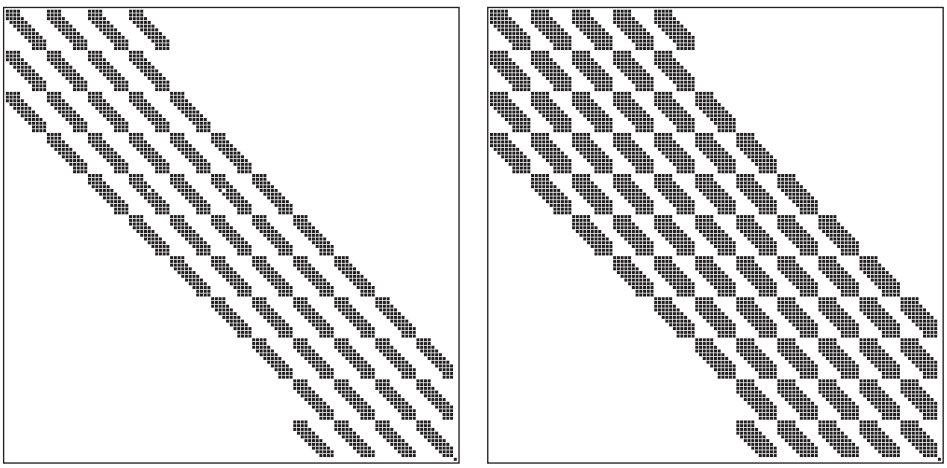

Figure 7. Sparsity pattern of $\mathscr{A}_A$ on a 13×13 uniform grid for splines of order $k = 4$, generated with the mass lumping approximation (left) and the approximate inverse of order 4 (right).

The semi-consistent method is now evaluated against the consistent method by solving numerically problem (43) with $\sigma = 1$, for the solution

$$\boldsymbol{v} = \mathbf{rot} \sin 4\pi x \sin 4\pi y, \quad p = \cos 4\pi x \cos 4\pi y$$

on a uniform distribution of knots. For comparison with results obtained in References [7, 8], the collocation points are set as the location of the maxima of the velocity B-splines. The maximum error on the first-component of the velocity u and the pressure, sampled on a 300×300 uniform grid, is reported in Figure 8. For splines of order 4 (Figure 8(a)), the use of an approximate inverse of order 4 maintains the order of accuracy of the consistent approximation, namely $O(N^{-4})$ for u and $O(N^{-2})$ for p. It is striking to observe that the magnitude of the error on p is almost identical for both schemes, while the velocity errors of the SCM scheme are only marginally higher. The latter is certainly the consequence of the inferior resolving power of the semi-consistent approximation that we observed in Section 4.1. As it would be expected, the 6th-order accuracy on u displayed in Figure 8(b) by the CM scheme with splines of order 6 is not recovered by the approximate inverse of order 4, and a fourth-order convergence rate is observed in this case.

It is valuable to compare the CPU cost required by the iterative solution of the CM and SCM equations. This comparison is restricted to the case $k = 4$, for which both methods display a similar asymptotic order of accuracy. To give a fair evaluation, both systems are solved by similar iterative techniques with the error tolerance (i.e. the l^2 norm of the discrete divergence (44)) set to $\varepsilon = 10^{-8}$, and are initialized with the initial guess set to zero.

Since the pressure operator of the CM discretization is dense and thus cannot be stored, the equations are solved by the Uzawa algorithm developed in References [7, 8], which is accelerated by the Bi-CGSTAB method (see e.g. References [12, 13]). The preconditioner of this system is $\mathscr{A}_L = \mathscr{D}\mathscr{M}_L^{-1}\tilde{\mathscr{D}}$, i.e. a SCM pressure operator where the lumping approximation is used. The use of this preconditioner has the effect to make the number of Uzawa iterations independent of the mesh size. Each step of the Uzawa algorithm requires inversions of the

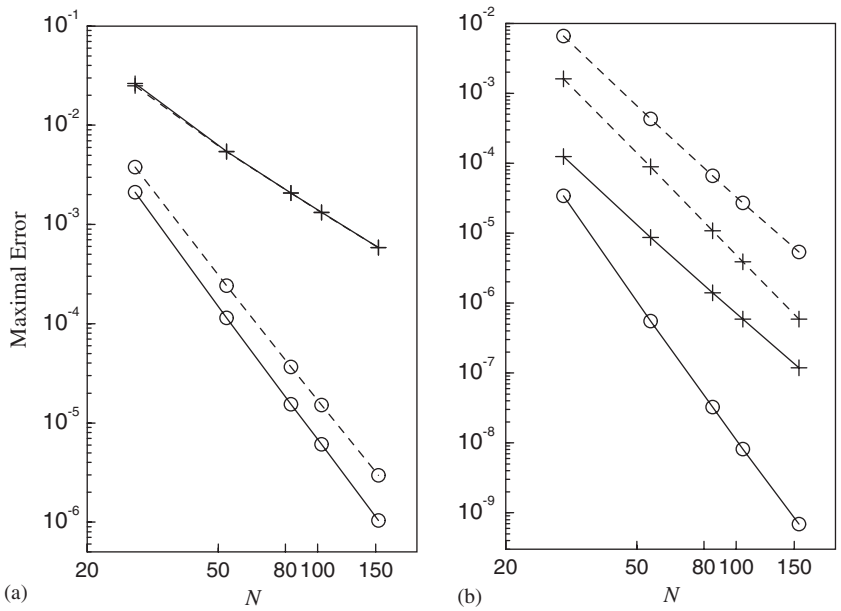

Figure 8. Solution of the Div–Grad problem with (*a*) splines of order 4, and (*b*) splines of order 6. Maximum error on u $\circ$—$\circ$, p +—+, CM approximation; maximum error on u $\circ$– – –$\circ$, p +– – –+, SCM approximation with approximate inverse of order 4.

consistent mass matrix and the preconditioner. These problems are respectively solved by a direct method and the Bi-CGSTAB algorithm with ILU(0) preconditioning.

The SCM system precludes the use of Uzawa iterations, resulting in a far less cumbersome solution procedure. The pressure equation (49) is solved with the same Bi-CGSTAB algorithm used for inverting the preconditioner $\mathscr{A}_L$ of the CM system. The velocity is then recovered by using Equation (46).

The computational efficiency of the methods is compared in Figure 9, where the CPU cost of the iterative solution of the linear systems is plotted against the maximal error obtained on the velocity and the pressure. The SCM method requires only a fraction of the CPU time of the CM method to reach the same level of accuracy: we observed here that the ratio is more than 25 for both velocity and pressure. Intuitively, these CPU savings can be understood when observing that the computational cost of the solution of the SCM pressure equation is roughly equivalent to a single preconditioner solve of the CM solution procedure. The ratio of savings is thus proportional to the number of Uzawa iterations required for convergence. Moreover, since it has been observed in References [7, 8] that the number of Uzawa iterations is independent of the grid size, but increases with the order of the B-splines, the SCM solution method would become more attractive as the order of the discretization is raised.

4.3. Navier–Stokes results

The SCM method is now applied to some benchmark Navier–Stokes applications. These important tests assess (a) the high spatial accuracy of the Navier–Stokes solver and (b) its

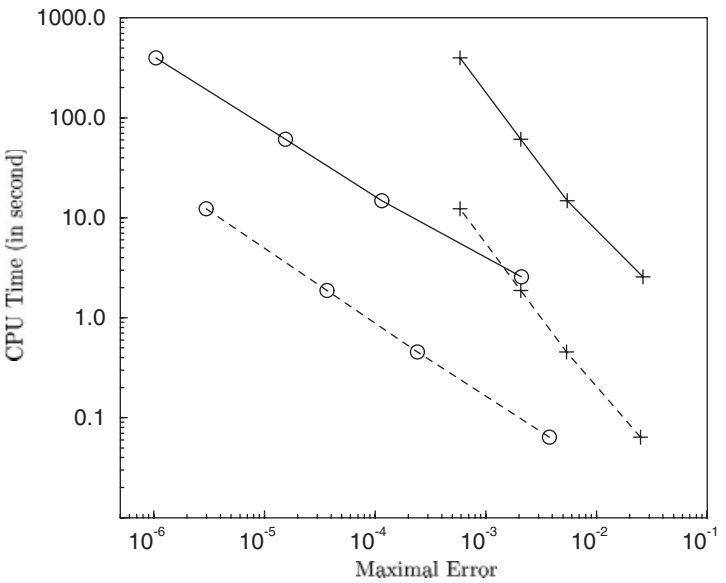

Figure 9. CPU time observed for the solution of the Div–Grad problem as a function of the maximal error for splines of order $k=4$: u ∘—∘, p +—+, CM approximation; u ∘– – –∘, p +– – –+, SCM approximation with approximate inverse of order 4.

robustness for unsteady computations. The time-integration is based on the semi-implicit scheme (16)–(17), with second-order backward-differentiation of the time-derivative and Adams–Bashforth discretization of the non-linear term written in the convective form. In association with the CM and SCM (with $\mathscr{M}_A^{-1}=\mathscr{M}_L^{-1}$) discretizations, this fractional-step scheme yields second-order time accuracy for both velocity and pressure [7, 8].

The spatial accuracy that is expected from the SCM scheme deserves some comments. When applied to differential problems, the collocation method we use, namely the smoothest spline collocation method, yields a suboptimal rate of convergence, which is generally $O(N^{d-k})$ for a d-order problem when B-splines of order k are employed (see e.g. References [7, 8, 29]). Since Equations (16) represents a Helmholtz problem for the provisional velocity, $O(N^{2-k})$ accuracy is expected for this quantity with velocity B-splines of order k. On the other hand, from the results of the previous section, it can be inferred that the SCM approximation for the projection step with an approximate inverse of order $k_1 \leqslant k$ yields $O(N^{-k_1})$ accuracy for the velocity and $O(N^{-\min\{k_1,k-2\}})$ accuracy for the pressure. As a result, the combination of both steps would yield a rate of convergence equal to $\min\{k_1, k-2\}$ for both velocity and pressure. Hence, in order to obtain a fourth-order accurate Navier–Stokes solver, B-splines of order $k=6$ are used, with the approximate inverse of order $k_1=4$ described in the previous sections. Note that a similar rate of convergence would have been obtained with the CM fractional step scheme using B-splines of the same order.

The first test concerns the validation of the spatial accuracy of the SCM scheme on a uniform grid. For this purpose, we consider the steady solution in $\Omega=]-1,1[^2$

$$\boldsymbol{v}=\mathbf{rot}\sin\pi x\sin\pi y, \quad p=\tfrac{1}{4}(\cos 2\pi x\cos 2\pi y)+10(x+y)$$

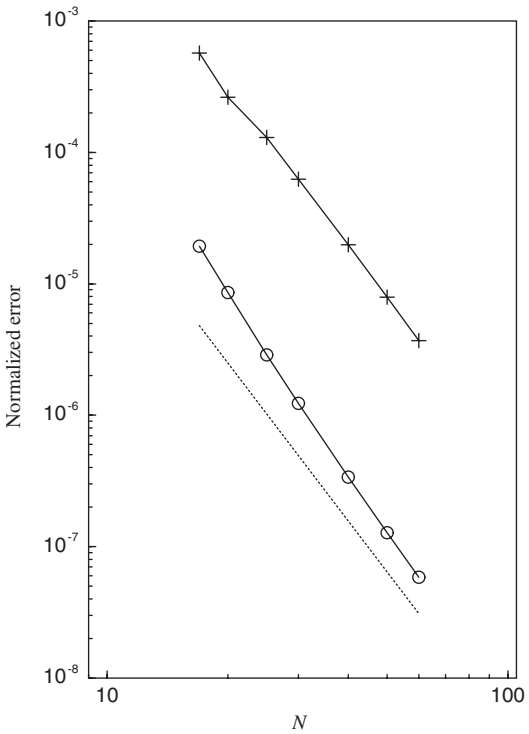

Figure 10. Spatial error on u (○——○) and p (+——+) obtained on the steady state solution of the Navier–Stokes equations; ·····: reference line of slope −4.

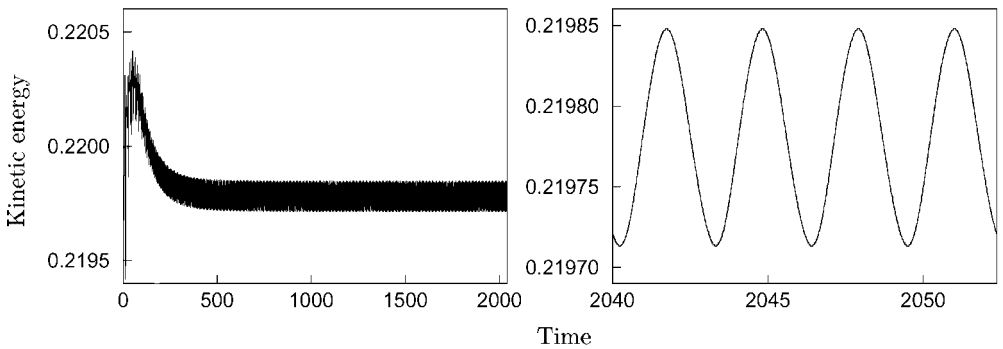

Figure 11. Time evolution of the kinetic energy in the driven cavity at $Re = 12000$ computed by the SCM fractional-step scheme on a 65×65 grid with $\Delta t = 5 \times 10^{-3}$.

from which the source term and the boundary conditions of the Navier–Stokes equations are defined. For the Reynolds number $Re = 100$, Figure 10 displays the normalized l^2 errors obtained when the steady-state is reached. This figure confirms that the method is indeed fourth-order accurate for both velocity and pressure.

The second test illustrates the accuracy and stability of the method for computing unsteady solutions. We consider the computation of the periodic flow in the regularized driven cavity at $Re = 12,000$, taking as a reference the spectral computation of Shen [30] performed with 65 Chebyshev polynomials in each direction and the time step $\Delta t = 5 \times 10^{-3}$. For comparison purpose, the SCM computation uses the same discretization parameters, with a similar grid refined near the boundary by a Chebyshev distribution of knots. The initial condition is defined as the steady flow at $Re = 10,000$. Figure 11 displays the time-evolution of the kinetic energy on nearly half a million time-steps. The periodic state is asymptotically reached with the same period $T = 3.085 \pm \Delta t$ measured in References [30]. This result shows the ability of the SCM method to conserve kinetic energy on a long time integration, and to reproduce spectral results with a similar coarse spatial resolution.

5. CONCLUDING REMARKS

The use of highly-accurate approximate inverses in association with the SCM fractional step scheme led to the development of a Navier–Stokes solver that preserves the accuracy of the B-spline in a cost-effective way. The high computational interest of the fractional step method is indeed recovered: the full decoupling of the velocity and pressure allows the solution to sparse elliptic problems only at each time-cycle.

The principle of the SCM scheme has a universal appeal that should not be restricted to the B-spline discretization used here: indeed, it has been initially proposed by Gresho and Chan [18] for the finite-element method in association to mass matrix lumping. In this paper, the introduction of sparse approximate inverses of the mass matrix has extended the latter scheme to higher-order discretizations. These approximate inverses have been constructed by application of local (or quasi-) interpolation schemes, which are mathematical tools borrowed from the spline interpolation theory. However, since their construction is based upon polynomial properties (see Equation (28)), we may believe that suitable sparse approximate inverses can be constructed for other piecewise polynomial bases such as spectral h/p elements [31], and alternate approximation methods such as the Galerkin method.

This SCM Navier–Stokes solver should be the building block for performing complex flow simulations with multivariate splines of wider flexibily than tensor-product B-splines, enabling approximations over arbitrary geometries and local refinement properties that are highly attractive for wall-bounded turbulent flow simulations. The development of accurate local interpolants and sparse approximate inverses represents a key element into this endeavor. In this respect, we refer to [32] where quasiinterpolation schemes are discussed for bivariate spline spaces over triangulations that display the aforementioned properties.

ACKNOWLEDGEMENTS

The author is grateful to Dr. K. Shariff for valuable discussions during the course of this work.

REFERENCES

1. Nicoud F, Baggett JS, Moin P, Cabot W. Large eddy simulation wall modelling based on suboptimal control theory and linear stochastic estimation. *Physics of Fluids* 2001; **13**:2968–2984.
2. Kravchenko AG, Moin P. Numerical studies of flow over a circular cylinder at $Re_d > 3900$. *Physics of Fluids* 2000; **12**:403–417.

3. Kravchenko AG, Moin P, Shariff KR. B-spline method and zonal grids for simulation of complex turbulent flows. *Journal of Computational Physics* 1999; **151**:757–789.
4. Shariff K, Moser RD. Two-dimensional mesh embedding for B-spline methods. *Journal of Computational Physics* 1998; **145**:471–488.
5. Kravchenko AG, Moin P. On the effect of numerical errors in large eddy simulations of turbulent flows. *Journal of Computational Physics* 1997; **131**:310–322.
6. Vasilyev OV. High order finite difference schemes on non-uniform meshes with good conservation properties. *Journal of Computational Physics* 2000; **157**:746–761.
7. Botella O. A velocity-pressure Navier–Stokes solver using a B-spline collocation method. *CTR Annual Research Briefs 1999*, Center for Turbulence Research, NASA Ames/Stanford Univ., 1999; 403–421 (http://ctr.stanford.edu/ResBriefs99/botella.pdf).
8. Botella O. On a collocation B-spline method for the solution of the Navier–Stokes equations. *Computers and Fluids* 2002; **31**:397–420.
9. Harlow FH, Welch JE. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surfaces. *Physics of Fluids* 1965; **8**:2181–2189.
10. Chorin A. Numerical simulation of the Navier–Stokes equations. *Mathematics of Computation* 1968; **22**: 745–762.
11. Temam R. Sur l'approximation de la solution des équations de Navier–Stokes par la méthode des pas fractionnaires II. *Archive for Rational Mechanics and Analysis* 1969; **32**:377–385.
12. Barrett R, Berry M, Chan TF, Demmel J, Donato JM, Dongarra J, Eijkhout V, Pozo R, Romine C, Van der Vorst H. *Templates for the Solution of Linear Systems*: *Building Blocks for Iterative Methods*. SIAM: Philadelphia, PA, 1994.
13. Saad Y. *Iterative Methods for Sparse Linear Systems*. PWS: Boston, MA, 1996.
14. Gresho PM, Sani RL. *Incompressible Flow and the Finite Element Method*. Wiley: Chichester, 1998.
15. Hughes TJR. *The Finite Element Method*: *Linear Static and Dynamic Finite Element Analysis*. Prentice-Hall: Englewood Cliffs, NJ, 1987.
16. Gresho PM, Lee RL, Sani RL. Advection-dominated flows, with emphasis on the consequence of mass lumping. In *Finite Elements in Fluids*, *Vol. 3*. Wiley: Chichester, 1978; 335–350.
17. Gresho PM. On the theory of semi-implicit projection methods for viscous incompressible flow and its implementation via a finite element method that also introduces a nearly consistent mass matrix. Part 1: Theory. *International Journal for Numerical Methods in Fluids* 1990; **11**:587–620.
18. Gresho PM, Chan ST. On the theory of semi-implicit projection methods for viscous incompressible flow and its implementation via a finite element method that also introduces a nearly consistent mass matrix. Part 2: Implementation. *International Journal for Numerical Methods in Fluids* 1990; **11**:621–659.
19. Benzi M, Tůma M. A comparative study of sparse approximate inverse preconditioners. *Applied Numerical Mathematics* 1999; **30**:305–340.
20. de Boor C. *A Practical Guide to Splines*. Springer: New York, NY, 1978.
21. de Boor C, Fix GJ. Spline approximation by quasiinterpolants. *Journal of Approximation Theory* 1973; **8**:19–45.
22. Lyche T, Schumaker LL. Local spline approximation methods. *Journal of Approximation Theory* 1975; **15**: 294–325.
23. Christara CC, Smith B. Multigrid and multilevel methods for quadratic spline collocation. *BIT* 1997; **37**(4): 781–803.
24. Guermond JL. Some implementations of projection methods for Navier–Stokes equations. *Modélisation Mathématique et Analyse Numérique* 1996; **30**:637–667.
25. Vichnevetsky R, Bowles JB. *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*. SIAM: Philadelphia, PA, 1982.
26. Lele SK. Compact finite difference schemes with spectral-like resolution. *Journal of Computational Physics* 1992; **103**:16–42.
27. Kwok WY, Moser RD, Jiménez J. A critical evaluation of the resolution properties of *B*-spline and compact finite difference methods. *Journal of Computational Physics* 2001; **174**:510–551.
28. Leonard BP. A stable and accurate convection modelling procedure based on quadratic upstream interpolation. *Computational Methods in Applied Mechanical Engineering* 1979; **19**:59–98.
29. Fairweather G, Meade D. A survey of spline collocation methods for the numerical solution of differential equations. In *Mathematics for Large Scale Computing*, Diaz JC (ed). Marcel Dekker: New York, NY, 1989; 297–341.
30. Shen J. Hopf bifurcation of the unsteady regularized driven cavity flow. *Journal of Computational Physics* 1991; **95**:228–245.
31. Karniadakis GE, Sherwin SJ. *Spectral/hp Element Methods for CFD*. Oxford University Press: New York, NY, 1999.
32. Lai M-J, Schumaker LL. On the approximation power of splines on triangulated quadrangulations. *SIAM Journal on Numerical Analysis* 1998; **36**:143–159.

