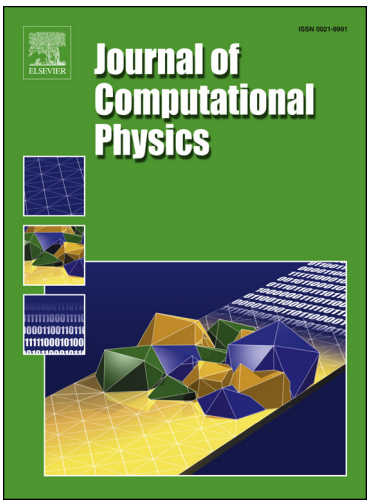
Accepted Manuscript



Optimised prefactored compact schemes for linear wave propagation phenomena

A. Rona, I. Spisso, E. Hall, M. Bernardini, S. Pirozzoli

PII: S0021-9991(16)30512-5
DOI: http://dx.doi.org/10.1016/j.jcp.2016.10.014
Reference: YJCPH 6892

To appear in: *Journal of Computational Physics*

Received date: 12 February 2016
Revised date: 3 October 2016
Accepted date: 5 October 2016

Please cite this article in press as: A. Rona et al., Optimised prefactored compact schemes for linear wave propagation phenomena, *J. Comput. Phys.* (2016), http://dx.doi.org/10.1016/j.jcp.2016.10.014

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Optimised prefactored compact schemes for linear wave propagation phenomena

A. Rona[a,*], I. Spisso[b], E. Hall[a], M. Bernardini[c], S. Pirozzoli[c]

[a] *Department of Engineering, University of Leicester, Leicester, LE1 7RH, England*
[b] *Cineca, via Magnanelli 6/3, 40033 Casalecchio di Reno, Italy*
[c] *Università degli Studi di Roma "La Sapienza", Dipartimento di Meccanica e Aeronautica, Via Eudossiana 18, 00184 Roma, Italy*

Abstract

A family of space- and time-optimised prefactored compact schemes are developed that minimize the computational cost for given levels of numerical error in wave propagation phenomena, with special reference to aerodynamic sound. This work extends the approach of Pirozzoli [1] to the MacCormack type prefactored compact high-order schemes developed by Hixon [2], in which their shorter Padé stencil from the prefactorization leads to a simpler enforcement of numerical boundary conditions. An explicit low-storage multi-step Runge-Kutta integration advances the states in time. Theoretical predictions for spatial and temporal error bounds are derived for the cost-optimised schemes and compared against benchmark schemes of current use in computational aeroacoustic applications in terms of computational cost for a given relative numerical error value. One- and two-dimensional test cases are presented to examine the effectiveness of the cost-optimised schemes for practical flow computations. An effectiveness up to about 50% higher than standard schemes is verified for the linear one-dimensional advection solver, which is a popular baseline solver kernel for computational physics problems. Substantial error reduction for a given cost is also obtained in the more complex case of a two-dimensional acoustic pulse propagation, provided the optimised schemes are made to operate close to their nominal design points.

Keywords: Computational aeroacoustics, Prefactored compact schemes, Optimised algorithms, Numerical dispersion, Numerical dissipation, Reduced numerical error

1. Introduction

1.1. Challenges in modelling wave generation and propagation phenomena

Models for the propagation of waves in a continuum are developed across the full spectrum of physical sciences, including aeroacoustics, where increasingly stringent aircraft noise

*Corresponding author. Tel.: +44 (0)116 2522510; fax: +44 (0)116 2522619.
Email addresses: ar45@leicester.ac.uk (A. Rona), i.spisso@cineca.it (I. Spisso), eh171@leicester.ac.uk (E. Hall), matteo.bernardini@uniroma1.it (M. Bernardini), sergio.pirozzoli@uniroma1.it (S. Pirozzoli)
URL: http://www.le.ac.uk/eg/ar45/ (A. Rona)

ACCEPTED MANUSCRIPT

regulations [3, 4] promote the development of accurate and affordable methods for predicting aerodynamically generated noise. Enhanced Computational Aeroacoustic (CAA) schemes are sought that can model sound generation and its propagation as part of the industrial design process, where predictions of accuracy compatible with the design specifications are required at an affordable cost, produced within specific time-constraints, earlier in the design process, using multi-processor computer clusters.

Computational aeroacoustic schemes that model sound generation and propagation with one simulation typically face a number of challenges. In aeroacoustics, modelling trailing edge noise is an example of where this approach is used. Small–scale flow structures, of the size of the boundary layer thickness, cross a trailing edge, where they experience large amplitude fluctuations in momentum. The interaction generates small amplitude acoustic pressure perturbations of long wavelength, compared to the boundary layer thickness. These radiate at the speed of sound, which is much larger than the flow structure mean convection speed.

In a direct computational aeroacoustic simulation of the trailing edge problem, the computational domain size scales with the long wavelength of the acoustic pressure perturbation whereas the spatial discretization scales with the small flow structures. This separation of length scales results in large sized computational meshes. Furthermore, supporting large amplitude localised momentum fluctuations and small amplitude acoustic perturbations in the same solution requires numerical schemes with low dispersion and low dissipation characteristics, to prevent the low–amplitude acoustic signal from being distorted and/or suppressed by the flow solver numerical viscosity.

1.2. Prefactored compact finite-difference schemes

Different aeroacoustic problems exhibit different flow physics. Therefore no single algorithm is available to model all problems with adequate resolution and accuracy. Low Mach number acoustically active flows typically involve capturing complex features that are nevertheless computationally smooth, that is, these features do not involve any sharp change in the flow state, such as from a shock. In these circumstances, it is computationally advantageous to use higher than second-order (high-order) schemes that can achieve exponential (e^{-aN}) convergence rates by increasing the scheme's order as opposed to algebraic (N^{-b}) convergence rates by increasing the spatial mesh refinement, where N is the number of degrees of freedom and a and b are positive real numbers. Several numerical methods for aeroacoustics have emerged in the last two decades [5, 6, 7] with attendant applications documented in four proceedings of the Computational Aeroacoustic (CAA) workshops on benchmark problems [8, 9, 10]. Lele [11] pioneered the use of Padé type compact and explicit optimised schemes in aeroacoustics. This work highlighted the requirement for special near-boundary treatment, driven by the longer finite-difference stencils used in these high-order methods. Hixon [2] introduced in 1996 a prefactorization method to reduce the non-dissipative central-difference stencil length of the compact schemes to two lower-order biased stencils which have easily soluble matrices of reduced size. The advantages of these schemes over traditional compact schemes arise from their reduced stencil length and the independent nature of the resultant factorized matrices. By reducing the stencil length of the compact schemes, the prefactorization method reduces the depth of the boundary frame over the perimeter of the computational domain where *ad–hoc* boundary stencils are required. This simplifies the specification of the boundary conditions [12]. Ashcroft and Zhang [13] extended the prefactorization method to a broader class of compact schemes using a more general derivation strategy, which combines Fourier analysis with the notion of a numeri-

ACCEPTED MANUSCRIPT

cal wavenumber. This class of optimised prefactored schemes exhibits better wave propagation characteristics than the standard prefactored compact ones.

1.3. Optimisation of the numerical methods

Several finite-difference explicit and compact methods are now available for solving wave propagation problems in a low Mach number flow. It is typical to involve a high performance computer cluster for such a computation, where the cost of the computation is important. It is therefore of interest to develop and implement an optimisation strategy for the computational cost, based on an acceptable level of numerical accuracy of the results. The issue of computational efficiency of finite-difference schemes has been investigated in detail by Colonius and Lele [7] and by Spisso and Rona [14]. These authors considered the dispersive and dissipative characteristics of several types of spatial discretization. The error associated with the approximate time integration is usually considered separately from the spatial error. Spatial and temporal errors combine in finite difference time–marching schemes to determine the overall numerical accuracy of the solution. Pirozzoli [1] developed in 2007 a general strategy for the analysis of finite-difference schemes for wave propagation problems, trying to involve time integration in the analysis in a natural way. The analysis of the global discretization error has shown the occurrence of two approximately independent sources of error, associated with the discretizations in space and in time. The performance improvement of the global scheme can be achieved by trying to separately minimize the two contributions. The analysis leads to logical and simple criteria for deriving optimised space- and time-discretization schemes, based on the concepts of spatial and temporal resolving efficiencies. Such efficiencies are achievable by a careful design of the space- and time-discretization schemes, as well as an appropriate choice of the grid spacing and of the time step. This analysis points towards substantial savings in computational resources.

1.4. Obtaining and testing optimised pre-factored compact schemes

The objectives of the present work are to extend the computational cost optimisation method of Pirozzoli [1] to prefactored compact schemes, assess by numerical experiment the actual computational cost of the newly developed optimised schemes, and verify that the desired level of numerical accuracy of the solution is achieved. This work advances the state of the art by presenting a stencil optimisation procedure for prefactored compact time-marching schemes that minimize the computational cost for a given numerical error level and spatial and temporal resolving efficiencies [15]. Furthermore, this work gives a first confirmation of the performance of spatial and temporal error estimators in prefactored compact finite-difference schemes coupled with the Runge-Kutta time integration by numerical experiment. Finally, this work provides a first verification of the predictive performance of a computational cost estimator of simple formulation, showing that this linearly relates to the actual computational time recorded in numerical experiments on a high performance computer. The impact of these three contributions is substantial. It enables tailoring the spatial and the temporal coefficients of the prefactored compact finite-difference time-marching schemes to obtain aeroacoustic predictions within computational cost and numerical error bounds. This contributes towards decreasing the uncertainty in computational aeroacoustic predictions for practical engineering and physical sciences applications, leading to a more optimal and reliable design solution.

1.5. Paper outline

The paper is organised as follows: in Section 2, the theory of the cost-optimisation of Pirozzoli [1], based on the optimisation of the computational cost for a given error level, is reviewed

ACCEPTED MANUSCRIPT

and the approximate decoupling of the spatial and temporal errors is introduced. The de-coupled errors are used in Section 3 to obtain a new family of prefactored cost-optimised schemes, with three different levels of numerical accuracy. In this section, the algebraic symmetry properties of the prefactorised stencils from Hixon [2] are used to define a new spatial differentiation for the cost-optimised compact schemes that is prefactored. The spatial differentiation is coupled with a Runge-Kutta time integration that is cost-optimised to a matching level of accuracy by the cost-optimisation process of Bernardini and Pirozzoli [16]. The theoretical performance of the optimised coupled schemes is then investigated for a range of cost–levels. Numerical tests in both one and two spatial dimensions are reported in Section 4, to compare the numerical accuracy achieved by the computational schemes against the design values from Section 3. Section 4 also addresses whether the analytic cost function used for the cost–optimisation is a good representation of CPU time recorded during the actual computations. Section 4 investigates the performance gain achieved by the optimised schemes versus their non–optimised counterparts. A discussion of the extension of the proposed schemes to nonlinear and higher dimensional problems, as well as to those with non-periodic boundary conditions, is given in Section 5 and, finally, conclusions from the current work and future perspectives are presented in Section 6.

2. Optimisation of finite-difference schemes for wave propagation phenomena

2.1. Compact finite difference schemes

By performing a Taylor's expansion of a smooth function $f : \mathbb{R} \to \mathbb{R}$, its derivatives can be shown to satisfy the following relations

$$\sum_{j=-P}^{Q} \alpha_j \frac{\partial f}{\partial x}(x + jh) = \frac{1}{h} \sum_{j=-R}^{S} a_j f(x + jh) + O(h^n), \tag{1}$$

where P, Q, R and S are non-negative integers, such that $(R + S) \geq 1$, $P \leq R$, $Q \leq S$ and $h > 0$. Consider now a set of uniformly spaced nodes indexed by i along the $\mathbb{R}$ axis, as shown in Figure 1, where $x_i = (i - 1)h$, for $1 \leq i \leq N$ and $h = L/(N - 1)$. Letting $f_i = f(x_i)$, then finite difference approximations f'_i of the first order derivatives $\frac{\partial f}{\partial x}(x_i)$ can be found using the $(R + S + 1)$ stencil:

$$\sum_{j=-P}^{Q} \alpha_j f'_{i+j} = \frac{1}{h} \sum_{j=-R}^{S} a_j f_{i+j}. \tag{2}$$

Selecting $P = Q = 0$ gives an explicit scheme, while other choices yield implicit schemes, otherwise known as Padé or compact schemes. Formally, from Eq. (1), the approximations f'_i have a truncation error of $O(h^n)$ and standard techniques seek to choose coefficients α_j and a_j that give the largest possible n for a given stencil width. For brevity, C$PQRS$ is the notation used to denote compact schemes. Of particular interest in CAA is the numerical error in the wave-like propagation of a single Fourier component of $f(x)$. A monochromatic wave of wavelength λ and wavenumber k has the scaled wavenumber $\kappa = kh$ with support $-\pi \leq \kappa \leq \pi$. Upon expanding f as a Fourier series, each Fourier coefficient satisfies Eq. (1) and, ignoring the contributions from the higher order terms, the scaled pseudo–wavenumber of the scheme Eq. (2) can be defined by

$$\bar{\kappa}(\kappa) = \frac{1}{\mathrm{i}} \frac{\sum_{j=-R}^{S} a_j \mathrm{e}^{\mathrm{i}j\kappa}}{\sum_{j=-P}^{Q} \alpha_j \mathrm{e}^{\mathrm{i}j\kappa}}. \tag{3}$$

ACCEPTED MANUSCRIPT

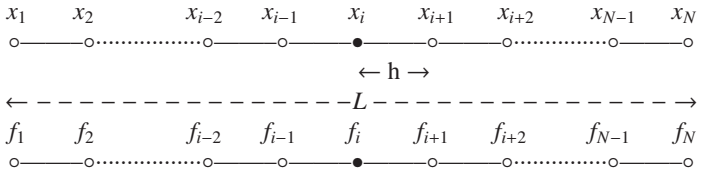


Figure 1: Variation of the discrete function $f_i = f(x_i)$ along a uniformly discretised streamwise length L.

If $\bar{\kappa} = \kappa$ could be achieved over the entire wavenumber range, then the finite difference estimate of f_i' by Eq. (2) would have spectral accuracy. In practice, the discrepancies between κ and $\bar{\kappa}$ increase as $\kappa \to \pi$, driven by the analytical form of the Padé approximation by which $\bar{\kappa} = 0$ at $\kappa = \pi$. $\bar{\kappa}$ is, in general, complex valued and its real part corresponds to a dispersive error, while its imaginary part is related to a dissipative error. In computational aeroacoustics, where disturbances are transmitted over large distances, it is important to minimise the dissipative error. Selecting $R = S$, $Q = P$ and letting a_j be antisymmetric and α_k be symmetric produces a scheme with zero dissipation. These central schemes are hereinafter denoted by CRQ. This desirable property is included in the current work by restricting the stencil optimisation to tridiagonal compact schemes ($P = Q = 1$) with a five point stencil ($R = S = 2$) *i.e.*, C12 schemes. In this case, the scaled pseudo–wavenumber is a real value given by

$$\bar{\kappa}(\kappa) = \frac{\sum_{j=1}^{2} 2a_j \sin(j\kappa)}{1 + 2\alpha_1 \cos(j\kappa)}. \tag{4}$$

This work shows the optimisation of the prefactored compact stencils on the one parameter family of C12 schemes given by

$$\begin{cases} \alpha_0 = 1, \\ a_0 = 0, \\ a_1 = -a_{-1} = \frac{1}{3}(\alpha_1 + 2), \\ a_2 = -a_{-2} = \frac{1}{12}(4\alpha_1 - 1). \end{cases} \tag{5}$$

The choice $\alpha_1 = 1/3$ yields the sixth–order C12 scheme. The analysis in this article is, in principle, extensible to more general C$PQRS$ compact schemes.

2.2. Runge–Kutta time stepping

The semi–discretisation of a general conservation law by means of finite differences gives rise to a non–autonomous system of ordinary differential equations of the form

$$\frac{\mathrm{d}\mathbf{U}}{\mathrm{d}t} = \mathbf{F}(\mathbf{U}(t), t), \tag{6}$$

$$\mathbf{U}(t_0) = \mathbf{U}^0, \tag{7}$$

where $\mathbf{U}$ represents the vector containing the solution values at spatial nodes. The solution can be time–marched from $t = t_n$ to $t = t_n + \Delta t$ by means of an explicit p–stage, two level Runge–Kutta scheme, which is formulated as

$$\mathbf{U}^{(0)} = \mathbf{U}^n, \tag{8}$$

$$\mathbf{U}^{(l)} = \mathbf{U}^n + \beta_l \Delta t \mathbf{F}(\mathbf{U}^{(l-1)}), \quad l = 1, \ldots, p, \tag{9}$$

$$\mathbf{U}^{n+1} = \mathbf{U}^{(p)}, \tag{10}$$

ACCEPTED MANUSCRIPT

where the coefficients β_l can be determined to achieve a given formal order of accuracy by means of a Taylor expansion, or, for example, to minimise dissipation and phase errors [17, 18]. Assuming $\mathbf{F}(\mathbf{U}) = A\mathbf{U}$, where A is linear, the scheme can be rewritten as

$$\mathbf{U}^{n+1} = \mathbf{U}^n + \sum_{j=1}^{p} (\Delta t)^p \gamma_j A^p \mathbf{U}^n, \tag{11}$$

where $\gamma_j = \Pi_{l=p-j+1}^{p} \beta_l$. This article uses the 4–stage Runge–Kutta scheme with $\gamma_1 = 1$, $\gamma_2 = 1/2$ and γ_3 and γ_4 as free parameters. Setting $\gamma_3 = 1/3!$ and $\gamma_4 = 1/4!$ yields a formally 4th order accurate scheme.

2.3. Linear advection equation analysis

2.3.1. The one dimensional single-scale problem

The one dimensional linear advection equation (LAE)

$$\frac{\partial u}{\partial t} + c\frac{\partial u}{\partial x} = 0, \quad u(x,0) = u_0(x), \tag{12}$$

where c is the wave speed, is a simple platform for developing and testing appropriate schemes for more complicated, multi–dimensional, multi–variable problems. Imposing the sinusoidal monochromatic initial condition $u_0(x) = \hat{u}_0 e^{ikx}$ on infinite and periodic domains gives rise to the solution $u(x,t) = \hat{u}_0 e^{i(-\omega t + kx)}$, where the angular frequency ω and the wavenumber k are related by the dispersion relation $\omega = ck$.

For the p–stage Runge–Kutta scheme defined in Eq. (11) and the compact scheme from Eq. (1), the amplification factor $r(\kappa, \sigma)$ from one time level to the next can be obtained [19] as

$$r(\kappa, \sigma) = 1 + \sum_{j=1}^{p} \gamma_j \left(-\mathrm{i}\sigma\bar{\kappa}(\kappa)\right)^j, \tag{13}$$

where $\sigma = c\Delta t/h$ is the Courant number and $j = 1, \ldots, p$. In the case of null spatial error, *i.e.*, if $\bar{\kappa} = \kappa$, the amplification factor is

$$r_t(z, \boldsymbol{\gamma}) = 1 + \sum_{j=1}^{p} \gamma_j (-\mathrm{i}z)^j, \tag{14}$$

where $z = \sigma\kappa$ and $\boldsymbol{\gamma} = \{\gamma_j\}_{j=1}^{p}$. The stability limit z_s is then given by the following condition:

$$z_s = \max\{z \,|\, |r_t(\bar{z}, \boldsymbol{\gamma})| \leq 1 \forall \bar{z}, 0 < \bar{z} \leq z\}. \tag{15}$$

Let $u_h(\cdot,\cdot)$ denote an approximation to $u(\cdot,\cdot)$. Following the work of Pirozzoli [1], the relative error in the L_2–norm across one wavelength is defined as

$$E = \frac{\|u_h(\cdot, T) - u(\cdot, T)\|_2}{\|u_0(\cdot)\|_2}, \tag{16}$$

where the norm of any scalar function w is defined as

$$\|w\|_2 = \left(\frac{1}{\lambda}\int_{x_0}^{x_0+\lambda} |w(x)|^2 \,\mathrm{d}x\right)^{1/2},$$

ACCEPTED MANUSCRIPT

and $\lambda = 2\pi/k$ is the wavelength. It can then be shown [1] that, to a good approximation,

$$E \approx (ckT)\,\frac{|r(k,\sigma) - \mathrm{e}^{-i\sigma\kappa}|}{\sigma\kappa} \equiv E_a, \tag{17}$$

which is accurate up to a term of order $E_a^2/2$. In the results that follow, $E_a \ll 1$, hence $E_a^2/2 \ll E_a \ll 1$. The error approximation E_a is likely to hold up to large κ values, at which numerical stability limits prevent the scheme from being operated.

2.3.2. Multi-dimensional single-scale problem

For a problem in d space dimensions, the cost C_d of its numerical solution, in terms of floating point operations, can be shown by dimensional arguments [7] to scale as

$$C_d \propto pN_{\mathrm{op}}\left(\frac{T}{\Delta t}\right)\left(\frac{L}{h}\right)^d, \tag{18}$$

where $(L/h)^d$ is the number of points in the domain, N_{op} is the required number of operations per mesh node that depends on the scheme used to approximate the spatial derivatives, p is the number of Runge-Kutta stages and $T/\Delta t$ is the number of time steps.

Consider, for instance, the cost of solving Eq. (12) in d dimensions using the spatial discretization of Fig. 1. For $L \gg h$, $N = (L+h)/h \approx L/h$ and $N^d \approx (L/h)^d$ is the number of u_i unknowns to be evaluated at time T. The solution is time-marched to time T in Δt time steps by Eqs. (8)-(10), which requires $T/\Delta t$ applications of the p-stage Runge-Kutta scheme of Section 2.2 per unknown. Evaluating the l-th Runge-Kutta stage in Eq. (9) requires d additions and 1 multiplication, or $d+1$ floating point operations (FLOPS) per unknown, for a given $\mathbf{F}(\mathbf{U})$. $\mathbf{F}(\mathbf{U})$ is estimated as cf_i', with f_i' from Eq. (2), in each dimension d. Equation (2) is a linear algebraic system of type $\mathbf{Ax} = \mathbf{b}$ with $\mathbf{A}$ banded. By the lower-upper-decomposition of $\mathbf{A}$, f_i' is computed by five additions and four multiplications, or nine FLOPS, per unknown away from the domain boundaries $i = 1$ and $i = N$, for $P = Q = 1$ and $R = S = 2$ [7]. Reducing $R = S$ to 1 lowers the FLOPS required to six [7]. The aggregate number of operations to compute one Runge-Kutta l stage per mesh node is $N_{\mathrm{op}} = (d+1) + d(9+1)$ with the C12 scheme. Therefore, the complexity of solving the linear system of equations, Eq. (12), in d dimensions by the C12 finite-difference scheme coupled to a p-stage Runge-Kutta integration is $p(11d+1)T\Delta t^{-1}N^d$ FLOPS, as represented by Eq. (18). The prefactorization of C12 in Section 3.2 reduces N_{op} but the complexity remains of the same order of N.

The low-storage Runge-Kutta method of Section 2.2, at each stage l, requires the storage of $O\left(3N^d\right)$ floating point numbers, which determines the baseline computer memory requirement. Additional memory storage is typically used in the parallel implementation of compact schemes, for instance by slab decomposition [20], compared to their equivalent explicit schemes.

By introducing the non-dimensional groups ckT and kL, which are, respectively, measures of the number of wavelengths travelled by a wave in time T and the number of wavelengths contained in the computational domain L, the normalised error e and the normalised cost c_d are obtained

$$e(\kappa,\sigma) \equiv \frac{E_a}{ckT} = \frac{|r(k,\sigma) - \mathrm{e}^{-i\sigma\kappa}|}{\sigma\kappa}, \tag{19}$$

$$c_d(\kappa,\sigma) \equiv \frac{C_d}{(ckT)(kL)^d} = pN_{\mathrm{op}}\frac{1}{\sigma\kappa^{d+1}}. \tag{20}$$

ACCEPTED MANUSCRIPT

The normalised error represents the numerical error incurred in advecting the wave $u(x,t)$ over one period. Similarly, the normalised cost represents the cost of advecting $u(x,t)$ in one dimension over the same period.

2.3.3. Multi-dimensional multi-scale problems

In practice, aeroacoustic signals are composed of a number of waves of differing wavelengths. It is therefore of interest to extend the error and cost definitions from Eq. (19) and Eq. (20) so they apply to broadband signals, supposing these have a spectrum of finite width $|k| \leq \hat{k}$ and varying propagation velocities $|c| \leq \hat{c}$. Let $\hat{\kappa} = \hat{k}h$, $\hat{\sigma} = \hat{c}\Delta t/h$, then the formulae for the normalised global error and the normalised global cost metric for a broadband signal can be defined as

$$\hat{e}(\hat{\kappa},\hat{\sigma}) = \frac{1}{\hat{\sigma}\hat{\kappa}} \max_{(\kappa,\sigma)\in[0,\hat{\kappa}]\times[0,\hat{\sigma}]} |r(k,\sigma) - \mathrm{e}^{-\mathrm{i}\sigma\kappa}|, \tag{21}$$

$$\hat{c}_d(\hat{\kappa},\hat{\sigma}) = pN_{\mathrm{op}} \frac{1}{\hat{\sigma}\hat{\kappa}^{d+1}}. \tag{22}$$

The definition of the global error given in Eq. (21) is an analytical estimator of the numerical error in a given simulation, based on the premise that all spectral scales in the wave propagation problem make equal contributions to the signal power. In real-world fluid mechanic applications, state variables have non-flat spectra and the smaller scales are usually less energetic than the larger scales, with the implication that the error estimator defined in Eq. (21) models a 'worst case scenario'.

2.4. Optimal performance for multi–scale problems

In this work, optimising the performance of a scheme is taken as minimising the computational cost for a given level of error, for given values of p and N_{op} and a given problem, *i.e.*, for specific values of the non-dimensional groups ckT and kL. To do this, a target level for the relative error ϵ is specified in the non-dimensional form

$$\hat{e}(\hat{\kappa},\hat{\sigma}) = \frac{\epsilon}{ckT} \equiv \tilde{\epsilon}, \tag{23}$$

and the optimisation consists of finding a pair of values $(\kappa^*(\tilde{\epsilon}), \sigma^*(\tilde{\epsilon}))$ that minimises the cost metric and satisfies the stability condition Eq. (15).

An illustration of this optimisation process is depicted in Figure 2(a), which shows iso–lines of the normalised global error $\hat{e}(\hat{\kappa},\hat{\sigma})$ and normalised cost $\hat{c}_1(\hat{\kappa},\hat{\sigma})$ in the wavenumber-Courant number $(\hat{\kappa},\hat{\sigma})$–plane, where the support is given by $[0 \leq \hat{\sigma} \leq \hat{\sigma}_{max}, 0 \leq \hat{\kappa} \leq \hat{\kappa}_{max}]$. $\hat{\sigma}_{max}$ and $\hat{\kappa}_{max}$ are the numerical stability limits of the specific scheme. Results are shown for the sixth–order compact spatial discretisation scheme coupled with the four–stage, fourth–order RK method, denoted by C12RK4, for which $\hat{\sigma}_{max} = 1.42$ and $\hat{\kappa}_{max} = \pi$. For the C12RK4 scheme, the iso–cost curves are convex whereas the iso–error curves are concave [1]. Under such a topology, the cost minima occur at the single point of tangency between the iso–lines and the circles in Figure 2(a) show the 'optimal' working conditions for a number of relative error values.

Let

$$\hat{e}(\hat{\kappa},\hat{\sigma}) \approx \max\left(\hat{e}_0(\hat{\kappa}), \hat{e}_t(\hat{z})\right) \equiv \hat{e}_a(\hat{\kappa},\hat{\sigma}), \tag{24}$$

ACCEPTED MANUSCRIPT

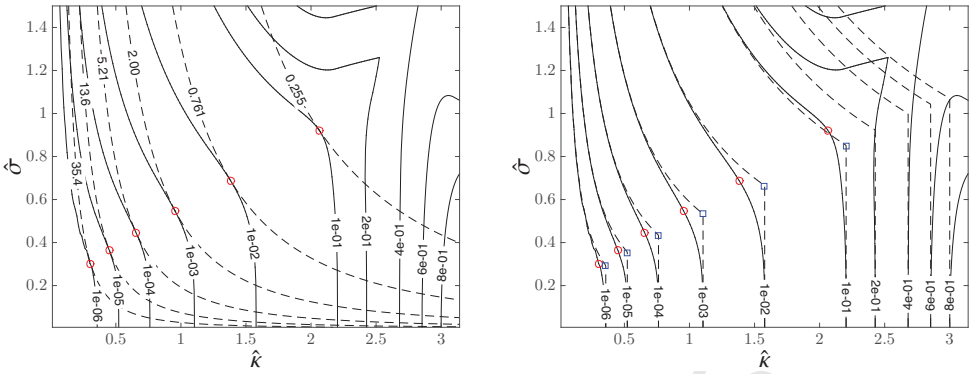


Figure 2: (a) Iso–lines of normalised 'global' error $\hat{e}(\hat{\kappa}, \hat{\sigma})$ estimated by Eq. (21) (solid lines) and of normalised cost $\hat{c}_1(\hat{\kappa}, \hat{\sigma})$ (dashed lines), for the C12RK4 scheme; the circles indicate 'optimal' working conditions. (b) Iso–lines of normalised 'global' error function $\hat{e}(\hat{\kappa}, \hat{\sigma})$ estimated by Eq. (21) (solid lines) and of its approximation $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ by Eq. (24) (dashed lines); the circles and squares indicate the respective optimised working conditions.

be an approximation [1] of $\hat{e}(\hat{\kappa}, \hat{\sigma})$ from Eq. (21) where $\hat{e}_0(\hat{\kappa})$ is the spatial error (assuming no temporal error) and $\hat{e}_t(\hat{z})$ is the temporal error (assuming no spatial error). $\hat{e}_0(\hat{\kappa})$ and $\hat{e}_t(\hat{z})$ are given by

$$\hat{e}_0(\hat{\kappa}) \equiv \frac{1}{\hat{\kappa}} \max_{0 \le \kappa \le \hat{\kappa}} |\bar{\kappa} - \kappa|, \quad \hat{e}_t(\hat{z}) \equiv \frac{1}{\hat{z}} \max_{0 \le z \le \hat{z}} \left| 1 + \sum_{j=1}^{p} \gamma_j (-\mathrm{i}z)^j - \mathrm{e}^{-\mathrm{i}z} \right|, \tag{25}$$

where $\hat{z} = \hat{\sigma}\hat{\kappa}$. $\hat{e}_a(\cdot, \cdot)$ enables a decoupling of the error into spatial and temporal contributions, which will lead to a simplification of the computations that follow. Figure 2(b) shows a comparison of the optimised working points for the C12RK4 scheme for the same levels of error as in Figure 2(a), when $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ is used in place of $\hat{e}(\kappa, \sigma)$. The dashed lines show the maximum of the spatial and temporal errors. The vertical segments are where the spatial error $\hat{e}_0(\hat{\kappa})$ is dominant and the dashed curves are where the temporal error $\hat{e}_t(\hat{z})$ is dominant. The circles correspond to the error $\hat{e}(\hat{\kappa}, \hat{\sigma})$ estimated by (21) and the squares to its approximation $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$. Table 1 shows the optimised values $(\hat{\kappa}^+, \hat{\sigma}^+)$ and $(\hat{\kappa}^*, \hat{\sigma}^*)$, corresponding to $(\hat{\kappa}, \hat{\sigma})$ from Eq. (21) and Eq. (24), respectively, and the relative distance between them defined by

$$D\left[(\hat{\kappa}^+, \hat{\sigma}^+), (\hat{\kappa}^*, \hat{\sigma}^*)\right] = \left[\left(\frac{\hat{\kappa}^+ - \hat{\kappa}^*}{\hat{\kappa}^+} \right)^2 + \left(\frac{\hat{\sigma}^+ - \hat{\sigma}^*}{\hat{\sigma}^+} \right)^2 \right]^{1/2}. \tag{26}$$

For the errors considered, there is a maximum of 17% difference in the location of the 'optimal' working conditions between the two error definitions. Whereas such a discrepancy has an appreciable magnitude, it still enables operating a finite difference time–marching scheme in the neighbourhood of its optimal design point, as verified in the propagation of a monochromatic wave [20]. This allows the splitting of the derivation of the cost-optimised coefficients into a spatial component and a temporal component.

ACCEPTED MANUSCRIPT

	$\hat{e}(\hat{\kappa},\hat{\sigma})$		$\hat{e}_a(\hat{\kappa},\hat{\sigma})$		
$\tilde{\epsilon}$	$\hat{\kappa}^+$	$\hat{\sigma}^+$	$\hat{\kappa}^*$	$\hat{\sigma}^*$	$D\left[(\hat{\kappa}^+,\hat{\sigma}^+),(\hat{\kappa}^*,\hat{\sigma}^*)\right]$
1×10^{-1}	2.064	0.922	2.203	0.849	0.10
1×10^{-2}	1.384	0.686	1.577	0.633	0.14
1×10^{-3}	0.954	0.548	1.103	0.534	0.16
1×10^{-4}	0.656	0.446	0.762	0.432	0.16
1×10^{-5}	0.449	0.365	0.522	0.354	0.17
1×10^{-6}	0.307	0.301	0.356	0.291	0.17

Table 1: Comparison between the 'optimal' pair $(\hat{\kappa}^+,\hat{\sigma}^+)$ based on the global error $\hat{e}(\hat{\kappa},\hat{\sigma})$ estimated by Eq. (21) and the optimised pair $(\hat{\kappa}^*,\hat{\sigma}^*)$ based on its approximation $\hat{e}_a(\hat{\kappa},\hat{\sigma})$ from Eq. (24), for the classical C12RK4 scheme.

3. Optimised compact, time-marching schemes and extension to prefactored compact schemes

3.1. Optimised compact schemes

This section describes the optimisation of compact finite difference schemes, by which $\hat{\kappa}$ is maximised for a given normalised error level $\tilde{\epsilon}$, assuming that the time integration is exact. This is equivalent to finding the maximum resolving efficiency $\hat{\kappa}_{\text{opt}}$ so that all wavenumbers $\hat{\kappa} \leq \hat{\kappa}_{\text{opt}}$ are resolved with an error less than $\tilde{\epsilon}$. For this purpose, the error $\hat{e}_0(\hat{\kappa})$ from Eq. (25) is used. The baseline spatial scheme is the compact C12 scheme of Eq. (5) with $\alpha_1 = 1/3$.

Figure 4 gives a graphical representation of how $\hat{e}_0(\hat{\kappa})$ varies with κ for this scheme by the difference between the exact value of κ, shown by the dotted line, and its modelled counterpart $\text{Re}(\bar{\kappa}(\kappa))$, shown by the dashed line. As the C12 scheme is compact, the Padé stencil imposes the constraint $\text{Re}(\bar{\kappa}(\kappa)) = 0$ at $\kappa = \pi$. The C12 scheme therefore exhibits a progressive departure from the exact relationship $\text{Re}(\bar{\kappa}(\kappa)) = \kappa$ as $\kappa \to \pi$ that generates the error $\hat{e}_0(\hat{\kappa})$.

The baseline compact C12 scheme of Eq. (5) is optimised by treating α_1 as a free parameter. The new class of optimised schemes will be denoted as C12–*n*, where *n* represents the exponent in the target $\tilde{\epsilon} = 10^{-n}$. Optimisation in this setting is a non–trivial matter. As it can be seen in Figure 3, the optimal value $\left(\alpha_1^n, \hat{\kappa}_{\text{opt}}^n\right)$ of each 10^{-n} error contour occurs at a corner point. The corner point for $n = 5$ is shown magnified in the inset of Figure 3. Near this corner point, for each α_1, there can be multiple values of $\hat{\kappa}$ which lie on the desired error contour. As such, a standard local minimisation routine will not necessarily locate the optimum. The procedure used for estimating $\left(\alpha_1^n, \hat{\kappa}_{\text{opt}}^n\right)$ is as follows: for a given α_1, the maximum $\hat{\kappa}$, which gives the prescribed level error, is found by utilising a number of local function solves starting at different initial guesses. A global pattern search algorithm, which does not require the objective function to be smooth, is then used in α_1, which is the search domain. The resulting optimised values $\left(\alpha_1^n, \hat{\kappa}_{\text{opt}}^n\right)$, for $n = 4, 5$ and 6 are shown as the open symbols in Figure 3 and are presented in Table 2. Table 2 shows that the optimised values for α_1 are above the 'standard' value of $1/3$ for all error levels, asymptoting toward $\alpha_1 = 1/3$ as the desired error level is reduced. As $\tilde{\epsilon} \to 0$ the asymptotic errors analysis holds true and the formal order of accuracy of the scheme controls the error. Hence, the optimised coefficient tends to the one which maximises the formal order of accuracy, *i.e.*, $\alpha_1 = 1/3$, which gives a sixth-order accurate scheme. The same argument can be extended to more general *CPQRS* compact schemes.

Table 2 shows the maximum resolving efficiency $\hat{\kappa}^n$ for a normalised error level $\tilde{\epsilon} = 10^{-n}$, when the non–optimised C12 scheme is used. Using the cost-optimised coefficients increases the spatial resolving efficiency by $\Delta\hat{\kappa}^n \sim 50\%$ over the range $10^{-6} \leq \epsilon \leq 10^{-4}$, indicating a gain in spatial resolving efficiency from the spatial cost-optimisation.

ACCEPTED MANUSCRIPT

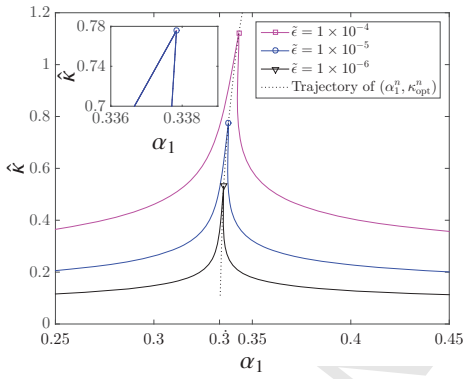


Figure 3: Iso–lines of 'global' spatial error $\hat{\epsilon}_0\,(\hat{\kappa})$ for the C12 scheme from Eq. (5) with varying α_1. The symbols indicate the optimised working conditions. The dotted line indicates the locus of optimised working conditions that asymptotes to $\alpha_1 = 1/3$ as $\hat{\kappa} \to 0$.

	C12–n			C12		RK4–n				RK4	
n	α_1^n	$\bar{\kappa}_{max}$	$\hat{\kappa}_{\text{opt}}^n$	$\hat{\kappa}$	$\Delta\hat{\kappa}^n\%$	γ_3^n	γ_4^n	z_s^n	$\hat{z}_{\text{opt}}^n$	$\hat{z}$	$\Delta\hat{z}^n\%$
4	0.34311986	2.03	1.121	0.762	47.1	0.16524207	0.04024863	2.8265	0.436	0.331	31.7
5	0.33783815	2.00	0.776	0.522	48.6	0.16610629	0.04111187	2.8281	0.272	0.186	46.2
6	0.33541923	1.99	0.533	0.357	49.3	0.16648569	0.04149236	2.8284	0.160	0.105	52.3

Table 2: Coefficients and resolving efficiencies of optimised spatial C12–n schemes and temporal RK4–n schemes. $\alpha = 1/3$, $\bar{\kappa}_{max} = 1.98$ for the standard sixth-order C12 scheme; $\gamma_3^{RK4} = 1/6$, $\gamma_4^{RK4} = 1/24$, $z_s = 2\sqrt{2}$ for the standard fourth-order RK4 scheme.

ACCEPTED MANUSCRIPT

3.2. Spatial differentiation by prefactored compact finite-difference schemes

In the C12 scheme of Eq. (5), the approximation of the derivatives at the nodal points requires the solution of a tridiagonal system of equations. In this section, following the methodology of Hixon [2], the tridiagonal system of equations is recast as two (upper and lower) bidiagonal systems. This gives a *prefactored* compact scheme. Let the derivative f_i' be split into a backward component $f_i'^B$ and a forward component $f_i'^F$ so that

$$f_i' = \frac{1}{2}\left(f_i'^F + f_i'^B\right). \tag{27}$$

Suppose that $f_i'^B$ and f'^F are found by using, respectively, the following backward and forward finite difference formulae:

$$\alpha_0^F f_i'^F + \alpha_1^F f_{i+1}'^F = \frac{1}{h}\left[a_{i-1}^F f_{i-1} + a_i^F f_i + a_{i+1}^F f_{i+1}\right], \tag{28}$$

$$\alpha_{-1}^B f_{i-1}'^B + \alpha_0^B f_i'^B = \frac{1}{h}\left[a_{i-1}^B f_{i-1} + a_i^B f_i + a_{i+1}^B f_{i+1}\right]. \tag{29}$$

The coefficients (α^F, α^B) and (a^F, a^B) in Eqs. (28) and (29) are determined so that the real (dispersive) components of the scaled pseudo–wavenumbers of the forward and backward stencils are equal to the scaled pseudo–wavenumbers of the original central scheme and the imaginary (dissipative) components of the scaled pseudo–wavenumbers are equal in magnitude and opposite in sign. Such schemes are termed *MacCormack* schemes [21]. The scaled pseudo–wavenumber of the generic C12 scheme of Eq. (4) is given by

$$\bar{\kappa}(\kappa) = \frac{\frac{2(\alpha_1+2)}{3(1+2\alpha_1)}\sin(\kappa) + \frac{(4\alpha_1-1)}{6(1+2\alpha_1)}\sin(2\kappa)}{\frac{1}{(1+2\alpha_1)} + \frac{2\alpha_1}{(1+2\alpha_1)}\cos(\kappa)}. \tag{30}$$

In a similar manner, the scaled pseudo–wavenumber of the generic forward operator $f_i'^F$ is

$$\mathrm{Re}\left(\bar{\kappa}^F(\kappa)\right) = \frac{\left(a_1^F\alpha_0^F - a_0^F\alpha_1^F - a_{-1}^F\alpha_0^F\right)\sin(\kappa) - a_{-1}^F\alpha_1^F\sin(2\kappa)}{(\alpha_0^F)^2 + (\alpha_1^F)^2 + 2\alpha_1^F\alpha_0^F\cos(\kappa)}, \tag{31}$$

$$\mathrm{Im}\left(\bar{\kappa}^F(\kappa)\right) = \frac{-\left(a_1^F\alpha_1^F + a_0^F\alpha_0^F\right) - \left(a_1^F\alpha_0^F + a_0^F\alpha_1^F + a_{-1}^F\alpha_0^F\right)\cos(\kappa) - a_{-1}^F\alpha_1^F\cos(2\kappa)}{(\alpha_0^F)^2 + (\alpha_1^F)^2 + 2\alpha_1^F\alpha_0^F\cos(\kappa)}, \tag{32}$$

and for the backward operator $f_i'^B$ is:

$$\mathrm{Re}\left(\bar{\kappa}^B(\kappa)\right) = \frac{\left(a_1^B\alpha_0^B + a_0^B\alpha_{-1}^B - a_{-1}^B\alpha_0^B\right)\sin(\kappa) + a_1^B\alpha_{-1}^B\sin(2\kappa)}{(\alpha_{-1}^B)^2 + (\alpha_0^B)^2 + 2\alpha_{-1}^B\alpha_0^B\cos(\kappa)}, \tag{33}$$

$$\mathrm{Im}\left(\bar{\kappa}^B(\kappa)\right) = \frac{-\left(a_0^B\alpha_0^B + a_{-1}^B\alpha_{-1}^B\right) - \left(a_1^B\alpha_0^B + a_0^B\alpha_{-1}^B + a_{-1}^B\alpha_0^B\right)\cos(\kappa) - a_1^B\alpha_{-1}^B\cos(2\kappa)}{(\alpha_{-1}^B)^2 + (\alpha_0^B)^2 + 2\alpha_{-1}^B\alpha_0^B\cos(\kappa)}. \tag{34}$$

Imposing the constraints

$$\alpha_j^F = \alpha_{-j}^B, \quad a_j^F = -a_{-j}^B, \quad -1 \le j \le 1$$

ACCEPTED MANUSCRIPT

ensures that the imaginary components of the forward and backward operators are equal in magnitude and opposite in sign and that the real components are equal to one another. Further, imposing

$$\sum_{j=-1}^{1} a_j^F = 0$$

guarantees that the spatial derivatives vanish in regions of zero gradient. Finally, by matching corresponding terms between Eq. (31) and Eq. (30) the following system of algebraic equations is obtained

$$\begin{cases} a_1^F\alpha_0^F - a_0^F\alpha_1^F - a_{-1}^F\alpha_0^F &= \frac{2(\alpha_1+2)}{3(1+2\alpha_1)}, \\ -a_{-1}^F\alpha_1^F &= \frac{4\alpha_1-1}{6(1+2\alpha_1)}, \\ (\alpha_0^F)^2 + (\alpha_1^F)^2 &= \frac{1}{(1+2\alpha_1)}, \\ 2\alpha_1^F\alpha_0^F &= \frac{2\alpha_1}{(1+2\alpha_1)}, \\ a_1^F + a_0^F + a_{-1}^F &= 0. \end{cases} \tag{35}$$

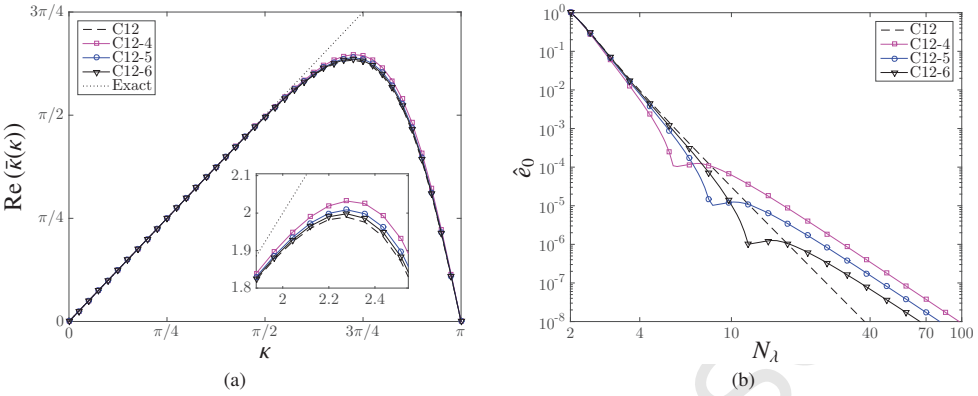
The quadratic terms mean that Eq. (35) has two solutions. The lower value solution for α_1^F is selected to minimize the ratio α_1^F/α_0^F, so that the influence of errors at the boundaries on the interior scheme is minimised, in problems that require closures at the computational domain boundaries. Substituting α_1^n from Section 3.1, Table 2, for α_1 in Eq. (35), the prefactored coefficients for the C12-*n* schemes are obtained. The resultant prefactored coefficients are listed in Table 3. It is noticeable that the standard C12 scheme appears to exhibit a greater central dominance in the stencil, as confirmed by the higher valued a_0^F coefficient. The coefficients in Table 3 together with Eqs. (27), (28) and (29) fully define the prefactored compact finite difference approximation f_i' of the first–order spatial derivatives $\partial f/\partial x$ in the domain interior, *i.e.*, $2 \le i \le N-1$.

Figure 4 shows the dispersive characteristics of the classical C12 and the cost-optimised C12-*n* schemes. Figure 4(a) shows the real component of the scaled pseudo-wavenumber $\mathrm{Re}\,(\bar{\kappa}\,(\kappa))$ of the compact scheme from Eq. (30) over the scaled wavenumber range $0 < \kappa \le \pi$. The dotted line shows the exact representation $\bar{\kappa}\,(\kappa) = \kappa$ and the lines with symbols the prefactored compact finite difference approximations using different levels of target error 10^{-n}. Table 2 and the insert in Fig. 4(a) show that the maximum value of the real part of the scaled pseudo-wavenumber $\bar{\kappa}_{max}$ does not vary significantly among the cost-optimised schemes and it is slightly higher for the C12-*n* schemes compared to the baseline scheme. Due to the symmetric properties of the MacCormack schemes [21], the real component of the scaled pseudo-wavenumber of the prefactored forward stencil from Eq. (31) is equal to the real component of the scaled pseudo-wavenumber of the prefactored backward stencil from Eq. (33) and is, by Eq. (27), also equal to the real part of the scaled pseudo-wavenumber of the original C12 scheme of Eq. (30), that is $\mathrm{Re}(\bar{\kappa}(\kappa)) = \mathrm{Re}(\bar{\kappa}^F(\kappa)) = \mathrm{Re}(\bar{\kappa}^B(\kappa))$. Hence, each line with symbols represents $\mathrm{Re}(\bar{\kappa}(\kappa)) = \mathrm{Re}(\bar{\kappa}^F(\kappa)) = \mathrm{Re}(\bar{\kappa}^B(\kappa))$ in Figure 4(a).

Figure 4(b) shows the spatial error $\hat{e}_0$ for the classical C12 and cost-optimised C12-*n* schemes versus the discrete number of points per wavelength $N_\lambda = 2\pi/\kappa$. The classical scheme C12 displays a straight line (constant logarithmic roll-off rate) along the whole wavenumber spectrum. The cost-optimised schemes C12-*n* feature local $\hat{e}_0$ minima of 10^{-n} for selected wavenumbers at which the spatial error $\hat{e}_0$ is reduced approximately by one order of magnitude compared to the C12 scheme. Similar spatial error-number of points per wavelengths $e_0 - N_\lambda$ plots with single and multiple $\hat{e}_0$ minima are reported in the literature [5, 11, 13, 22] of optimised finite-difference schemes.

The two imaginary components of scaled pseudo-wavenumber of the prefactored forward and

ACCEPTED MANUSCRIPT



(a) (b)

Figure 4: Dispersive characteristics of the prefactored classical C12 and cost-optimised C12–4, C12–5 and C12–6 schemes. (a) Real component of the prefactored compact scheme from Eq. (30). (b) Spatial error $\hat{e}_0$ versus number of points-per-wavelength N_λ.

	C12	C12–4	C12–5	C12–6
α_1^F	0.27639320	0.28431973	0.28002945	0.27807444
α_0^F	0.72360679	0.71568026	0.71997054	0.72192555
a_1^F	0.87939886	0.87051424	0.87520493	0.87743432
a_0^F	-0.758797731	-0.74102848	-0.75040987	-0.75486864
a_{-1}^F	-0.120601132	-0.12948575	-0.12479506	-0.12256567

Table 3: Prefactored spatial discretization coefficients of the classical C12 and optimised C12–n schemes.

backward stencils, respectively from Eqs. (32) and (34), are equal in magnitude and opposite in sign and they cancel each other out by the summation in Eq. (27), to give a dissipation-free scheme.

3.3. Cost-optimised Runge-Kutta time integration

The optimisation approach of Section 3.1 is herein extended to the Runge–Kutta time integration schemes of Eqs. (8–10). Assuming a zero spatial differentiation error, values for the coefficients γ_j are sought that maximise the temporal resolving efficiency $\hat{z}_{\mathrm{opt}}(\tilde{\epsilon})$ for a given value of the normalised error $\tilde{\epsilon}$

$$\hat{z}_{\mathrm{opt}} = \max\{\tilde{z} : \hat{e}_t(\tilde{z}, \boldsymbol{\gamma}) \leq \tilde{\epsilon}\},$$

under the following stability constraint

$$\zeta \hat{z}_{\mathrm{opt}}(\tilde{\epsilon}) \leq z_s, \tag{36}$$

where z_s is the stability limit defined in Eq. (15). The factor $\zeta \geq 1.1$ has been introduced to guarantee a greater stability margin beyond the range of well resolved angular frequencies z. The four–stage, fourth–order RK scheme is selected as the baseline temporal solver. To preserve formal second–order time integration accuracy, let $\gamma_1 = 1$ and $\gamma_2 = 1/2$. This leaves two free parameters γ_3, γ_4 to be determined from $\zeta \hat{z}_{\mathrm{opt}}$.

ACCEPTED MANUSCRIPT

Scheme	$\hat{e}_a$	σ_{max}	$\hat{c}_1^*$	$\hat{c}_1^*$(C12RK4)	$\Delta\hat{c}_1(\%)$	$\hat{e}_a$(C12RK4)	$\Delta\hat{e}_a$
C12RK4–4	10^{-4}	1.392	73.65	142.67	48.45	4.98×10^{-4}	79.92
C12RK4–5	10^{-5}	1.414	170.63	370.19	53.91	6.48×10^{-5}	84.57
C12RK4–6	10^{-6}	1.421	422.36	963.55	56.17	7.28×10^{-6}	86.26

Table 4: Theoretical performance of cost-optimised schemes C12RK4–n for different target errors in one dimensional space and comparison with the standard C12RK4 scheme. $\sigma_{max} = 1.428$ for the C12RK4 scheme.

To find optimal values for differing levels of target error $\tilde{\epsilon} = 1 \times 10^{-n}$, a global pattern search algorithm similar to the one in Section 3.1 is used, in which an additional nonlinear constraint is incorporated to satisfy Eq. (36). In this case, $\zeta = 1.1$ was found heuristically to provide a good balance between performance and stability. The results are plotted in Figure 5(a). These show that the cost–optimised time integration schemes have smaller values of coefficients γ_3 and γ_4 than the classical fourth–order accurate RK scheme, indicated by the discontinuous lines. As the level of normalised target error $\tilde{\epsilon} = 1 \times 10^{-n}$ decreases, the corresponding cost–optimised coefficients γ_3^n and γ_4^n, shown by the lines with symbols, tend asymptotically to γ_3^{RK4} and γ_4^{RK4}, respectively. The reason for this is the same as for the spatial scheme. For $\tilde{\epsilon} \to 0$, the formal order of accuracy of the scheme determines the error and the optimised coefficients tend to the ones that maximise the highest formal order of accuracy, which is four in this case. For finite values of $\tilde{\epsilon}$, by optimising γ_3^n and γ_4^n, Figure 5(a) shows that the RK4—–n schemes consistently achieve a higher temporal resolving efficiency $\hat{z}^n_{\mathrm{opt}}$ than the resolving efficiency $\hat{z}^{\mathrm{RK4}} = \max\left[\tilde{z} : \tilde{e}_t\left(\tilde{z}, \boldsymbol{\gamma}\right) \leq \tilde{\epsilon}\right]$, with $\gamma_j = 1/j!$, achieved by the classical RK4 scheme for the same target error $\tilde{\epsilon} = 1 \times 10^{-n}$.
Figure 5(b) shows the stability footprints from the Von-Neumann analysis [19] for the classical RK4 and the temporal cost-optimised RK4–n, $n = 4, 5, 6$ schemes. The stability footprints are the locus of points in the $\mathbb{C}$ complex plane where the amplification factor from Eq. (14) is equal to unity $|r_t = 1|$, in case of null spatial error. It is shown that the stability footprints for the temporal cost-optimised schemes do not differ too much from the classical non-optimised RK4 scheme. The temporal cost-optimised schemes have a slightly larger footprint than the classical one in the left half-side of the Laplace plane. The larger increment is obtained for $n = 4$.
The coefficients of the optimised RK4–n schemes γ_3^n, γ_4^n, their respective temporal resolving efficiencies $\hat{z}^n$ and the stability limits z_s^n from Eq. (15) are listed in the Table 2. The stability limits z_s^n of the temporal cost-optimised schemes, which are the intercepts of the stability footprints with the Im(z)-axis of the complex plane in Fig. 5(b), are marginally lower than the corresponding limit of the classical RK4 scheme, which is $z_s = 2\sqrt{2}$. Using the cost–optimised coefficients increases the temporal resolving efficiencies between approximately 30% and 50% over the range $10^{-6} \leq \tilde{\epsilon} \leq 10^{-4}$, indicating a gain in temporal resolving efficiency from the temporal cost-optimisation.

3.4. Theoretical performance of the cost–optimised schemes

Sections 3.2 and 3.3 considered the potential benefits of cost optimisation in solving ordinary differential equations in space and in time, respectively. This section explores combining C12–n and RK–n schemes in a single numerical procedure, denoted C12RK4–n, to solve the partial differential equation Eq. (12). The limitation on the maximum value of the Courant number $\sigma_{\max}$

ACCEPTED MANUSCRIPT

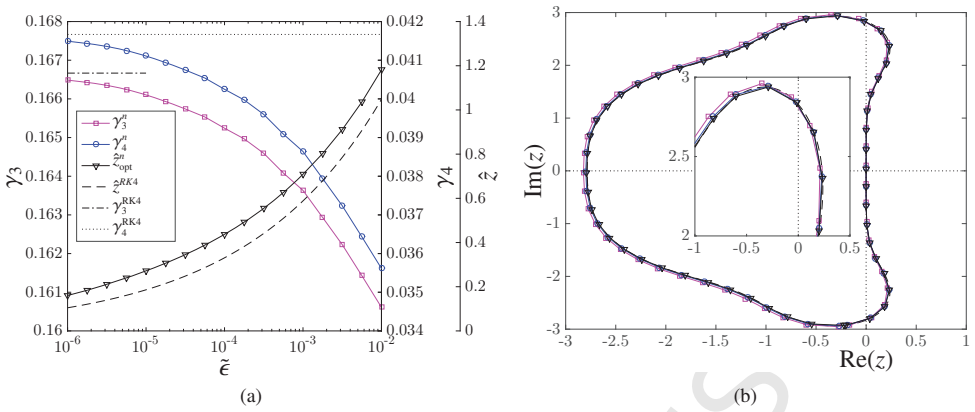


Figure 5: (a) γ_3 and γ_4 coefficients and corresponding temporal resolving efficiency $\hat{z}$ for the classical RK4 and the cost-optimised RK4—n schemes for a range of target error levels $\tilde{\epsilon} = 1 \times 10^{-n}$. (b) Stability footprints for the classical and cost-optimised RK4 time integration schemes. Line patterns as in Fig. 4.

for the class of cost-optimised C12RK4–n schemes is given by

$$\sigma_{max} = \frac{z_s}{\max\limits_{\kappa\in(0,\pi)} \bar{\kappa}\,(\kappa)}, \tag{37}$$

which depends on both the spatial and temporal discretizations. Table 4 reports the σ_{max} values for $n = 4, 5, 6$.

Figure 2(a) has shown that, given a target level of normalised error $\hat{e}(\hat{\kappa}, \hat{\sigma})$, it is possible to determine the optimal working conditions $(\hat{\kappa}^+, \hat{\sigma}^+)$ that give the smallest normalised cost $\hat{c}_1(\hat{\kappa}, \hat{\sigma})$ and Figure 2(b) has shown that the normalised approximate error $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ can be used in place of $\hat{e}(\hat{\kappa}, \hat{\sigma})$ for computing the estimations $(\hat{\kappa}^*, \hat{\sigma}^*)$ of $(\hat{\kappa}^+, \hat{\sigma}^+)$, which are listed in Table 1. Following the same procedure, the normalised cost $\hat{c}_1(\hat{\kappa}^*, \hat{\sigma}^*)$ of operating the C12RK4 and the C12RK4–n schemes for different levels of normalised approximate error $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ were determined and the results reported in Fig. 6. For each C12RK4–n scheme, the symbols denotes the normalised cost $\hat{c}_1(\hat{\kappa}^*, \hat{\sigma}^*)$ of operating this scheme at the normalised approximate error $\hat{e}_a(\hat{\kappa}^*, \hat{\sigma}^*) = 10^{-n}$, which is the design error level of the C12RK4–n scheme. Figure 6 confirms that, for each C12RK4–n scheme, the design level of error 10^{-n} of its spatial differentiation C12–n and temporal integration RK4–n constituents is retained once the spatial differentiation and temporal integration are coupled together in one partial differential equation solver.

Whereas it is encouraging to notice that the C12RK4–n schemes retain their design error level based on the approximation $\hat{e}_a(\hat{\kappa}^*, \hat{\sigma}^*)$ at their design point, it is of interest to discuss the error trend away from the $\hat{e}_a = 10^{-n}$ level, in comparison with the conventional C12RK4 scheme. All C12RK4–n schemes exhibit a rate of convergence higher than the C12RK4 scheme up to their respective design error levels. Thereafter, the C12RK4–n schemes exhibit a $\hat{e}_a$ plateau with increasing $\hat{c}_1$, followed by a convergence rate that is lower than that of C12RK4. The $\hat{e}_a$ trend is monotonically decreasing for all schemes. The variable error roll-off of the C12RK4–n schemes offers some advantages and limitations in computational physics applications. The monotonicity of the numerical error indicates that the scheme can be operated with confidence in wavenumber bandwidth limited problems knowing the design error will not be exceeded. This indicates that

ACCEPTED MANUSCRIPT

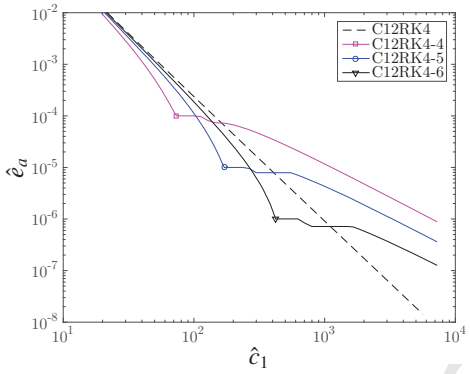


Figure 6: Error estimated by Eq. (24) as a function of $\hat{c}_1$ for C12RK4–n schemes, one-dimensional implementation. Symbols indicate $(\hat{\kappa}^*, \hat{\sigma}^*)$ values.

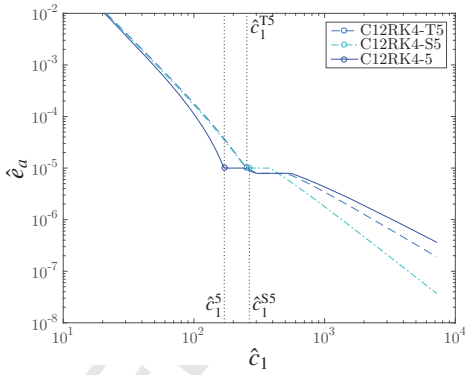


Figure 7: Comparison of the error estimated by Eq. (24) among the C12RK4–5, C12RK4–T5 and C12RK4–S5 schemes. Symbols indicate $(\hat{\kappa}^*, \hat{\sigma}^*)$ values.

each optimised scheme could be expected to perform well when operated close to its design point. The variable error roll-off rate, however, prevents the implementation of classical roll-off rate checks for mesh convergence computations, as constant error roll-off rates exhibit only at a normalised cost $\hat{c}_1$ about two orders of magnitude higher than the scheme's optimal operating point. This cost difference is likely to amount to an onerous mesh resolution in a practical computation, the rate of convergence of which requires testing by alternative means.

Table 4 shows the computational cost $\hat{c}_1^*$ required by each of the C12RK4–n schemes to achieve their design level of error $\hat{e}_a = 10^{-n}$. This is then compared with the corresponding cost $\hat{c}_1^*$(C12RK4) required by the standard C12RK4 scheme to obtain the same level of error; the percentage improvement $\Delta\hat{c}_1$, based on the cost $\hat{c}_1^*$(C12RK4) of the classical scheme, is between 48% and 56% and increases with decreasing error level. In addition, the smallest error $\hat{e}_a$(C12RK4) achieved by the standard scheme for the estimated computational costs $\hat{c}_1^*$ is reported in Table 4 and the percentage numerical error reduction $\Delta\hat{e}_a$, based on the error $\hat{e}_a$(C12RK4) of the classical scheme, is seen to be in the region of 80% for all the optimised schemes.

ACCEPTED MANUSCRIPT

Let space–only optimised schemes be denoted by C12RK4–Sn, where the standard RK4 scheme is used with the space optimised schemes of Section 3.1, and time–only optimised schemes be denoted by C12RK4–Tn, where the standard C12 spatial discretisation is combined with the temporal optimised schemes of Section 3.3; consistently, n represents the exponent of the error level. Figure 7 shows a comparison between the fully–optimised C12RK4–5 scheme and the C12RK4–S5 and C12RK4–T5 ones. The symbols and the corresponding vertical dotted lines show the smallest normalised cost $\hat{c}_1$ required by each scheme to give a normalised error $\hat{e}_a = 10^{-5}$. The values of $\hat{c}_1$ are $\hat{c}_1^5 = 170.63$, $\hat{c}_1^{S5} = 265.52$ and $\hat{c}_1^{T5} = 254.11$, respectively for the C12RK4–5, C12RK4–S5 and C12RK4–T5 schemes. To achieve an error of 10^{-5}, the fully optimised C12RK4–5 scheme, represented by a dashed line with circles, shows a cost advantage of 32.9% and 35.7% over the C12RK4–T5 and C12RK4–S5 schemes, respectively. Similar results hold for other error levels n, but these are omitted from Figure 7 for clarity. These results highlight the advantages that can be achieved by using a scheme fully optimised in both space and time.

4. Numerical Experiments

This section presents a series of numerical examples to highlight the improvement in computational efficiency that is attainable by using cost–optimised prefactored schemes. The classical C12RK4 and the cost-optimised C12RK4–n scheme of Secs. 3.2 and 3.3 are used to solve the governing equations. Periodic boundary conditions are used at the computational domain boundaries. The numerical experiments reported in this work were performed on the High Performance Computer cluster Alice at the University of Leicester. The cluster is an assembly of 208 pairs of 2.6 GHz CPUs with 64 GB RAM per node pair. The tests were run as scalar computations on one dedicated computer core with its own dedicated memory, isolated from the remainder of the hardware to prevent other jobs running concurrently on the cluster from affecting the assessment of the computational cost.

4.1. Polychromatic wave

In the first numerical example, the one-dimensional LAE (Eq. (12)) is solved on the interval $(0, 1)$ with $c = 1$, periodic boundary conditions $u(0, t) = u(1, t)$ at $t > 0$, and the initial condition $u(x, 0) = \sum_{j=1}^{4} \sin(2^{(j+1)}\pi x)$. The solution is a polychromatic wave with $\hat{k} = 32\pi$. For a given theoretical cost level $\hat{c}_1$, a simulation is run with the optimal Courant numbers $\hat{\sigma}^*$ and the optimal pseudo-wavenumbers $\hat{\kappa}^*$ found in Section 3.4. Once an approximate solution $u_h(x, t)$ has been computed, the normalised relative numerical error $\bar{e}$ is calculated. Here, $\bar{e}$ is given by

$$\bar{e}^2 = \frac{1}{(c\hat{k}T)^2} \frac{\sum_{i=1}^{N-1}(u_h(x_i, T) - u(x_i, T))^2}{\sum_{i=1}^{N-1}(u(x_i, 0))^2}, \tag{38}$$

where N is the number of nodes in the computational domain. $\bar{e}$ represents an approximation to $E/(c\hat{k}T)$ due to there being only a finite number of points at which the solution is available. However, tests where $\bar{e}$ was replaced by an error computed using a coarse grained solution yielded negligible changes to the results, except where very few points per wavelength were utilised. These results indicate that $\bar{e}$ is not affected by spurious numerical integration effects.

Figure 8(a) and Figure 8(b) show, respectively, the iso-lines of the theoretical global approximate error function $\hat{e}_a(\cdot, \cdot)$ and the iso-lines of the numerical error $\bar{e}(\cdot, \cdot)$ for the C12RK4–4 scheme with $(\hat{\kappa}, \hat{\sigma}) \in [0, 1.5] \times [0, 1]$. The iso-lines of the numerical error $\bar{e}(\cdot, \cdot)$ reported in Fig. 8(b) were

ACCEPTED MANUSCRIPT

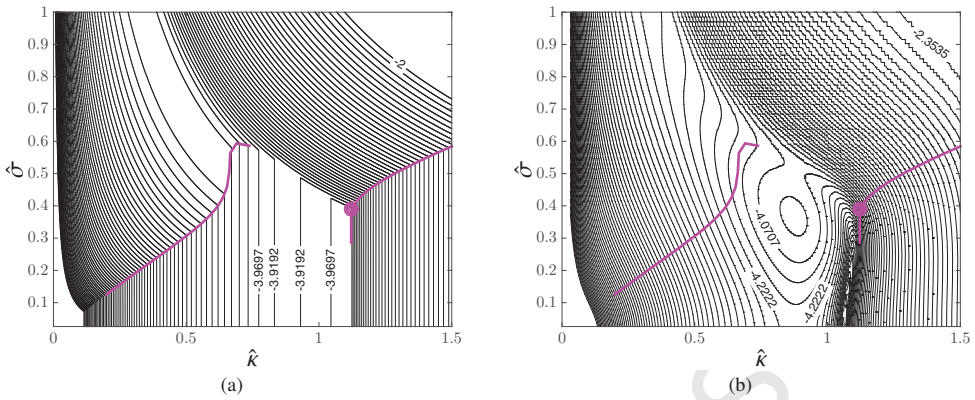


Figure 8: Polychromatic wave: (a) Iso-lines of global approximate error $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ for the C12RK4–4 scheme and (b) iso-lines of polychromatic wave numerical error. The iso-lines are equally spaced on a log-scale and the levels shown are the exponent of 10. Bold (magenta) lines indicate the locus of optimal (κ^*, σ^*) for the C12RK4–4 scheme, while the open circle corresponds to $(\hat{\kappa}^4_{\text{opt}}, \hat{\sigma}^4_{\text{opt}})$.

calculated according to Eq. (38), with $c = 1$, $\hat{k} = 32\pi$, $T = 1$ and the non-uniform $(\hat{\kappa}, \hat{\sigma})$ discretisation used by Spisso [20]. The locus of optimal points $(\hat{\kappa}^*, \hat{\sigma}^*)$ and the design point $(\hat{\kappa}^4_{\text{opt}}, \hat{\sigma}^4_{\text{opt}})$ for the C12RK4–4 scheme are overlaid on both figures by using, respectively, the thick (magenta) line and the filled (magenta) circle. The *a posteriori* computed numerical error $\bar{e}(\hat{\kappa}, \hat{\sigma})$ of Eq. (38) is dependent on the wavenumber spectrum of the problem being solved. Still, the comparison between Figure 8(a) and Figure 8(b) shows that there is a strong correlation between the theoretical error $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ and the computed numerical error $\bar{e}(\hat{\kappa}, \hat{\sigma})$ for this polychromatic wave propagation problem.

There are two branches of the locus of optimal points $(\hat{\kappa}^*, \hat{\sigma}^*)$. The design point $(\hat{\kappa}^4_{\text{opt}}, \hat{\sigma}^4_{\text{opt}})$ lies on the right hand branch. The discontinuity in the locus of optimal points is due to the non-monotonicity of the approximate spatial error function $\hat{e}_{\text{a}}(\cdot, \cdot)$ as $\hat{\kappa}$ increases from 0.75 to 1.2 for a fixed $\hat{\sigma} < 0.58$, as shown in Figure 8(a) by the contour error labels over the range $0.75 < \hat{\kappa} < 1.2$. The non-monotonic error variation is lower in magnitude than the error change outside this region, as indicated by the packing of the contour lines in Figure 8(a), so that this region can be regarded as featuring an error plateau. In practice, by increasing the cost above that given by the design point, the theoretical error initially plateaus, until the jump from the right to the left branch occurs. Thereafter, it monotonically decays with the reduction of both $\hat{\kappa}$ and $\hat{\sigma}$ as shown in Figure 6.

Figure 8(b) follows essentially the same main trend as Fig. 8(a). Interestingly, for $(\hat{\kappa}, \hat{\sigma}) \in [1, 1.2] \times [0, 0.5]$ there is a valley where the computed numerical error is significantly lower than the theoretical predictions. The presence of the valley is due to the actual numerical error $\bar{e}$ being the sum of local errors associated with the different harmonics which comprise the solution, while the error estimate $\hat{e}_a$ is the maximum error over the entire wavenumber range using the worst case assumptions of Eq. (25).

Figure 9(a) plots the numerical error $\bar{e}$ against cost $\hat{c}_1$, using the same optimal conditions $(\hat{\kappa}^*, \hat{\sigma}^*)$ as in Figure 6, for the optimised schemes C12RK4–n (lines with symbols) and the standard C12RK4 scheme (solid lines), at a non-dimensional time $T = 1$ and overlays the theoretical

ACCEPTED MANUSCRIPT

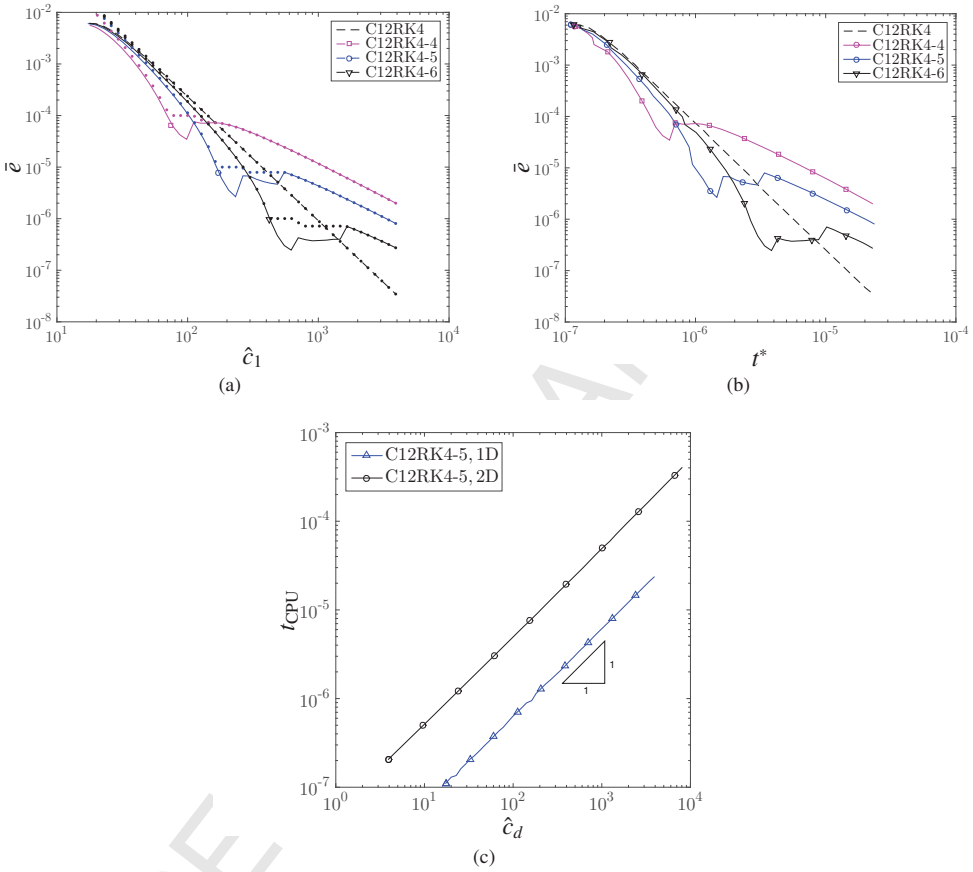


Figure 9: Polychromatic wave: (a) Computed numerical error (lines with symbols) as a function of one-dimensional cost function $\hat{c}_1$, overlaid with theoretical predictions (dotted lines). (b) Numerical error as a function of normalised CPU time t_{CPU} for the classical C12RK4 and optimised C12RK4–n schemes. (c) Normalised CPU time t_{CPU} compared with theoretical cost, in one- and two-space dimensions.

ACCEPTED MANUSCRIPT

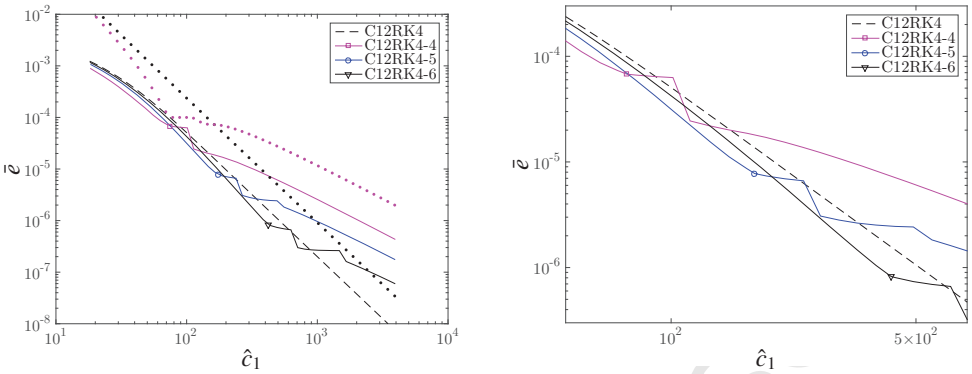


Figure 10: One-dimensional Gaussian pulse: (a) Numerical error as a function of one-dimensional cost function $\hat{c}_1$ with error estimates $\hat{e}_a$ overlaid (dotted lines) and (b) detail in the vicinity of the C12RK4–4, C12RK4–5 and C12RK4–6 design points.

results (dotted lines) from Figure 8(a). The normalised numerical errors $\bar{e}\,(\hat{\kappa}^*, \hat{\sigma}^*)$ computed *a posteriori* compare very favourably with the approximate error estimates $\hat{e}_a\,(\hat{\kappa}^*, \hat{\sigma}^*)$. Each of the optimised schemes gives a significant computational cost reduction over the C12RK4 scheme close to its design level of error, as indicated by the open symbols lying left of the dashed line, although the maximum cost reduction is seen at error levels slightly below the design level of error. The approximate error $\hat{e}_a\,(\hat{\kappa}^*, \hat{\sigma}^*)$ in Figure 9(a) is shown to be consistently of the same order of magnitude as the *a posteriori* numerical error $\bar{e}\,(\hat{\kappa}^*, \hat{\sigma}^*)$ over the normalised cost range $2 \times 10^1 \leq \hat{c}_1 \leq 4 \times 10^3$. More specifically $\hat{e}_a \geq \bar{e}$. This is a desirable feature, as it indicates that $\hat{e}_a$ can be used as an upper bound indicator for the normalised error for the application of a cost optimised prefactored compact scheme to multi-scale linear advection problems. It implies that the numerical solution will not exceed an acceptable level of numerical error $\bar{e} = 10^{-n}$ that is arbitrarily set by the numerical modeller selecting a specific C12RK4–n scheme. By construction, from Eq. (21), $\hat{e}$ and its approximation $\hat{e}_a$ represent the maximum normalised error over the wavenumber and Courant number range $[0, \hat{\kappa}] \times [0, \hat{\sigma}]$. This implicitly attributes the same importance to all the (κ, σ) scales over this range [1], whereas the most energetic scales in a wave propagation problem may differ from the ones associated to the normalised error maximum $\hat{e}$ from Eq. (21), which typically lie towards the upper limit $\hat{\kappa}$ of the wavenumber range, as shown in Figure 4(a). For instance, in the propagation of broadband noise, the sound intensity spectrum decays at higher wavenumbers and frequencies so that propagation errors that occur at these scales have a modest impact on the overall sound pressure level prediction. $\hat{e}$ and its approximation $\hat{e}_a$ therefore represent the worst case scenario of the most energetic scale in a multi-scale wave propagation problem being also the scale of maximum normalised propagation error $\hat{e}$ in Eq. (21), from which $\hat{e}_a \geq \bar{e}$.

Figure 9(a) also shows that, for each of the C12RK4–n schemes, there is a discrepancy between the numerical errors and the theoretical predictions in Figure 9(a) near its design level of error $\tilde{\epsilon} = 10^{-n}$. In this region, the theoretical error $\hat{e}_a$ is larger than the numerical error $\bar{e}$. This trend is explainable by considering the right-hand branch of the $(\hat{\kappa}^*, \hat{\sigma}^*)$ locus in Figure 8(b), moving from right to left. Using the C12RK4–4 scheme near to its design point, the locus of

ACCEPTED MANUSCRIPT

Scheme	$\bar{e}$	$\hat{c}_1^*$	$\hat{c}_1^*$(C12RK4)	$\Delta\hat{c}_1(\%)$	$\bar{e}$(C12RK4)	$\Delta\bar{e}(\%)$
C12RK4–4	6.443×10^{-5}	74.376	170.56	56.39	4.706×10^{-4}	86.31
C12RK4–5	7.702×10^{-6}	171.7987	412.19	58.32	6.332×10^{-5}	87.84
C12RK4–6	9.565×10^{-7}	424.136	981.3	56.78	7.191×10^{-6}	86.71

Table 5: Polychromatic wave: performance of cost-optimised schemes at their design points $(\hat{\kappa}^n_{\text{opt}}, \hat{\sigma}^n_{\text{opt}})$ and comparison with the standard C12RK4 scheme. $\Delta\hat{c}_1(\%)$ and $\Delta e(\%)$ indicate the percentage cost and error reduction with respect to the C12RK4 scheme.

Scheme	$\bar{e}$	$\hat{c}_1^*$	$\hat{c}_1^*$(C12RK4)	$\Delta\hat{c}_1(\%)$	$\bar{e}$(C12RK4)	$\Delta\bar{e}(\%)$
C12RK4–4	6.784×10^{-5}	74.677	88.29	15.42	9.954×10^{-5}	31.85
C12RK4–5	7.7647×10^{-6}	173.227	218.78	20.82	1.365×10^{-5}	43.12
C12RK4–6	8.2525×10^{-7}	424.245	554.45	23.48	1.573×10^{-6}	47.54

Table 6: One-dimensional Gaussian pulse: performance of cost-optimised schemes at their operating points $(\hat{\kappa}^n_{\text{opt}}, \hat{\sigma}^n_{\text{opt}})$ and comparison with standard C12RK4 scheme. $\Delta\hat{c}_1(\%)$ and $\Delta\bar{e}(\%)$ indicate the percentage cost and error reduction with respect to the C12RK4 scheme.

optimal points begins to enter the valley of low error, thus $\bar{e}$ diverges from $\hat{e}_a$ due to this valley not being present in Figure 8(a). As the locus of points 'climbs' back out of the valley, $\bar{e}$ increases with cost until once again $\hat{e}_a \approx \bar{e}$. A similar discussion applies to both the C12RK4–5 and C12RK4–6 curves in Figure 9(a).

Table 5 reports the performance of the C12RK4–n schemes benchmarked against the C12RK4 scheme at their respective design points and confirms quantitatively that $\bar{e} \simeq \hat{e}_a$. For each optimised scheme operating at its design point, the numerical error $\bar{e}$ is slightly below the corresponding design level of error $\tilde{\epsilon}$. Table 5 also reports the percentage cost reduction $\Delta\hat{c}_1$ over the standard C12RK4 scheme in achieving the same numerical error level $\bar{e}$ attained by each optimised scheme operated at its design point. This cost reduction varies from 50% to 60% and it increases with n. Similarly, the percentage error reduction $\Delta\bar{e}$ achieved by each optimised scheme over the C12RK4 scheme for the same cost $\hat{c}_1^*$ is over 80%.

Figure 9(b) shows the normalised numerical error $\bar{e}$ against the normalised computational time t_{CPU}, at the same $(\hat{\kappa}^*, \hat{\sigma}^*)$ as in Figure 9(a). Here, to be consistent with Eq. (20), $t_{\text{CPU}} = T_{\text{CPU}}/((c\hat{k}T)(\hat{k}L))$, where T_{CPU} is the CPU time used by the simulation in seconds. Computations were carried out on a machine with an Intel Xeon Ivy Bridge CPU running at 2.6GHz and the computational time was averaged over 100 runs per $(\hat{\kappa}^*, \hat{\sigma}^*)$ pair to ensure the consistency of the results. Qualitatively, Figure 9(b) is very similar to Figure 9(a). This indicates that $\hat{c}_1$ is a good pseudo-variable for the normalised CPU time; this is confirmed in Figure 9(c), where the normalised CPU time, t_{CPU}, is plotted against $\hat{c}_1$ for the C12RK4–5 scheme and shows $\hat{c}_1 \propto t_{\text{CPU}}$. The computations were repeated for the C12RK4–4 and C12RK4–5 schemes and the constant of proportionality was found to be the same for all three schemes. Using the $t_{\text{CPU}} \propto \hat{c}_1$ relationship from Figure 9(c), each optimised scheme offers the same (percentage) reduction in CPU time over the C12RK4 scheme as the $\Delta\hat{c}(\%)$ reported in Table 5.

4.2. One-dimensional Gaussian pulse

The second sample application of the C12RK4–n schemes is derived from Test case B from [16]. A numerical solution is sought for the one-dimensional LAE equation Eq. (12) over the domain $[-100, 100]$, with $c = 1$, periodic boundary conditions, and the initial condition

ACCEPTED MANUSCRIPT

Scheme	$\tilde{\epsilon}$	$\hat{c}_2^*$	$\hat{c}_2^*$(C12RK4)	$\Delta\hat{c}_2(\%)$	$\hat{e}_a$(C12RK4)	$\Delta\hat{e}_a(\%)$
C12RK4–4	10^{-4}	65.69	187.19	64.91	6.215×10^{-4}	83.91
C12RK4–5	10^{-5}	219.98	708.56	68.95	7.557×10^{-5}	86.77
C12RK4–6	10^{-6}	792.86	2699.30	70.63	8.238×10^{-6}	87.86

Table 7: Theoretical performance of cost-optimised schemes for different target errors in two dimensional space.

$u(x,0) = \frac{1}{2}\mathrm{e}^{-(x/3)^2}$. The solution has a broadband Fourier decomposition and provides a good test of the C12RK4–n schemes applied to a broadband signal. The characteristic length scale for this problem is $\lambda = 6$, which corresponds to $\hat{k} = \pi/3$. For a given cost level $\hat{c}_1$, the optimal Courant number $\hat{\sigma}^*$ and the optimal pseudo-wavenumber $\hat{\kappa}^*$ have been selected from Figure 6 for $\hat{e}_a = 10^{-4}$, 10^{-5} and 10^{-6}. These set the spatial and temporal discretisations (h and Δt) for the application of the cost optimised schemes. Figure 10(a) plots the normalised numerical error $\bar{e}$, computed using Eq. (38), against cost $\hat{c}_1$ for each of the optimised schemes C12RK4–n and the standard C12RK4 scheme at the final time $T = 100$. In Figure 10(a), all optimal operating points of the C12RK4–n schemes, denoted by the open symbols, lie to the left of the dashed line, therefore, numerical solutions with the same level of normalised numerical error $\bar{e}$ can be obtained for a lower computational cost than by using the C12RK4 scheme. In this case, the reduction in computational cost $\hat{c}_1$ obtained by using the optimised schemes is not as large as in the polychromatic wave test case, but there is still a clear advantage in using the optimised schemes at their respective design points. This is better appreciated in the enlargement of Figure 10(b) in the vicinity of $\bar{e} = 1 \times 10^{-6}$ and $\bar{e} = 1 \times 10^{-5}$. Table 6 shows a comparison between the optimised schemes C12RK4–n, $n = 4, 5, 6$, operating at their design points and the standard C12RK4 scheme. Table 6 confirms that, when operating at their design points, all optimised schemes require a 15% to 25% lower cost than the C12RK4 scheme to achieve the same normalised numerical error $\bar{e}$. Similarly, when the optimised schemes are operated at their design point, they offer between a 30% and 50% reduction in normalised numerical error for the same normalised cost $\hat{c}_1^*$ when compared with the C12RK4 scheme.

Notwithstanding the positive outcome from this broadband wave propagation test, which points to a consistent performance advantage of the C12RK4–n schemes over the C12RK4 scheme, this performance advantage is significantly lower than that predicted by the theory in Table 4. The background of this lower performance gain is explored by reporting in Figure 10(a), by the dotted lines, the approximate normalised error $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ versus the normalised cost $\hat{c}_1$ for the C12RK4–6 and C12RK4 schemes. Plots of the approximate normalised error for the other optimised schemes are omitted for clarity. Recall that Figure 4(a) indicates an increase in dispersion with increasing wavenumber. Given that the Gaussian pulse has a definite roll-off at high wavenumbers, $\hat{e}_a(\hat{\kappa}, \hat{\sigma})$ overestimates $\bar{e}$, as shown by the dotted curves lying above the continuous and discontinuous line in Figure 10(a). The approximate errors can be considered to be upper bound estimates for $\bar{e}$ because $\hat{e}_a$ assumes the error is generated at all wavenumbers $\leq \hat{\kappa}$ whereas, in a physical problem, some wavenumbers may have comparatively negligible energy levels and make a negligible contribution to the actual error $\bar{e}$. Given that the baseline *a posteriori* normalised numerical error is lower than $\hat{e}_a$, a more modest reduction in cost $\hat{c}_1$ to achieve $\bar{e} = 10^{-n}$ appears to be a reasonable outcome by the use of a C12RK4–n scheme.

ACCEPTED MANUSCRIPT

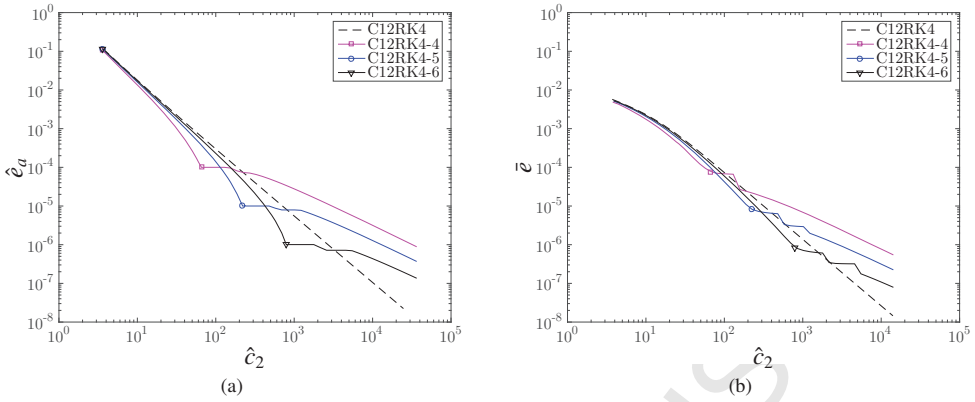
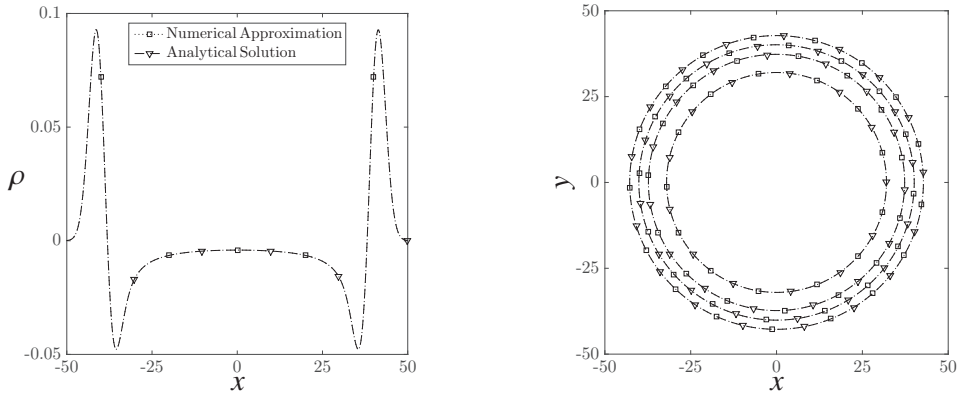


Figure 11: Two-dimensional Gaussian pulse: (a) Theoretical errors as a function of $\hat{c}_2$ for space-time optimised schemes. (b) Numerical errors as a function of cost $\hat{c}_2$.



(a) Non-dimensional density perturbation ρ along the y=0 line.

(b) Contours of non-dimensional density perturbation, contour levels: -0.025, 0.075. Line patterns as in Figure 12(a).

Figure 12: Propagation of a two-dimensional acoustic pulse on an unbounded domain at non-dimensional $T = 40$, fixed $\sigma = 0.05$.

Scheme	$\bar{e}$	$\hat{c}_2^*$	$\hat{c}_2^*$(C12RK4)	$\Delta\hat{c}_2(\%)$	$\bar{e}^*$(C12RK4)	$\Delta\bar{e}(\%)$
C12RK4–4	7.581×10^{-5}	66.93	97.85	31.61	1.449×10^{-4}	47.67
C12RK4–5	8.297×10^{-6}	221.72	352.74	37.14	1.850×10^{-5}	55.16
C12RK4–6	8.540×10^{-7}	799.93	1317.42	39.28	2.020×10^{-6}	57.73

Table 8: Two-dimensional Gaussian pulse: performance of cost-optimised schemes at their design point ($\hat{\kappa}^n_{\text{opt}}, \hat{\sigma}^n_{\text{opt}}$) and comparison with the standard C12RK4 scheme. $\Delta\hat{c}_2(\%)$ and $\Delta\bar{e}(\%)$ indicate the percentage cost and error reduction with respect to the C12RK4 scheme.

ACCEPTED MANUSCRIPT

4.3. Linearised Euler Equations

The last test case is the application of the same optimised schemes from Section 2 to a broad-band two-dimensional linear wave propagation problem, which is governed by the Linearised Euler Equations (LEE) in two-dimensions. The two-dimensional normalised LEE equations, in non-conservative form, are

$$\frac{\partial U}{\partial t} + A_0 \frac{\partial U}{\partial x} + B_0 \frac{\partial U}{\partial y} = 0, \tag{39}$$

where

$$U = \begin{bmatrix} \rho \\ u \\ v \\ p \end{bmatrix}, \quad A_0 = \begin{bmatrix} M_x & 1 & 0 & 0 \\ 0 & M_x & 0 & 1 \\ 0 & 0 & M_x & 0 \\ 0 & 1 & 0 & M_x \end{bmatrix}, \quad B_0 = \begin{bmatrix} M_y & 0 & 1 & 0 \\ 0 & M_y & 0 & 0 \\ 0 & 0 & M_y & 1 \\ 0 & 1 & 0 & M_y \end{bmatrix}, \tag{40}$$

with M_x the Mach number in the x-direction and M_y the Mach number in the y-direction. In this example $M_x = M_y = 0$ and Eq. (39) is solved in the non-dimensional spatial domain $(-100, 100)^2$ with the initial conditions

$$U_0 = \begin{bmatrix} \mathrm{e}^{-\ln(2)(x^2+y^2)/9} \\ 0 \\ 0 \\ \mathrm{e}^{-\ln(2)(x^2+y^2)/9} \end{bmatrix} \tag{41}$$

that represent an initially quiescent medium perturbed by a two-dimensional Gaussian pulse. $M_x = M_y = 0$ is used for consistency with the numerical example presented in [1]. Periodic boundary conditions are applied at the computational domain perimeter and the solution is time-advanced to the non-dimensional time $T = 40$ at which the pulse, which travels at the speed of sound, has yet to reach the perimeter, which therefore remains unperturbed. The initial condition of Eq. (41) has characteristic wavenumber $\hat{k} = \pi (\ln 2)^{1/2} /3$.

The procedure of Section 2 is used to estimate the optimal values $(\hat{\kappa}^*, \hat{\sigma}^*)$ which give the smallest error estimate $\hat{e}$, from Eq. (21), for a given cost estimate $\hat{c}_2$, from Eq. (22), for the C12RK4 and C12RK4–n, $n = 4, 5, 6$ schemes. Figure 11(a) shows the resulting cost-error trends, using the same notation as Figure 6. The open symbols represent the estimated optimal operating points $(\hat{\kappa}_{\mathrm{opt}}, \hat{\sigma}_{\mathrm{opt}})$ of each scheme and indicate a good cost-advantage potential where the C12RK4–n schemes work at these spatial and temporal discretization level. A quantitative comparison in the predicted performance between the standard C12RK4 scheme and each of the cost-optimised C12RK4–n schemes is presented in Table 7, for the optimal values $(\hat{\kappa}_{\mathrm{opt}}, \hat{\sigma}_{\mathrm{opt}})$. All optimised schemes are predicted to provide more than 60% reduction in cost $\Delta\hat{c}_2$ to achieve the same magnitude of normalised numerical error as the C12RK4 scheme. Furthermore, all optimised schemes are predicted to give about an 80% reduction in normalised numerical error $\Delta\hat{e}_a$ compared to the C12RK4 scheme for the same computational cost estimate $\hat{c}_2^*$. Comparing $\Delta\hat{c}_d$ and $\Delta\hat{e}_a$ values between Tables 4 and 7 shows that these error and cost reductions from using C12RK4–n schemes at their respective optimal operating points are predicted to increase in two-dimensional applications.

These cost and error reduction estimates are verified *a posteriori* by solving numerically the two-dimensional Gaussian pulse propagation problem of Eqns. (39), (41). The baseline solution is provided by the C12RK4 scheme used on a uniform spatial mesh with $h = 0.25$ in both coordinate directions and $\sigma = 0.05$. Figures 12(a) and 12(b) show a comparison between the numerical approximation and the analytical solution of the two-dimensional propagation of the

ACCEPTED MANUSCRIPT

acoustic pulse at the non-dimensional computational time $T = 40$. There is no appreciable azimuthal distortion of the wave, showing that the numerical solution does not suffer from any appreciable degradation from the wave front being not-orthogonal to the mesh.

Equivalently, solutions are obtained using the C12RK4–n schemes with spatial and temporal discretizations corresponding to their $(\hat{\kappa}^n_{\text{opt}}, \hat{\sigma}^n_{\text{opt}})$ design points. From these solutions, the *a posteriori* normalised numerical error $\bar{e}$ at $T = 40$ is computed as

$$\bar{e}^2 = \frac{1}{(c\hat{k}T)^2} \frac{\sum_{j=1}^{N-1} \sum_{i=1}^{N-1} (\rho_h(x_i, y_j, T) - \rho(x_i, y_j, T))^2}{\sum_{j=1}^{N-1} \sum_{i=1}^{N-1} (\rho(x_i, y_j, 0))^2}, \tag{42}$$

where N is the number of nodes in each coordinate direction and each (x_i, y_j) is a coordinate of a node in the mesh. The normalised numerical error computed *a posteriori* is shown in Figure 11(b). Each component of the solution is a broadband signal that spans the full wavenumber spectrum. Thus, similarly to the one-dimensional Gaussian pulse propagation problem of Section 4.2, the numerical results from the cost-optimised schemes do not show the same reduction in cost over the C12RK4 scheme as the theoretical results. Nonetheless, qualitatively, the numerical errors show a similar trend to the theoretical results in Figure 11(a). At its design point $(\hat{\kappa}^n_{\text{opt}}, \hat{\sigma}^n_{\text{opt}})$, each C12RK4–$n$ scheme shows a reduction in cost compared with the C12RK4 scheme and with the other C12RK4–n schemes to achieve the same level of normalised numerical error $\bar{e}$. Indeed, Table 8 shows that each cost-optimised scheme gives a reduction in normalised cost $\Delta\hat{c}_2$ of over 30%, when operating at its design point $(\hat{\kappa}^n_{\text{opt}}, \hat{\sigma}^n_{\text{opt}})$, compared to the standard C12RK4 scheme. In addition, the optimised schemes operated at their design points give approximately a 50% reduction in error $\Delta\bar{e}$ over the standard C12RK4 scheme for the same cost $\hat{c}^*_2$(C12RK4).

Finally, Figure 9(c) confirms that in two-dimensions the cost function $\hat{c}_2$ is a good pseudo-variable for the normalised CPU-time $t_{\text{CPU}} = T_{\text{CPU}}/((c\hat{k}T)(\hat{k}L)^2)$.

5. Applicability of Cost-Optimised Schemes to Three-dimensional and to Non-linear Problems

The present work has dealt only with the optimisation of interior compact finite difference schemes and the understanding of their wave propagation properties in unbounded, periodic, one- and two-dimensional domains. Most real flow physics problems are further complicated by boundary condition effects, three-dimensionality, and non-linearity.

A preliminary assessment of the impact on the stability and accuracy of the prefactored cost-optimised schemes when coupled with high-order boundary closures is given in [20]. The baseline solver coupled with high-order boundary closures is shown to be stable [20]. The L_2-norm error between the analytical reference solution and the numerical prediction obtained from the C12RK4 scheme using a CFL number of $O\left(10^{-2}\right)$ was shown to match the sixth-order roll-off predicted by the truncation error of Eq. (1). Future efforts will focus on designing suitable boundary closures for the optimised schemes.

Preliminary solutions of three-dimensional linear wave propagation problems have been obtained using the C12RK4 scheme. The results show that the C12RK4 scheme coupled with a buffer zone at the through-flow boundaries is able to model the rotor-stator interaction in a three-dimensional turbomachinery geometry, producing engineering accurate aeroacoustic predictions [23].

The proposed methods are expected to encounter some difficulties when applied to multi-dimensional systems of equations with differing wave speeds. In such cases, the overall solution

ACCEPTED MANUSCRIPT

error, in a linear setting, can be estimated from a weighted average of the errors associated with each wavenumber.

It is expected that the present analysis may also give some benefits for weakly non-linear problems, because the leading order effects due to variable wave speed and multiple flow scales are naturally accounted for. A rigorous assessment of this statement is made difficult by the lack of appropriate explicit closed-form solutions and further efforts are required in these directions [1]. Non-linear problems can lead to the wavenumber spectrum of the solution varying with time. As such, it is worth considering an adaptive algorithm that can detect the changes in wavenumber and modify the mesh spacing and time steps accordingly. In conclusion, for real flow physics problems, the cost-optimised schemes are expected to retain some advantage over the non-optimised ones, even though the actual improvement may be less than the theoretical prediction reported in Section 6.

6. Conclusion

A class of optimised prefactored schemes have been developed based on fully spatio-temporal *a priori* error estimates. The following procedure can be outlined in order to achieve the best performance for a given physical problem in which a typical maximum wavelength (and associated wavenumber $\hat{k}$), a typical maximum propagation speed $\hat{c}$, and final time T can be identified: i) given the allowed error tolerance ϵ, the corresponding normalized error level is determined $\tilde{\epsilon} = \epsilon/(\hat{c}\,\hat{k}\,T)$; ii) the most efficient prefactored space- and time- discretization scheme is selected among the optimised ones herein proposed upon inspection of the respective cost vs. error plots, as given for instance in Fig. 2; iii) for the selected scheme, optimal values of the scaled wavenumber ($\hat{\kappa}^*(\tilde{\epsilon})$) and of the Courant number ($\hat{\sigma}^*(\tilde{\epsilon})$) are determined; iv) optimal values of the grid spacing and of the time step are finally determined according to $h^*(\tilde{\epsilon}) = \hat{\kappa}^*(\tilde{\epsilon})/\hat{k}, \quad \Delta t^*(\tilde{\epsilon}) = \hat{\sigma}^*(\tilde{\epsilon})\, h^*(\tilde{\epsilon})/\hat{c}$.

A theoretical performance analysis shows the clear superiority of optimised schemes over classical CAA schemes, especially if high-fidelity approximations are sought in multiple space dimensions. The proposed cost reduction is, according to Eq. (22), more effective in higher dimensions, due to the exponent of $\hat{\kappa}$ being $d + 1$. The theory suggests that a cost reduction of over 75% is possible in three dimensions when using the C12RK4-n schemes, with $n = 4, 5, 6$, compared with a maximum of 56% and 70% in one- and two-dimensions, respectively. Numerical experiments convincingly support the above statements, showing a cost reduction for a given error level of $O(50\%)$ for one-dimensional monochromatic wave propagation and of about 20% for a two-dimensional acoustic pulse propagation. This feature, coupled with simplicity in the enforcement of numerical boundary conditions imparted by pre-factorization, makes the present class of schemes especially promising for carrying out computations of long-range acoustic propagation. It is important to note that the favourable properties of space-and-time optimised schemes depend sensitively on the choice of the mesh size and time step, which should be sufficiently close to the nominal optimal point of operation. It is an open issue whether the same features also carry on to the fully non-linear regime (i.e. turbulent flow) and to flow cases in which no *a priori* estimate is given for the bulk flow properties (i.e. length and time scales).

7. Acknowledgements

This research project has been supported by a Marie Curie Early Stage Research Training Fellowship of the European Community's Sixth Framework Programme under contract number

ACCEPTED MANUSCRIPT

MEST CT 2005 020301.

References

[1] S. Pirozzoli, Performance analysis and optimization of finite-difference schemes for wave propagation problems, Journal of Computational Physics 222 (2007) 809–831.

[2] R. Hixon, Prefactored small-stencil compact schemes, Journal of Computational Physics 165 (2) (2000) 522–41.

[3] G. UK, The Future of Air Transport, White Paper and the Civil Aviation Bill, Tech. Rep., UK Department of Transport, 2003.

[4] ACARE, The Strategic Research Agenda SRA-1, SRA-2 and the 2008 Addendum to the Strategic Research Agenda, Tech. Rep., http://www.acare4europe.org/, 2009.

[5] W. D. Rock, W. Desmet, M. Baelmans, P. Sas (Eds.), An overview of high-order finite difference schemes for computational aeroacoustics, Proceedings of the 2004 International Conference on Noise and Vibration Engineering, ISMA, Leuven, Belgium, 2004.

[6] K. A. Kurbatskii, R. R. Mankbadi, Review of computational aeroacoustics algorithms, International Journal of Computational Fluid Dynamics 18 (6) (2004) 533–546.

[7] T. Colonius, S. Lele, Computational aeroacoustics: progress on nonlinear problems of sound generation, Progress in Aerospace Sciences 40 (2004) 365–416.

[8] J. Hardin, J. R. Ristorcelli, C. Tam (Eds.), ICASE/LaRC Workshop on Benchmark Problems in Computational Aeroacoustics (CAA), NASA Conference Publication 3300, Hampton, Virginia, 1995.

[9] J. Hardin, C. Tam (Eds.), Second Computational Aeroacoustics (CAA) Workshop on Benchmark Problems, NASA Conference Publication, Hampton, Virginia, Langley Research Center, 1996.

[10] J. Hardin, D. Huff, C. Tam (Eds.), Third Computational Aeroacoustics (CAA) Workshop on Benchmark Problems, NASA Conference Publication 2000-209790, Cleveland, Ohio, Langley Research Center, 2000.

[11] S. K. Lele, Compact finite difference schemes with spectral-like resolution, Journal of Computational Physics 103 (1) (1992) 16–42.

[12] R. Hixon, On Increasing the accuracy of MacCormack schemes for Aeroacoustic Applications, Nasa Contract Report 202311 ICOMP-96-11, NASA, 1996.

[13] G. Ashcroft, X. Zhang, Optimized prefactored compact schemes, Journal of Computational Physics 190 (2) (2003) 459–477.

[14] I. Spisso, A. Rona, A selective overview of high-order finite difference schemes for aeroacoustic applications, in: International Conference on Sound and Vibration14th, Cairns, Australia, 2007.

[15] A. Rona, I. Spisso, M. Bernardini, S. Pirozzoli, Comparison of optimized high-order finite-difference compact scheme for computational aeroacoustics, Conference paper 2009-0498, Aerospace Sciences Meeting and Exhibit47th, Orlando, Florida, 2009.

[16] M. Bernardini, S. Pirozzoli, A general strategy for the optimization of Runge-Kutta schemes for wave propagation phenomena, Journal of Computational Physics 228 (2009) 4182–4199.

[17] C. Bogey, C. Bailly, A family of low dispersive and low dissipative explicit schemes for flow and noise computations, Journal of Computational Physics 194 (1) (2004) 194–214.

[18] F. Q. Hu, M. Y. Hussaini, J. L. Manthey, Low-dissipation and low-dispersion Runge-Kutta schemes for computational acoustics, Journal of Computational Physics 124 (1) (1996) 177–191.

[19] C. Hirsch, Numerical Computation of Internal and External flow, vol. 1, Fundamental of Computational Fluid Dynamics, New York, second edn., 2007.

[20] I. Spisso, Development of a Prefactored High-Order Compact Scheme for Low-Speed Aeroacoustics, Ph.D. thesis, University of Leicester, optional, optional, 2013.

[21] R. Hixon, E. Turkel, Compact Implicit MacCormack-Type Schemes with High Accuracy, Journal of Computational Physics 158 (1) (2000) 51–70.

[22] C. K. W. Tam, J. C. Web, Dispersion-Relation-Preserving finite difference schemes for Computational Acoustics, Journal of Computational Physics 107 (1993) 262–281.

[23] P. Ghillani, Aeroacustic simulation of a linear cascade by a prefactored compact scheme, Ph.D. thesis, University of Leicester, Leicester, UK, 2012.

ACCEPTED MANUSCRIPT