BIT manuscript No.
(will be inserted by the editor)

Faster SDC convergence on non-equidistant grids by DIRK sweeps*

Martin Weiser

the date of receipt and acceptance should be inserted later

Abstract Spectral deferred correction methods for solving stiff ODEs are known to converge reasonably fast towards the collocation limit solution on equidistant grids, but show a less favourable contraction on non-equidistant grids such as Radau-IIa points. We interprete SDC methods as fixed point iterations for the collocation system and propose new DIRK-type sweeps for stiff problems based on purely linear algebraic considerations. Good convergence is recovered also on non-equidistant grids. The properties of different variants are explored on a couple of numerical examples.

Keywords spectral deferred correction · diagonally implicit Runge-Kutta · contraction rate

Mathematics Subject Classification (2000) 65L06 · 65M20 · 65M70

1 Introduction

Spectral deferred correction methods (SDC) for solving ODEs are iterative schemes based on approximately integrating defect equations with simple low order methods. They have been introduced by Dutt, Greengard and Rokhlin [9] as a more directly derived variant of the classical iterated defect correction methods [11, 23]. One of the main differences is the

* Supported by the DFG Research Center MATHEON "Mathematics for key technologies" in Berlin, Projects A17 and F9. A preprint is available as ZIB Report 13-30. This version has been accepted for BIT Numerical Mathematics at 2014-09-11. The final publication is available at Springer via http://dx.doi.org/10.1007/s10543-014-0540-y.

M. Weiser
Zuse Institute Berlin
Takustr. 7
14195 Berlin
Germany
Tel.: +49-30-84185-170
Fax: +49-30-84185-107
E-mail: weiser@zib.de

derivation in terms of the Picard integral equation instead of the differential equation itself. A very similar approach, but staying closer to the classical defect correction structure, has been suggested by Auzinger et al. [1–3].

SDC methods can be and have been interpreted as quite a number of different methods. For a fixed number of sweeps, they are foremost Runge-Kutta methods, the properties of which, such as order or accuracy and stability domains, have been studied extensively [7,9]. Applied to nonlinear differential equations, they can be seen as Newton-like methods for solving the collocation system. Applied to the linearized system, they form preconditioners for Krylov methods [5,14,15] or fixed point iterations in their own right [11]. Recently, SDC methods have also been used to construct efficient time-parallel solvers [6,10].

The interpretation of SDC methods as linear fixed point iterations will be presumed here. Convergence results for deferred correction methods based on the contractivity of linear fixed point iterations are scattered throughout the literature, we mention just [2,11, 14,15,21]. While this perspective has been used to analyze various SDC methods, it does not appear to have been used for the construction of efficient SDC variants. This is the aim pursued in this work. The ultimate motivation and intended application is the application to reaction-diffusion equations such as cardiac excitation [4], the properties of which are to be taken into account. In particular we restrict the attention to negative real eigenvalues of the Jacobian and to the occurence of very stiff components. Moreover, we anticipate the use of iterative solvers for the implicit basic steps and neglect the potential advantage of reusing factorizations of the Jacobian as considered, e.g., in [2]. In Section 2 we will introduce the notation of SDC methods used here from the perspective of linear algebra. This will be exploited in Sections 3 and 4 for the construction of specialized correction sweeps with diagonally implicit Runge-Kutta (DIRK) structure. Stability and accuracy domains of the resulting methods are experimentally investigated in Section 5. Finally, numerical experiments are performed at three different ODE systems with different properties in Section 6, illustrating the different convergence properties of the constructed methods.

2 A linear algebra view on SDC convergence

First we will recollect spectral deferred correction methods in two variants, mainly to introduce a consistent notation. The first one is based on a direct defect equation formulation, close in spirit to the "classical" defect correction methods, and the second one relies on the equivalent Picard equation as introduced in [9]. Subsequently, we will interprete SDC methods as fixed point iterations for solving linear collocation equations and investigate their convergence properties.

2.1 Spectral deferred correction methods

We consider approximate solutions of the initial value problem

$$\dot{y}(t) = f(y(t), t), \quad y(0) = y_0 \tag{2.1}$$

on the interval $[0, \tau]$.

Spectral differentiation: DSDC. With an approximate polynomial solution $y^0 \in \mathbb{P}_n$ at hand, the error $\delta = y - y^0$ satisfies the defect equation

$$\dot{\delta}(t) = f(y(t), t) - \dot{y}^0(t) = f(y^0(t) + \delta(t), t) - \dot{y}^0(t), \quad \delta(0) = 0, \tag{2.2}$$

which can be approximately solved by integration with a simple time stepping scheme on a time grid $0 = t_0 < t_1 < \cdots < t_n \le \tau$. Popular choices are equidistant nodes, or Gauß and Radau points. Let $\tau_i = t_i - t_{i-1}$ for $i = 1, \ldots, n$. Using the implicit Euler scheme, one obtains for $\delta_i \approx \delta(t_i)$ and $y_i^0 = y^0(t_i)$

$$\delta_i = \delta_{i-1} + \tau_i \left(f(y_i^0 + \delta_i, t_i) - \dot{y}_i^0 \right), \quad i = 1, \ldots, n, \tag{2.3}$$

starting at $\delta_0 = 0$. Linearization around y^0, i.e. substituting $f(y_i^0 + \delta_i, t_i)$ by $f(y_i^0, t_i) + f'(y_i^0, t_i)\delta_i$ yields the linearly implict scheme

$$(I - \tau_i f'(y_i^0, t_i))\delta_i = \delta_{i-1} + \tau_i \left(f(y_i^0, t_i) - \dot{y}_i^0 \right), \quad i = 1, \ldots, n. \tag{2.4}$$

The derivatives $\dot{y}_i^0$ can be obtained as a linear combination of the values y_i^0,

$$\dot{y}_i^0 = \tau^{-1} \sum_{j=1}^{n} D_{ij}^d (y_j^0 - y_0), \quad i = 1, \ldots, n,$$

where the coefficients $D_{ij}^d = L_j'(t_i)$ of the spectral differentiation matrix $D^d \in \mathbb{R}^{n \times n}$ are given in terms of the Lagrange polynomials $L_j \in \mathbb{P}_n$ with respect to the normalized time grid $0, t_1/\tau, \ldots, t_n/\tau$.

A usually better approximation of the solution is then $y^1 \in \mathbb{P}_n$ given by

$$y^1(t) = y^0(t) + \sum_{i=1}^{n} \delta_i L_i(t/\tau), \quad \text{i.e.} \quad y_i^1 = y_i^0 + \delta_i. \tag{2.5}$$

Of course, this correction scheme can be iterated. If this fixed-point iteration converges, the limit value $y^* \in \mathbb{P}_n$ satisfies the collocation conditions

$$\dot{y}^*(t_i) = f(y^*(t_i), t_i) \text{ for } i = 1, \ldots, n, \quad y^*(0) = y_0, \tag{2.6}$$

and is therefore the solution of an implicit Runge-Kutta method of collocation type.

Note that the numerical differentiation realized by D^d is ill-conditioned, which, according to [9], is one of the reasons why the DSDC approaches are unpopular and should not be used for large n. It is included here to highlight the structural symmetry of differential and quadrature-based SDC methods in this and the next section.

Spectral quadrature: QSDC. The defect equation (2.2) is easily tranformed into the equivalent Picard equation

$$\delta(t) = \int_{s=0}^{t} \left(f(y^0(s)+\delta(s),s) - \dot y^0(s)\right) ds. \tag{2.7}$$

Following the derivation in [14], we obtain

$$\begin{aligned}\delta(t_i) = \delta(t_{i-1}) &+ \int_{s=t_{i-1}}^{t_i} \left(f(y^0(s)+\delta(s),s) - f(y^0(s),s)\right) ds \\ &+ \int_{s=t_{i-1}}^{t_i} f(y^0(s),s)\, ds - (y_i^0 - y_{i-1}^0), \quad i=1,\dots,n.\end{aligned}$$

Approximating the first integral by a simple numerical quadrature, e.g., right-looking rectangular rule, and the second one by the canonical quadrature rule on the nodes $t_1,\dots,t_n$, exact for polynomials of degree at most $n-1$, one obtains

$$\delta_i = \delta_{i-1} + \tau_i \big(f(y_i^0+\delta_i,t_i) - f(y_i^0,t_i)\big) + \tau_i \sum_{j=1}^{n} S_{ij}^q f(y_j^0,t_j) - (y_i^0 - y_{i-1}^0), \tag{2.8}$$

again starting at $\delta_0 = 0$. Linearization around y^0 yields the linearly implicit scheme

$$(I - \tau_i f'(y_i^0,t_i))\delta_i = \delta_{i-1} + \tau_i \sum_{j=1}^{n} S_{ij}^q f(y_j^0,t_j) - (y_i^0 - y_{i-1}^0). \tag{2.9}$$

An improved approximation is obtained as before by (2.5). As is apparent from (2.8), any fixed point satisfies the collocation conditions (2.6) as long as the quadrature S^q is exact for polynomials of degree up to $n-1$, which is the case for

$$S_{ij}^q = \frac{\tau}{\tau_i} \int_{t=t_{i-1}/\tau}^{t_i/\tau} \hat L_j(t)\, dt, \quad i,j=1,\dots,n,$$

in terms of the Lagrange polynomials $\hat L_j \in \mathbb{P}_{n-1}$ on the normalized grid $t_1/\tau,\dots,t_n/\tau$.

The quadrature based formulation has been introduced in [9], and named “spectral deferred correction” method.

Remark 2.1 In favor of a simpler presentation in terms of Lagrange interpolation, the setting is deliberately restricted to collocation limit schemes where the left interval end point t_0 is not included. This is the case, e.g., for Radau and Gauss points. Aiming at stiff problems with real spectrum, L-stable Radau collocation is the structurally most appropriate choice, such that the restriction is no significant limitation. Collocation schemes including t_0 can be treated similarly, using Hermite interpolation.

2.2 SDC on Dahlquist's equation

Many properties of time stepping schemes are already visible when applied to the simple, linear test equation

$$\dot{y} = \lambda y, \quad y(0) = 1. \tag{2.10}$$

In the following, we will apply the SDC methods to (2.10) and start developments from there. The implicit Euler DSDC variants (2.3) and (2.4) both read

$$\delta_i - \delta_{i-1} - \tau_i \lambda \delta_i = \tau_i \left(\lambda y_i^0 - \tau^{-1} \sum_{j=1}^{n} D_{ij}^d (y_j^0 - 1) \right).$$

Multiplication by τ/τ_i yields

$$\frac{\tau}{\tau_i}(\delta_i - \delta_{i-1}) - \tau\lambda\delta_i = \tau\lambda y_i^0 - \sum_{j=1}^{n} D_{ij}^d (y_j^0 - 1),$$

or, in matrix form with $\delta = (\delta_1, \ldots, \delta_n)^T$,

$$(\hat{D}_E - zI)\delta = -(D^d - zI)y^0 + D^d \mathbf{1}, \tag{2.11}$$

with $z = \tau\lambda$ and lower bidiagonal approximate differentiation matrix

$$(\hat{D}_E)_{ij} = \frac{\tau}{\tau_i}(\delta_{i,j} - \delta_{i,j-1}), \quad i, j = 1, \ldots, n,$$

where $\delta_{i,j}$ is the Kronecker-δ, realized by the implicit Euler method. The QSDC variants (2.8) and (2.9) both read

$$\delta_i - \delta_{i-1} - \tau_i \lambda \delta_i = \tau_i \sum_{j=1}^{n} S_{ij}^q \lambda y_j^0 - (y_i^0 - y_{i-1}^0).$$

Again, multiplication by τ/τ_i yields

$$\frac{\tau}{\tau_i}(\delta_i - \delta_{i-1}) - z\delta_i = z \sum_{j=1}^{n} S_{ij}^q y_j^0 - \frac{\tau}{\tau_i}(y_i^0 - y_{i-1}^0),$$

or, in matrix form,

$$(\hat{D}_E - zI)\delta = -(\hat{D}_E - zS^q)y^0 + \hat{D}_E \mathbf{1}. \tag{2.12}$$

Comparing (2.11) and (2.12) reveals that both SDC variants differ only in the matrices building up the right hand sides from values of y^0. In joint notation, they can be written as

$$(\hat{D} - z\hat{S})\delta = -(D - zS)y^0 + D\mathbf{1} \tag{2.13}$$

with $\hat{D}, \hat{S}, D, S$ as given in Tab. 2.1. Note that the spectral differentiation and integration matrices D, S appearing in the right hand side are "exact" up to the collocation error, whereas $\hat{D}, \hat{S}$ represent the lower order implicit Euler basic scheme and take the role of approximate differentiation and integration matrices, respectively. The convergence of the

method	$\hat{D}$	$\hat{S}$	D	S
DSDC	$\hat{D}_E$	I	D^d	I
QSDC	$\hat{D}_E$	I	$\hat{D}_E$	S^q

Table 2.1 Spectral differentiation and integration matrices arising in implicit Euler based DSDC and QSDC methods (2.13).

corresponding fixed point iteration

$$y^{k+1} = y^k + (\hat{D} - z\hat{S})^{-1}\left(-(D - zS)y^k + y_0 D\mathbf{1}\right) \tag{2.14}$$

towards the collocation solution depends only on the properties of the iteration matrix

$$G(z) = I - (\hat{D} - z\hat{S})^{-1}(D - zS). \tag{2.15}$$

Its properties have been studied in some detail in [14] for the QSDC case.

As the motivating interest is in reaction-diffusion equations exhibiting a real spectrum of the Jacobian with dominant large negative eigenvalues, we restrict our attention to $0 > z \in \mathbb{R}$. The two limit cases $z \to 0$ and $z \to -\infty$ are important for convergence of non-stiff and very stiff components, respectively. The case $z = \mathcal{O}(1)$ is important for spectra spread out over the whole negative real line, and is considered in Section 4.

Case $z \to 0$. This case is the limit of $\tau \to 0$ for a fixed λ and thus determines the non-stiff convergence of the integrator. The rule of thumb is that with a first-order basic scheme, each SDC sweep increases the order by one until the convergence order of the collocation discretization is reached (higher order basic schemes have been considered for increasing the order by more than one in each sweep, see [12,13,18,22]). This implies that the contraction factor of the SDC iteration is $\mathcal{O}(-z)$, which requires a vanishing spectral radius $\rho(G(0)) = 0$. This is automatically satisfied in the quadrature-based SDC formulation due to $\hat{D} = D$, but in general not by the differentiation-based, see Fig. 2.1. This is a striking point in favor of QSDC for non-stiff problems.

It has since long been known, however, that on *equidistant* grids, the differential variant exhibits good contraction properties as well [11]: $\rho(G(0)) = 0$ holds due to $G(0)$ being nilpotent. This implies that, at least asymptotically, one order per sweep is gained.

Case $z \to -\infty$. This case is the limit of $|\lambda| \gg \tau^{-1}$ for moderate step sizes τ. It is usually encountered in differential-algebraic equations, in problems with a pronounced scale separation between non-stiff and highly stiff components, and penalty treatment of constraints, e.g., Dirichlet boundaries in parabolic PDE problems. Here we have $G(z) \to I - \hat{S}^{-1}S =: G(-\infty)$. A rapid convergence of the SDC iteration, in particular the aim of L-stability or at least a vanishing stability function $R(\infty) = 0$ independently of y^0, requires $\rho(G(-\infty)) = 0$, which is automatically satisfied by the differentiation-based SDC formulation due to $\hat{S} = S$, but in general not by the quadrature-based. In fact, stagnation of QSDC iterations has been observed for differential-algebraic problems [15]. This is a striking point in favor of DSDC for stiff problems. In linear autonomous problems, this is of somewhat less importance as transient components are nevertheless damped out, only much slower than they should. In

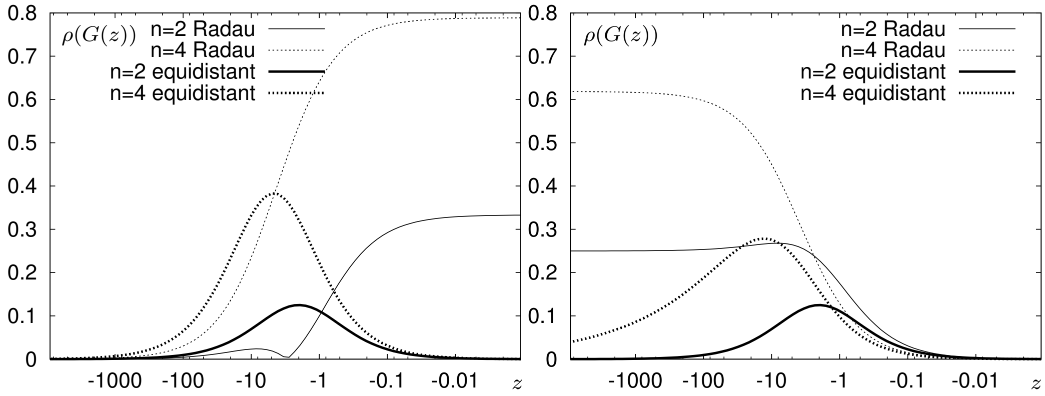

Fig. 2.1 Asymptotic contraction factor $\rho(G(z))$ of implicit Euler SDC iterations on equidistant and Radau points for Dahlquist's equation versus z. Left: DSDC. Right: QSDC.

nonlinear or non-autonomous problems, these transient errors in stiff components can spill over into non-stiff components and result in an order reduction.

Note that when the provisional solution y^0 is computed by an L-stable scheme, QSDC methods not including t_0 as collocation point can be L-stable despite $\rho(G(-\infty)) > 0$ [17]. This is essentially the same as starting with constant $y^0 = y_0$ and performing one DSDC sweep before starting the QSDC iteration. As the DSDC iteration matrix satisfies $\rho(G(-\infty))$ by construction, very stiff error components are immediately eliminated.

Again, QSDC exhibits an improved convergence on equidistant grids, where $\rho(G(-\infty)) = 0$ holds, see Fig. 2.1.

Remark 2.2 The compact matrix notation used in this section for scalar ODEs can be directly extended to ODE systems by using Kronecker products, see, e.g., [5].

3 Nilpotent DIRK sweeps by LU decomposition

Usually, SDC methods are designed by choosing a simple basic integration scheme (often explicit, or linearly or fully implicit Euler) for the defect equations (2.2) or (2.7), which realizes an SDC iteration with lower triangular approximate differentiation and integration matrices $\hat{D} = \hat{D}_E$ and $\hat{S} = I$.

The linear algebra perspective offers a different approach: We can choose the matrices $\hat{D}$ and $\hat{S}$ first, and afterwards interprete the resulting SDC sweep as a time stepping scheme. Staying in the DSDC ($\hat{S} = S$) and QSDC ($\hat{D} = D$) frameworks for their remarkable limit-behavour discussed before, we are free to choose either $\hat{D}$ or $\hat{S}$. A lower triangular shape of the approximate matrices $\hat{D}$ and $\hat{S}$, with nonvanishing diagonal entries for implicit schemes, will retain the sweep structure of the SDC iteration and hence allow an efficient implementation with only one system solve in each step of the sweep. The SDC sweep now reads

$$\left(\hat{D}_{ii} - z\hat{S}_{ii}\right)\delta_i^k = -\sum_{j=1}^{i-1}(\hat{D}_{ij} - z\hat{S}_{ij})\delta_j^k - \sum_{j=1}^{n}\left(D_{ij}(y_j^k - y_0) - zS_{ij}y_j^k\right), \qquad (3.1)$$

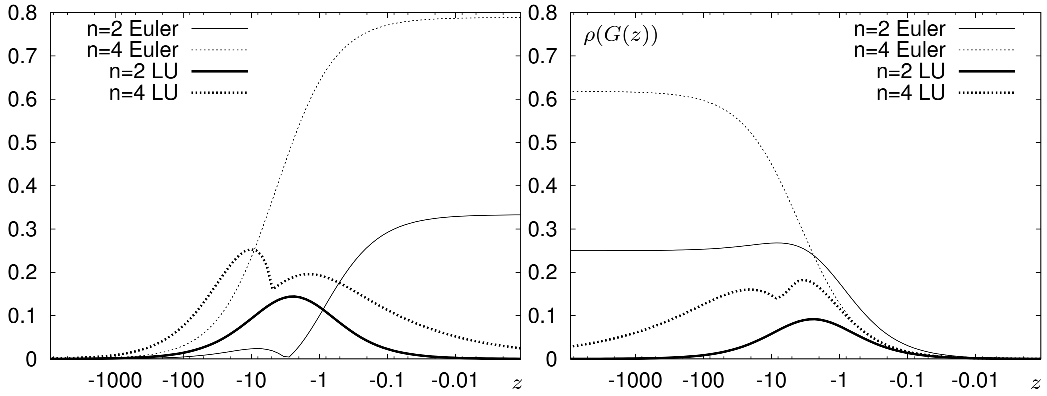

Fig. 3.1 Asymptotic contraction factor $\rho(G(z))$ of Euler- and LU-based SDC iterations on n Radau IIa points for Dahlquist's equation versus z. Left: DSDC. Right: QSDC.

which is but a minor modification of the simple linearly implicit Euler sweep. In other words, an SDC sweep is no longer a sequence of simple basic schemes, but one step of a diagonally implicit Runge-Kutta method.

Now assume that $G(z) = I - \hat{A}^{-1}A$ for some $z \in [-\infty, 0]$, and we want to enforce fast asymptotic convergence, i.e. $\rho(G(z)) = 0$. One way to achieve this is to select $\hat{A}$ based on an LU decomposition of A.

Lemma 3.1 *Let $G = I - \hat{A}^{-1}A$ and $A^T = LU$ with lower triangular L with diagonal entries all 1 and U being upper tiangular. Then $\hat{A} = U^T$ implies $\rho(G) = 0$.*

Proof We have $G = I - U^{-T}U^TL^T = I - L^T$, which is a strictly upper triangular matrix and hence nilpotent.

For the QSDC method with $\rho(G(0)) = 0$ already satisfied, we will enforce $\rho(G(\infty)) = 0$ and choose $\hat{S} = U^T$ with $S^T = LU$. The good convergence properties for $z = 0$ will not be affected in any way, as $\hat{S}$ plays no role in the limit case. Analogously, for the DSDC method, we will enforce $\rho(G(0)) = 0$ and choose $\hat{D} = U^T$ with $D^T = LU$. Again, the good convergence properties for $z \to -\infty$ will not be affected. The improvement over the standard approach on non-equidistant grids shown in Fig. 3.1 is quite pronounced. For QSDC methods, the contraction factors are an almost uniform improvement even over the case of equidistant collocation nodes. Nevertheless, for DSDC the convergence order does not increase by one for each sweep, but only asymptotically every n sweeps.

Theorem 3.1 *Assume that $D^T = LU$ and $\hat{D} = U^T$ in the DSDC method. Then, for $G(z)$ from* (2.15)*, there are constants $c < \infty$ and $\gamma > 0$ such that $\rho(G(z)) \le c|z|^{1/n}$ for all $|z| \le \gamma$.*

Proof Due to $\hat{S} = S = I$ in DSDC, the iteration matrix is

$$G(z) = I - (U^T - zI)^{-1}(U^TL^T - zI)$$

with

$$(U^T - zI)^{-1} = \left(I + \sum_{k=1}^{\infty} (zU^{-T})^k \right) U^{-T}$$

if $|z| \le \gamma := \min(\frac{1}{2}\|U^{-T}\|^{-1}, 1)$. Thus,

$$\begin{aligned} G(z) &= I - \left(I + \sum_{k=1}^{\infty} (zU^{-T})^k \right) (L^T - zU^{-T}) \\ &= \underbrace{I - L^T}_{=:N} + z \underbrace{\left(U^{-T} - U^{-T} \sum_{k=0}^{\infty} (zU^{-T})^k (L^T - zU^{-T}) \right)}_{=:A} \end{aligned}$$

holds with N being nilpotent of order n and A bounded independently of $|z| \le \gamma$ due to $\|zU^{-T}\| \le 1/2$. As $\rho(G(z)) = \lim_{k\to\infty} \|G(z)^k\|^{1/k}$, we consider

$$G(z)^k = (N + zA)^k,$$

which is a sum of 2^k products of factors N and zA. Grouping the terms by the number of factors zA, we obtain $\binom{k}{i}$ terms consisting of i factors zA and $k - i$ factors N each. For $i \ge k/n$ these terms are bounded by

$$|z|^i \|A\|^i \|N\|^{k-i} \le |z|^{k/n} \max\{\|A\|, \|N\|\}^k$$

due to $|z| \le 1$. Otherwise there is at least one sequence of n factors of N, such that the term vanishes. With $c = 2\max\{\|A\|, \|N\|\}$ we obtain the estimate

$$\|G(z)^k\|^{1/k} \le \left(2^k |z|^{k/n} \frac{c^k}{2^k} \right)^{1/k} = c|z|^{1/n},$$

which completes the proof.

The analogous result $\rho(G(z)) = \mathcal{O}(|z|^{-1/n})$ for $z \to -\infty$ holds for QSDC methods.

Remark 3.1 An LU decomposition of D or S, respectively, need not exist without pivoting. While pivoting can in principle considered here as well, the corresponding permutations would modify the sweep structure, that no longer runs simply forward in time. In favour of a traditional SDC sweep structure, we omit pivoting here.

4 Direct optimization of DIRK sweeps

The ad-hoc approach in the previous section has led to a considerable improvement on nonuniform grids despite its simple, almost explicitly given approximation matrices. Yet we can explore the design space of possible choices of lower triangular $\hat{D}$ and $\hat{S}$ more comprehensively while restricting the setup to the Dahlquist test equation. We will discuss interesting quantities as objectives for optimizing the approximation matrices in the following subsection. In particular, the spectral radius of G for the limit cases may be of less interest than its norm, or some other quantity. Next, we will explore possibilities for choosing design variables. In particular we need not restrict the discussion to just $\hat{D}$ and $\hat{S}$, as we can use different approximate matrices in each sweep.

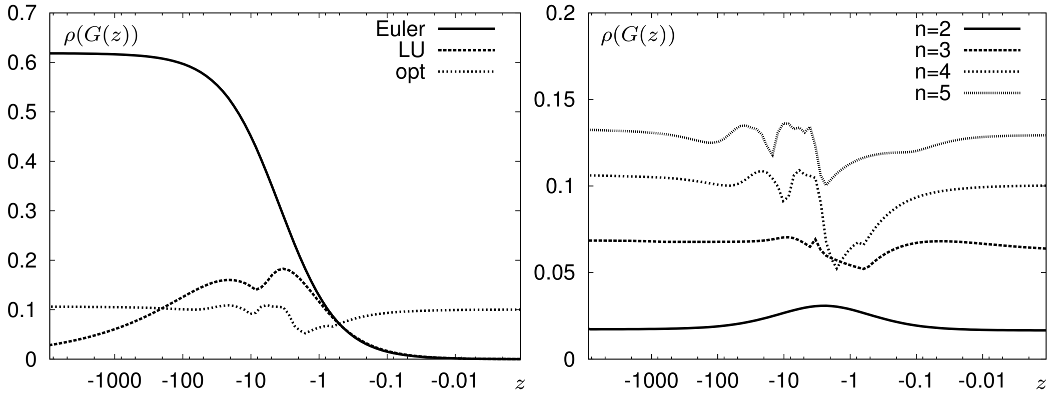

Fig. 4.1 Spectral radius $\rho(G(z))$ versus z for QSDC sweeps applied to the Dahlquist equation. Left: Radau(4) QSDC sweeps based on implicit Euler, LU-DIRK, and numerically optimized matrices $\hat{D}, \hat{S}$ for $w \equiv 1$. Right: Optimized Radau(n) sweeps for $w \equiv 1$ and different numbers n of collocation points.

4.1 Objectives

"Fast convergence" of SDC methods can mean quite a number of different things in practice, such as asymptotic convergence rate, good error reduction in the first few iterates, error reduction in the whole time interval or only at its end, and so on. Here we will formulate a number of reasonable criteria for optimizing $\hat{D}$ and $\hat{S}$, and investigate their properties. The optimization objectives are formulated in terms of the iteration matrix $G(z)$ from (2.15).

Spectral radius. The spectral radius $\rho(G(z))$ of the iteration matrix determines the asymptotic convergence rate, and is therefoere relevant when performing many SDC sweeps. The LU-based choices of $\hat{D}$ and $\hat{S}$ above guarantee $\rho(G(z)) = 0$ for the limit cases $z = 0$ and $z \to -\infty$. As is apparent from Fig. 3.1, intermediate values of z, which are bound to occur in parabolic problems with sufficiently fine spatial grid, experience a worse error reduction. In order to reduce those error components faster, we may choose the SDC matrices such that the maximal contraction factor is minimized:

$$J(\hat{D}, \hat{S}) = \max_{z \le 0} \rho(G(z)) \tag{4.1}$$

This choice sacrifices the good properties at $z = 0$. As $\rho(G(z)) = \mathcal{O}(z)$ for $z \to 0$ no longer holds, we cannot expect to gain one order of convergence per SDC sweep. As opposed to an error reduction by the factor of $\mathcal{O}(\tau)$ we have to settle for a reduction factor of approximately $\rho(G(0)) > 0$. The contraction behavior in terms of z can be adjusted by introducing a weight function $w(z) \ge 0$:

$$J(\hat{D}, \hat{S}) = \max_{z \le 0} w(z)\rho(G(z))$$

With $w(z) = 1 - z^{-1}$ we enforce $\rho(G(0)) = 0$ again. Note that this does not yet guarantee one order of convergence improvement per SDC sweep (compare Theorem 3.1) above.

In Fig. 4.1, resulting spectral radii are shown versus z for different numbers of Radau-IIa collocation points and constant weight function $w \equiv 1$. Optimization of $\hat{D}$ and $\hat{S}$ were

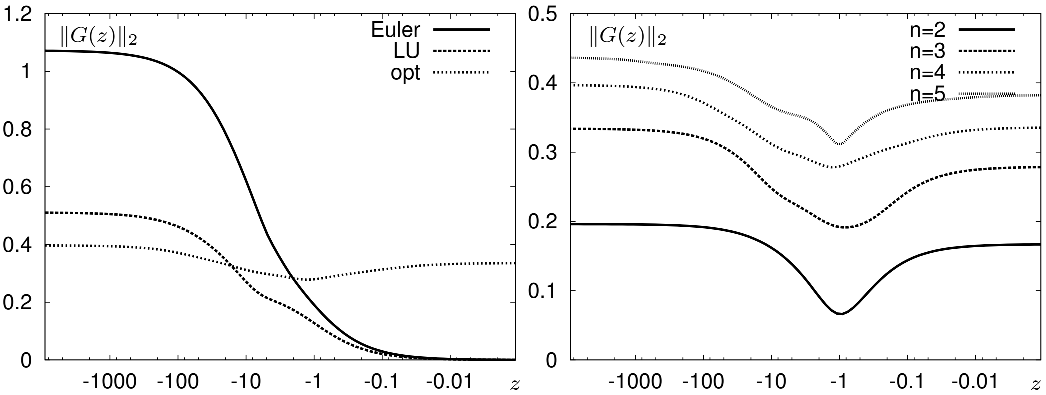

Fig. 4.2 Norm $\|G(z)\|_2$ versus z for QSDC sweeps applied to the Dahlquist equation. Left: Radau(4) QSDC sweeps based on implicit Euler, LU-DIRK, and numerically optimized matrices $\hat{D}, \hat{S}$ for $w \equiv 1$. Right: Optimized Radau(n) sweeps for $w \equiv 1$ and different numbers n of collocation points.

perfomed using a very simple-minded SQP scheme with numerical differentiation and using the LU-based matrices $\hat{D}_{\text{LU}}$ and $\hat{S}_{\text{LU}}$ as initial values. The maximum in (4.1) has been approximated by the l^p norm $\|\cdot\|_{64}$ evaluated on a logarithmic grid with 100 points on $[10^{-4}, 10^4]$. No global optimization has been attempted, such that the results shown are likely to be somewhat less than optimal. Nevertheless, a reduction of the worst case spectral radius compared to LU-based QSDC is observed, at the cost of worse contraction in the limits $z \to 0$ and $z \to -\infty$.

Pre-asymptotic contraction factors. Spectral deferred correction methods are particularly attractive if they lead to efficient time stepping schemes with *few* SDC sweeps, in which case not the spectral radius but the norm of G determines the error reduction. Hence a reasonable optimization criterion would be

$$J(\hat{D}, \hat{S}) = \max_{z \le 0} w(z) \|G(z)\| \tag{4.2}$$

with any appropriate matrix norm $\|\cdot\|$. In Fig. 4.2, the resulting error reduction factors $\|G(z)\|_2$ are shown versus z for weight function $w \equiv 1$. Again we observe a significant improvement of LU-based SDC sweeps over standard Euler sweeps, and a further moderate improvement by numerically optimized sweeps, again at the cost of worse contraction for $z \to 0$. Notice that for $z \to -\infty$, the LU-based sweeps do not lead to $\|G(z)\|_2 \to 0$, as $G(-\infty)$ is nilpotent of order n, but not zero.

Final time error. While local error reduction is a worthwhile goal, the overall quality of the computed solution will hinge on the global error transport. In this conception, the relevant property of G is the error at the end of the time interval. Whenever the last collocation point is at the end of the interval, e.g., with Radau points, this error is determined by the last row of G only, which would then to be minimized:

$$J(\hat{D}, \hat{S}) = \max_{z \le 0} w(z) \|e_n^T G(z)\| \tag{4.3}$$

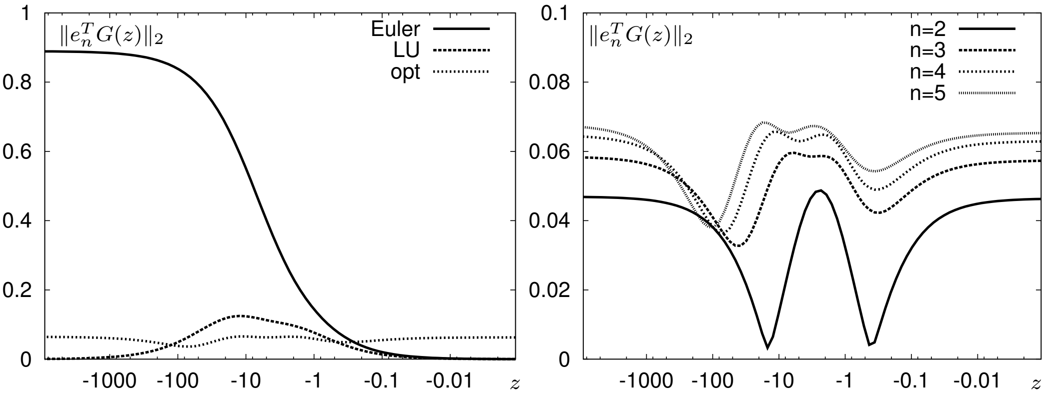

Fig. 4.3 Norm $\|e_n^T G(z)\|_2$ at final time versus z for QSDC sweeps applied to the Dahlquist equation. Left: Radau(4) QSDC sweeps based on implicit Euler, LU-DIRK, and numerically optimized matrices $\hat{D}, \hat{S}$ for $w \equiv 1$. Right: Optimized Radau(n) sweeps for $w \equiv 1$ and different numbers n of collocation points.

Here, e_n denotes the n-th unit vector in $\mathbb{R}^n$. For different collocation points, e.g., Gauß points, a suitable linear combination of the rows of G has to be used instead.

In Fig. 4.3, the resulting error reduction factors $\|e_n^T G(z)\|_2$ are shown versus z for weight function $w \equiv 1$. A similar pattern as for the spectral radius emerges: LU-DIRK leads to a dramatic improvement over Euler-based SDC. In particular, for $z \to -\infty$ the reduction factor approaches zero, as the last line of $G(-\infty)$ is exactly zero due to its strictly upper triangular shape. Compared to that, a modest worst case improvement is achieved by numerical optimization, again sacrificing the good properties in the limit cases $z \to 0$ and $z \to -\infty$.

Sweep blocks. Already for the purpose of time error estimation, at least two SDC sweeps will be performed. If we intend to apply a certain number of m sweeps, the relevant error reduction is given by $G(z)^m$ instead of $G(z)$, which may lead to different optimal values of $\hat{D}$ and $\hat{S}$. Hence we may want to look at the average reduction factors

$$J(\hat{D},\hat{S}) = \max_{z\le 0} w(z)\|G(z)^m\|^{1/m} \quad \text{or} \quad J(\hat{D},\hat{S}) = \max_{z\le 0} w(z)\|e_n^T G(z)^m\|^{1/m}$$

instead of (4.2) and (4.3), respectively.

The results shown in Fig. 4.4 show a significant improvement versus the single-sweep optimization for the norm objective (4.2), but a moderate deterioration for the final time objective (4.3). Both approach the spectral radius result for increasing m, which is to be expected due to $\rho(G) = \lim_{m\to\infty} \|G^m\|^{1/m} \le \|G\|$. The deterioration in the final time objective is due to the fact that large errors remaining at times $t_1, \ldots, t_{n-1}$ have an impact on the error at t_n in the next iteration and cannot be ignored. In this aspect, the objective (4.3) is too optimistic, as submultiplicativity $\|e_n^T G^m\| \le \|e_n^T G\|^m$ usually does not hold.

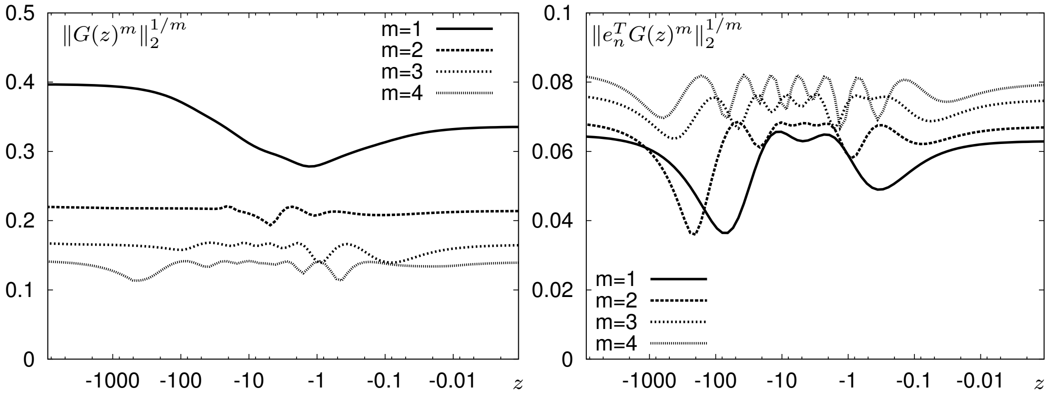

Fig. 4.4 Averaged error reduction per sweep versus z for Radau(4) QSDC sweeps with numerically optimized matrices $\hat{D}, \hat{S}$ for $w \equiv 1$. Left: Norm $\|G(z)^m\|_2^{1/m}$. Right: Final time error $\|e_n^T G(z)^m\|_2^{1/m}$.

4.2 Design variables.

Up to now we followed the strategy to design a single optimal SDC sweep that is to be applied a sufficient number of times. Here we will explore SDC designs with more freedom and flexibility, all of which can be combined with any of the optimization objectives discussed above.

Flexible sweep blocks. Always applying m SDC sweeps at a time offers the freedom to choose different approximate matrices in each of the m sweeps, leading to optimization variables $\hat{D}_j, \hat{S}_j$, $j = 1, \dots, k$ corresponding to iteration matrices $G_j(z)$. The objective will then be

$$J(\hat{D}_1, \hat{S}_1, \dots \hat{D}_m, \hat{S}_m) = \max_{z \le 0} w(z) \Big\| e_n^T \prod_{j=1}^{m} G_j(z) \Big\|^{1/m}. \tag{4.4}$$

As visible in Fig. 4.5, to be compared directly with Fig. 4.4, this richer design space allows a further improvement over the optimization of single sweeps for multiple iterations.

Greedy sweeps. Applying always a fixed number m of SDC sweeps at a time may incur inefficiencies, in particular for larger m, when one wants to control the number of iterations such that a certain accuracy is achieved. A greedy style choice of $\hat{D}_m, \hat{S}_m$ allows to terminate the SDC iteration after each sweep. The locally optimal choice of approximate matrices is then given by

$$J(\hat{D}_m, \hat{S}_m) = \max_{z \le 0} w(z) \Big\| e_n^T \prod_{j=1}^{m} G_j(z) \Big\|^{1/m},$$

where the optimization is performed sequentially for $m = 1, \dots$.

The resulting average contraction factors are shown in Fig. 4.6. As expected, the improvement over single-sweep optimization $m = 1$ is less pronounced than in the cases where the number m of sweeps is known beforehand (Figs. 4.4 and 4.5). This is the price to pay for the flexibility to terminate the SDC iteration at any m. Given that the reduction factor

$\|G_2(z)G_1(z)\|^{1/2}$ visible in Fig. 4.5 is already better than the greedy approach up to $m = 8$ and virtually as good as any block scheme for larger m, the greedy style of choosing $\hat{D}_j, \hat{S}_j$ is probably not worthwhile.

5 Stability and accuracy domains

Even though the intended application to reaction-diffusion system fixes the focus on real λ in the Dahlquist equation (2.10), we briefly explore the stability and accuracy properties of different SDC variants for complex λ as well. In Fig. 5.1 we show the domains of stability for Euler-QSDC (top) and LU-QSDC (bottom), both for simultaneously increasing number of collocation points n and SDC iteration count k (left, $k = n$) and fixed number of collocation

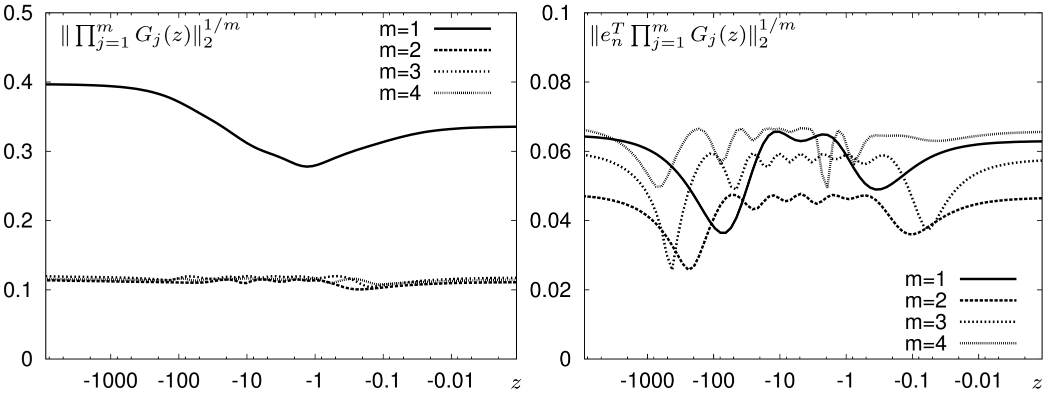

Fig. 4.5 Averaged error reduction per sweep versus z for m-block Radau(4) QSDC sweeps with numerically optimized matrices $\hat{D}_j(m), \hat{S}_j(m)$ for $w \equiv 1$. Left: Norm $\|\prod_{j=1}^m G_j(z)\|_2^{1/m}$. Right: Final time error $\|e_n^T \prod_{j=1}^m G_j(z)\|_2^{1/m}$.

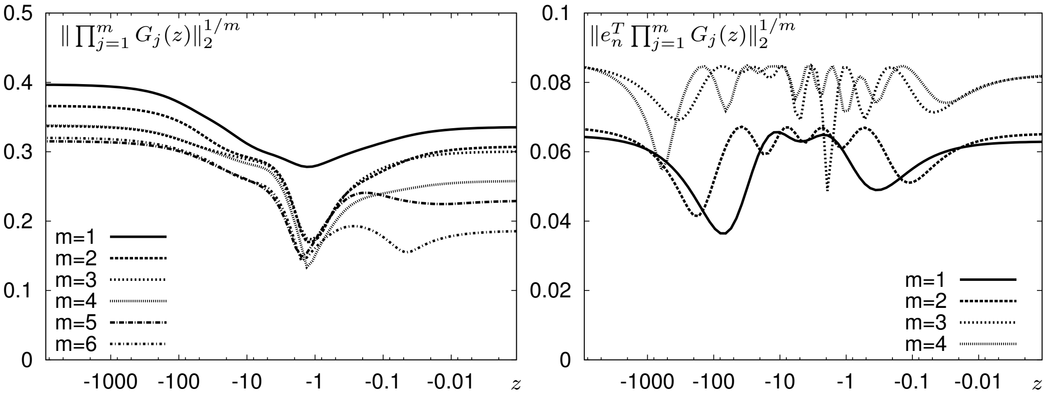

Fig. 4.6 Averaged error reduction per sweep versus z for nested Radau(4) QSDC sweeps with numerically optimized matrices $\hat{D}_j, \hat{S}_j$ for $w \equiv 1$. Left: Norm $\|\prod_{j=1}^m G_j(z)\|_2^{1/m}$. Right: Final time error $\|e_n^T \prod_{j=1}^m G_j(z)\|_2^{1/m}$.

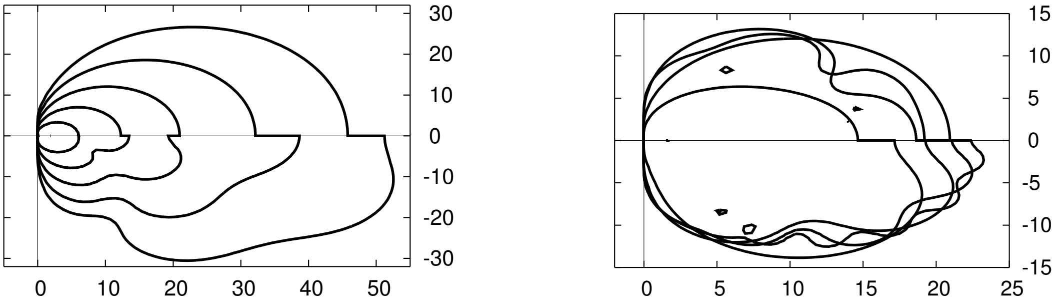

Fig. 5.1 Stability domains of QSDC methods based on Euler steps (top) and LU-DIRK steps (bottom) Left: $n = k = 2, \dots, 6$. Right: $n = 4, k = 2, 4, 8, 16$.

points $n = 4$ with increasing iteration count k. The stability domain is the subset of the complex plane, where for the rational approximation R of the exponential function realized by the QSDC method $|R(z)| \le 1$ holds. Note that the stability domain is outside the plotted curves.

None of the QSDC methods tested is A-stable. The LU-based methods are $A(\alpha)$-stable with $\alpha \approx 89.7$, while the Euler-based methods achieve $\alpha \approx 89.9$. Of course, as the iterations converge towards the L-stable underlying Radau collocation scheme, these angles will increase towards 90 degrees. For larger positive real values of z, the SDC convergence towards the collocation scheme appears to be rather slow or even nonexistent.

Accuracy domains for the same methods are shown in Fig. 5.2. The accuracy domain for $\epsilon = 10^{-4}$ is the subset of the complex plane for which $|R(z) - e^z| \le \epsilon$, which is inside the plotted curves.

No significant difference between the Euler- and LU-based QSDC schemes is apparent here. The slight differences can be attributed to the somewhat slower convergence of the Euler-based SDC iterations, as they resemble longer a sequence of Euler steps and approach the properties of the underlying Radau collocation method later.

Stability and accuracy domains of directly optimized QSDC methods are shown in Figs. 5.3 and 5.4, using the flexible sweep block design for a block size $m = 2$ with spectral radius objective or final time error objective (4.4), respectively. Even though these methods are of first order only if a limited number of iterations is performed, the accuracy domains

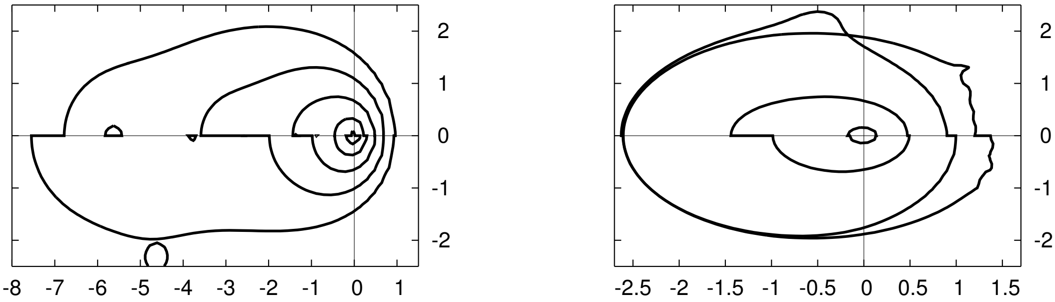

Fig. 5.2 Accuracy domains of QSDC methods based on Euler steps (top) and LU-DIRK steps (bottom) for an error of 10^{-4}. Left: $n = k = 2, \dots, 6$. Right: $n = 4, k = 2, 4, 8, 16$.

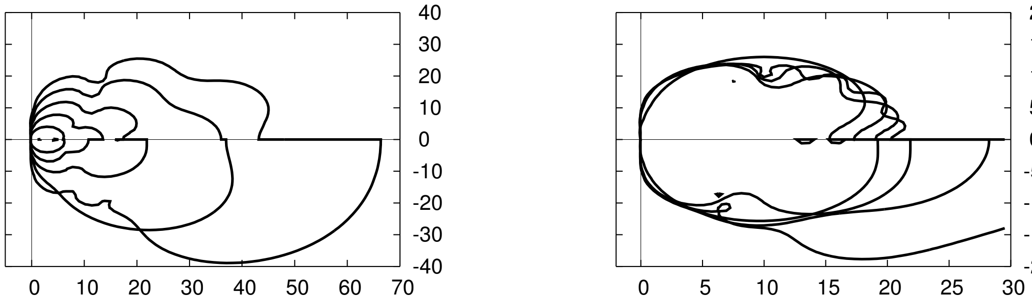

Fig. 5.3 Stability domains of QSDC methods optimized for spectral radius (top) and error at terminal time (bottom) Left: $n = k = 2, \ldots, 6$. Right: $n = 4, k = 2, 4, 8, 16$. Compare with Fig. 5.1.

for $\epsilon = 10^{-4}$ tend to be somewhat larger than for the Euler and LU variants. This holds in particular for the final time error objective, since the accuracy domain quantifies the error at the final time of the SDC step only. We may notice that the accuracy domains are particularly large for an even number of iterations, which corresponds to the iteration block size used for optimizing the QSDC matrices.

Both variants are $A(\alpha)$-stable with $\alpha \approx 89.1$ for the spectral radius objective and $\alpha \approx 89.8$ for the final time error objective.

We may expect the linear algebra construction of QSDC matrices $\hat{S}$, though done here for the negative real axis only, to work fairly well also for general systems.

6 Numerical examples

Up to now we have designed SDC methods to work in some sense optimal on the Dahlquist equation, in the expectation that their good properties translate to more complex systems of nonlinear or nonautonomous ODEs, DAEs, and PDEs. In this section, we will apply those SDC methods to a couple of examples, comparing and interpreting the results. In particular we will check the conjectures that $\|e_n^T G(z)\|$ is more relevant for the global error transport

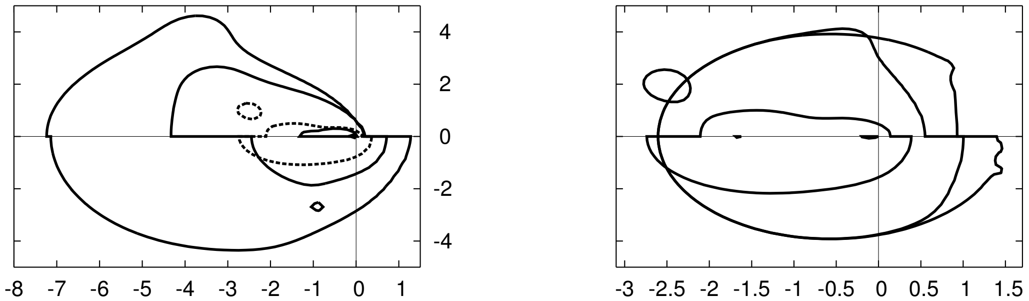

Fig. 5.4 Accuracy domains of QSDC methods optimized for spectral radius (top) and error at terminal time (bottom) for an error of 10^{-4}. Left: $n = k = 2, \ldots, 6$. Right: $n = 4, k = 2, 4, 8, 16$. Compare with Fig. 5.2.

than the other error reduction measures, and that $\|G(z)\|$ or $\|G(z)^k\|$ is more relevant for integrators using a small number of sweeps than $\rho(G(z))$.

Extending the DIRK sweep structure from the Dahlquist test equation (3.1) to general ODEs is straightforward. The DIRK sweep based on lower triangular approximate differentiation and integration matrices $\hat{D}$ and $\hat{S}$ reads

$$\begin{aligned}\left(\hat{D}_{ii} - \tau f'(y_i^k, t_i)\hat{S}_{ii}\right) \delta_i^k \\ &= -\sum_{j=1}^{i-1} (\hat{D}_{ij} - \tau f'(y_j^k, t_j)\hat{S}_{ij})\delta_j^k - \sum_{j=1}^{n} \left(D_{ij}(y_j^k - y_0) - \tau S_{ij} f(y_j^k, t_j)\right).\end{aligned}$$

Equally straightforward is the concatenation of several SDC steps of size τ each to cover a fixed integration interval $[0, T]$, such that the convergence behaviour for different step sizes τ can be examined. For simplicity, no adaptive step size selection is performed.

6.1 Prothero-Robinson example

We consider the nonautonomous generalization of Dahlquist's test equation due to [20],

$$\dot{y} = \lambda(y - g(t)) + \dot{g}(t), \quad y(0) = g(0), \tag{6.1}$$

with $g = \sin$ on the time interval $[0, 1]$. The exact solution is $\sin(t)$. We consider a "mildly stiff" setting with $\lambda = -10^3$, that shows a transition towards non-stiff behavior for reasonably small step size τ.

Understanding SDC convergence. Compared to the limiting collocation method, SDC methods with a fixed number k of sweeps show a more complex convergence behavior, see Fig. 6.1. From a theoretical point of view, the deviation between the SDC solution y_τ and the exact solution y in the mildly stiff example should exhibit a "hump" when plotted versus the step size τ, as the LU based SDC contraction rate given by the spectral radius of G is larger for $z = \tau\lambda$ in the range $[-10, -1]$ than in the limit cases $z \to 0$ and $z \to -\infty$ (see Fig. 3.1 and [11]). In particular, the rule of thumb "one order per sweep" is only reached asymptotically for very small time steps, as then $z = \tau\lambda$ has crossed the hump and only then the contraction factor $\rho(G(z))$ approaches zero. In this regime, a k-sweep QSDC method can even show a convergence rate k that exceeds the order of the underlying collocation scheme. This convergence behavior is actually observed in numerical computations and can be expected to arise in Euler sweep SDC methods on equidistant grids as well.

It is apparent from Fig. 6.1 and the considerations above that the notion of "order" is insufficient to describe the error behavior of SDC methods properly in the practically relevant pre-asymptotic step size and accuracy regions. In particular the numerical estimation of convergence order can yield quite arbitrary results depending on the chosen step sizes, and is therefore to be interpreted with utmost care.

Comparison of objectives and designs. Next we will compare different types of DIRK schemes. As the convergence of LU based QSDC is impeded by the “hump” around $z \approx -2$, we try to have a “flat” contraction factor and therefore choose the weight function $w \equiv 1$. Comparing Fig. 6.2 left with Fig. 6.1 right one can indeed observe both the absence of a “hump” and the worse behaviour for $\tau \to 0$, as the SDC contraction factor is no longer $\mathcal{O}(|z|)$. Overall, the error is somewhat smaller than with LU based SDC sweeps.

Limiting the presentation to an intermediate value of $k = 5$ sweeps on Radau(3) nodes, we compare different sweep types. In Fig. 6.2 right the effect of increasing m in optimizing for $\|G(\cdot)^m\|_{L^\infty[-\infty,0]}$ is shown. A significant improvement of up to two orders of magnitude is achieved moving from $m = 1$ to $m = 4$ and above in a wide range of $z < -1$. Taking more than one sweep iteration into account when optimizing DIRK sweeps appears to be quite beneficial.

Next we compare different sweep constructions, namely standard implicit Euler, LU based sweeps, and directly optimized DIRK sweeps according to the criteria (4.1), (4.2), and (4.3), respectively, in Fig. 6.3. The most striking improvement is that of DIRK sweeps over standard implicit Euler sweeps, which are inefficient for stiff problems. Only for $z > -1$, i.e. for non-stiff problems, the Euler sweep is competitive. We can observe that all optimization criteria give a moderate overall improvement over the LU based SDC sweeps. The rather large variation in the error curves are probably due to “accidentally accurate” results, which can occur quite frequently because of the low dimension of the ODE (6.1).

From all the error plots it is clear that no single SDC method with fixed number of sweeps is an efficient method over the whole range of step sizes. Instead, an adaptive selection of the number of sweeps is indispensable for robust efficiency and time error estimation.

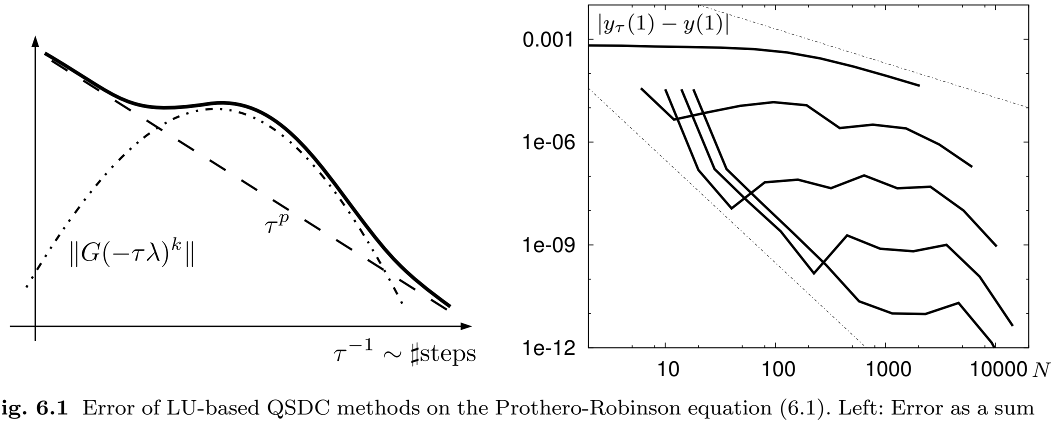

Fig. 6.1 Error of LU-based QSDC methods on the Prothero-Robinson equation (6.1). Left: Error as a sum of collocation time discretization error and SDC iteration error. Right: Actual observed error-work behavior on Radau(3) for $k = 1, 3, 5, 7, 9$ sweeps and different time step sizes: final time error $|y_\tau(1) - y(1)|$ vs. total number $N = k/\tau$ of sweeps. Limiting lines are N^{-1} and N^{-3}.

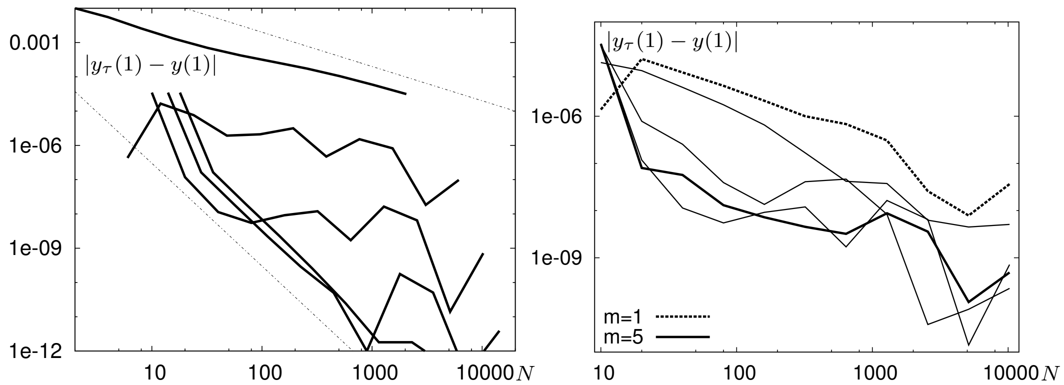

Fig. 6.2 Final time error vs. total number N of sweeps on Radau(3) of QSDC methods on the Prothero-Robinson equation (6.1). Left: $k = 1, 3, 5, 7, 9$ DIRK sweeps optimized for $\|G(\cdot)^4\|_{L^\infty[-\infty,0]}$. Limiting lines are N^{-1} and N^{-3}. Right: $k = 5$ DIRK sweeps optimized for $\|G(\cdot)^m\|_{L^\infty[-\infty,0]}$, $m = 1, \dots, 5$.

6.2 Vienna equation

A more challenging example is the autonomous nonlinear system for $y = [y_1, y_2]^T$

$$\begin{aligned} \dot{y}_1 &= -y_2 + \lambda y_1(\|y\|_2^2 - 1) \\ \dot{y}_2 &= y_1 + 3\lambda y_2(\|y\|_2^2 - 1) \end{aligned} \tag{6.2}$$

with initial value $y(0) = [1, 0]^T$, $\lambda = -10^5$, exact solution $[\cos t, \sin t]^T$, and final time $T = 3$ studied in [3]. The system exhibits one stiff and one non-stiff component, which continuously change their direction. The varying eigenvectors of the Jacobian make this problem considerably harder to solve than the previous one, in particular using the linearly implicit Euler method as a basic scheme. In fact, a significant error reduction is observed

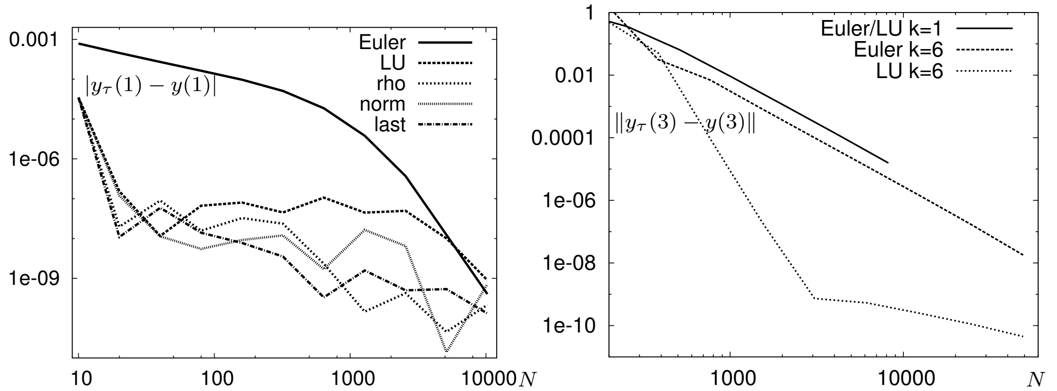

Fig. 6.3 Final time error vs. total number N of sweeps on Radau(3) of QSDC methods. Left: Prothero-Robinson equation (6.1) with Euler, LU based, and directly optimized DIRK sweeps for $\|\rho(G(\cdot))\|_{L^\infty[-\infty,0]}$, $\|G(\cdot)^4\|_{L^\infty[-\infty,0]}$, and $\|e_3^T G(\cdot)^4\|_{L^\infty[-\infty,0]}$. $k = 5$ sweeps. Right: Vienna equation (6.2) with standard Euler and LU based sweeps for $k = 1$ and $k = 6$ sweeps.

with either method only for step sizes $\tau < 10^{-2}$. A fully implicit Euler basic scheme gives better results [3, Table 6], but even then Newton's method for computing the implicit Euler steps is reported to diverge for step sizes larger than 0.05. To overcome this problem, a separation of the components by a QR factorization within SDC has been suggested in [3] but will not be pursued here.

In Fig. 6.3, right, the error at T is shown versus the total number of sweeps for Euler- and LU-based QSDC. The late onset of linearly implict Euler error reduction requires to use small time steps for which the Radau(3) collocation error is already in the order of 10^{-13}. The total error is dominated by the SDC iteration error. The Jacobian has two real eigenvalues, one close to zero, the other one in the order of $\lambda = -10^5$. With time steps larger than 10^{-4} and $|z| > 10$, the Euler QSDC suffers from its large contraction factor, whereas LU-based QSDC converges rather quickly towards the Radau(3) solution.

6.3 Nonautonomous heat equation

The linear but nonautonomous heat equation

$$\dot{u} = \Delta u + xe^{-t}, \quad x \in]0,1[, \quad u(0) = u(1) = 0, \tag{6.3}$$

from [19] is considered on $t \in [0, 0.1]$. The "saw-tooth" shaped source term xe^{-t} introduces a long tail of high-frequency modes in every time step, such that essentially the whole negative real axis covered by the spectrum of the Jacobian is excited. Hence one expects a significant order reduction as the eigenvalues cross the hump of LU based SDC for $\tau \to 0$. In fact, the Rosenbrock scheme GRK4T [16] shows an effective order of 3.25 instead of the nominal consistency order 5 for similar reasons [8, Chap. 9]. An equidistant finite difference discretization of size h leads to discrete eigenvalues roughly covering the range $[-h^{-2}, 0]$. Unless the usual stability constraint $\tau < ch^2$ for parabolic problems is satisfied, meaning that all eigenvalues have crossed the critical range around one, the SDC time stepping schemes can be expected to behave as for the continuous problem. Consequently, a spatial discretization with $h = 10^{-2}$ is chosen for time step sizes $\tau > 10^{-4}$.

In Fig. 6.4 left we observe that for a small number k of sweeps, the standard implicit Euler QSDC method converges quite rapidly, and results in a time stepping scheme of effective estimated order 2.3. The fast error reduction during the first few sweeps is due to the elimination of dominant low-frequent error modes with small $|z|$. From about four sweeps on, these modes are sufficiently reduced for the high-frequent error modes with small initial amplitude to dominate the total error. Consequently, the large spectral radius of Euler QSDC effectively prevents significant further progress.

In contrast, the convergence of LU based QSDC is steady due to the spectral radius being significantly smaller than one. The price to pay is a somewhat larger error for smaller sweep numbers $k \leq 4$. Ultimately, all QSDC methods with $\hat{D} = D$ and fixed k achieve an effective converge order of 2.3.

In Fig. 6.4 right, we explore the impact of the weight function w in the final time objective (4.3), which turned out to give the best results among the different optimization criteria in this example. As the excitation modes due to the nonautonomous source term are of amplitude $(-z)^{-1/2}$, we choose this power of z as weight function w in case A, and for comparison $w = 1$ as case B. In case A, the weight function singularity implies

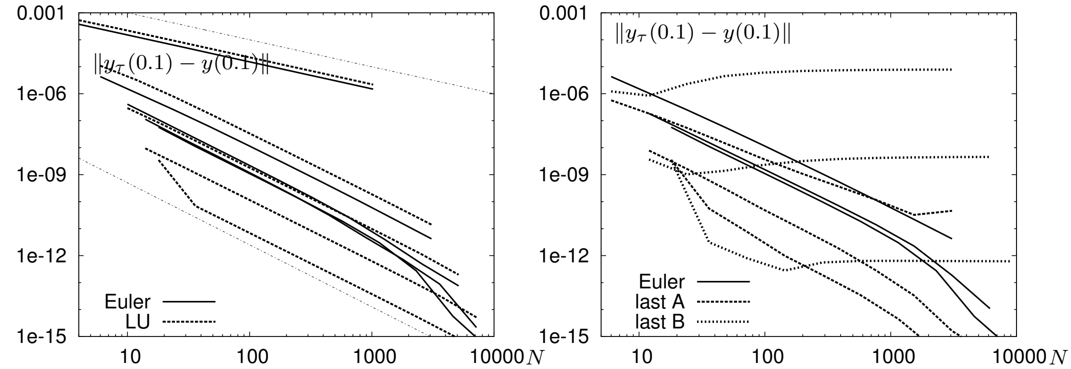

Fig. 6.4 Final time error vs. total number N of sweeps taken by QSDC methods for the nonautonomous heat equation (6.3) on Radau(5). Left: LU and Euler sweeps $k = 1, 3, 5, 7, 9$. Limiting lines are N^{-1} and $N^{-2.3}$. Right: Euler and DIRK sweeps optimized for (4.3) at $k = 3, 6, 9$ sweeps. A is with weight function $w = (-z)^{-1/2}$, B with $w = 1$.

$\|e_n^T G(\cdot)^m\|^{1/m}_{L^\infty[-\infty,0]} \to 0$ for $z \to 0$, such that the order should increase with sweep count k. This is actually achieved again up to the effective order 2.3, with a slight improvement compared to the LU based sweeps. In case B, the contraction factor does not vanish for $z \to 0$, such that errors in low-frequent modes are retained to some extent. As those are not damped out quickly by (6.3), they add up, and a step size reduction does not lead to an error reduction. The convergence order with respect to time step size τ is therefore 0. Nevertheless, a somewhat faster SDC iteration convergence can be achieved, such that for an appropriate combination of sweep number k and step size τ, even better results are obtained.

This raises the question which combinations of sweep number k and step size τ are most efficient in terms of accuracy (measured as final time error ϵ) per work (total number $N = 0.1k/\tau$ of sweeps). Obviously, there cannot be a single optimal combination, since different accuracy requirements will lead to different choices. To examine this, we vary both $\tau \in [10^{-3}, 0.1]$ and $k \in \{1, \dots, 35\}$, and look at the Pareto front of efficient (k, τ) combinations. Borrowing terminology from multicriteria optimization, a point (N, ϵ) is said to dominate a different point (N', ϵ'), if $N \le N'$ and $\epsilon \le \epsilon'$. For an SDC method (here to be interpreted as a scheme where every parameter except for the the sweep number k and the step size τ is fixed a priori), the Pareto front is the boundary of the dominated set, i.e. all points (N, ϵ) for which a combination (k, τ) exists such that the resulting SDC effort $N_{k,\tau}$ and error $\epsilon_{k,\ tau}$ dominate (N, k).

In Fig. 6.5, left, those Pareto fronts are shown. The significant gain of improved DIRK sweeps over Euler sweeps is again apparent, as are the smaller differences between the DIRK sweeps. Optimizing the sweep for criterion (4.3) with weight function $w = 1$ appears to perform best in this example. This is due to the smaller number of sweeps it requires for a certain error reduction, which can be observed also in Fig. 6.5, right, where the Pareto-optimal number of sweeps is shown versus the number of time steps $0.1/\tau = N/k$. Obviously, the stage order 7 of the underlying Radau(5) collocation method can be recovered, with only a logarithmic efficiency loss, if a sufficiently large number of SDC sweeps is performed.

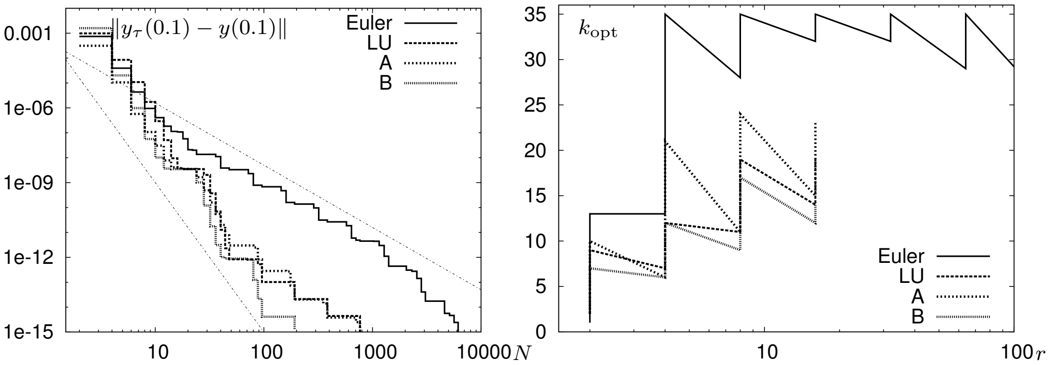

Fig. 6.5 Left: Pareto fronts of final time error vs. total number $N = 0.1k_{\text{opt}}/\tau$ of sweeps taken by QSDC methods for the nonautonomous heat equation (6.3) on Radau(5). $k = 1, \dots, 12$ sweeps are used. Limiting lines are $N^{-2.5}$ and N^{-6}. Right: most efficient number $k_{\text{opt}} \in \{1, \dots, 35\}$ of sweeps in each step for given number $r = 0.1/\tau$ of time steps.

Conclusion

The focus on SDC contraction makes it comparatively easy to construct efficient SDC sweeps by simple linear algebra means, either by the generic LU approach or direct numerical optimization. The thus constructed schemes often outperform standard Euler based sweeps by a significant factor. In particular they are able to recover the better convergence properties for stiff problems that are observed on equidistant grids also on non-equidistant grids such as Radau-IIa. The selection of a subset of the complex plane over which to optimize the SDC contraction allows to tailor the schemes to particular problem classes, such as the stiff reaction-diffusion equations considered here. The drawback is, of course, that the resulting schemes are no more general-purpose methods, and probably give worse results on problems they are not designed for.

Optimizing for flexible sweep blocks of size $m = 2$ according to the average norm objective $\max_z w(z)\| \prod_{j=1}^{m} G_j(z)\|^{1/m}\|$ with $w \equiv 1$ for problems with many stiff error components or $w(z) = 1/|z|$ for mildly stiff problems appears to be the most promising approach.

The convergence behavior of those methods can be quite complex. The notion of "order" in time step size τ is clearly insufficient to describe the error-work relationship in a meaningful way, which makes an adaptive choice of both, step size and sweep number, highly desirable.

Acknowledgement. The author is indebted to Bodo Erdmann for many fruitful discussions and thorough computational assistance. Partial funding by the DFG Research Center MATHEON "Mathematics for key technologies", projects A17 and F9, is gratefully acknowledged.

References

1. W. Auzinger, H. Hofstätter, O. Koch, W. Kreuzer, and E. Weinmüller. Superconvergent defect correction algorithms. *WSEAS Transactions on Systems*, 4:1378–1383, 2004.

2. W. Auzinger, H. Hofstätter, W. Kreuzer, and E. Weinmüller. Modified defect correction algorithms for ODEs. Part I: General theory. *Numer. Algorithms*, 36(2):135–156, 2004.
3. W. Auzinger, H. Hofstätter, W. Kreuzer, and E. Weinmüller. Modified defect correction algorithms for ODEs. Part II: Stiff initial value problems. *Numer. Algorithms*, 40(3):285–303, 2005.
4. M.M. Bowen. *A Spectral Deferred Correction Method for Solving Cardiac Models.* PhD thesis, Duke University, 2011.
5. S. Bu, J. Huang, and M.L. Minion. Semi-implicit Krylov deferred correction methods for differential algebraic equations. *Math. Comput.*, 81:2127–2157, 2012.
6. A.J. Christlieb, C.B. Macdonald, and B.W. Ong. Parallel high-order integrators. *SIAM J. Sci. Comput.*, 32(2):818–835, 2010.
7. A.J. Christlieb, B.W. Ong, and J.-M. Qiu. Comments on high order integrators embedded within integral deferred correction methods. *Comm. Appl. Math. Comput. Sci*, 4(1):27–56, 2009.
8. P. Deuflhard and M. Weiser. *Adaptive Numerical Solution of PDEs.* de Gruyter, 2012.
9. A. Dutt, L. Greengard, and V. Rokhlin. Spectral deferred correction methods for ordinary differential equations. *BIT*, 40(2):241–266, 2000.
10. M. Emmett and M.L. Minion. Toward an efficient parallel in time method for partial differential equations. *Comm. App. Math. and Comp. Sci.*, 7(1):105–132, 2012.
11. R. Frank and C.W. Überhuber. Iterated defect correction for the efficient solution of stiff systems of ordinary differential equations. *BIT*, 17:146–159, 1977.
12. B. Gustafsson and W. Kress. Deferred correction methods for initial value problems. *BIT*, 41(5):986–995, 2001.
13. A.C. Hansen and J. Strain. On the order of deferred correction. *Appl. Num. Math.*, 61:961–973, 2011.
14. J. Huang, J. Jia, and M.L. Minion. Accelerating the convergence of spectral deferred correction methods. *J. Comp. Phys.*, 214(2):633–656, 2006.
15. J. Huang, J. Jia, and M.L. Minion. Arbitrary order Krylov deferred correction methods for differential algebraic equations. *J. Comp. Phys.*, 221(2):739–760, 2007.
16. P. Kaps and P. Rentrop. Generalized Runge-Kutta methods of order four with stepsize control for stiff ordinary differential equations. *Numer. Math.*, 33:55–68, 1979.
17. A.T. Layton and M.L. Minion. Implications of the choice of quadrature nodes for Picard integral deferred corrections methods for ordinary differential equations. *BIT*, 45:341–373, 2005.
18. A.T. Layton and M.L. Minion. Implications of the choice of predictors for semi-implicit Picard integral deferred correction methods. *Comm. App. Math. and Comp. Sci.*, 2(1):1–34, 2007.
19. A. Ostermann and M. Roche. Rosenbrock methods for partial differential equations and fractional order of convergence. *SIAM J. Numer. Anal.*, 30(4):1084–1098, 1993.
20. A. Prothero and A. Robinson. On the stability and accuracy of one-step methods for solving stiff systems of ordinary differential equations. *Math. Comput.*, 28:145–162, 1974.
21. K.H. Schild. Gaussian collocation via defect correction. *Numer. Math.*, 58:369–386, 1990.
22. T. Tang, H. Xie, and X. Yin. High-order convergence of spectral deferred correction methods on general quadrature nodes. *J. Sci. Comput.*, pages 1–13, 2012.
23. P.E. Zadunaisky. On the estimation of errors propagated in the numerical integration of ODEs. *Numer. Math.*, 27:21–39, 1976.