Numerical Integration of Positive Linear Differential-Algebraic Systems [‡]

A.K. Baum [*] V. Mehrmann [†]

Abstract

In the simulation of dynamical processes in economy, social sciences, biology or chemistry, the analyzed values often represent nonnegative quantities like the amount of goods or individuals or the density of a chemical or biological species. Such systems are typically described by positive ordinary differential equations (ODEs) that have a non-negative solution for every non-negative initial value. Besides positivity, these processes often are subject to algebraic constraints that result from conservation laws, limitation of resources, or balance conditions and thus the models are differential-algebraic equations (DAEs). In this work, we present conditions under which both these properties, the positivity as well as the algebraic constraints, are preserved in the numerical simulation by Runge-Kutta or multistep discretization methods. Using a decomposition approach, we separate the dynamic and the algebraic equations of a given linear, positive DAE to give positivity preserving conditions for each part separately. For the dynamic part, we generalize the results for positive ODEs to DAEs using the solution representation via Drazin inverses. For the algebraic part, we use the consistency conditions of the discretization method to derive conditions under which this part of the approximation overestimates the exact solution and thus is non-negative. We test these conditions for some common Runge-Kutta and multistep methods and observe that none of these methods is suitable to solve positive higher index DAEs in a proper way.

2011 Mathematics Subject Classification: 65L80, 65L06, 15A16, 15B48
Key words. Differential-algebraic equation, positive system, Runge-Kutta method, multistep method, positivity preserving discretization, Z-matrix, M-matrix, stability function, Drazin inverse

1 Introduction

We consider the numerical solution of initial value problems for linear differential-algebraic equations (DAEs) with constant coefficients

$$E\dot{x} = Ax + f, \qquad x(t_0) = x_0, \tag{1}$$

where $E, A \in \mathbb{R}^{n\times n}$. We assume that the system is positive, i. e., that every solution $x(t)$ of (1) that starts with an element-wise non-negative initial value x_0 stays non-negative for all times $t \geq t_0$. Positive systems arise in every application in which $x(t)$ models a quantity that does not take negative values, like e.g. the concentration of chemical and biological species or the amount of goods and individuals in economic and social sciences. Examples are Leontief- and Leslie-Models, [13, 14], compartment-models [2, 4, 10, 13, 35] or semi-discretized advection-diffusion equations [1, 2, 3, 5, 11, 13, 23, 24, 25, 32, 33, 34]. Besides positivity, these processes usually have to satisfy

[‡]Supported by DFG Research Center Matheon, *Mathematics for Key Technologies* in Berlin
[*]Institut für Mathematik, MA 4-5, TU Berlin, Str. des 17. Juni 136, 10623 Berlin, Germany. email: baum@math.tu-berlin.de.
[†]Institut für Mathematik, MA 4-5, TU Berlin, Str. des 17. Juni 136, 10623 Berlin, Germany. email: mehrmann@math.tu-berlin.de. Supported by ERC Advanced Grant, MODSIMCONMP.

additional algebraic conditions resulting from limitation of resources, conservation laws or balance conditions and which extend the differential system by accessory algebraic equations.

The description of the system by an input-output map or its linearization around equilibrium solutions then leads to linear models of the type (1). In order to obtain a physically meaningful simulation of such processes, one has to assure that the numerical approximations satisfy the algebraic constraints. For unconstrained problems, i. e., problems in which E is invertible, this topic is well studied, see e. g.. [20, 21, 22, 24] and the references therein. For systems in which additional conditions like balance equations or conservation laws lead to DAEs this question has not been analyzed so far.

The paper is organized as follows. In section 2, the basic concepts for the analysis of linear, time-invariant DAEs and their discretization by one- and multistep methods are introduced. In section 3, the main result for positivity preserving discretizations of ODEs is presented. In section 4, these results are generalized to DAEs, first for the dynamic components in 4.4, then for the algebraic part in 4.5. In section 4.6 an example is presented.

2 Preliminaries

A matrix $A \in \mathbb{R}^{n\times n}$ is called *-Z-matrix*, if there exists $\mu > 0$, such that $A \geq -\mu I$, where the inequality is considered entry-wise. If the eigenvalues λ of A additionally satisfy $\max_{\lambda\in\sigma(A)} |\mu + \lambda| \leq \mu$, where $\sigma(A)$ denotes *the spectrum of* A, then A is called *-M-matrix*. If this inequality is strict, then A is called *nonsingular -M-matrix* and $(-A)^{-1}$ is non-negative, see [19].

A matrix pair (E, A), $E, A \in \mathbb{R}^{n\times n}$ is called *regular*, if there exist nonsingular transformation matrices $S, T \in \mathbb{R}^{n\times n}$ such that (E, A) can be written in Kronecker canonical form

$$(E,\ A) \quad = \quad \left(S \begin{bmatrix} I_d & 0 \\ 0 & N_a \end{bmatrix} T,\ S \begin{bmatrix} J_d & 0 \\ 0 & I_a \end{bmatrix} T \right),$$

where $J_d \in \mathbb{R}^{d\times d}$ is a matrix in Jordan canonical form [15], associated with the *finite generalized eigenvalues* $\lambda \in \sigma_{fin}(E, A)$, i. e., those $\lambda \in \mathbb{C}$, for which there exists $0 \neq v \in \mathbb{R}^n$ with $\lambda Ev = Av$, and $N_a \in \mathbb{R}^{a\times a}$ is a nilpotent matrix in Jordan canonical form associated with the *infinite eigenvalues* of E, A.

For a regular pair (E, A), the *index of nilpotency of* N_a, i. e., the smallest $\nu \in \mathbb{N}_0$ with $N_a^\nu = 0$, $N_a^{\nu-1} \neq 0$, is called the *(Kronecker) index* $ind(E, A) := \nu$ *of* (E, A). Note that this definition is independent of the choice of the transformations S and T, see [27].

In order to give an explicit solution representation of (1) for singular $E \in \mathbb{R}^{n\times n}$, we introduce the *Drazin inverse* E^D, that is characterized by the following properties, see e. g. [9].

$$(i)\ E^D E = EE^D, \quad (ii)\ E^D EE^D = E^D, \quad (iii)\ E^D E^{\nu+1} = E^\nu,\ \nu = \mathrm{ind}(E),$$

where the *index* of a *single matrix* E is defined by $\mathrm{ind}(E) := \mathrm{ind}(E, I_n)$. For every matrix $E \in \mathbb{R}^{n\times n}$, there exists a unique Drazin inverse and if E is nonsingular, then $E^D = E^{-1}$, see [9].

For a regular pair (E, A) the product $E^D E$ is a projector onto that part of E that corresponds to the differential equations of the DAE (1), see [9]. The complementary projection $P_\infty := I_n - E^D E$, respectively, picks out the algebraic equations.

Theorem 2.1 (See e. g. [27])**.** *Let $E, A \in \mathbb{R}^{n\times n}$ be a regular, commuting matrix pair with $ind(E, A) = \nu$, $\nu \in \mathbb{N}_0$ and let $f \in C^\nu(\mathbb{R}, \mathbb{R}^n)$. The initial value problem $E\dot{x} = Ax + f$, $x(t_0) = x_0$ is uniquely solvable, if there exists $v_0 \in \mathbb{R}^n$ with*

$$x_0 \quad = \quad E^D Ev_0 - P_\infty \sum_{\ell=0}^{\nu-1} (A^D E)^\ell A^D f^{(\ell)}(t_0).$$

If this is the case, then the solution is given by

$$x(t) \quad = \quad e^{(t-t_0)E^D A} E^D Ev_0 + \int_{t_0}^{t} e^{(t-s)E^D A} E^D f(s)\ ds - P_\infty \sum_{\ell=0}^{\nu-1} (A^D E)^\ell A^D f^{(\ell)}(t),$$

for $t \geq t_0$.

The commutativity assumption in Theorem 2.1 is not a restriction because any DAE with a regular matrix pair (E, A) can be transformed into an equivalent system with commuting matrix pair. This is done by a transformation of $E\dot{x} = Ax + f$ from the left with $(E - \lambda A)^{-1}$, where $\lambda \in \mathbb{C}$ is such that $\det(E - \lambda A) \neq 0$, see [8].

The crucial point in our analysis will be the decomposition of the solution x into its differential and algebraic components, $x = x_d + x_a$ with $x_d = E^D E x$ and $x_a = P_\infty x$.

The differential part of the solution can then be written as

$$x_d = E^D E x(t) \quad = \quad e^{(t-t_0)E^D A} E^D E v_0 + \int_{t_0}^{t} e^{(t-s)E^D A} E^D f(s)\, ds$$

for $t \geq t_0$, by applying the projection $E^D E$ to the solution $x(t)$ and using the commutativity and that $(E^D E)^2 = E^D E$. Note that this corresponds to the solution of the ordinary differential equation (ODE) $\dot{x}_d = E^D A x_d + E^D f$ with initial condition $x_d(t_0) = E^D E v_0$. The algebraic part is given by

$$x_a = P_\infty x(t) \quad = \quad -P_\infty \sum_{\ell=0}^{\nu-1} (A^D E)^\ell A^D f^{(\ell)}(t),$$

for $t \geq t_0$, which is completely determined by the inhomogeneity f and its derivatives $f^{(1)}, ..., f^{(\nu-1)}$. In particular, for the initial value x_0, we get the consistency condition

$$P_\infty x_0 = -P_\infty \sum_{\ell=0}^{\nu-1} (A^D E)^\ell A^D f^{(\ell)}(t_0).$$

Thus, the dynamic behavior of the solution is determined by the matrix $E^D A$, which corresponds to the Jordan matrix J_d associated with the generalized finite eigenvalues and vanishes on the algebraic components. The constrained components are fixed by the function f and its derivatives, and the matrix $P_\infty A^D E$, which corresponds to the nilpotent matrix N_a on $\mathrm{im}(P_\infty)$, cp. [39].

For the numerical solution of (1), we denote by x_N the approximation of the exact solution $x(t)$ at time $t = t_N$. As discretization schemes, we consider implicit Runge-Kutta methods and linear multistep methods.

For a given initial value x_0, an Implicit Runge-Kutta method has the form, see e. g. [7, 17]

$$x_{N+1} \quad = \quad x_N + \tau \sum_{i=1}^{s} \beta_i \dot{X}_{j,i}, \;\; N \in \mathbb{N}, \tag{2a}$$

$$E\dot{X}_{N,i} \quad = \quad Ax_j + \tau \sum_{j=1}^{s} \alpha_{ij} A\dot{X}_{N,j} + f(t_N + \tau\gamma_i), \text{ for } i = 1, ..., s, \tag{2b}$$

where $(\mathcal{A}, \beta, \gamma)$ with coefficients $\mathcal{A} := [\alpha_{ij}]$, $\beta := [\beta_i]$, and $\gamma := [\gamma_i]$, $i, j, = 1, .., s$.

For given initial values $x_0, ..., x_k$, a linear multistep method is given by , see e. g. [12, 17],

$$\sum_{j=0}^{k} \alpha_{k-j} E x_{N-j} \quad = \quad \tau \sum_{j=0}^{k} \beta_{k-j} A x_{N-j} + \tau \sum_{j=0}^{k} \beta_{k-j} f(t_{N-j}), \tag{3}$$

where (α, β) with coefficients $\alpha := [\alpha_j]$, $\beta := [\beta_j]$, $j = 1, .., k$.

For Runge-Kutta-Methods, formula (2b) can be rewritten as

$$\begin{aligned} x_{N+1} \quad &= \big(I_n + (\tau\beta^T \otimes I_n)(I_s \otimes E - \tau\mathcal{A} \otimes A)^{-1}(\mathbf{1} \otimes A)\big) x_N \\ &\quad + \sum_{i=1}^{s} (\tau\beta^T \otimes I_n)(I_s \otimes E - \tau\mathcal{A} \otimes A)^{-1}(e_i \otimes I_n) f(t_N + \tau\gamma_i), \end{aligned} \tag{4}$$

where $\otimes$ denotes the *Kronecker product* defined by $A \otimes B := [a_{ij}B] \in \mathbb{R}^{rk \times ml}$ for $A \in \mathbb{R}^{r\times m}, B \in \mathbb{R}^{k\times l}$, see [29]. Throughout the paper $e_i \in \mathbb{R}^s$, $j = 1, ..., s$, denotes the i-th unit vector and $\mathbf{1} := [1, ..., 1]^T$ of appropriate size. If (E, A) is a regular matrix pair and the coefficient matrix $\mathcal{A}$ is nonsingular, then the inverse in (4) is always well defined for small $\tau > 0$.

For multistep methods, formula (3) can be rewritten as

$$x_N = \sum_{j=0}^{k-1} (\alpha_k E - \beta_k \tau A)^{-1}(\beta_j \tau A - \alpha_j E)x_{N-k+j} + \tau \sum_{j=0}^{k} \beta_j (\alpha_k E - \beta_k \tau A)^{-1} f(t_{N-k+j}), \quad (5)$$

and this is well defined for small $\tau > 0$ if $\alpha_k \neq 0$ and if (E, A) is a regular matrix pair.

Thus, for linear problems of the form (1), a Runge-Kutta or multistep discretization corresponds to a linear iteration of the previous values $x_{N-1}, ..., x_{N-k}$. The properties of these iterations are determined by the iteration matrices, i. e., the matrices acting on the previous values x_N or $x_{N-k}, ..., x_{N-1}$, respectively, and the corresponding inhomogeneities involving the input function f. In the following, we refer to these matrices for Runge-Kutta methods by

$$\begin{aligned} \mathcal{R}(E, \tau A) &:= I_n + (\tau \beta^T \otimes I_n)(I_s \otimes E - \tau \mathcal{A} \otimes A)^{-1}(\mathbf{1} \otimes A), \\ \mathcal{Q}_i(E, \tau A) &:= (\tau \beta^T \otimes I_n)(I_s \otimes E - \tau \mathcal{A} \otimes A)^{-1}(e_i \otimes I_n),\ i = 1, ..., s \end{aligned}$$

and for multistep methods by

$$\begin{aligned} r_j(E, \tau A) &:= -(\alpha_k E - \beta_k \tau A)^{-1}(\alpha_j E - \beta_j \tau A),\ j = 0, ..., k-1, \\ q_j(E, \tau A) &:= \tau \beta_j (\alpha_k E - \beta_k \tau A)^{-1},\ j = 0, ..., k-1. \end{aligned}$$

If a Runge-Kutta or multistep method is applied to the scalar problem $\dot{x} = \lambda x$, $\lambda \in \mathbb{R}$, then $\mathcal{R}$ and r_j correspond to rational functions in λ. These are called the *stability functions* of the method. In the following discussion, we will mostly consider the discretization of the scaled system $\hat{E}\dot{x} = \hat{A}x + \hat{f}$, where

$$\hat{E} := (\hat{\lambda} E - A)^{-1} E, \quad (6a)$$

$$\hat{A} := (\hat{\lambda} E - A)^{-1} A, \quad (6b)$$

$$\hat{f} := (\hat{\lambda} E - A)^{-1} f. \quad (6c)$$

Here it is important to note that the iteration matrices as well as the inhomogeneities are invariant against this scaling, i. e.,

$$\mathcal{R}(\hat{E}, \tau\hat{A}) = \mathcal{R}(E, \tau A), \quad \mathcal{Q}_i(\hat{E}, \tau\hat{A}) = \mathcal{Q}_i(E, \tau A),$$

and

$$r_j(\hat{E}, \tau\hat{A}) = r_j(E, \tau A), \quad q_j(\hat{E}, \tau\hat{A}) = q_j(E, \tau A),$$

which immediately follows from

$$\begin{aligned} \mathcal{R}(\hat{E}, \tau\hat{A}) &= I_n + (\tau\beta^T \otimes I_n)(I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})^{-1}(\mathbf{1} \otimes \hat{A}) \\ &= I_n + (\tau\beta^T \otimes I_n)(I_s \otimes (\hat{\lambda}E - A)^{-1}E - \tau\mathcal{A} \otimes (\hat{\lambda}E - A)^{-1}A)^{-1}(\mathbf{1} \otimes (\hat{\lambda}E - A)^{-1}A) \\ &= I_n + (\tau\beta^T \otimes I_n)(I_s \otimes (\hat{\lambda}E - A))(I_s \otimes E - \tau\mathcal{A} \otimes A)^{-1}(I_s \otimes (\hat{\lambda}E - A))^{-1}(\mathbf{1} \otimes A) \\ &= I_n + (\tau\beta^T \otimes I_n)(I_s \otimes E - \tau\mathcal{A} \otimes A)^{-1}(\mathbf{1} \otimes A) \\ &= \mathcal{R}(E, \tau A), \end{aligned}$$

and similar computations for the invariance properties of the other functions.

We say that a Runge-Kutta or multistep method is *positive* for problem (1), if it provides nonnegative approximations x_N, $N \in \mathbb{N}$, for every set of non-negative starting values x_0 or $x_0, ..., x_k$. Considering the initial values e_i or 0, we immediately see that the iterations (4) and (5) are positive, if and only if the iteration matrices and the inhomogeneous part are elementwise nonnegative, see e. g. [13].

3 Positivity preservation for ODEs

For ODEs, i. e., where $E = I$, positivity is characterized by the sign pattern of the system matrix A. Since the continuous evolution operator e^{tA} is nonnegative for every $t \geq 0$, if and only if $A \in \mathbb{R}^{n\times n}$ is a -Z-matrix, see e. g. [38], the linear problem $\dot{x} = Ax + f$ is positive, if and only if A is a -Z-matrix and $f \geq 0$, see, e. g. [13, 24].

To obtain a positive discretization of such problems, it has been shown in [6] that it is sufficient that the stepsizes for the discretization method are bounded by the *radius of absolute monotonicity*, which is defined as the largest real number $\gamma_+ \geq 0$, such that the stability functions are *absolutely monotonic* on $[-\gamma_+, 0]$, i. e., the stability functions and all its derivatives are nonnegative and have no poles in $[-\gamma_+, 0]$.

For problems where the system matrix is a -M-matrix, then one obtains the following condition for a positive discretization, see e. g. [6] or [24]. In preparation of the differential-algebraic case, we present this theorem with a proof in our notation.

Theorem 3.1. *Consider a positive system $\dot{x} = Ax + f$ for which A is a -M-matrix with $\mu > 0$.*

1. *A Runge-Kutta method with coefficients $(\mathcal{A}, \beta, \gamma)$ and stability function $\mathcal{R}(z) := 1 + z\beta^T(I_s - z\mathcal{A})^{-1}\mathbf{1}$, that is absolutely monotonic on $[-\gamma_+, 0]$, $\gamma_+ \geq 0$, is positive for $\dot{x} = Ax$ if the stepsize τ satisfies $0 < \tau \leq \frac{\gamma_+}{\mu}$.*

   Moreover, if additionally $\mathcal{Q}_i(z) := \beta^T(I_s - z\mathcal{A})^{-1}e_i$ is absolutely monotonic on $[-\gamma_+, 0]$ for $i = 1, .., s$, then the method is also positive for $\dot{x} = Ax + f$ provided that $0 < \tau \leq \frac{\gamma_+}{\mu}$.

2. *A multistep method with coefficients (α, β) for which the stability functions $r_{k-j}(z) := -\frac{\alpha_{k-j} - \beta_{k-j}z}{\alpha_k - \beta_k z}$, $j = 1, .., k$, are absolutely monotonic on $[-\gamma_+, 0]$, $\gamma_+ \geq 0$, is positive on $\dot{x} = Ax$ if the stepsize τ satisfies $0 < \tau \leq \frac{\gamma_+}{\mu}$.*

   Moreover, if additionally $\frac{\beta_k}{\alpha_k} \geq 0$ and $\frac{\beta_{k-1}}{\alpha_k} \geq 0$ holds for $j = 1, .., k$, then the method is also positive for $\dot{x} = Ax + f$ provided that $0 < \tau \leq \frac{\gamma_+}{\mu}$.

Proof. The basic idea for the proof is to consider the iteration matrices and inhomogeneities of a discretization method as matrix valued functions and show their nonnegativity by identifying them with a nonnegative Taylor expansion.

1. For a Runge-Kutta method, we can expand the stability functions in a Taylor series centered in $-\tau\mu I_n$ and obtain

$$\mathcal{R}(\tau A) \;=\; \sum_{k=0}^{\infty} \frac{1}{k!}\mathcal{R}^{(k)}(-\tau\mu I_n)\ (\tau B)^k, \tag{7}$$

and

$$\mathcal{Q}_i(\tau A) \;=\; \sum_{k=0}^{\infty} \frac{1}{k!}\mathcal{Q}_i^{(k)}(-\tau\mu I_n)\ (\tau B)^k, \tag{8}$$

where $B := \mu I_n + A$ is nonnegative, since A is a -M-matrix. Using the structure of $\mathcal{R}(-\tau\mu I_n)$ and $\mathcal{Q}_i(-\tau\mu)$, the coefficients in these expansions can be written as

$$\mathcal{R}(-\tau\mu I_n) \;=\; I_n - (\tau\mu\beta^T \otimes I_n)\ \left(I_s \otimes I_n + \tau\mu\mathcal{A} \otimes I_n\right)^{-1}\ (e_s \otimes I_n) \;=\; \mathcal{R}(-\tau\mu)I_n,$$

and

$$\mathcal{Q}_i(-\tau\mu I_n) \;=\; (\beta^T \otimes I_n)(I_s \otimes I_n + \tau\mu A \otimes I_n)^{-1}(\mathbf{1} \otimes I_n) \;=\; \mathcal{Q}_i(-\tau\mu)I_n,$$

respectively, such that (7) and (8) correspond to matrix power series with scalar coefficients. These series converge, if the spectral radius of B satisfies $\rho(\tau B) < r_{\mathcal{R}}$ and $\rho(\tau B) < r_{\mathcal{Q}_i}$, where $r_{\mathcal{R}}, r_{\mathcal{Q}_i} \geq 0$, respectively, denote the radius of convergence of the scalar power series $\sum_{k=0}^{\infty} \frac{1}{k!}\mathcal{R}^{(k)}(-\tau\mu)\xi^k$

and $\sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{Q}_i^{(k)}(-\tau\mu)\xi^k$ for $\xi \in \mathbb{C}$, $i = 1, ..., s$. Since A is a -M-matrix, $\rho(\tau B)$ is bounded by $\rho(\tau B) = \tau \max_{\lambda \in \sigma(A)} |\mu + \lambda| \leq \tau\mu$ and this means that the expansions (7) and (8) are valid for every $\tau > 0$ with $\tau\mu < r_{\mathcal{R}}$ and $\tau\mu < r_{\mathcal{Q}_i}$, $i = 1, ..., s$, respectively.

To show the positivity of these expansions, we note that $B \geq 0$ holds by construction and $\mathcal{R}^{(k)}(-\tau\mu)$ and $\mathcal{Q}_i^{(k)}(-\tau\mu)$ are nonnegative if $-\tau\mu \in [\gamma_+, 0]$ by the absolute monotonicity assumption. The absolute monotonicity further implies that $\mathcal{R}^{(k)}(-\tau\mu)$, $\mathcal{Q}_i^{(k)}(-\tau\mu)$ have no poles in $[\gamma_+, 0]$, i. e., it follows that $\gamma_+ \leq r_{\mathcal{R}}, r_{\mathcal{Q}_i}$ and we obtain a nonnegative Taylor expansion of $\mathcal{R}(\tau A)$ and $\mathcal{Q}_i(\tau A)$ for every τ with $\tau\mu < \gamma_+$.

2. For multistep methods, we apply an analogous argument. For $r_{k-j}(\tau A)$, the absolute monotonicity on $[-\gamma_+, 0]$ implies the existence of a nonnegative expansion $\sum_{k=0}^{\infty} \frac{1}{k!} r_{k-j}^{(k)}(-\tau\mu I_n)\ (\tau B)^k$ for every $\tau > 0$ with $\tau\mu < \gamma_+$. The $q_{k-j}(\tau A)$ are nonnegative for any $\tau > 0$, if $\beta_{k-j} \geq 0$ and $\frac{\beta_k}{\alpha_k} \geq 0$, because the matrix $(I_n - \frac{\beta_k}{\alpha_k}\tau A)$ is a nonsingular M-matrix, if A is a -M-matrix. □

After this brief recollection of the results for ODEs, in the next section we extend these results to DAEs.

4 Positivity preservation for DAEs

4.1 Positivity concepts for matrix pairs

In order to generalize Theorem 3.1 to DAEs, we first give a positivity characterization for the continuous problem (1) in the case of singular E. We recall the result given in [39] in a way that admits a uniform description of ODEs and DAEs.

Definition 4.1. *A matrix pair (E, A) in $\mathbb{R}^{n\times n}$ is called* Z-matrix pair*, if $E^D E \geq 0$ and if there exists $\mu > 0$, such that $E^D A \leq \mu E^D E$. Furthermore, (E, A) is a* -Z-matrix pair *if $(E, -A)$ is a Z-matrix pair.*

Lemma 4.1 ([39])**.** *Let (E, A) be a regular, commuting matrix pair with $E, A \in \mathbb{R}^{n\times n}$. Then $e^{tE^D A}E^D E \geq 0$ for $t \geq 0$ if and only if (E, A) is a -Z-matrix pair.*

Proof. "⇒": Suppose that $e^{tE^D A}E^D E \geq 0$ for all $t \geq 0$. For small $t \geq 0$, this implies that $(I_n + tE^D A)E^D E \geq 0$ by expanding $e^{tE^D A}E^D E$ in $t = 0$ via a Taylor series, i. e., $E^D E + tE^D A \geq 0$ by the commutativity of E, A and the properties of the Drazin inverse. This implies that $E^D E \geq 0$, because otherwise $E^D E + tE^D A$ would become negative.

Furthermore, for indices (i, j) with $[E^D E]_{ij} \neq 0$ it follows that $\frac{(E^D A)_{ij}}{(E^D E)_{ij}} \geq -\frac{1}{t}$, whereas $(E^D A)_{ij} \geq 0$ holds for those indices i, j with $[E^D E]_{ij} = 0$. This means that $E^D A \geq -\mu E^D E$ for every $\mu \geq |\min_{[E^D E]_{ij}\neq 0} \frac{(E^D A)_{ij}}{(E^D E)_{ij}}|$.

"⇐": Let (E, A) be a -Z-matrix pair and $\tau \geq 0$. Since $(E^D E)^N = E^D E$ for $N \in \mathbb{N}\setminus\{0\}$, setting $t = N\tau$, we can write $e^{tE^D A}E^D E$ as $e^{\tau E^D A}\big(E^D E\big)^N$, and since $E^D E \geq 0$, this iteration has the same sign as $e^{\tau E^D A}E^D E$. Thus, it is sufficient to prove that $e^{\tau E^D A}E^D E \geq 0$ for $\tau \geq 0$ sufficiently small. But for small τ, we again may use the approximation $e^{\tau EA}E^D E \approx (I_n + \tau E^D A)E^D E$, which is nonnegative if $\tau \leq \frac{1}{\mu}$, since (E, A) is a -Z-matrix pair with $\mu > 0$. □

Using Lemma 4.1, the positivity of a DAE can be characterized in the same manner as for ODEs, if the nonnegativity of the algebraic components in (1) is ensured for all $t \geq 0$.

Lemma 4.2 ([39])**.** *Let (E, A) with $E, A \in \mathbb{R}^{n\times n}$ be a regular, commuting matrix pair with $ind(E, A) = \nu$, $\nu \in \mathbb{N}_0$, $f \in C^\nu(\mathbb{R}, \mathbb{R}^n)$. If $E^D E \geq 0$ and $P_\infty \sum_{\ell=0}^{\nu-1}(A^D E)^\ell A^D f^{(\ell)} \geq 0$ for $t \geq 0$, then $E\dot{x} = Ax + f$ is positive if and only if (E, A) is a -Z-matrix pair and $E^D f \geq 0$ for $t \geq 0$.*

Proof. "⇒": Suppose that $E\dot{x} = Ax + f$ is positive. For homogeneous problems, every $v_0 \in \mathbb{R}^n$ defines a consistent initial value $x_0 = E^D E v_0$ and $x_0 \geq 0$ if $v_0 \geq 0$, because $E^D E$ is nonnegative.

Therefore, we can choose, in particular, $v_{0,i} = e_i$ to obtain nonnegative initial values $x_{0,i} = E^D E v_{0,i}$ with associated solutions

$$x_i \quad = \quad e^{(t-t_0)E^D A} E^D E v_{0,i} \quad i = 1, ..., n.$$

Since these solutions are nonnegative by the positivity assumption, this implies that $e^{(t-t_0)E^D A} E^D E \geq 0$ for $t \geq t_0$ and by Lemma 4.1 this means that (E, A) is a -Z-matrix pair.

To prove that $E^D f \geq 0$, we consider the special solution

$$x(t) \quad = \quad \int_{t_0}^{t} e^{(t-s)E^D A} E^D f(s) ds - P_\infty \sum_{\ell=0}^{\nu-1} (A^D E)^\ell A^D f^{(\ell)}(t) \geq 0$$

that is associated with the initial condition $x_0 = -\sum_{\ell=0}^{\nu-1} P_\infty (A^D E)^\ell A^D f^{(\ell)}(t_0)$. This means that $E^D E x_0 = 0$ and $x_0 \geq 0$ and, since $E^D E \geq 0$, the projection $E^D E x(t) = \int_{t_0}^{t} e^{(t-s)E^D A} E^D f(s) ds$ is nonnegative as well. By the monotonicity of the integral, this implies that $e^{(t-s)E^D A} E^D f(s) \geq 0$ holds for $t \geq s$ and considering this for $t = s$, we get that $E^D f(t) \geq 0$ holds for $t \geq t_0$.

"$\Leftarrow$": Suppose that (E, A) is a -Z-matrix pair and that $E^D f(t) \geq 0$ for $t \geq 0$. Then $e^{(t-t_0)E^D A} E^D E \geq 0$ holds for $t \geq t_0$ by Lemma 4.1 and the homogeneous solution $x(t) = e^{(t-t_0)E^D A} E^D E v_0$ is nonnegative for every consistent initial value $x_0 = E^D E v_0 \geq 0$.

For the solution of the inhomogeneous problem, the monotonicity of the integral preserves the nonnegativity of $E^D f$ and $e^{(t-t_0)E^D A} E^D E$, which means that $\int_{t_0}^{t} e^{(t-s)E^D A} E^D f(s) ds \geq 0$ holds for $t \geq t_0$.

The algebraic parts of the solution, i. e., $P_\infty \sum_{\ell=0}^{\nu-1} (A^D E)^\ell A^D f^{(\ell)}(t)$, are nonnegative by assumption. □

Note that it is not necessary to assume that $t \geq 0$. This assumption is simply made to enlarge the class of admissible inhomogeneities for positive DAEs. If a specific application requires that $t \in [a, b]$ for $a, b \in \mathbb{R}$, then Theorem 4.2 can be applied by assuming that $P_\infty \sum_{\ell=0}^{\nu-1} (A^D E)^\ell A^D f^{(\ell)}(t) \geq 0$ holds for any $t \in [a, b]$.

4.2 Results on Drazin inverses

To give conditions for a positive discretization of DAEs we need some results about Drazin inverses.

Lemma 4.3. *Let $E \in \mathbb{R}^{n \times n}$. The projections $E^D E$ and $P_\infty = I_n - E^D E$ are invariant under the application of the Drazin inverse, i. e., $E^D E = \left(E^D E\right)^D$ and $P_\infty = P_\infty^D$.*

Proof. We prove the assertion by verifying the properties of the Drazin inverse. (i) The commutativity of $(E^D E)^D$ and $E^D E$ is an immediate result of the commutativity of E and E^D, i. e.,

$$(E^D E)^D E^D E = E^D E E^D E^D.$$

(ii) Since $(E^D E)^2 = E^D E$, we immediately get that

$$(E^D E)^D (E^D E)(E^D E)^D = (E^D E)^3 = E^D E = (E^D E)^D.$$

(iii) If the Jordan canonical form of E is given by $E = T^{-1} \left[\begin{smallmatrix} J_E & 0 \\ 0 & N_E \end{smallmatrix}\right] T$ with J_E nonsingular and N_E nilpotent, then $E^D = T^{-1} \left[\begin{smallmatrix} J^{-1} & 0 \\ 0 & 0 \end{smallmatrix}\right] T$ and $E^D E = T^{-1} \left[\begin{smallmatrix} I & 0 \\ 0 & 0 \end{smallmatrix}\right] T$. This means that the index of $E^D E$ is one and hence

$$(E^D E)^D (E^D E)^2 = E^D E E^D E = E^D E.$$

The proof for P_∞ is analogous. □

We also need the Drazin inverses of matrix products.

Lemma 4.4. *Let $E, C, T \in \mathbb{R}^{n \times n}$.*

1. *If C is nonsingular and $EC = CE$, then $(EC)^D = C^{-1}E^D = E^DC^{-1}$.*

2. *If T nonsingular, then $(T^{-1}ET)^D = T^{-1}E^DT$.*

Proof. We again verify the assertions by checking the characteristic properties of the Drazin inverse.

1. (i) Note that if E and C commute, then E, C, E^D and C^{-1} all commute as well, cp. [27] and with $C^D = C^{-1}$, we immediately obtain

$$(EC)^D(EC) = E^DC^{-1}EC = ECE^DC^{-1} = (EC)(EC)^D.$$

(ii) In the same manner, we verify that

$$(EC)^D(EC)(EC)^D = E^DC^{-1}ECE^DC^{-1} = C^{-1}E^DEE^D = C^{-1}E^D = (EC)^D.$$

(iii) For the third property, we note that if E and C commute, then $(EC)^k = E^kC^k$ holds for every $k \in \mathbb{N}$. Then, by the regularity of C, it follows that $\ker((EC)^k) = \ker(E^k)$, because $(EC)^kv = C^kE^kv = 0$ holds for $v \neq 0$ if and only if $E^kv = 0$, i. e., $v \in \ker(E^k)$.

Note further, that $\mathrm{ind}(E)$ can be characterized as that number $\nu \in \mathbb{N}$, for which $\ker(E^\nu) = \ker(E^{\nu+1})$, but $\ker(E^{\nu-1}) \neq \ker(E^\nu)$. This can be seen by considering the Jordan canonical form of E, i. e., $E = T^{-1}\left[\begin{smallmatrix} J_E & 0 \\ 0 & N_E \end{smallmatrix}\right]T$ with J_E nonsingular and N_E nilpotent with index $\nu = \mathrm{ind}(E)$. Then $E^k = T^{-1}\left[\begin{smallmatrix} J_E^k & 0 \\ 0 & N_E^k \end{smallmatrix}\right]T$ and $\ker(E^k) = \{x \in \mathbb{R}^n | x = T^{-1}\left[\begin{smallmatrix} 0 \\ \tilde{x}_2 \end{smallmatrix}\right], \tilde{x}_2 \in \ker(N^k)\}$. If $k \geq \nu$, then $N^k = 0$, i. e., $\ker(E^k)$ does not dependent of k anymore, while $\dim(\ker(E^k)) < \dim(\ker(E^{k+1}))$ if $k < \nu$. So if $\ker((EC)^k) = \ker(E^k)$ for $k \in \mathbb{N}$, then $\ker((EC)^k) = \ker((EC)^{k+1})$ for $k \geq \nu$ and $\ker((EC)^k) \neq \ker(E^{k+1}C)$ for $k < \nu$ and it follows that $\mathrm{ind}(EC) = \mathrm{ind}(E)$. With this, we immediately obtain

$$(EC)^D(EC)^{\nu+1} = E^DC^{-1}E^{\nu+1}C^{\nu+1} = C^\nu E^\nu = (EC)^\nu.$$

2. (i) As a direct result of the commutativity of E and E^D and the nonsingularity of T, we obtain

$$(T^{-1}ET)^D(T^{-1}ET) = T^{-1}E^DET = T^{-1}EE^DT = (T^{-1}ET)(T^{-1}ET)^D.$$

(ii) From $E^DEE^D = E^D$ we immediately obtain that

$$(T^{-1}ET)^D(T^{-1}ET)(T^{-1}ET)^D = T^{-1}E^DEE^DT = T^{-1}E^DT = (T^{-1}ET)^D.$$

(iii) Note that $(T^{-1}ET)^k = T^{-1}E^kT$ holds for every $k \in \mathbb{N}$. By the nonsingularity of T it follows that $\ker((T^{-1}ET)^k) = \ker(E^k)$, because $(T^{-1}ET)^kv = T^{-1}E^kTv = 0$ holds for $v \neq 0$ if and only if $E^kTv = 0$, i. e., $Tv \in \ker(E^k)$. Then, $\ker((T^{-1}E^\nu T)^k) = \ker(E^\nu) = \ker(E^{\nu+1}) = \ker((T^{-1}E^{\nu+1}T)^k)$, but $\ker((T^{-1}E^{\nu-1}T)^k) = \ker(E^{\nu-1}) \neq \ker(E^\nu) = \ker((T^{-1}E^\nu T)^k)$ using the index characterization given in part 1., i. e., $\mathrm{ind}(T^{-1}ET) = \mathrm{ind}(E)$. With this, we immediately obtain

$$(T^{-1}ET)^D(T^{-1}ET)^{\nu+1} = T^{-1}E^DE^{\nu+1}T = T^{-1}E^\nu T = (T^{-1}ET)^\nu.$$

□

4.3 Kronecker products and Drazin inverses

In this subsection we summarize properties of the Kronecker product, see [28], as they are needed in the upcoming discussion. Let A, B, C, D be matrices, such that the products AC and BD exist, then the products of $(A \otimes B)$ and $(C \otimes D)$ is given by $(A \otimes B)(C \otimes D) = AC \otimes BD$, see [28]. If A, C and B, D commute respectively, this immediately implies that the Kronecker products commute as well, i. e., $(A \otimes B)(C \otimes D) = (C \otimes D)(A \otimes B)$. The factors in $A \otimes B$ can be reversed by the perfect shuffle matrices Π_r and Π_c, i. e., $\Pi_z^T(A \otimes B)\Pi_r = B \otimes A$. If $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{m\times m}$ are nonsingular, then the inverse of the Kronecker-Product is given by $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$. The next Lemma shows that the same result holds for the Drazin inverse of a Kronecker Product.

Lemma 4.5. *Let $E \in \mathbb{R}^{n\times n}$ and $D \in \mathbb{R}^{m\times m}$, then $(E \otimes D)^D = E^D \otimes D^D$.*

Proof. We verify the assertion by checking the characteristic properties of a Drazin inverse. (i) The commutativity follows directly from the commutativity of E and D with their respective Drazin inverses, i. e.,

$$(E \otimes D)^D(E \otimes D) = E^D E \otimes D^D D = EE^D \otimes DD^D = (E \otimes D)(E \otimes D)^D.$$

(ii) The second property of the Drazin inverse is implied by the corresponding property of E and D, i. e.,

$$(E \otimes D)^D(E \otimes D)(E \otimes D)^D = E^D E E^D \otimes D^D D D^D = E^D \otimes D^D = (E \otimes D)^D.$$

(iii) To prove the third property, we first determine the index of $E \otimes D$ by the same argument as in Lemma 4.4. The kernel of $E \otimes D$ is given by

$$\ker(E \otimes D) \;=\; \{v \otimes w \mid v \in \ker(E),\ w \in \mathbb{R}^m\} \cup \{v \otimes w \mid v \in \mathbb{R}^n,\ w \in \ker(D)\}.$$

More exactly, transforming $E \otimes D$ to Jordan canonical form, cp. [28], i. e.,

$$E \otimes D \;=\; T^{-1} \otimes S^{-1} \begin{bmatrix} J_E & 0 \\ 0 & N_E \end{bmatrix} \otimes \begin{bmatrix} J_D & 0 \\ 0 & N_D \end{bmatrix} T \otimes S$$

where J_E, J_D are nonsingular and N_E, N_D are nilpotent with index $\nu_E = \operatorname{ind}(E)$ and $\nu_D = \operatorname{ind}(D)$, respectively, then

$$E^k \otimes D^k \;=\; T^{-1} \otimes S^{-1} \begin{bmatrix} J_E^k & 0 \\ 0 & N_E^k \end{bmatrix} \otimes \begin{bmatrix} J_D^k & 0 \\ 0 & N_D^k \end{bmatrix} T \otimes S.$$

Thus,

$$\begin{aligned} \ker(E^k \otimes D^k) \;=\; & \{v \otimes w \mid v = T^{-1}\left[\begin{smallmatrix} 0 \\ \tilde{v}_2 \end{smallmatrix}\right], \tilde{v}_2 \in \ker(N_E^k), w \in \mathbb{R}^m\} \\ & \cup \{v \otimes w \mid w = S^{-1}\left[\begin{smallmatrix} 0 \\ \tilde{w}_2 \end{smallmatrix}\right], \tilde{w}_2 \in \ker(N_D^k), v \in \mathbb{R}^n\}. \end{aligned}$$

This means that $k = \max\{\nu_E, \nu_E\}$ is the smallest number, for which $\ker(E^k \otimes D^k)$ does not depend on k anymore, i. e., $\operatorname{ind}(E \otimes D) = \max\{\nu_E, \nu_E\}$. Setting $\nu := \max\{\nu_E, \nu_E\} = \operatorname{ind}(E \otimes D)$, the third property is verified by

$$(E \otimes D)^D(E \otimes D)^{\nu+1} = E^D E^{\nu+1} \otimes D^D D^{\nu+1} = E^\nu \otimes D^\nu = (E \otimes D)^\nu.$$

□

4.4 Positivity preservation for the differential part

With the results of the previous section, we now discuss the positive discretization of the dynamic parts of a DAE. We first note that scaling $E\dot{x} = Ax + f$ by $(\hat{\lambda}E - A)^{-1}$ not only yields commutative system matrices, but also provides a simultaneous similarity transformation of E, A that separates the dynamic and algebraic components. More exactly, if $E = S\left[\begin{smallmatrix} I_d & 0 \\ 0 & N_a \end{smallmatrix}\right]T$, $A = S\left[\begin{smallmatrix} J & 0 \\ 0 & I_a \end{smallmatrix}\right]T$ denotes the Kronecker canonical form, then $\hat{E}, \hat{A}$ from (6) can be written as, see [39],

$$\hat{E} = T^{-1} \begin{bmatrix} \Lambda_d^{-1} & 0 \\ 0 & \Lambda_a^{-1} N_a \end{bmatrix} T, \qquad \hat{A} = T^{-1} \begin{bmatrix} \Lambda_d^{-1} J & 0 \\ 0 & \Lambda_a^{-1} \end{bmatrix} T, \tag{9}$$

where $\Lambda_d := (I_d - \hat{\lambda} J_d)$ and $\Lambda_a := (I_a - \hat{\lambda} N_a)$.

The Drazin inverses of $\hat{E}, \hat{A}$ are then given by

$$\hat{E}^D = T^{-1} \begin{bmatrix} J_d^D \Lambda_d & 0 \\ 0 & 0 \end{bmatrix} T, \qquad \hat{A}^D = T^{-1} \begin{bmatrix} \Lambda_d & 0 \\ 0 & \Lambda_a \end{bmatrix} T$$

from which we obtain

$$\hat{E}^D\hat{E} = T^{-1}\begin{bmatrix} I_d & 0 \\ 0 & 0 \end{bmatrix} T, \qquad \hat{E}^D\hat{A} = T^{-1}\begin{bmatrix} J_d & 0 \\ 0 & 0 \end{bmatrix} T. \tag{10}$$

With Lemma 4.4, we note that

$$\hat{E}^D\hat{E} = \big((\hat{\lambda}E - A)^{-1}E\big)^D(\hat{\lambda}E - A)^{-1}E = E^D(\hat{\lambda}E - A)(\hat{\lambda}E - A)^{-1}E = E^DE$$

as well as

$$\hat{E}^D\hat{A} = \big((\hat{\lambda}E - A)^{-1}E\big)^D(\hat{\lambda}E - A)^{-1}A = E^D(\hat{\lambda}E - A)(\hat{\lambda}E - A)^{-1}A = E^DA,$$

i. e., the products in (10) are independent of the chosen $\hat{\lambda} \in \mathbb{C}$. In particular, this implies that the positivity of $E\dot{x} = Ax + f$ is not destroyed by scaling with $(\hat{\lambda}E - A)^{-1}$, since Lemma 4.2 only makes assumptions on the products E^DE, A^DE and $(A^DE)^pA^Df^{(p)}$.

With these transformations, we reformulate the projected inverses in (4) and (5) in terms of the Drazin inverse.

Lemma 4.6. *Let (E, A) be a regular matrix pair with $E, A \in \mathbb{R}^{n\times n}$ and ind$(E, A) = \nu$, and let $\hat{E}, \hat{A}$ be defined by* (6)*.*

1. *If $\mathcal{A} \in \mathbb{R}^{s\times s}$, then*

$$\begin{aligned}(I_s \otimes \hat{E}^D\hat{E})(I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})^{-1} &= \big((I_s \otimes \hat{E}^D\hat{E})(I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})\big)^D \\ &= (I_s \otimes \hat{E}^D)(I_s \otimes \hat{E}^D\hat{E} - \tau\mathcal{A} \otimes \hat{E}^D\hat{A})^D.\end{aligned}$$

2. *If $\alpha_k, \beta_k \neq 0$, then*

$$\hat{E}^D\hat{E}(\alpha_k\hat{E} - \beta_k\tau\hat{A})^{-1} = \big(\hat{E}^D\hat{E}(\alpha_k\hat{E} - \beta_k\tau\hat{A})\big)^D = \hat{E}^D\big(\alpha_k\hat{E} - \beta_k\tau\hat{A}\big)^D.$$

Proof. 1. To prove the first identity, we note that $\hat{E}^D\hat{E} = E^DE$ and Lemma 4.5 imply that $(I_s \otimes \hat{E}^D\hat{E}) = (I_s \otimes \hat{E}^D\hat{E})^D$. Since $(I_s \otimes \hat{E}^D\hat{E})$ and $(I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})$ commute, we then get from Lemma 4.4 that

$$\begin{aligned}(I_s \otimes \hat{E}^D\hat{E})\ (I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})^{-1} &= \Big((I_s \otimes \hat{E}^D\hat{E})(I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})\Big)^D \\ &= \Big((I_s \otimes \hat{E})(I_s \otimes \hat{E}^D\hat{E} - \tau\mathcal{A} \otimes \hat{E}^D\hat{A})\Big)^D.\end{aligned}$$

For the second identity, we exploit that the differential and algebraic components in $(\hat{E}, \hat{A})$ can be separated by the similarity transformation (9). Inserting this into $((I_s \otimes \hat{E})(I_s \otimes \hat{E}^D\hat{E} - \tau\mathcal{A} \otimes \hat{E}^D\hat{A}))^D$, we obtain

$$\begin{aligned}&\big((I_s \otimes \hat{E})(I_s \otimes \hat{E}^D\hat{E} - \tau\mathcal{A} \otimes \hat{E}^D\hat{A})\big)^D \\ &\quad= \left((I_s \otimes T^{-1})\Big(I_s \otimes \begin{bmatrix} \Lambda_d & 0 \\ 0 & \Lambda_aN_a \end{bmatrix}\Big)\Big(I_s \otimes \begin{bmatrix} I_d & 0 \\ 0 & 0 \end{bmatrix} - \tau\mathcal{A} \otimes \begin{bmatrix} J_d & 0 \\ 0 & 0 \end{bmatrix}\Big)(I_s \otimes T)\right)^D\end{aligned}$$

and permuting the Kronecker factors by the perfect shuffle matrix Π, we get

$$\begin{aligned}&\left((I_s \otimes T^{-1})\Pi^T\Big(\begin{bmatrix} \Lambda_d & 0 \\ 0 & \Lambda_aN_a \end{bmatrix} \otimes I_s\Big)\ \Big(\begin{bmatrix} I_d & 0 \\ 0 & 0 \end{bmatrix} \otimes I_s - \begin{bmatrix} J_d & 0 \\ 0 & 0 \end{bmatrix} \otimes \tau\mathcal{A}\Big)\Pi(I_s \otimes T)\right)^D \\ &\quad= \left((I_s \otimes T^{-1})\Pi^T\begin{bmatrix} (\Lambda_d \otimes I_s)(I_d \otimes I_s - J_d \otimes \tau\mathcal{A}) & 0 \\ 0 & 0 \end{bmatrix}\Pi(I_s \otimes T)\right)^D.\end{aligned}$$

Applying Lemma 4.5, this can be written as

$$\begin{aligned}&\left((I_s\otimes T^{-1})\Pi^T\begin{bmatrix}(\Lambda_d\otimes I_s) & (I_d\otimes I_s - J_d\otimes\tau\mathcal{A}) & 0\\ & 0 & 0\end{bmatrix}\Pi(I_s\otimes T)\right)^D\\ &\quad = (I_s\otimes T^{-1})\Pi^T\begin{bmatrix}(\Lambda_d\otimes I_s) & (I_d\otimes I_s - J_d\otimes\tau\mathcal{A}) & 0\\ & 0 & 0\end{bmatrix}^D\Pi(I_s\otimes T). \qquad (11)\end{aligned}$$

Since $\Lambda_d\otimes I_s$ and $I_d\otimes I_s - J_d\otimes\tau\mathcal{A}$ are both nonsingular, the Drazin inverse can be written as

$$\begin{aligned}\begin{bmatrix}(\Lambda_d\otimes I_s) & (I_d\otimes I_s - J_d\otimes\tau\mathcal{A}) & 0\\ & 0 & 0\end{bmatrix}^D &= \begin{bmatrix}\Big((\Lambda_d\otimes I_s)(I_d\otimes I_s - J_d\otimes\tau\mathcal{A})\Big)^{-1} & 0\\ & 0\end{bmatrix}\\ &= \begin{bmatrix}\Lambda_d^{-1}\otimes I_s & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}I_d\otimes I_s - J_d\otimes\tau\mathcal{A} & 0\\ & 0\end{bmatrix}^D,\end{aligned}$$

which can easily be seen by checking the properties of the Drazin inverse. Inserting this into (11), we thus get

$$\begin{aligned}&\Big((I_s\otimes\hat{E})(I_s\otimes\hat{E}^D\hat{E} - \tau\mathcal{A}\otimes\hat{E}^D\hat{A})\Big)^D\\ &\quad = (I_s\otimes T^{-1})\Pi^T\begin{bmatrix}\Lambda_d^{-1}\otimes I_s & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}I_d\otimes I_s - J_d\otimes\tau\mathcal{A} & 0\\ & 0\end{bmatrix}^D\Pi(I_s\otimes T).\end{aligned}$$

Since

$$\begin{aligned}(I_s\otimes T^{-1})\Pi^T\begin{bmatrix}\Lambda_d^{-1}\otimes I_s & 0\\ 0 & 0\end{bmatrix}\Pi(I_s\otimes T) &= (I_s\otimes T^{-1})\begin{bmatrix}I_s\otimes\Lambda_d^{-1} & 0\\ 0 & 0\end{bmatrix}(I_s\otimes T)\\ &= I_s\otimes\Big(T^{-1}\begin{bmatrix}\Lambda_d^{-1} & 0\\ 0 & 0\end{bmatrix}T\Big) = I_n\otimes\hat{E}^D,\end{aligned}$$

and

$$\begin{aligned}&(I_s\otimes T^{-1})\Pi^T\begin{bmatrix}I_d\otimes I_s - J_d\otimes\tau\mathcal{A} & 0\\ & 0\end{bmatrix}^D\Pi(I_s\otimes T)\\ &\quad = \Big((I_s\otimes T^{-1})\Pi^T\begin{bmatrix}I_d\otimes I_s - J_d\otimes\tau\mathcal{A} & 0\\ & 0\end{bmatrix}\Pi(I_s\otimes T)\Big)^D\\ &\quad = \Big((I_s\otimes T^{-1})\begin{bmatrix}I_s\otimes I_d & 0\\ 0 & 0\end{bmatrix} - \begin{bmatrix}\tau\mathcal{A}\otimes J_d & 0\\ 0 & 0\end{bmatrix}(I_s\otimes T)\Big)^D\\ &\quad = \Big(I_n\otimes T^{-1}\begin{bmatrix}I_d & 0\\ 0 & 0\end{bmatrix}T - \tau\mathcal{A}\otimes T^{-1}\begin{bmatrix}J_d & 0\\ 0 & 0\end{bmatrix}T\Big)^D\\ &\quad = \Big(I_s\otimes\hat{E}^D\hat{E} - \tau\mathcal{A}\otimes\hat{E}^D\hat{A}\Big)^D,\end{aligned}$$

we finally obtain that

$$\Big((I_s\otimes\hat{E})(I_s\otimes\hat{E}^D\hat{E} - \tau\mathcal{A}\otimes\hat{E}^D\hat{A})\Big)^D = I_n\otimes\hat{E}^D\Big(I_s\otimes\hat{E}^D\hat{E} - \tau\mathcal{A}\otimes\hat{E}^D\hat{A}\Big)^D.$$

Note that in general for $E, C\in\mathbb{R}^{n\times n}$ it is not true that $(EC)^D = E^DC^D$, see [37], but in our particular case, we obtain this by exploiting the block diagonal structure given for $\hat{E}$ and $\hat{A}$ by (9).

2. With

$$\hat{E}^D\hat{E}(\alpha_k\hat{E} - \beta_k\tau\hat{A})^{-1} = \frac{1}{\alpha_k}\hat{E}^D\hat{E}(\hat{E} - \tau\frac{\beta_k}{\alpha_k}\hat{A})^{-1},$$

the assertion immediately follows from part 1. with $s = 1$ and $\mathcal{A} := \frac{\beta_k}{\alpha_k}$. □

As for ODEs, we now restrict the class of considered problems regarding the eigenvalues and introduce the notion of M-matrix pairs.

Definition 4.2. *A matrix pair (E, A) in $\mathbb{R}^{n\times n}$ is called* M-matrix pair*, if (E, A) is a Z-matrix pair and $\max_{\lambda\in\sigma_{fin}(E,A)} |\mu - \lambda| \leq \mu$. Moreover, (E, A) in $\mathbb{R}^{n\times n}$ is called* -M-matrix pair*, if $(E, -A)$ is an M-matrix pair.*

If the inequality on the finite eigenvalues in Definition 4.2 is strict, then (E, A) is called a *strict M-matrix pair* and we can generalize the property of a nonnegative inverse in terms of the Drazin inverse.

Lemma 4.7. *If (E, A) is a regular, commuting, strict M-matrix pair in $\mathbb{R}^{n\times n}$, then $(E^DA)^D \geq 0$.*

Proof. If (E, A) is a strict M-matrix pair, then there exists $\mu > 0$, such that $\mu E^DE - E^DA \geq 0$ and $\max_{\lambda\in\sigma_{fin}(E,A)} |\mu - \lambda| < \mu$. This means that the matrix $B := \mu E^DE - E^DA$ is nonnegative and satisfies $\rho(B) < \mu$, such that for $\mu \neq 0$ we can expand $(\mu I_n - B)^{-1}$ into a nonnegative Neumann series, i. e., $(\mu I_n - B)^{-1} = \sum_{k=0}^{\infty} \frac{1}{\mu^{k+1}} B^k \geq 0$. With $E^DE \geq 0$, the projection $E^DE(\mu I_n - B)^{-1}$ is nonnegative as well and, since E^DE is invariant under the Drazin inverse, we obtain

$$0 \leq (E^DE)^D\big(\mu I_n - (\mu E^DE - E^DA)\big)^{-1} = \big(E^DE(\mu I_n - \mu E^DE + E^DA)\big)^D = \mu\big(E^DA\big)^D.$$

□

As for single matrices, we can shift a -M-matrix pair to an M-matrix pair.

Lemma 4.8. *If (E, A) is -M-matrix pair with $\mu > 0$, then $(E, \kappa E - A)$ is a strict M-matrix pair with $\mu + \kappa$ for any $\kappa > 0$.*

Proof. If (E, A) is a -M-matrix pair, then there exists $\mu > 0$, such that $\kappa E^DE - E^DA \leq \mu E^DE + \kappa E^DE$ holds for arbitrary $\kappa > 0$, i. e., $E^D(\kappa E - A) \leq (\kappa + \mu)E^DE$. Furthermore, we have $\max_{\lambda\in\sigma_{fin}(E,A)} |\mu+\lambda| \leq \mu$, such that $\max_{\lambda\in\sigma_{fin}(E,A)} |\mu+\lambda| < \mu+\kappa$ holds for every $\kappa > 0$. Hence, $(E, \kappa E - A)$ is a strict M-matrix pair with $\mu + \kappa > 0$ for every $\kappa > 0$. □

With these results, we can present conditions for the positivity of the differential part of a Runge-Kutta or multistep discretization corresponding to the discretization of $\dot{x}_d = E^DAx_d + E^Df(t)$ with initial condition $x_d(t_0) = E^DEv_0$.

Theorem 4.1. *Let $E\dot{x} = Ax + f$ be a positive DAE with regular matrix pair (E, A), with $\nu = ind(E, A)$ and let $f \in \mathcal{C}^\nu(\mathbb{R}, \mathbb{R}^n)$. If (E, A) is a -M-matrix pair with $\mu > 0$, then the following assertions hold:*

1. *A Runge-Kutta method with coefficients $(\mathcal{A}, \beta, \gamma)$, $\mathcal{A}$ nonsingular and stability function $\mathcal{R}_d(z) := 1 + z\beta^T(I_s - z\mathcal{A})^{-1}\mathbf{1}$ that is absolutely monotonic on $[-\gamma_+, 0]$ for $\gamma_+ \geq 0$, is positive for $E^DE\dot{x} = E^DAx$ if the stepsize satisfies $0 < \tau \leq \frac{\gamma_+}{\mu}$.*

   Moreover, if additionally $\mathcal{Q}_{d,i}(z) := \beta^T(I_s - z\mathcal{A})^{-1}e_i$, is absolutely monotonic on $[-\gamma_+, 0]$ for $\gamma_+ \geq 0$, $i = 1, ..., s$ then the method is positive for $E^DE\dot{x} = E^DAx + E^Df$ provided that $0 < \tau \leq \frac{\gamma_+}{\mu}$.

2. *A multistep method with coefficients (α, β), $\alpha_k, \beta_k \neq 0$, for which the stability functions $r_{d,j}(z) := -\frac{\alpha_j - \beta_j z}{\alpha_k - \beta_k z}$, $j = 0, ..., k-1$, are absolutely monotonic on $[-\gamma_+, 0]$ for $\gamma_+ \geq 0$, is positive for $E^DE\dot{x} = E^DAx$ if the stepsize satisfies $0 < \tau \leq \frac{\gamma_+}{\mu}$.*

   Moreover, if additionally $\frac{\beta_j}{\alpha_k}, \frac{\beta_k}{\alpha_k}, -\frac{\beta_j}{\beta_k} \geq 0$ holds for $j = 0, ..., k-1$, then the method is positive for $E^DE\dot{x} = E^DAx + E^Df$ provided that $0 < \tau \leq \frac{\gamma_+}{\mu}$.

Proof. As in the proof of Theorem 3.1, we consider the iteration matrices as rational functions and show their nonnegativity via Taylor expansion. Due to the singularity of E, we have to formulate these arguments in terms of the Drazin inverse.

1. As for the exact solution, the differential part of a Runge-Kutta discretization of $E\dot{x} = Ax+f$ is obtained by the projection of (4) by E^DE, i. e., by

$$E^DEx_{N+1} = E^DE\mathcal{R}(E,\tau A)x_N + \sum_{i=0}^{s} E^DE\mathcal{Q}_i(E,\tau A)f(t_N+\tau\gamma_i),$$

where

$$\mathcal{R}(E,\tau A) \;:=\; I_n + (\tau\beta^T\otimes I_n)(I_s\otimes E - \tau\mathcal{A}\otimes A)^{-1}(\mathbf{1}\otimes A) \tag{12}$$

and

$$\mathcal{Q}_i(E,\tau A) \;:=\; (\tau\beta^T\otimes I_n)(I_s\otimes E - \tau\mathcal{A}\otimes A)^{-1}(e_i\otimes I_n). \tag{13}$$

Using Lemma 4.6 and the identity $\hat{E}^D\hat{E}\mathcal{R}(\hat{E},\tau\hat{A}) = E^DE\mathcal{R}(E,\tau A)$, the projection of $\mathcal{R}(E,\tau A)$ can be written as

$$\begin{aligned} E^DE\mathcal{R}(E,\tau A) &= E^DE + (\tau\beta^T\otimes A)(I_s\otimes E^DE)\,(I_s\otimes E - \tau\mathcal{A}\otimes A)^{-1}(\mathbf{1}\otimes E^DE) \\ &= E^DE + (\tau\beta^T\otimes E^DA)\big(I_s\otimes E^DE - \tau\mathcal{A}\otimes E^DA\big)^D(\mathbf{1}\otimes E^DE). \end{aligned}$$

Note that this term corresponds to the iteration matrix of a Runge-Kutta method applied to the ODE $\dot{x}_d = E^DAx_d + E^Df$ with initial condition $x_d(t_0) = E^DEx_0$. Therefore, the projected iteration matrix is a rational function that is evaluated in τE^DA. Denoting this function by $\mathcal{R}_d(\tau E^DA) := E^DE\mathcal{R}(E,\tau A)$ and defining the matrix $B := \mu E^DE + E^DA$, we can expand this function $\mathcal{R}_d(\tau E^DA)$ into a Taylor series centered in $-\tau\mu E^DE$, i. e.,

$$\mathcal{R}_d(\tau E^DA) \;=\; \sum_{k=0}^{\infty}\frac{1}{k!}\mathcal{R}_d^{(k)}(-\tau\mu E^DE)\,(\tau B)^k. \tag{14}$$

As for ODEs, the expansion (14) corresponds to a matrix power series with scalar coefficients, since $\mathcal{R}(-\tau\mu E^DE)$ can be reduced to a scalar function scaling the projection E^DE, i. e.,

$$\begin{aligned} \mathcal{R}_d(-\tau\mu E^DE) &= E^DE - (\tau\mu\beta^T\otimes E^DE)\Big(I_s\otimes E^DE + \tau\mu\mathcal{A}\otimes E^DE\Big)^D(\mathbf{1}\otimes I_n) \\ &= E^DE - (\tau\mu\beta^T\otimes E^DE)\Big((I_s+\tau\mu\mathcal{A})^{-1}\otimes\big(E^DE\big)^D\Big)\,(\mathbf{1}\otimes I_n) \\ &= (1-\tau\mu\beta^T(I_s+\tau\mu\mathcal{A})^{-1}e_s)\otimes E^DE \\ &= \mathcal{R}_d(-\tau\mu)\;E^DE. \end{aligned}$$

Inserting this into (14) and noting that $E^DEB = E^DE(\mu E^DE + E^DA) = B$, we obtain the power series

$$\sum_{k=0}^{\infty}\frac{1}{k!}\mathcal{R}_d^{(k)}(-\tau\mu E^DE)\,(\tau B)^k = \sum_{k=0}^{\infty}\frac{1}{k!}\mathcal{R}_d^{(k)}(-\tau\mu)\;E^DE\,(\tau B)^k = \sum_{k=0}^{\infty}\frac{1}{k!}\mathcal{R}_d^{(k)}(-\tau\mu)\,(\tau B)^k.$$

This series converges if $\rho(\tau B) < r$, where $r \geq 0$ denotes the radius of convergence of the scalar series $\sum_{k=0}^{\infty}\frac{\mathcal{R}_d^{(k)}(-\tau\mu)}{k!}\xi^k$, $\xi\in\mathbb{C}$. But from the -M-matrix pair property of (E,A), we know that $\rho(B) = \max_{\lambda\in\sigma_{fin}(E,A)}|\mu+\lambda| \;\leq\; \mu$, i. e., $\rho(\tau B) < r$ holds if $\tau\mu \leq \gamma_+$ and $\mathcal{R}$ is absolutely monotonic on $[-\gamma_+, 0]$. Since B is nonnegative if (E,A) is an -M-matrix pair, and $\mathcal{R}_d^{(k)}(-\tau\mu)\geq 0$ holds by the absolutely monotonicity assumption, we thus have shown that $\mathcal{R}_d(\tau E^DA)\geq 0$ if $\tau\mu\leq\gamma_+$. The nonnegativity of the inhomogeneity is proved in a similar way by considering

$E^D E\mathcal{Q}_i(E,\tau A)$ for each $i = 1,...,s$ as rational function in $\tau E^D A$. We define $\mathcal{Q}_{d,i}(\tau E^D A) := E^D E\mathcal{Q}_i(E,\tau A)$, where

$$E^D E\mathcal{Q}_i(E,\tau A) \quad = \quad (\beta^T \otimes E^D)(I_s \otimes E^D - \tau\mathcal{A} \otimes E^D A)^D (e_i \otimes E^D E)$$

is obtained in the same way as $E^D E\mathcal{R}(E,\tau A)$, using Lemma 4.6 and the identity $\hat{E}^D \hat{E}\mathcal{Q}_i(\hat{E},\tau\hat{A}) = E^D E\mathcal{Q}_i(E,\tau A)$. As before, this corresponds to the inhomogeneity of a Runge-Kutta method applied to $\dot{x}_d = E^D A x_d + E^D f$. The expansion of $\mathcal{Q}_{d,i}(\tau E^D A)$ in $-\tau\mu E^D E$ is given by

$$\mathcal{Q}_{d,i}(\tau E^D A) = \sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{Q}_{d,i}^{(k)}(-\tau\mu E^D E)(\tau B)^k = \sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{Q}_{d,i}^{(k)}(-\tau\mu)(\tau B)^k E^D,$$

since $\mathcal{Q}_{d,i}(-\tau\mu E^D E)$ is reduced to a scalar function scaling the Drazin inverse E^D, i. e.,

$$\begin{aligned}
\mathcal{Q}_i(-\tau\mu E^D E) &= (\beta^T \otimes E^D)(I_s \otimes E^D E - \tau\mathcal{A} \otimes E^D E)^D (e_i \otimes E^D E) \\
&= E^D(\beta^T \otimes E^D E)\big((I_s + \tau\mu\mathcal{A}) \otimes E^D E\big)^D (e_i \otimes E^D E) \\
&= E^D(\beta^T (I_s + \tau\mu\mathcal{A})^{-1} e_i) \otimes E^D E \\
&= \mathcal{Q}_i(-\tau\mu) E^D,
\end{aligned}$$

where we have used that $(\beta^T \otimes E^D) = (1 \otimes E^D)(\beta^T \otimes I_n) = E^D(\beta^T \otimes I_n)$ and $E^D E E^D = E^D$. The convergence and nonnegativity of the power series $\sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{Q}_{d,i}^{(k)}(-\tau\mu)(\tau B)^k$ is shown in the same way as for $\mathcal{R}_d(\tau E^D A)$ via the absolute monotonicity of $\mathcal{Q}_{d,i}(z)$, $i = 1,...,s$. Thus, the inhomogeneity of the Runge-Kutta iteration is given by

$$\sum_{i=0}^{s} E^D E\mathcal{Q}_i(E,\tau A) f(t_N + \tau\gamma_i) \quad = \quad \sum_{i=0}^{s} \Big(\sum_{k=0}^{\infty} \frac{1}{k!} \mathcal{Q}_{d,i}^{(k)}(-\tau\mu)(\tau B)^k \Big) E^D f(t_N + \tau\gamma_i)$$

and since $E^D f(t)$ is nonnegative for $t \geq 0$ for a positive DAE, this is nonnegative for every $t_n \geq 0$ as well.

2. For a multistep method, the differential part of the iteration (5) is given by

$$E^D E x_N = \sum_{j=0}^{k-1} E^D E r_j(E,\tau A) x_{N-k+j} + \sum_{j=0}^{k} E^D E q_j(E,\tau A) f(t_{N-k+j}),$$

with $r_j(E,\tau A) = -(\alpha_k E - \beta_k \tau A)^{-1}(\alpha_j E - \beta_j \tau A)$, and $q_j(E,\tau A) := \tau\beta_j(\alpha_k E - \beta_k \tau A)\big)^{-1}$ for $j = 0,...,k-1$.

As before, using Lemma 4.6 and the identity $\hat{E}^D \hat{E} r_j(\hat{E},\tau\hat{A}) = E^D E r(E,\tau A)$, the projection of $r_j(E,\tau A)$ onto $\mathrm{im}(E^D E)$ can be written as

$$E^D E r_j(E,\tau A) = -\big(\alpha_j E^D E - \beta_j \tau E^D A\big)\big(\alpha_k E^D E - \beta_k \tau E^D A\big)^D$$

and again, this corresponds to the multistep discretization of $E^D E\dot{x} = E^D A x + E^D f$. Considering this as rational function $r_{d,j}(\tau E^D A) := E^D E r_j(E,\tau A)$ and expanding it in $-\tau\mu E^D E$, we obtain

$$E^D E r_j(E,\tau A) = \sum_{l=0}^{\infty} \frac{1}{l!} r_{d,j}^{(l)}(-\tau\mu E^D E)\ (\tau B)^l,$$

where $B = \mu E^D E + E^D A$. Using the identity

$$r_j(-\tau\mu E^D E) = -\frac{\alpha_j - \beta_j \tau\mu}{\alpha_k - \beta_k \tau\mu} E^D E,$$

this expansion corresponds to the matrix power series $\sum_{l=0}^{\infty} \frac{r_j^{(l)}(-\tau\mu)}{l!}(\tau B)^l$ and thus, as in the case of Runge-Kutta methods, the convergence and non-negativity of this series is implied by the

absolute monotonicity of $r_{d,j}$ on $[-\gamma_+, 0]$. For $q_{d,j}(E, \tau A)$, we note that the projected inverse can be written as

$$E^D E(\alpha_k E - \beta_k \tau A)^{-1} = \left(E^D (E - \frac{\beta_k}{\alpha_k}\tau A)\right)^D E^D$$

and this is nonnegative for $\tau > 0$ if $\frac{\beta_k}{\alpha_k} > 0$, because (E, A) is an M-matrix pair and thus $(E, E - \frac{\beta_k}{\alpha_k}\tau A)$ is a strict M-matrix pair that leads to a nonnegative Drazin inverse, see Lemma 4.7. Since $E^D f(t)$ is nonnegative for $t \geq 0$ by the positivity assumption, we thus get that $E^D E q_j(E, \tau A) f(t_{N-k+j})$ is nonnegative if $\frac{\beta_j}{\alpha_k} \geq 0$. □

If E is nonsingular, then the conditions of Theorem 4.1 correspond to those given in Theorem 3.1.

In Table 1 the absolute monotonicity radius of some commonly used Runge-Kutta and multistep methods is presented. The values for the Radau and Lobatto methods are taken from [16], where the absolute monotonicity radius is computed for general Padé approximations. In [16], it is also shown that for even stage numbers $s = 2l$, $l \in \mathbb{N}$, the absolute monotonicity radius is zero and $\gamma_+ = \infty$ is achieved only for methods of convergence order $p = 1$, see also [6]. Note that by the assumptions of Theorem 4.1, we are restricted to implicit methods.

To explain the failure of BDF methods in preserving positivity, we note that their stability functions are given by $r_j(z) = -\frac{\alpha_j}{\alpha_k - z}$. Since $\alpha_k > 0$ for $k = 1, ..., 6$, the denominator is nonnegative for any $z \leq 0$ and the sign of r_j in $\mathbb{R}^-$ is determined by the corresponding coefficient α_j. But, since for a given step number k, the α_j have an alternating sign pattern, it follows that $\gamma_+ = 0$ for every BDF method with $k \geq 2$.

In [16] and [31] the optimal values, i. e., the maximal admissible monotonicity radii for one- and multistep methods of a given order of convergence, are presented in Table 2.

	$s=1$	$s=2$	$s=3$	$s=4$	$s=5$	$s=6$
Radau IIA	$\gamma_+ = \infty$	$\gamma_+ = 0$	$\gamma_+ = 1.7034$	$\gamma_+ = 0$	$\gamma_+ = 1.7940$	$\gamma_+ = 0$
Lobatto IIIC		$\gamma_+ = 0$	$\gamma_+ = 1.1954$	$\gamma_+ = 0$	$\gamma_+ = 1.4242$	$\gamma_+ = 0$
	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
BDF	$\gamma_+ = 0$	$\gamma_+ = 0$	$\gamma_+ = 0$	$\gamma_+ = 0$	$\gamma_+ = 0$	$\gamma_+ = 0$

Table 1: Absolute monotonicity radius of Radau IIA, Lobatto IIIC and BDF methods with stage number s and step number k.

Further references on this topic are [16, 18, 26, 36] and [30, 31], where the absolute monotonicity radius is discussed in the context of contractivity preserving one- and multistep methods, respectively.

4.5 Positivity preservation for the algebraic part

Having characterized the positivity of the differential part of a DAE discretization, we now turn to the algebraic components. Here, the approach is to find conditions under which the considered method overestimates the exact solution, such that $P_\infty x_N \geq P_\infty x(t_N)$ holds for every $t_N \geq t_0$. Since $P_\infty x(t_N)$ is nonnegative if the DAE is positive, this implies that $P_\infty x_N \geq 0$. This estimate certainly holds, if the local discretization error is nonnegative in every step, so we will first compute an explicit expression of the local discretization error and then analyze conditions under which this error is nonnegative. This requires the following results.

Lemma 4.9 (See e. g. [38]). *Let $B \in \mathbb{R}^{n\times n}$, then $(I_n - B)^{-1}$ exists and $(I_n - B)^{-1} = \sum_{l=0}^{\infty} B^l$ if and only if $\rho(B) < 1$.*

Lemma 4.10 (See e. g. [27]). *Let (E, A) be a regular, commuting matrix pair in $\mathbb{R}^{n\times n}$, then $P_\infty A^D A = P_\infty$.*

Optimal onestep method	$s=3$	$s=1$	$s=2$	$s=3$	$s=4$
$p=2$				$\gamma_+ = 0.703$	
$p=3$	$\gamma_+ = 0.703$	$\gamma_+ = 3.6085$	$\gamma_+ = 2.732$	$\gamma_+ = 1.2906$	$\gamma_+ = 0.7035$

Optimal multistep method	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$
$p=2$	$\gamma_+ = 2$	$\gamma_+ = 2$	$\gamma_+ = 2$	$\gamma_+ = 2$	$\gamma_+ = 2$
$p=3$	$\gamma_+ = 1.225$	$\gamma_+ = 1.1.572$	$\gamma_+ = 1.703$	$\gamma_+ = 1.772$	$\gamma_+ = 1.815$
$p=4$	$\gamma_+ = 1.2432$	$\gamma_+ = 1.2432$	$\gamma_+ = 1.2432$	$\gamma_+ = 1.2432$	$\gamma_+ = 1.2432$
$p=6$	$\gamma_+ = 0.9053$	$\gamma_+ = 0.9053$	$\gamma_+ = 0.9053$	$\gamma_+ = 0.9053$	$\gamma_+ = 0.9053$

Table 2: Optimal absolute monotonicity radius of one- and multistep methods of convergence order p.

With this, we can expand the inverses occurring in Runge-Kutta and multistep iterations. For brevity, we set $P_N := P_\infty A^D E$ and $\hat{P_N} := P_\infty \hat{A}^D \hat{E}$.

Lemma 4.11. *Let (E,A) be a regular matrix pair in $\mathbb{R}^{n\times n}$ with ind$(E,A)=\nu$ and let $\hat{E},\hat{A}$ be given by* (6).

1. Let $(\mathcal{A},\beta,\gamma)$ be the coefficients of a Runge-Kutta method with $\mathcal{A}\in\mathbb{R}^{s\times s}$ nonsingular, then

$$(I_s\otimes P_\infty)\big(I_s\otimes\hat{E}-\tau\mathcal{A}\otimes\hat{A}\big)^{-1} = -\sum_{\ell=0}^{\nu-1}(\tau\mathcal{A})^{-(\ell+1)}\otimes \hat{P_N}^{\ell}\hat{A}^D.$$

2. Let (α_k,β_k) be the coefficients of a multistep method with $\alpha_k,\beta_k\neq 0$, then

$$P_\infty(\alpha_k\hat{E}-\beta_k\tau\hat{A})^{-1} = -\frac{1}{\alpha_k}\sum_{\ell=0}^{\nu-1}\big(\frac{\alpha_k}{\tau\beta_k}\big)^{\ell+1}\hat{P_N}^{\ell}\hat{A}^D.$$

Proof. 1. As in the proof of Theorem 4.1, we rewrite the projected inverse in terms of the Drazin inverse. Using the identities $P_\infty^D = P_\infty$, Lemma 4.3, Lemma 4.10, and $P_\infty\hat{A}^D\hat{A}=P_\infty$, we obtain that

$$\begin{aligned}(I_s\otimes P_\infty)\big(I_s\otimes\hat{E}-\tau\mathcal{A}\otimes\hat{A}\big)^{-1} &= \big((I_s\otimes P_\infty\hat{A}^D\hat{A})(I_s\otimes\hat{E}-\tau\mathcal{A}\otimes\hat{A})\big)^D\\ &= \big((I_s\otimes P_\infty\hat{A})\big(I_s\otimes P_N-\tau\mathcal{A}\otimes P_\infty\big)\big)^D\\ &= (I_s\otimes P_\infty\hat{A}^D)(I_s\otimes P_N-\tau\mathcal{A}\otimes P_\infty)^D. \qquad (15)\end{aligned}$$

To filter out those components of $(I_s\otimes\hat{P_N}-\tau\mathcal{A}\otimes P_\infty)^D$ lying in the image of P_∞, we again use the decomposition of $\hat{E},\hat{A}$ into finite and infinite eigenspaces. By (6), it holds that $P_\infty = T^{-1}\left[\begin{smallmatrix}0&0\\0&I_a\end{smallmatrix}\right]T$ and $\hat{P_N} = T^{-1}\left[\begin{smallmatrix}0&0\\0&N\end{smallmatrix}\right]T$, and using Lemma 4.4, this implies

$$\begin{aligned}(I_s\otimes P_N-\tau\mathcal{A}\otimes P_\infty)^D &= \Big(I_s\otimes T^{-1}\begin{bmatrix}0&0\\0&N\end{bmatrix}T-\tau\mathcal{A}\otimes T^{-1}\begin{bmatrix}0&0\\0&I_a\end{bmatrix}T\Big)^D\\ &= (I_s\otimes T^{-1})\Big(I_s\otimes\begin{bmatrix}0&0\\0&N_a\end{bmatrix}-\tau\mathcal{A}\otimes\begin{bmatrix}0&0\\0&I_a\end{bmatrix}\Big)^D(I_s\otimes T).\end{aligned}$$

To separate the differential and algebraic components, we apply the perfect shuffle permutation Π, and we obtain

$$\begin{aligned}\Big(I_s\otimes P_N-\tau\mathcal{A}\otimes P_\infty\Big)^D &= (I_s\otimes T^{-1})\Big(\Pi^T\Big(\begin{bmatrix}0&0\\0&N_a\end{bmatrix}\otimes I_s-\begin{bmatrix}0&0\\0&I_a\end{bmatrix}\otimes\tau\mathcal{A}\Big)\Pi\Big)^D(I_s\otimes T)\\ &= (I_s\otimes T^{-1})\;\Pi^T\begin{bmatrix}0&\\&\big(N_a\otimes I_s-I_a\otimes\tau\mathcal{A}\big)^D\end{bmatrix}\Pi(I_s\otimes T). \qquad (16)\end{aligned}$$

We then apply Lemma 4.9 and expand $\big(N_a\otimes I_s-I_a\otimes\tau\mathcal{A}\big)^D$ in a power series. Since $\mathcal{A}$ is nonsingular, this can be written as

$$\big(N_a\otimes I_s-I_a\otimes\tau\mathcal{A}\big)^D = -(I_n\otimes(\tau\mathcal{A})^{-1})\big(I_a\otimes I_n-N_a\otimes(\tau\mathcal{A})^{-1}\big)^D.$$

The eigenvalues of $N_a \otimes (\tau\mathcal{A})^{-1}$ are given by $\frac{\eta_\vartheta}{\tau\sigma_i}$, $\vartheta = 1, ..., a$, $i = 1, ..., s$, where η_ϑ and σ_l denote the eigenvalues of N and $\mathcal{A}$, respectively. But since N_a is nilpotent, i. e., $\eta_\vartheta = 0$, $\vartheta = 1, ..., a$, it follows that $\rho(N_a \otimes (\tau\mathcal{A})^{-1}) = 0$. By Lemma 4.9, this implies that $I_a \otimes I_n - N_a \otimes (\tau\mathcal{A})^{-1}$ is nonsingular and we get

$$\big(I_a \otimes I_n - N_a \otimes (\tau\mathcal{A})^{-1}\big)^{-1} = \sum_{\ell=0}^{\nu-1} N_a^\ell \otimes (\tau\mathcal{A})^{-\ell}$$

by the nilpotency of N_a. Thus, we have

$$\big(N_a \otimes I_s - I_a \otimes \tau\mathcal{A}\big)^D \;=\; (I_n \otimes (\tau\mathcal{A})^{-1}) \sum_{\ell=0}^{\nu-1} N_a^\ell \otimes (\tau\mathcal{A})^{-\ell} = \sum_{\ell=0}^{\nu-1} N_a^\ell \otimes (\tau\mathcal{A})^{-(\ell+1)}$$

and inserting this into (16) we get

$$\begin{aligned}
\Big(I_s \otimes P_N - \tau\mathcal{A} \otimes P_\infty\Big)^D &= -\sum_{\ell=0}^{\nu-1} (I_s \otimes T^{-1})\Pi^T \Big(\begin{bmatrix} 0 & 0 \\ 0 & N_a^\ell \end{bmatrix} \otimes \begin{bmatrix} 0 & 0 \\ 0 & (\tau\mathcal{A})^{-(\ell+1)} \end{bmatrix} \Big) \Pi (I_s \otimes T) \\
&= -\sum_{\ell=0}^{\nu-1} (I_s \otimes T^{-1}) \Big(\begin{bmatrix} 0 & 0 \\ 0 & (\tau\mathcal{A})^{-(\ell+1)} \end{bmatrix} \otimes \begin{bmatrix} 0 & 0 \\ 0 & N_a^\ell \end{bmatrix} \Big) (I_s \otimes T). \\
&= -\sum_{\ell=0}^{\nu-1} (\tau\mathcal{A})^{-(\ell+1)} \otimes T^{-1} \begin{bmatrix} 0 & 0 \\ 0 & N_a^\ell \end{bmatrix} T.
\end{aligned}$$

Since $T^{-1}\left[\begin{smallmatrix} 0 & 0 \\ 0 & N_a^\ell \end{smallmatrix}\right] T = T^{-1}\left[\begin{smallmatrix} 0 & 0 \\ 0 & N_a \end{smallmatrix}\right] T^\ell T^{-\ell} \left[\begin{smallmatrix} 0 & 0 \\ 0 & N_a \end{smallmatrix}\right] T = P_N^\ell$, it follows that

$$(I_s \otimes P_N - \tau\mathcal{A} \otimes P_\infty)^D = -\sum_{\ell=0}^{\nu-1} (\tau\mathcal{A})^{-(\ell+1)} \otimes P_N^\ell,$$

and inserting this into (15), we obtain the desired expansion, i. e.,

$$(I_s \otimes P_\infty)\big(I_s \otimes E - \tau\mathcal{A} \otimes A\big)^{-1} = -(I_s \otimes A^D)(I_s \otimes P_N - \tau\mathcal{A} \otimes P_\infty)^D = -\sum_{\ell=0}^{\nu-1} (\tau\mathcal{A})^{-(\ell+1)} \otimes P_N^\ell \hat{A}^D.$$

2. Using

$$\begin{aligned}
P_\infty(\alpha_k\hat{E} - \beta_k\tau\hat{A})^{-1} &= \big(P_\infty \hat{A}^D \hat{A}\big(\alpha_k\hat{E} - \beta_k\tau\hat{A}\big)\big)^D \\
&= \big(\hat{A}\big(\alpha_k P_\infty \hat{A}^D \hat{E} - \beta_k\tau P_\infty \hat{A}^D \hat{A}\big)\big)^D \\
&= \frac{1}{\alpha_k} P_\infty \hat{A}^D \big(\hat{P_N} - \tau\frac{\beta_k}{\alpha_k} P_\infty\big)^D, \qquad (17)
\end{aligned}$$

we see that (17) corresponds to (15) for $s = 1$ and $\mathcal{A} = \frac{\beta_k}{\alpha_k}$. Thus, we immediately obtain the expansion

$$P_\infty(\alpha_k\hat{E} - \beta_k\tau\hat{A})^{-1} = -\frac{1}{\alpha_k} \sum_{\ell=0}^{\nu-1} (\frac{\alpha_k}{\tau\beta_k})^{\ell+1} \hat{P_N}^\ell \hat{A}^D.$$

□

We make use of the following result.

Theorem 4.2 ([27]). *Consider the DAE $E\dot{x} = Ax + f$, where (E, A) is a regular, commuting matrix pair with ind$(E, A) = \nu$ and $f \in \mathcal{C}^\nu(\mathbb{R}, \mathbb{R}^n)$.*

1. *Let $(\mathcal{A},\beta,\gamma)$ be the coefficients of a Runge-Kutta method with nonsingular $\mathcal{A}$. If there exist $\kappa_0,...,\kappa_{\nu-1}\in\mathbb{N}$, such that $(\mathcal{A},\beta,\gamma)$ satisfy*

   (i) $\beta^T\mathcal{A}^{-m}\mathbf{1}=\frac{\beta^T\mathcal{A}^{-(\ell+1)}\gamma^{\ell+1-m}}{(\ell+1-m)!}$ *for* $m=1,...,\ell$,

   (ii) $\beta^T\mathcal{A}^{-(\ell+1)}\gamma^m=\frac{m!}{(m-\ell)!}$ *for* $m=\ell+1,...,\kappa_\ell$,

   for $\ell=0,...,\nu-1$, where $\gamma^m=[\gamma_1^m...\gamma_s^m]$, then the method is consistent and the local error satisfies

$$x(t_{N+1})-x_{N+1}=\mathcal{O}(\tau^{\kappa_0+1})+...+\mathcal{O}(\tau^{\kappa_{\nu-1}-\nu+2})$$

   provided that $x(t_N)=x_N$.

2. *Let (α,β) be the coefficients of a multistep method with $\alpha_k,\beta_k\neq 0$. If there exists $p\in\mathbb{N}$, such that (α,β) satisfy $\sum_{j=0}^k j^m\alpha_j=m\sum_{j=0}^k j^{m-1}\beta_j$ for $m=0,...,p$, then the method is consistent of order p, i. e., the local error satisfies*

$$x(t_{N+1})-x_{N+1}=\mathcal{O}(\tau^{p+1})$$

   provided that $x(t_{N-k+j})=x_{N-k+j}$ holds for $j=0,...,k-1$.

For the convergence analysis, we further assume that $p<\kappa_\ell$ and $|1-\beta^T\mathcal{A}^{-1}\mathbf{1}|<1$, cp. [27]. We use Theorem 4.2 and Lemma 4.11 to compute an explicit representation of the local error in the algebraic part.

For this we denote by $R_\kappa f_N(\tau)$ the remainder of the Taylor expansion $f(t_N+\tau)=\sum_{\imath=0}^{\kappa}\frac{\tau^\imath}{\imath!}f_N^{(\imath)}+R_\kappa\bar{f}_N(\tau)$ and write for the projected inhomogeneity $\bar{f}:=P_\infty A^D f$.

Lemma 4.12. *Consider a DAE $E\dot{x}=Ax+f$, where (E,A) is a regular, commuting matrix pair in $\mathbb{R}^{n\times n}$ with ind$(E,A)=\nu$ and $f\in\mathcal{C}^\nu(\mathbb{R},\mathbb{R}^n)$.*

1. *Let $(\mathcal{A},\beta,\gamma)$, with $\mathcal{A}$ nonsingular be the coefficient of a Runge-Kutta method that is consistent with exponents $\kappa_0,...,\kappa_{\nu-1}\in\mathbb{N}$. If $P_\infty x(t_N)=P_\infty x_N$ holds for some $N\in\mathbb{N}$, then the new approximation $P_\infty x_{N+1}$ is given by*

$$P_\infty x_{N+1}=-\sum_{\ell=0}^{\nu-1}P_N^\ell\sum_{\imath=0}^{\kappa_\ell-\ell}\frac{\tau^\imath}{\imath!}\bar{f}_N^{(\ell+\imath)}-\sum_{\ell=0}^{\nu-1}P_N^\ell\sum_{i=1}^{s}\frac{\beta^T\mathcal{A}^{-(\ell+1)}e_i}{\tau^\ell}R_{\kappa_\ell}\bar{f}_N(\tau\gamma_i)$$

   and the local error $P_\infty\epsilon_{loc}=P_\infty x_{N+1}-P_\infty x(t_{N+1})$ is given by

$$P_\infty\epsilon_{loc}=\Big(\sum_{i=1}^{s}\frac{\beta^T\mathcal{A}^{-(\ell+1)}e_i}{\tau^\ell}R_{\kappa_\ell}\bar{f}(\tau\gamma_i)-R_{\kappa_\ell-\ell}\bar{f}^{(\ell)}(\tau)\Big).$$

2. *Let (α,β), $\alpha_k,\beta_k\neq 0$, denote a multistep method that is consistent of order $p\in\mathbb{N}$. If $P_\infty x(t_{N-k+j})=P_\infty x_{N-k+j}$ holds for $j=0,...,k-1$ and some $N\in\mathbb{N}$, then the new approximation $P_\infty x_N$ is given by*

$$\begin{aligned}P_\infty x_N &= -\sum_{\ell=0}^{\nu-1}P_N^\ell\bar{f}_n^{(\ell)}\\ &- \frac{1}{\tau\beta_k}\sum_{\ell=0}^{\nu-1}P_N^\ell\sum_{\imath=0}^{\ell-1}(\frac{\alpha_k}{\tau\beta_k})^{\ell-\imath-1}\sum_{j=0}^{k}\Big(\alpha_jR_{p-\imath}\bar{f}_{N-k}^{(\imath)}(j\tau)-\tau\beta_jR_{p-\imath-1}\bar{f}_{N-k}^{(\imath+1)}(j\tau)\Big)\end{aligned}$$

   and the local error $P_\infty\epsilon_{loc}=P_\infty x_N-P_\infty x(t_N)$ satisfies

$$\epsilon_{loc}=-\frac{1}{\tau\beta_k}\sum_{\ell=0}^{\nu-1}P_N^\ell\sum_{\imath=0}^{\ell-1}(\frac{\alpha_k}{\tau\beta_k})^{\ell-\imath-1}\sum_{j=0}^{k}\Big(\alpha_jR_{p-\imath}\bar{f}_{N-k}^{(\imath)}(j\tau)-\tau\beta_jR_{p-\imath-1}\bar{f}_{N-k}^{(\imath+1)}(j\tau)\Big).$$

Proof. In order to apply Lemma 4.11, we consider the scaled DAE $\hat{E}\dot{x} = \hat{A} + \hat{f}$ with $\hat{E}, \hat{A}, \hat{f}$ as in (6)

1. The algebraic part of a Runge-Kutta discretization (4) is given by

$$P_\infty x_{N+1} = P_\infty \mathcal{R}(\hat{E}, \tau\hat{A}) x_N + \sum_{i=0}^{s} P_\infty \mathcal{Q}_i(\hat{E}, \tau\hat{A}) f_{N,i}.$$

Using the expansion provided by Lemma 4.11, the iteration matrix can be written as

$$\begin{aligned} P_\infty \mathcal{R}(\hat{E}, \tau\hat{A}) &= P_\infty + (\tau\beta^T \otimes I_n)(I_s \otimes P_\infty)(I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})^{-1}(\mathbf{1} \otimes \hat{A}) \\ &= P_\infty + (\tau\beta^T \otimes I_n)\big(- \sum_{\ell=0}^{\nu-1} (\tau\mathcal{A})^{-(\ell+1)} \otimes \hat{P_N}^{\ell} \hat{A}^D\big)(\mathbf{1} \otimes \hat{A}) \\ &= P_\infty - \sum_{\ell=0}^{\nu-1} \frac{\beta^T \mathcal{A}^{-(\ell+1)} \mathbf{1}}{\tau^\ell} \hat{P_N}^{\ell}, \end{aligned} \tag{18}$$

since $P_\infty \hat{P_N}^{\ell} \hat{A}^D \hat{A} = \hat{P_N}^{\ell}$. In the same manner, we obtain

$$\begin{aligned} P_\infty \mathcal{Q}_i(\hat{E}, \tau\hat{A}) &= (\tau\beta^T \otimes P_\infty)(I_s \otimes P_\infty)(I_s \otimes \hat{E} - \tau\mathcal{A} \otimes \hat{A})^{-1}(e_i \otimes P_\infty \hat{A}^D) \\ &= -\sum_{\ell=0}^{\nu-1} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} \hat{P_N}^{\ell} \hat{A}^D, \end{aligned}$$

such that the inhomogeneous part is given by

$$\sum_{i=0}^{s} P_\infty \mathcal{Q}_i(\hat{E}, \tau\hat{A}) f_{N,i} = -\sum_{\ell=0}^{\nu-1} \sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} \hat{P_N}^{\ell} \bar{f}_{N,i}. \tag{19}$$

If we insert (18) and (19) into the projected iteration and assume consistent initial values, i. e., $P_\infty x_N = -\sum_{m=0}^{\nu-1} \hat{P_N}^{m} \bar{f}_N^{(m)}$, then the new approximation is given by

$$P_\infty x_{N+1} = -\sum_{m=0}^{\nu-1} \hat{P_N}^{m} \bar{f}_N^{(m)} + \sum_{\ell,m=0}^{\nu-1} \frac{\beta^T \mathcal{A}^{-(\ell+1)} \mathbf{1}}{\tau^\ell} \hat{P_N}^{\ell+m} \bar{f}_N^{(m)} - \sum_{\ell=0}^{\nu-1} \sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} \hat{P_N}^{\ell} \bar{f}_{N,i}.$$

In the proof of Lemma 4.11 it was shown that $\hat{P_N}^{j} = 0$ for $j \geq \nu$, such that by reordering the sums in powers of P_N, we obtain

$$P_\infty x_{N+1} = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^{\ell} \Big(\bar{f}_N^{(\ell)} + \sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} \bar{f}_{N,i} - \sum_{m=0}^{\ell} \frac{\beta^T \mathcal{A}^{-(\ell+1-m)} \mathbf{1}}{\tau^{\ell-m}} \bar{f}_N^{(m)} \Big).$$

If f is sufficiently smooth, then we can expand $\bar{f}_{N,i} = \bar{f}(t_N + \gamma_i \tau)$ in a Taylor series centered in t_N, i. e.,

$$\begin{aligned} P_\infty x_{N+1} = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^{\ell} \Bigg(& \bar{f}_N^{(\ell)} + \sum_{\imath=0}^{\kappa_\ell} \frac{\beta^T \mathcal{A}^{-(\ell+1)} \gamma^\imath}{\imath! \tau^{\ell-\imath}} \bar{f}_N^{(\imath)} + \sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} R_{\kappa_\ell} \bar{f}_N(\tau\gamma_i) \\ & - \sum_{m=0}^{\ell} \frac{\beta^T \mathcal{A}^{-(\ell+1-m)} \mathbf{1}}{\tau^{\ell-m}} \bar{f}_N^{(m)} \Bigg), \end{aligned}$$

where again $\gamma := \sum_{i=1}^{s} \gamma_i e_i$. Since the Runge-Kutta method with coefficients $(\mathcal{A}, \beta, \gamma)$ is consistent with exponents $\kappa_0, ..., \kappa_{\nu-1}$, it follows that $\beta^T \mathcal{A}^{-(\ell+1-m)} \mathbf{1} = \frac{\beta^T \mathcal{A}^{-(\ell+1)} \gamma^l}{l!}$ for $m = 1, ..., \ell$ and we can combine the first and the last sum, i. e.,

$$P_\infty x_{N+1} = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^{\ell} \Big(\bar{f}_N^{(\ell)} + \sum_{\imath=\ell+1}^{\kappa_\ell} \frac{\beta^T \mathcal{A}^{-(\ell+1)} \gamma^\imath}{\imath! \tau^{\ell-\imath}} \bar{f}_N^{(\imath)} + \sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} R_{\kappa_\ell} \bar{f}_N(\tau\gamma_i) \Big).$$

By assumption (ii) of Theorem 4.2, the coefficients $(\mathcal{A}, \beta, \gamma)$ furthermore satisfy $\beta^T \mathcal{A}^{-(\ell+1)} \gamma^\imath = \frac{\imath!}{(\imath-\ell)!}$ for $\imath = \ell+1, ..., \kappa_\ell$, which implies that

$$P_\infty x_{N+1} = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^\ell \Big(\bar{f}_N^{(\ell)} + \sum_{\imath=\ell+1}^{\kappa_\ell} \frac{\tau^{\imath-\ell}}{(\imath-\ell)!} \bar{f}_N^{(\imath)} + \sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} R_{\kappa_\ell} \bar{f}_N(\tau\gamma_i) \Big).$$

Shifting the summation index, the new approximation is thus given by

$$P_\infty x_{N+1} = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^\ell \sum_{\imath=0}^{\kappa_\ell-\ell} \frac{\tau^\imath}{\imath!} \bar{f}_N^{(\ell+\imath)} - \sum_{\ell=0}^{\nu-1} \hat{P_N}^\ell \sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} R_{\kappa_\ell} \bar{f}_N(\tau\gamma_i).$$

For the local error, we need to compare this expression with the exact solution $P_\infty x(t_{N+1}) = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^\ell \bar{f}_{N+1}^{(\ell)}$. Expanding each $\bar{f}_{N+1}^{(\ell)}$ in a Taylor series, $P_\infty x(t_{N+1})$ is given by

$$P_\infty x(t_{N+1}) = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^\ell \sum_{\imath=0}^{\kappa_\ell-\ell} \frac{\tau^\imath}{\imath!} \bar{f}_N^{(\ell+\imath)} - \sum_{\ell=0}^{\nu-1} \hat{P_N}^\ell R_{\kappa_\ell-\ell} \bar{f}_N^{(\ell)}(\tau),$$

and we obtain the local error as

$$\epsilon_{loc} = -\sum_{\ell=0}^{\nu-1} \hat{P_N}^\ell \Big(\sum_{i=1}^{s} \frac{\beta^T \mathcal{A}^{-(\ell+1)} e_i}{\tau^\ell} R_{\kappa_\ell} \bar{f}(\tau\gamma_i) - R_{\kappa_\ell-\ell} \bar{f}^{(\ell)}(\tau) \Big).$$

2. For a multistep iteration, the algebraic part of (5) is given by

$$P_\infty x_N = P_\infty \sum_{j=0}^{k-1} r_j x_{N-k+j} + P_\infty \sum_{j=0}^{k} q_j f_{N-k+j}.$$

Using Lemma 4.11, the projected iteration matrix can be written as

$$\begin{aligned} P_\infty r_j(\hat{E}, \tau\hat{A}) &= P_\infty (\alpha_k I_n - \beta_k \tau \hat{A})^{-1} (\beta_j \tau \hat{A} - \alpha_j \hat{E}) \\ &= \big(-\frac{1}{\alpha_k} \sum_{\ell=0}^{\nu-1} \big(\frac{\alpha_k}{\tau\beta_k}\big)^{\ell+1} \hat{P_N}^\ell \hat{A}^D \big) (\beta_j \tau \hat{A} - \alpha_j \hat{E}) \\ &= -\frac{1}{\alpha_k} \sum_{\ell=0}^{\nu-1} \big(\frac{\alpha_k}{\tau\beta_k}\big)^{\ell+1} (\beta_j \tau \hat{P_N}^\ell \hat{A}^D \hat{A} - \alpha_j \hat{P_N}^\ell \hat{P_N}) \\ &= \frac{1}{\tau\beta_k} \sum_{\ell=0}^{\nu-1} \big(\frac{\alpha_k}{\tau\beta_k}\big)^{\ell} \hat{P_N}^\ell \ (\alpha_j \hat{P_N} - \beta_j \tau P_\infty), \end{aligned} \tag{20}$$

and for the inhomogeneities we obtain

$$\begin{aligned} P_\infty q_j(\hat{E}, \tau\hat{A}) &= \tau\beta_j P_\infty \ (\alpha_k \hat{E} - \beta_k \tau \hat{A})^{-1}, \\ &= -\frac{\tau\beta_j}{\alpha_k} \sum_{\ell=0}^{\nu-1} \big(\frac{\alpha_k}{\tau\beta_k}\big)^{\ell+1} P_N^\ell \hat{A}^D \\ &= -\frac{\beta_j}{\beta_k} \sum_{\ell=0}^{\nu-1} \big(\frac{\alpha_k}{\tau\beta_k}\big)^{\ell} \hat{P_N}^\ell \ \hat{A}^D. \end{aligned} \tag{21}$$

Inserting (20) and (21) into the projected iteration, we get

$$\begin{aligned} P_\infty x_N = &\Big(\frac{1}{\tau\beta_k} \sum_{\ell=0}^{\nu-1} (\frac{\alpha_k}{\tau\beta_k})^\ell \sum_{j=0}^{k-1} (\alpha_j P_N - \beta_j \tau P_\infty) P_N^\ell \Big) \ x_{N-k+j} \\ &-\frac{1}{\beta_k} P_\infty \sum_{\ell=0}^{\nu-1} (\frac{\alpha_k}{\tau\beta_k})^\ell \sum_{j=0}^{k} \beta_j \hat{P_N}^\ell \bar{f}_{N-k+j}. \end{aligned}$$

If we assume consistent initial values, i. e., $P_\infty x_{N-k+j} = -\sum_{m=0}^{\nu-1} P_N^m \bar{f}_{N-k+j}^{(m)}$, this implies

$$\begin{aligned} P_\infty x_N &= \frac{1}{\tau\beta_k} P_\infty \sum_{\ell,m=0}^{\nu-1} (\frac{\alpha_k}{\tau\beta_k})^\ell \sum_{j=0}^{k-1} (\alpha_j \hat{P_N} - \beta_j \tau P_\infty) \hat{P_N}^{\ell+m} \bar{f}_{N-k+j}^{(m)} \\ &\quad - \frac{1}{\beta_k} P_\infty \sum_{\ell=0}^{\nu-1} (\frac{\alpha_k}{\tau\beta_k})^\ell \sum_{j=0}^{k} \beta_j \hat{P_N}^\ell \bar{f}_{N-k+j}. \end{aligned}$$

Reordering the terms and sorting in powers of P_N gives

$$\begin{aligned} P_\infty x_N &= P_\infty \sum_{\ell=0}^{\nu-1} \sum_{m=0}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^m \sum_{j=0}^{k-1} \frac{\beta_j}{\beta_k} \hat{P_N}^\ell \bar{f}_{N-k+j}^{(\ell-m)} \\ &\quad - \frac{1}{\tau} P_\infty \sum_{\ell=0}^{\nu-1} \sum_{m=0}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^m \sum_{j=0}^{k-1} \frac{\alpha_j}{\beta_k} \hat{P_N}^{\ell+1} \bar{f}_{N-k+j}^{(\ell-m)} \\ &\quad - P_\infty \sum_{\ell=0}^{\nu-1} (\frac{\alpha_k}{\tau\beta_k})^\ell \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \hat{P_N}^\ell \bar{f}_{N-k+j}. \end{aligned}$$

Transforming summation indices, and combining the first and last double sum, noting that for $\ell = 0$, the corresponding terms add up to $-P_\infty \bar{f}_n$, this is equivalent to

$$\begin{aligned} P_\infty x_N &= -P_\infty \bar{f}_N - P_\infty \sum_{\ell=1}^{\nu-1} \hat{P_N}^\ell \left(- \sum_{m=0}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^m \sum_{j=0}^{k-1} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)} \right. \\ &\quad \left. + \frac{1}{\tau} \sum_{m=1}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^{m-1} \sum_{j=0}^{k-1} \frac{\alpha_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)} + (\frac{\alpha_k}{\tau\beta_k})^\ell \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j} \right). \end{aligned}$$

Observing that for $m = 1, ..., \ell$, several summands cancel each other out, we may write this as

$$\begin{aligned} P_\infty x_N &= -P_\infty \bar{f}_N - P_\infty \sum_{\ell=1}^{\nu-1} \hat{P_N}^\ell \left(- \sum_{j=0}^{k-1} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell)} - \sum_{m=1}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^m \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)} \right. \\ &\quad \left. + \frac{1}{\tau} \sum_{m=1}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^{m-1} \sum_{j=0}^{k} \frac{\alpha_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)} + (\frac{\alpha_k}{\tau\beta_k})^\ell \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j} \right), \end{aligned}$$

and by combining further terms we obtain

$$\begin{aligned} P_\infty x_N &= -P_\infty \bar{f}_N - P_\infty \sum_{\ell=1}^{\nu-1} \hat{P_N}^\ell \left(- \sum_{j=0}^{k-1} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell)} \right. \\ &\quad \left. - \sum_{m=1}^{\ell-1} (\frac{\alpha_k}{\tau\beta_k})^m \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)} + \frac{1}{\tau} \sum_{m=1}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^{m-1} \sum_{j=0}^{k} \frac{\alpha_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)} \right). \end{aligned}$$

To compare with the exact solution, we consider the difference

$$\delta_\ell := - \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell)} - \sum_{m=1}^{\ell-1} (\frac{\alpha_k}{\tau\beta_k})^m \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)} + \frac{1}{\tau} \sum_{m=1}^{\ell} (\frac{\alpha_k}{\tau\beta_k})^{m-1} \sum_{j=0}^{k} \frac{\alpha_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell-m)}$$

so that the new approximation $P_\infty x_N$ is then given by

$$\begin{aligned} P_\infty x_N &= -P_\infty \bar{f}_N - P_\infty \sum_{\ell=1}^{\nu-1} P_N^\ell \left(-\sum_{j=0}^{k-1} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell)} + \sum_{j=0}^{k} \frac{\beta_j}{\beta_k} \bar{f}_{N-k+j}^{(\ell)} + \delta_\ell \right) \\ &= -P_\infty \sum_{\ell=0}^{\nu-1} P_N^\ell \bar{f}_n^{(\ell)} + \delta_\ell. \end{aligned}$$

For convenience, we multiply δ_ℓ by $\tau^\ell \beta_k$ and combine the first two sums, i. e.,

$$\tau^\ell \beta_k \delta_\ell = -\sum_{m=0}^{\ell-1} \tau^{\ell-m} (\frac{\alpha_k}{\beta_k})^m \sum_{j=0}^{k} \beta_j \bar{f}_{N-k+j}^{(\ell-m)} + \sum_{m=0}^{\ell-1} \tau^{\ell-m-1} (\frac{\alpha_k}{\beta_k})^m \sum_{j=0}^{k} \alpha_j \bar{f}_{N-k+j}^{(\ell-m-1)},$$

such that, by combining the sums and transforming the summation indices, we get

$$\tau^\ell \beta_k \delta_\ell = \sum_{m=0}^{\ell-1} \tau^m (\frac{\alpha_k}{\beta_k})^{\ell-m-1} \sum_{j=0}^{k} \Big(\alpha_j \bar{f}_{N-k+j}^{(m)} - \tau \beta_j \bar{f}_{N-k+j}^{(m+1)} \Big).$$

If $\bar{f}$ is sufficiently smooth, then we can expand each derivative into a Taylor series centered in t_{N-k}, such that for equidistant stepsizes $t_{N-k+j} = t_{N-k} + j\tau$, we obtain

$$\begin{aligned} \tau^\ell \beta_k \delta_\ell &= \sum_{m=0}^{\ell-1} \tau^m (\frac{\alpha_k}{\beta_k})^{\ell-m-1} \sum_{j=0}^{k} \left(\alpha_j \sum_{\imath=0}^{p-m} \frac{(j\tau)^\imath}{\imath!} \bar{f}_{N-k}^{(m+\imath)} - \tau \beta_j \sum_{\imath=0}^{p-m-1} \frac{(j\tau)^\imath}{\imath!} \bar{f}_{N-k}^{(m+\imath+1)} \right) \\ &\quad + \sum_{m=0}^{\ell-1} \tau^m (\frac{\alpha_k}{\beta_k})^{\ell-m-1} \sum_{j=0}^{k} \Big(\alpha_j R_{p-m} \bar{f}_{N-k}^{(m)}(j\tau) - \tau \beta_j R_{p-m-1} \bar{f}_{N-k}^{(m+1)}(j\tau) \Big), \end{aligned}$$

or by sorting the terms in orders of $\bar{f}^{(m+\imath)}$,

$$\begin{aligned} \tau^\ell \beta_k \delta_\ell &= \sum_{m=0}^{\ell-1} \tau^m (\frac{\alpha_k}{\beta_k})^{\ell-m-1} \sum_{j=0}^{k} \left(\alpha_j \bar{f}_{N-k}^{(\imath)} + \sum_{\imath=1}^{p-m} \frac{\tau^\imath}{\imath!} \big(j^\imath \alpha_j - \imath j^{\imath-1} \beta_j \big) \bar{f}_{N-k}^{(m+\imath)} \right) \\ &\quad + \sum_{m=0}^{\ell-1} \tau^m (\frac{\alpha_k}{\beta_k})^{\ell-m-1} \sum_{j=0}^{k} \Big(\alpha_j R_{p-m} \bar{f}_{N-k}^{(m)}(j\tau) - \tau \beta_j R_{p-m-1} \bar{f}_{N-k}^{(m+1)}(j\tau) \Big). \end{aligned}$$

If the method is consistent of order p, this reduces to

$$\tau^\ell \beta_k \delta_\ell = \sum_{m=0}^{\ell-1} \tau^m (\frac{\alpha_k}{\beta_k})^{\ell-m-1} \sum_{j=0}^{k} \Big(\alpha_j R_{p-m} \bar{f}_{n-k}^{(m)}(j\tau) - \tau \beta_j R_{p-m-1} \bar{f}_{N-k}^{(m+1)}(j\tau) \Big)$$

and the new approximation is given by

$$\begin{aligned} P_\infty x_N &= -P_\infty \sum_{\ell=0}^{\nu-1} P_N^\ell \bar{f}_N^{(\ell)} \\ &\quad - \frac{1}{\tau \beta_k} \sum_{\ell=0}^{\nu-1} P_N^\ell \sum_{m=0}^{\ell-1} (\frac{\alpha_k}{\tau \beta_k})^{\ell-m-1} \sum_{j=0}^{k} \Big(\alpha_j R_{p-m} \bar{f}_{N-k}^{(m)}(j\tau) - \tau \beta_j R_{p-m-1} \bar{f}_{N-k}^{(m+1)}(j\tau) \Big). \end{aligned}$$

Compared with the exact solution $P_\infty x(t_N) = -P_\infty \sum_{\ell=0}^{\nu-1} P_N^\ell \bar{f}_N^{(\ell)}$, this yields the local error

$$\epsilon_{loc} = -\frac{1}{\tau \beta_k} \sum_{\ell=0}^{\nu-1} P_N^\ell \sum_{m=0}^{\ell-1} (\frac{\alpha_k}{\tau \beta_k})^{\ell-m-1} \sum_{j=0}^{k} \Big(\alpha_j R_{p-m} \bar{f}_{n-k}^{(m)}(j\tau) - \tau \beta_j R_{p-\imath-1} \bar{f}_{N-k}^{(m+1)}(j\tau) \Big).$$

□

It follows that for Runge-Kutta methods the derivatives $\hat{P_N}^{\ell}\hat{A}^D f^{(\ell)}(t)$ occurring in the exact solution $P_\infty x(t)$ are approximated up to the order κ_ℓ, respectively, whereas for multistep methods, all derivatives are approximated up to the same order p. We further note that for multistep methods $\epsilon_{loc} = 0$ holds if $\nu \leq 1$, i. e., multistep methods yield an exact discretization of the algebraic part of DAEs with index at most one.

In order to estimate the sign of the local error, we restrict the class of considered problems to those that admit an explicit expression of the remainders.

Theorem 4.3. *Let $E\dot{x} = Ax + f$ be a positive DAE with regular, commuting matrix pair (E, A) in $\mathbb{R}^{n\times n}$, ind$(E, A) = \nu$ and let $P_\infty P_N^\ell A^D f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^n)$ and $-P_\infty P_N^\ell \bar{f}^{(\iota)}(t) \geq 0$ for $t \geq 0$.*

1. *Let $(\mathcal{A}, \beta, \gamma)$ with $\mathcal{A}$ nonsingular be the coefficients of a Runge-Kutta method that is consistent with exponents $\kappa_0, ..., \kappa_{\nu-1} \in \mathbb{N}$. If $\beta^T \mathcal{A}^{-(\ell+1)}\gamma^m \geq \frac{m!}{(m-\ell)!}$ holds for $m = \kappa_\ell + 1, ..., \infty$ and $\ell = 0, ..., \nu - 1$, then the method is positive for the system $P_N\dot{x} = P_\infty x + P_\infty A^D f$ for every $\tau > 0$.*

2. *Let (α, β), $\alpha_k, \beta_k \neq 0$ be the coefficients of a multistep method that is consistent of order $p \in \mathbb{N}$. If $\sum_{j=0}^k j^m\alpha_j \geq \sum_{j=0}^k m j^{m-1}\beta_j$ holds for $m = p + 1, ..., \infty$, then the method is positive for the system $P_N\dot{x} = P_\infty x + P_\infty A^D f$ for every $\tau > 0$.*

Proof. 1. If $\bar{f} \in C^\infty(\mathbb{R}, \mathbb{R}^n)$, then the remainders in the local error of a Runge-Kutta method are given by

$$R_{\kappa_\ell}\bar{f}_N^{(\ell)}(\gamma_i\tau) = \sum_{m=\kappa_\ell+1}^{\infty} \frac{(\gamma_i\tau)^m}{m!}\bar{f}_N^{(m)}, \qquad R_{\kappa_\ell-\ell}\bar{f}_N^{(\ell)}(\tau) \ = \sum_{m=\kappa_\ell-\ell+1}^{\infty} \frac{\tau^m}{m!}\bar{f}_n^{(\ell+m)} \tag{22}$$

and thus

$$\begin{aligned}
\epsilon_{loc} &= -\sum_{\ell=0}^{\nu-1} P_N^\ell \Big(\sum_{i=1}^{s} \frac{\beta^T\mathcal{A}^{-(\ell+1)}e_i}{\tau^\ell} \sum_{m=\kappa_\ell+1}^{\infty} \frac{(\gamma_i\tau)^m}{m!}\bar{f}_N^{(m)} - \sum_{m=\kappa_\ell+1-\ell}^{\infty} \frac{\tau^m}{m!}\bar{f}_N^{(\ell+m)} \Big) \\
&= -\sum_{\ell=0}^{\nu-1} P_N^\ell \Big(\frac{\beta^T\mathcal{A}^{-(\ell+1)}\gamma^m}{\tau^\ell} \sum_{m=\kappa_\ell+1}^{\infty} \frac{\tau^m}{m!}\bar{f}_N^{(m)} - \sum_{m=\kappa_\ell+1}^{\infty} \frac{\tau^{m-\ell}}{(m-\ell)!}\bar{f}_N^{(m)} \Big) \\
&= -\sum_{\ell=0}^{\nu-1} \frac{1}{\tau^\ell} P_N^\ell \sum_{m=\kappa_\ell+1}^{\infty} \frac{\tau^m}{m!} \Big(\beta^T\mathcal{A}^{-(\ell+1)}\gamma^m - \frac{m!}{(m-\ell)!} \Big) \bar{f}_N^{(m)}.
\end{aligned}$$

As we see from Tables 4 and 3, the values of the consistency exponents κ_ℓ differ considerably in size, so we need to prove the nonnegativity for each ℓ separately, i. e., we need to show that

$$-\sum_{m=\kappa_\ell+1}^{\infty} \frac{\tau^m}{m!} \Big(\beta^T\mathcal{A}^{-(\ell+1)}\gamma^m - \frac{m!}{(m-\ell)!} \Big) P_N^\ell \bar{f}_N^{(m)} \quad \geq \quad 0$$

for $m = \kappa_\ell + 1, ..., \infty$ and $\ell = 0, ..., \nu - 1$. This certainly is satisfied for $t_N \geq 0$ and $\tau > 0$, if $-P_N^\ell \bar{f}^{(m)}(t) \geq 0$ and $\sum_{i=1}^s \beta^T\mathcal{A}^{-(\ell+1)}\gamma^m \geq \frac{m!}{(m-\ell)!}$.

2. If $\bar{f} \in C^\infty(\mathbb{R})$, then the remainders of a multistep method in the local error are given by

$$R_{p-\iota}\bar{f}_{N-k}^{(\iota)}(j\tau) = \sum_{m=p-\iota+1}^{\infty} \frac{(j\tau)^m}{m!}\bar{f}_{N-k}^{(\iota+m)}, \qquad R_{p-\iota-1}\bar{f}_{n-k}^{(\iota+1)}(j\tau) \ = \sum_{m=p-\iota}^{\infty} \frac{(j\tau)^m}{m!}\bar{f}_{N-k}^{(\iota+m+1)}.$$

The local error is then given by

$$\begin{aligned}
\epsilon_{loc} &= -\frac{1}{\tau\beta_k} \sum_{\ell=0}^{\nu-1} P_N^\ell \sum_{\iota=0}^{\ell-1} \Big(\frac{\alpha_k}{\tau\beta_k}\Big)^{\ell-\iota-1} \sum_{j=0}^{k} \alpha_j \sum_{m=p-\iota+1}^{\infty} \frac{(j\tau)^m}{m!}\bar{f}_{N-k}^{(\iota+m)} - \tau\beta_j \sum_{m=p-\iota+1}^{\infty} \frac{\tau^{m-1}}{(m-1)!}\bar{f}_{N-k}^{(\iota+m)} \\
&= -\frac{1}{\tau\beta_k} \sum_{\ell=0}^{\nu-1} P_N^\ell \sum_{\iota=0}^{\ell-1} \Big(\frac{\alpha_k}{\tau\beta_k}\Big)^{\ell-\iota-1} \sum_{m=p-\iota+1}^{\infty} \frac{\tau^m}{m!} \sum_{j=0}^{k} \big(j^m\alpha_j - j^{m-1}m\beta_j \big) \bar{f}_{N-k}^{(\iota+m)}.
\end{aligned}$$

and by a transformation of the summation indices, we obtain

$$\epsilon_{loc} = -\frac{1}{\tau\beta_k}\sum_{\ell=0}^{\nu-1} P_N^\ell \sum_{\imath=0}^{\ell-1}(\frac{\alpha_k}{\tau\beta_k})^{\ell-\imath-1} \sum_{m=p+1}^{\infty}\frac{\tau^{m-\imath}}{(m-\imath)!}\sum_{j=0}^{k}\left(j^{m-\imath}\alpha_j - j^{m-\imath-1}(m-\imath)\beta_j\right)\bar{f}_{N-k}^{(m)}.$$

This is nonnegative, if $-P_\infty P_N^\ell \bar{f}^{(m)}(t) \geq 0$ holds for $t \geq 0$ and if the coefficients of the multistep method satisfy $\alpha_k, \beta_k > 0$ and $\sum_{j=0}^k j^m\alpha_j \geq \sum_{j=0}^k m j^{m-1}\beta_j$ for $m = p+1, ..., \infty$, $\ell = 1, ..., \nu - 1$. □

This analysis shows that for smooth constraints, the approximation $P_\infty x_{N+1}$ overestimates the exact solution $P_\infty x(t)$, if the derivatives $-\hat{P_N}^\ell \hat{A}^D f^{(\kappa_\ell+1)}$ are absolutely monotonic for $t \geq 0$ and the applied method overestimates the consistency conditions of Theorem 4.2.

The assumptions of Theorem 4.2 are analyzed for Radau-IIA and Lobatto-IIIC methods of stage order $s = 2, 3$ in Tables 4 and 3. As one can see, none of these schemes meets the positivity condition for higher index problems. But at least for DAEs of index at most one, a positive discretization is possible, because these methods are stiffly accurate, i. e., their coefficients satisfy $\beta^T\mathbf{1} = 1$ and $\beta^T = e_s^T\mathcal{A}$ and thus $\kappa_0 = \infty$.

Likewise, the most common multistep schemes, like e. g. BDF methods, do not satisfy these conditions for higher index problems as one can see from Table 5. Only for problems of index at most one, a positive discretization again is possible, since in this case multistep methods provide an exact discretization of the algebraic components.

We summarize these observations in the next Corollary.

Corollary 4.1. *Consider a positive DAE $E\dot{x} = Ax + f$ with regular, commuting pair (E, A) with $ind(E, A) = 1$ and let $P_\infty A^D f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^n)$.*

1. *Let $(\mathcal{A}, \beta, \gamma)$, $\mathcal{A}$ nonsingular, be the coefficients of a consistent, stiffly accurate Runge-Kutta method with $\gamma = \mathcal{A}\mathbf{1}$, then the discretization of $P_N\dot{x} = P_\infty x + P_\infty A^D f$ is positive for every $\tau > 0$.*

2. *Let (α, β), $\alpha_k, \beta_k > 0$ be the coefficients of a multistep method, then the discretization of $P_N\dot{x} = P_\infty x + P_\infty A^D f$ is positive for every $\tau > 0$.*

Proof. 1. By assumption, the coefficients $(\mathcal{A}, \beta, \gamma)$ of the Runge-Kutta method satisfy $\beta^T\mathbf{1} = 1$, $\beta^T = e_s^T\mathcal{A}$ and $1 = \beta^T e = e_s^T\mathcal{A}e = \gamma_s$. With $e_s^T\gamma^m = 1 = e_s^T\mathbf{1}$, we get $\beta^T\mathcal{A}^{-1}\gamma^m = \beta^T\mathcal{A}^{-1}\mathbf{1}$, which means that $\kappa_0 = \infty$, i. e., for $\ell = 0$, the consistency condition (ii) of Theorem 4.2 holds for arbitrary $m \geq 0$. Then, for a DAE of index at most one, the algebraic part of a Runge-Kutta discretization (4) is given by $P_\infty x_{N+1} = -P_\infty \sum_{m=0}^\infty \frac{\tau^m}{m!}\bar{f}_N^{(m)}$, which means that $P_\infty x_{N+1} = P_\infty x(t_{N+1}) = -P_\infty \bar{f}_{N+1}$ if $P_\infty A^D f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^n)$. But this is nonnegative by assumption.

2. If $\mathrm{ind}(E, A) = 1$, then the algebraic part of the discretization of is given by $P_\infty x_N = -P_\infty \bar{f}_N = P_\infty x(t_{N+1})$ and this is nonnegative if $E\dot{x} = Ax + f$ is positive. □

For DAEs of higher index than one, we now analyze the necessity of the conditions of Theorem 4.3. We denote by $\Pi^r(\mathbb{R}, \mathbb{R}^n) := \{\sum_{\vartheta=0}^r c_\vartheta t^\vartheta | c \in \mathbb{R}^n, t \in \mathbb{R}\}$ the set of vector valued polynomials of maximal degree r.

Lemma 4.13. *Consider a positive DAE $E\dot{x} = Ax + f$ with a regular and commuting matrix pair (E, A), $ind(E, A) = \nu$ and let $P_\infty A^D f \in \Pi^m(\mathbb{R}, \mathbb{R}^n)$.*

1. *Let $(\mathcal{A}, \beta, \gamma)$, with $\mathcal{A}$ nonsingular be the coefficients of a Runge-Kutta method that is consistent with exponents $\kappa_0, ..., \kappa_{\nu-1} \in \mathbb{N}$. If $P_\infty x_{N+1} \geq P_\infty x(t_{N+1})$ holds for $N \in \mathbb{N}$ with $t_{N+1} \geq 0$, then the coefficients $(\mathcal{A}, \beta, \gamma)$ satisfy*

$$\beta^T\mathcal{A}^{-(\ell_{\min}+1)}\gamma^{\kappa_{\ell_{\min}}+1} - \frac{(\kappa_{\ell_{\min}}+1)!}{(\kappa_{\ell_{\min}}+1-\ell_{\min})!} \geq 0$$

where $\kappa_{\ell_{\min}} := \min\{\kappa_0, ..., \kappa_{\nu-1}\}$.

2. *Let (α,β), $\alpha_k,\beta_k\neq 0$ be the coefficients of a multistep method that is consistent of order $p\in\mathbb{N}$. If $P_\infty x_{N+1}\geq P_\infty x(t_{N+1})$ holds for $N\in\mathbb{N}$ with $t_{N+1}\geq 0$, then the coefficients (α,β) satisfy*

$$\sum_{j=0}^{k} j^{p+1}\alpha_j \geq (p+1)\sum_{j=0}^{k} j^p\beta_j.$$

Proof. 1. If $P_\infty x_{N+1}\geq P_\infty x(t_{n+1})$ holds for every $t_{n+1}\in\mathbb{R}_+$, then the algebraic components of the local error $\epsilon_{loc}=P_\infty x_{N+1}-P_\infty x(t_{n+1})$, as it is defined in Theorem 4.12 are nonnegative for $t_{N+1}\geq 0$ and $\tau>0$. For $P_\infty A^D f=\sum_{\vartheta=0}^{r} c_\vartheta t^\vartheta\in\Pi^r(\mathbb{R},\mathbb{R}^n)$, the local error ϵ_{loc} is given by

$$\epsilon_{loc} \;=\; -\sum_{\ell=0}^{\nu-1}\sum_{m=\kappa_\ell+1}^{\infty}\frac{\tau^m}{m!}\vartheta_{\ell,m}\sum_{\vartheta=0}^{r-m}\frac{(\vartheta+m)!}{\vartheta!}P_\infty P_N^\ell c_{\vartheta+m}t_N^\vartheta,$$

where $\vartheta_{\ell,m}:=\beta^T\mathcal{A}^{-(\ell+1)}\gamma^m-\frac{m!}{(m-\ell)!}$. With $\kappa_{\ell_{\min}}:=\min\{\kappa_0,...,\kappa_{\nu-1}\}$, then the nonnegativity of ϵ_{loc} implies that

$$-\sum_{\ell=0}^{\nu-1}\sum_{m=\kappa_\ell+1}^{\infty}\frac{\tau^{m-\kappa_{\ell_{\min}}-1}}{m!}\vartheta_{\ell,m}\sum_{\vartheta=0}^{r-m}\frac{(\vartheta+m)!}{\vartheta!}P_\infty P_N^\ell c_{\vartheta+m}t_N^\vartheta \;\geq\; 0 \tag{23}$$

for $t_{N+1}\geq 0$, $\tau>0$ as well and considering this for small τ, we get

$$-\frac{\vartheta_{\ell_{\min},\kappa_{\ell_{\min}}+1}}{(\kappa_{\ell_{\min}}+1)!}\sum_{\vartheta=0}^{r-\kappa_{\ell_{\min}}-1}\frac{(\vartheta+\kappa_{\ell_{\min}}+1)!}{\vartheta!}P_\infty P_N^{\ell_{\min}}c_{\vartheta+\kappa_{\ell_{\min}}+1}t_N^\vartheta \;\geq\; 0.$$

Since $\vartheta_{\ell_{\min},\kappa_{\ell_{\min}}+1}$ is constant in time, this implies that $\sum_{\vartheta=0}^{r-\kappa_{\ell_{\min}}-1}\frac{(\vartheta+\kappa_{\ell_{\min}}+1)!}{\vartheta!}P_\infty P_N^{\ell_{\min}}c_{\vartheta+\kappa_{\ell_{\min}}+1}t_N^\vartheta$ has constant sign on $\mathbb{R}_+$. Thus, to prove the assertion $\vartheta_{\ell_{\min},\kappa_{\ell_{\min}}+1}\geq 0$, it is sufficient to show that this term is nonnegative for some $t\geq 0$.

By the positivity assumption, we know that $P_\infty x(t)=-P_\infty\sum_{\ell=0}^{\nu-1}P_N^\ell\bar{f}^\ell(t)$ is nonnegative for $t\geq 0$. Sorting with respect to powers of t, we obtain

$$\begin{aligned}P_\infty x(t) \;=\; &-\sum_{\vartheta=0}^{r-\nu+1}\Big(\sum_{\ell=0}^{\nu-1}\frac{(\vartheta+\ell)!}{\vartheta!}P_\infty P_N^\ell c_{\vartheta+\ell}\Big)t^\vartheta\\ &-\sum_{\vartheta=0}^{\nu-2}\Big(\frac{(r+\vartheta-\nu+2)!}{\vartheta!}P_\infty P_N^\vartheta a_{r+\vartheta-\nu+2}\Big)t^{r-\nu+2}-...-P_\infty c_r t^r\end{aligned}$$

and considering the limit $t\to\infty$, the positivity implies that

$$\begin{aligned}0 \;\leq\; &\lim_{t\to\infty}\Big(\big(-\sum_{\vartheta=0}^{r-\nu+1}\big(\sum_{\ell=0}^{\nu-1}\frac{(\vartheta+\ell)!}{\vartheta!}P_\infty P_N^\ell c_{\vartheta+\ell}\big)t^\vartheta\\ &-\sum_{\vartheta=0}^{\nu-2}\big(\frac{(r+\vartheta-\nu+2)!}{\vartheta!}P_\infty P_N^\vartheta c_{r+\vartheta-\nu+2}\big)t^{r-\nu+2}-...-P_\infty c_r t^r\big)\Big)\\ =\; &-P_\infty c_r.\end{aligned}$$

The same argument then shows that $-P_\infty\sum_{\ell=0}^{\nu-1}P_N^\ell\sum_{\vartheta=0}^{r-\kappa_\ell-1}\frac{(\vartheta+\kappa_\ell+1)!}{\vartheta!}c_{\vartheta+\kappa_1+1}t_N^\vartheta$ is nonnegative for large $t_N\geq 0$ and thus for any $t_N\geq 0$ by the constant sign condition. Thus, we have proved that

$$\vartheta_{\ell_{\min},\kappa_{\ell_{\min}}+1} \;=\; \beta^T\mathcal{A}^{-(\ell_{\min}+1)}\gamma^{\kappa_{\ell_{\min}}+1}-\frac{(\kappa_{\ell_{\min}}+1)!}{(\kappa_{\ell_{\min}}+1-\ell_{\min})!}\geq 0.$$

κ_ℓ	Lobatto-IIIC, $s=2$			
$\kappa_0=\infty$	0	0	0	$m=1,2,3$
$\kappa_1=1$	-1	-2	-3	$m=2,3,4$
$\kappa_2=2$	-6	-12	-20	$m=3,4,5$
$\kappa_3=3$	-26	-62	-122	$m=4,5,6$

κ_ℓ	Lobatto-IIIC, $s=3$			
$\kappa_0=\infty$	0	0	0	$m=1,2,3$
$\kappa_1=2$	-0.5	-1.25	-2.125	$m=3,4,5$
$\kappa_2=2$	-3	-8.5	-16.25	$m=3,4,5$
$\kappa_3=3$	-28.5	-65.25	-125.625	$m=4,5,6$

Table 3: Positivity conditions of Theorem 4.3 for the Lobatto-IIIC methods.

κ_ℓ	Radau-IIA, $s=2$			
$\kappa_0=\infty$	0	0	0	$m=1,2,3$
$\kappa_1=2$	-0.$\bar{6}$	-1.$\bar{5}$	-2.5185	$m=3,4,5$
$\kappa_2=2$	-2.$\bar{6}$	-8.$\bar{2}$	-16.0741	$m=3,4,5$
$\kappa_3=3$	-23.$\bar{5}$	-59.1852	-119.0617	$m=4,5,6$

κ_ℓ	Radau-IIA, $s=3$			
$\kappa_0=\infty$	0	0	0	$m=1,2,3$
$\kappa_1=3$	-0.3	-0.84	-1.542	$m=4,5,6$
$\kappa_2=3$	-2.7	-8.46	-16.998	$m=4,5,6$
$\kappa_3=3$	-13.5	-45.9	-103.47	$m=4,5,6$

Table 4: Positivity conditions of Theorem 4.3 for the Radau-IIA methods.

For multistep methods, if $P_\infty x_{N+1} \geq P_\infty x(t_{N+1})$ for every $t_{N+1} \geq 0$, then the algebraic components of the local error $\epsilon_{loc} = P_\infty x_{N+1} - P_\infty x(t_{N+1})$ are nonnegative for $t_{N+1} \geq 0$ and $\tau > 0$. For $P_\infty A^D f \in \Pi^r(\mathbb{R}, \mathbb{R}^n)$, i. e., $P_\infty A^D f(t) = \sum_{\vartheta=0}^{r} c_\vartheta t^\vartheta$, then the local error ϵ_{loc} of the multistep method is given by, cp. Theorem 4.12,

$$\epsilon_{loc} \quad = \quad -\frac{1}{\beta_k} \sum_{\ell=0}^{\nu-1} \sum_{\imath=0}^{\ell-1} \sum_{m=p+1}^{\infty} (\frac{\alpha_k}{\beta_k})^{\ell-1} \frac{\tau^{m-\ell}}{(m-\imath)!} \varphi_{m-\imath} \sum_{\vartheta=0}^{r-m-\imath} \frac{(\vartheta+m-\imath)!}{\vartheta!} P_\infty P_N^\ell c_{\vartheta+m-\imath} t_{N-k}^\vartheta,$$

where we use the abbreviation $\varphi_m := \sum_{j=0}^{k} \left(j^m \alpha_j - j^{m-1} m \beta_j \right)$. The smallest power of τ occurring in ϵ_{loc} is attained for $\ell = \nu - 1$ and $m = p + 1$. Multiplication of ϵ_{loc} by $\tau^{p-\nu+2}$ and considering this for small $\tau > 0$, the non-negativity assumption implies that

$$-\frac{1}{\beta_k} (\frac{\alpha_k}{\beta_k})^{\nu-2} \sum_{\imath=0}^{\nu-2} \frac{\varphi_{p-\imath+1}}{(p-\imath+1)!} \; P_N^{\nu-1} \sum_{\vartheta=0}^{r-p-1-\imath} \frac{(\vartheta+p+1-\imath)!}{\vartheta!} P_\infty P_N^\ell c_{\vartheta+p+1-\imath} t_{N-k}^\vartheta \quad \geq \quad 0.$$

As we have shown in the proof of Lemma 4.13, the positivity of $E\dot{x} = Ax + f$ implies that $-P_N^{\nu-1} \bar{f}_{N-k}^{(p+1)}$ is nonnegative for large $t \geq 0$. This implies that $-P_N^{\nu-1} \bar{f}_{N-k}^{(p+1)}$ is nonnegative for every $t \geq 0$, since $\varphi_{p-\imath+1}$ is constant in time. With $\alpha_k, \beta_k > 0$, this means that

$$\sum_{\imath=0}^{\nu-2} \frac{\varphi_{p-\imath+1}}{(p-\imath+1)!} \quad \geq \quad 0$$

holds and for $\nu \geq 2$, we have proved our assertion, i. e., that $\varphi_{p+1} \geq 0$. □

$k=p$	BDF, $k=3,4$		
$k=p=3$	-6	-54	$\ell=4,5$
$k=p=4$	-24	-336	$\ell=5,6$

Table 5: Positivity conditions of Theorem 4.3 for the BDF methods.

4.6 Example

To illustrate the conditions and implications of Theorem 4.1 and 4.3, we consider the matrix pair (E, A) with

$$E = \begin{bmatrix} -e_{11} & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

where $e_{11} > 0$. This pair (E, A) is regular, since $\det(\lambda E - A) \neq 0$, E, A commute and the index is $\nu = 3$. The Drazin inverses of the system matrices are given by

$$E^D = \begin{bmatrix} -\frac{1}{e_{11}} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A^D = A,$$

which yield the projections

$$E^D E = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad P_\infty = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the products

$$E^D A = \begin{bmatrix} -\frac{1}{e_{11}} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad A^D E = \begin{bmatrix} -e_{11} & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

For the algebraic part of the solution, we further need the products

$$P_\infty A^D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}, \qquad P_\infty A^D (A^D E) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$P_\infty A^D (A^D E)^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

As inhomogeneity, we consider the function

$$f = \begin{bmatrix} -1 & 0 & 0 & 0 & (t-t_0-0.1)^2 & \frac{1}{100}(t+0.3)^{-2} & \frac{1}{100}(t+0.1)^{-2} \end{bmatrix}^T,$$

whose derivatives are given by

$$\begin{aligned} f^{'} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 2(t-t_0-0.1) & -\frac{1}{50}(t+0.3)^{-3} & -\frac{1}{50}(t+0.1)^{-3} \end{bmatrix}^T, \\ f^{'''} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 2 & \frac{3}{50}(t+0.3)^{-4} & \frac{3}{50}(t+0.1)^{-4} \end{bmatrix}^T. \end{aligned}$$

For the exact solution, we first note that

$$\exp\left((t-t_0)\begin{bmatrix} -\frac{1}{e_{11}} & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}\right) = \begin{bmatrix} e^{-\frac{t-t_0}{e_{11}}} & * & * & * & 0 & 0 & 0 \\ 0 & e^{-(t-t_0)} & * & * & 0 & 0 & 0 \\ 0 & 0 & e^{-(t-t_0)} & * & 0 & 0 & 0 \\ 0 & 0 & 0 & e^{-(t-t_0)} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

such that

$$\int_{t_0}^{t} e^{(t-t_0)E^D A} E^D f ds = \begin{bmatrix} \int_{t_0}^{t} \frac{1}{e_{11}} e^{-\frac{t-s}{e_{11}}} ds \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -e^{-\frac{t-s}{e_{11}}}|_{t_0}^{t} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} (e^{-\frac{t-t_0}{e_{11}}} - 1) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

For the algebraic part,we compute

$$P_\infty A^D f = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -(t-t_0-0.1)^2 \\ -\frac{1}{100(t-t_0+0.3)^2} \\ -\frac{1}{100(t-t_0+0.1)^2} \end{bmatrix}, \quad P_\infty A^D (A^D E) f^{'} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\frac{1}{50(t-t_0+0.3)^3} \\ -\frac{1}{50(t-t_0+0.1)^3} \\ 0 \end{bmatrix}, \quad P_\infty (A^D E)^2 A^D f^{''} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\frac{3}{50(t-t_0+0.1)^4} \\ 0 \\ 0 \end{bmatrix}. \quad (24)$$

In conclusion, the exact solution is given by

$$x(t) = \begin{bmatrix} e^{-\frac{t-t_0}{e_{11}}} v_1 + * + e^{\frac{t-t_0}{e_{11}}} - 1 \\ e^{-(t-t_0)} v_2 + * \\ e^{-(t-t_0)} v_3 + * \\ e^{-(t-t_0)} v_4 \\ (t-t_0-0.1)^2 + \frac{1}{50(t-t_0+0.3)^3} + \frac{3}{50(t-t_0+0.1)^4} \\ \frac{1}{100(t-t_0+0.3)^2} + \frac{1}{50(t-t_0+0.1)^3} \\ \frac{1}{100(t-t_0+0.1)^2} \end{bmatrix}$$

where $[v_1, v_2, v_3, v_4]^T \in \mathbb{R}^4$ denotes the initial value for the dynamic components.

For the positivity, we check the conditions of Theorem 4.2. We observe that $E^D E \geq 0$ and $E^D A \geq -\mu E^D A$ for $\mu \geq \max\{\frac{1}{e_{11}}, 1\}$. Furthermore, it holds that $E^D f = [\frac{1}{e_{11}}\ 0\ 0\ 0\ 0\ 0\ 0] \geq 0$, which means that the differential components $E^D E x(t)$ are nonnegative. For the algebraic part, we see that $P_\infty x(t) = -P_\infty \sum_{\ell=0}^{2} (A^D E)^\ell A^D f^{(\ell)}(t)$ is nonnegative for $t \geq t_0$. In conclusion, the system $E\dot{x} = Ax + f$ is positive.

Now, we analyze the conditions under which this property is preserved in a numerical approximation. For the differential part, we apply Theorem 4.1 and check if (E, A) is a -M-matrix pair. The finite eigenvalues of (E, A) are given by $\sigma_{fin}(E, A) = \{-e_{11}, -1\}$, so it holds that $|\mu + \rho_{fin}| \leq \mu$ for $\mu \geq \frac{\rho_{fin}}{2}$ where $\rho_{fin} := \max\{\lambda \,|\, \lambda \in \sigma_{fin}(E, A)\}$. Since the inequality $E^D A \geq -\mu E^D E$ must be satisfied as well, we get that (E, A) is a -M-matrix pair with $\mu \geq \max\{\frac{1}{e_{11}}, 1\}$. For a given discretization method with absolute monotonicity radius γ_+, this means that the stepsize τ must be chosen according to

$$\tau \leq \frac{\gamma_+}{\max\{\frac{1}{e_{11}}, 1\}} \tag{25}$$

to ensure a nonnegative discretization.

But the threshold (25) is not strict, as we see from Table 7 and Table 8, where the projected iteration matrix $E^D E R(\tau E^D A)$ is computed for $e_{11} = 0.1, 1$ and the stepsize, where it loses its nonnegativity.

For $e_{11} = 0.1$, it holds that $\mu \geq 10$, such that for the 3-stage Radau-IIA and Lobatto-IIIC method, the stepsize is restricted by $\tau \leq 0.17940$ and $\tau \leq 0.11954$ respectively, cp. Table 1. But, as we see from Table 7 and Table 8, τ may be chosen more than $16\times$ larger than this threshold for the Radau method and still $3.5\times$ larger for the Lobatto method.

For $e_{11} = 1$, it holds that $\mu \geq 1$, so the stepsize for these methods is restricted by $\tau \leq 1.7940$ and $\tau \leq 1.1954$, respectively. From Table 7 and Table 8, we see that this threshold may be exceeded by more than $2.5\times$ for the Radau method and almost by $3.5\times$ for the Lobatto method.

For the 2-stage Radau-IIA and Lobatto-IIIC methods, this effect is even more remarkable, since these methods have the absolute monotonicity radius $\gamma_+ = 0$, cp. Table 1. But as it is shown in Table 6 for the 2-stage Radau-IIA method, the iteration matrix is nonnegative for stepsizes up to $\tau = 0.3$ if $e_{11} = 1$ and up to $\tau = 3$ if $e_{11} = 1$. For the Lobatto-III method, the iteration matrix stays nonnegative for any reasonable stepsize.

These observations can be explained as follows. The nonnegative matrix $B = \mu E^D E + E^D A$ used in the proof of Theorem 4.1 is given by

$$B_1 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{if } e_{11} = 1$$

and by

$$B_{0.1} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 9 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 9 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{if } e_{11} = 0.1.$$

The powers of these matrices are given by

$$\tilde{B}_1^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \tilde{B}_1^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \tilde{B}_1^\ell = 0, \ \ell = 4, 5, ...$$

and by

$$\tilde{B}_{0.1}^2 = \begin{bmatrix} 0 & 9 & 1 & 0 \\ 0 & 81 & 18 & 1 \\ 0 & 0 & 81 & 18 \\ 0 & 0 & 0 & 81 \end{bmatrix}, \qquad B_{0.1}^\ell = \begin{bmatrix} 0 & \ell 9^{\ell-1} & (\ell-1)9^{\ell-2} & (\ell-2)9^{\ell-3} \\ 0 & 9^\ell & \ell 9^{\ell-1} & 9[B_{0.1}^{\ell-1}]_{2,4} + [B_{0.1}^{\ell-1}]_{3,4} \\ 0 & 0 & 9^\ell & \ell 9^{\ell-1} \\ 0 & 0 & 0 & 9^\ell \end{bmatrix}, \quad \ell = 3, 4, ...,$$

where for convenience, we have dropped the columns and rows belonging to the algebraic components. Then, the expansion of $E^D ER(\tau E^D A)$ is given by

$$E^D ER(\tau E^D A) \quad = \quad \begin{bmatrix} R(-\tau) & R^{'}(-\tau) & R^{''}(-\tau) & R^{'''}(-\tau) & 0 & 0 & 0 \\ 0 & R(-\tau) & R^{'}(-\tau) & R^{''}(-\tau) & 0 & 0 & 0 \\ 0 & 0 & R(-\tau) & R^{'}(-\tau) & 0 & 0 & 0 \\ 0 & 0 & 0 & R(-\tau) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{if } e_{11} = 1$$

and by

$$E^D ER(\tau E^D A) \quad = \quad \begin{bmatrix} R(-10\tau) & b_1 & b_2 & b_3 & 0 & 0 & 0 \\ 0 & R(-10\tau) & b_1 & * & 0 & 0 & 0 \\ 0 & 0 & R(-10\tau) & b_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & R(-10\tau) & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \text{if } e_{11} = 0.1$$

where

$$b_1 := \sum_{\ell=1}^{\infty} R^{(\ell)}(-10\tau)9^{\ell-1}, \quad b_2 := \sum_{\ell=2}^{\infty} R^{(\ell)}(-10\tau)(\ell-1)9^{\ell-2}, \quad b_3 := \sum_{\ell=3}^{\infty} R^{(\ell)}(-10\tau)(\ell-2)9^{\ell-3}.$$

For the 2-stage Radau-IIA method and the 3-stage Lobatto-IIIC, the stability function $R(z)$ changes sign before its derivatives do, see Table 9 and Table 10. Thus, the iteration matrix becomes negative for $\tau \geq -3$ and $\tau \geq -4$ on the diagonal if $e_{11} = 1$ and in the first diagonal entry if $e_{11} = 0.1$.

For the 2-stage Lobatto-IIIC method none of the occurring derivatives $R(-\tau)$, $R^{'}(z)$, $R^{''}(z)$, $R^{'''}(z)$ change their sign on the negative real axis, see Table 10, which is why $E^D ER(\tau E^D A)$ is nonnegative for any stepsize $\tau > 0$ if $e_{11} = 1$. For $e_{11} = 0.1$, where higher derivatives of $R(z)$ occur in b_1, b_2, b_3, potential negative terms seem to be balanced by the positive initial terms.

For the 3-stage Radau-IIA method, $R^{'}(z)$ first changes sign, see Table 9. Therefore, the iteration matrix gets negative in the super-diagonal for $\tau = -4.8084$ if $e_{11} = 1$. For $e_{11} = 0.1$, it is also the first super-diagonal entry that loses its nonnegativity. But since this entry includes higher derivatives as well, the stepsize for the nonnegativity loss occurs is larger than the expected multiple $\tau = -0.48084$.

For the positive discretization of the algebraic components, we note that the inhomogeneity f and its derivatives satisfy the conditions of Theorem 4.3, i.e., $-P_\infty A^D f$, $-P_\infty A^D (A^D E) f^{'}$,

	$E^D ER(\tau E^D A)$, $e_{11} = 0.1$		$E^D ER(\tau E^D A)$, $e_{11} = 0.1$
$\tau = 0.29$	$\begin{bmatrix} 0.0077 & 0.8228 & 0.1500 & 0.0179 \\ 0 & 0.7482 & 0.2172 & 0.0311 \\ 0 & 0 & 0.7482 & 0.2172 \\ 0 & 0 & 0 & 0.7482 \end{bmatrix}$	$\tau = 2.99$	$\begin{bmatrix} 0.0007 & 0.2231 & 0.2473 & 0.2001 \\ 0 & 0.0007 & 0.2231 & 0.2473 \\ 0 & 0 & 0.0007 & 0.2231 \\ 0 & 0 & 0 & 0.0007 \end{bmatrix}$
$\tau = 0.3$	$\begin{bmatrix} 0 & 0.8230 & 0.1558 & 0.0193 \\ 0 & 0.7407 & 0.2225 & 0.0330 \\ 0 & 0 & 0.7407 & 0.2225 \\ 0 & 0 & 0 & 0.7407 \end{bmatrix}$	$\tau = 3.00$	$\begin{bmatrix} 0.0000 & 0.2222 & 0.2469 & 0.2003 \\ 0 & 0.0000 & 0.2222 & 0.2469 \\ 0 & 0 & 0.0000 & 0.2222 \\ 0 & 0 & 0 & 0.0000 \end{bmatrix}$
$\tau = 0.31$	$\begin{bmatrix} -0.0071 & 0.8228 & 0.1616 & 0.0208 \\ 0 & 0.7334 & 0.2277 & 0.0348 \\ 0 & 0 & 0.7334 & 0.2277 \\ 0 & 0 & 0 & 0.733 \end{bmatrix}$	$\tau = 3.01$	$\begin{bmatrix} -0.0007 & 0.2213 & 0.2466 & 0.2004 \\ 0 & -0.0007 & 0.2213 & 0.2466 \\ 0 & 0 & -0.0007 & 0.2213 \\ 0 & 0 & 0 & -0.0007 \end{bmatrix}$

Table 6: Differential part of the iteration matrix for Radau-IIA, s=2, and $e_{11} = 0.1, 1$.

	$E^D ER(\tau E^D A)$, $e_{11} = 0.1$		$E^D ER(\tau E^D A)$, $e_{11} = 1$
$\tau = 2.94$	$\begin{bmatrix} 0.0568 & 0.0004 & 0.1568 & 0.2502 \\ 0 & 0.0571 & 0.1412 & 0.2409 \\ 0 & 0 & 0.0571 & 0.1412 \\ 0 & 0 & 0 & 0.0571 \end{bmatrix}$	$\tau = 4.798$	$\begin{bmatrix} 0.0253 & 0.0004 & 0.1015 & 0.1740 \\ 0 & 0.0253 & 0.0004 & 0.1015 \\ 0 & 0 & 0.0253 & 0.0004 \\ 0 & 0 & 0 & 0.0253 \end{bmatrix}$
$\tau = 2.95$	$\begin{bmatrix} 0.0567 & 0.0000 & 0.1556 & 0.2496 \\ 0 & 0.0567 & 0.1400 & 0.2402 \\ 0 & 0 & 0.0567 & 0.1400 \\ 0 & 0 & 0 & 0.0567 \end{bmatrix}$	$\tau = 4.808$	$\begin{bmatrix} 0.0253 & 0.0000 & 0.1008 & 0.1735 \\ 0 & 0.0253 & 0.0000 & 0.1008 \\ 0 & 0 & 0.0253 & 0.0000 \\ 0 & 0 & 0 & 0.0253 \end{bmatrix}$
$\tau = 2.96$	$\begin{bmatrix} 0.0566 & -0.0005 & 0.1544 & 0.2490 \\ 0 & 0.0562 & 0.1389 & 0.2395 \\ 0 & 0 & 0.0562 & 0.1389 \\ 0 & 0 & 0 & 0.0562 \end{bmatrix}$	$\tau = 4.818$	$\begin{bmatrix} 0.0253 & -0.0004 & 0.1002 & 0.1730 \\ 0 & 0.0253 & -0.0004 & 0.1002 \\ 0 & 0 & 0.0253 & -0.0004 \\ 0 & 0 & 0 & 0.0253 \end{bmatrix}$

Table 7: Differential part of the iteration matrix for Radau-IIA, s=3, and $e_{11} = 0.1, 1$.

	$E^D ER(\tau E^D A)$, $e_{11} = 0.1$		$E^D ER(\tau E^D A)$, $e_{11} = 1$
$\tau = 0.39$	$\begin{bmatrix} 0.0025 & 0.7495 & 0.2102 & 0.0338 \\ 0 & 0.6770 & 0.2641 & 0.0514 \\ 0 & 0 & 0.6770 & 0.2641 \\ 0 & 0 & 0 & 0.6770 \end{bmatrix}$	$\tau = 3.99$	$\begin{bmatrix} 0.0002 & 0.0944 & 0.1676 & 0.1923 \\ 0 & 0.0002 & 0.0944 & 0.1676 \\ 0 & 0 & 0.0002 & 0.0944 \\ 0 & 0 & 0 & 0.0002 \end{bmatrix}$
$\tau = 0.4$	$\begin{bmatrix} 0 & 0.7448 & 0.2152 & 0.0356 \\ 0 & 0.6703 & 0.2682 & 0.0535 \\ 0 & 0 & 0.6703 & 0.2682 \\ 0 & 0 & 0 & 0.6703 \end{bmatrix}$	$\tau = 4$	$\begin{bmatrix} 0 & 0.0938 & 0.1670 & 0.1920 \\ 0 & 0 & 0.0938 & 0.1670 \\ 0 & 0 & 0 & 0.0938 \\ 0 & 0 & 0 & 0 \end{bmatrix}$
$\tau = 0.41$	$\begin{bmatrix} -0.0022 & 0.7399 & 0.2202 & 0.0374 \\ 0 & 0.6636 & 0.2722 & 0.0557 \\ 0 & 0 & 0.6636 & 0.2722 \\ 0 & 0 & 0 & 0.6636 \end{bmatrix}$	$\tau = 4.01$	$\begin{bmatrix} -0.0002 & 0.0932 & 0.1664 & 0.1916 \\ 0 & -0.0002 & 0.0932 & 0.1664 \\ 0 & 0 & -0.0002 & 0.0932 \\ 0 & 0 & 0 & -0.0002 \end{bmatrix}$

Table 8: Differential part of the iteration matrix for Lobatto-IIIC, s=3, and $e_{11} = 0.1, 1$.

Radau-IIA, $s=2$	$R^{(\ell)}(z)\geq 0$	Lobatto-IIIC, $s=2$	$R^{(\ell)}(z)\geq 0$
$R(z)=\frac{2z+6}{z^2-4z+6}$	$z\in[-3,0]$	$R(z)=\frac{2}{z^2-2z+2}$	$z\in[-\infty,0]$
$R^{'}(z)=-\frac{2z^2+12z-48}{(z^2-4z+6)^2}$	$z\in[-8.7446,0]$	$R^{'}(z)=-\frac{4(z-1)}{(z^2-2z+2)^2}$	$z\in[-\infty,0]$
$R^{''}(z)=\frac{4z^3+36z^2-264z+312}{(z^2-4z+6)^3}$	$z\in[-14.0807,0]$	$R^{''}(z)=\frac{4(3z^2-6z+2)}{(z^2-2z+2)^3}$	$z\in[-\infty,0]$
$R^{'''}(z)=-\frac{12z^4+144z^3-1536z^2+3552z-2160}{(z^2-4z+6)^4}$	$z\in[-19.4064,0]$	$R^{'''}(z)=-\frac{48z(z-1)(z-2)}{(z^2-2z+2)^4}$	$z\in[-\infty,0]$

Table 9: Stability function of the Radau-IIA and Lobatto-IIIC method, $s=2$.

Radau-IIA, $s=3$	$R^{(\ell)}(z)\geq 0$
$R(z)=-\frac{3z^2+24z+60}{z^3-9z^2+36z-60}$	$z\in[-\infty,0]$
$R^{'}(z)=\frac{3z^4+48z^3-144z^2-720z+3600}{(z^3-9z^2+36z-60)^2}$	$z\in[-4.8084,0]$
$R^{''}(z)=-\frac{9z^6+117z^5-1116z^4-2196z^3+49680z^2-172800z+216000}{(z^3-9z^2+36z-60)^3}$	$z\in[-10.4362,0]$
$R^{'''}(z)=\frac{27z^8+468z^7-7605z^6+1728z^5+482328z^4-3438720z^3+11113200z^2-18144000z+12960000}{(z^3-9z^2+36z-60)^4}$	$z\in[-6.8370,0]$
Lobatto-IIIC, $s=3$	$R^{(\ell)}(z)\geq 0$
$R(z)=-\frac{6z+24}{z^3-6z^2+18z-24}$	$z\in[-4,0]$
$R^{'}(z)=-\frac{12z^3+36z^2-288z+576}{(z^3-6z^2+18z-24)^2}$	$z\in[-7.2345,0]$
$R^{''}(z)=\frac{36z^5+72z^4-2088z^3+9504z^2-17280z+13824}{(z^3-6z^2+18z-24)^3}$	$z\in[-10.4362,0]$
$R^{'''}(z)=-\frac{-126z^7+684z^6+1512z^5-30024z^4+149256z^3-386208z^2+528768z-290304}{(z^3-6z^2+18z-24)^4}$	$z\in[-6.8370,0]$

Table 10: Stability function of the Radau-IIA and Lobatto-IIIC method, $s=3$.

$-P_\infty(A^DE)^2A^Df^{''}$ are each nonnegative, cp. (24). But, none of the tested methods satisfies the assertion $\beta^T\mathcal{A}^{-(\ell+1)}\gamma^m-\frac{m!}{(m-\ell)!}\geq 0$ for $\ell=0,1,2$ and $m=\ell+1,\ell+2,...$, cp. Table 4 and Table 3.

Nonetheless, a positive discretization is possible for sufficiently small stepsizes, which can be seen in Table 11. There, the stepsizes $\tau>0$ are presented for which each of the algebraic components x_5, x_6, x_7 are still nonnegative. For the computation, we note that the stepsize was increased in steps of 0.01 until $\tau=5$.

Again, the 2-stage Lobatto method admits the largest stepsizes for a positive discretization, even though it is the method that underestimates the exact solution the most, cp. Table 3.

The other methods either fail on the first algebraic component x_5, the 3-stage Radau-IIA method, or on the second component x_6, the 3-stage Lobatto-IIIC method or on both like the 2-stage Radau-IIA method.

But, since the tested methods are stiffly accurate, all of them yield a positive discretization of the last component x_7, which does not involve any derivatives.

Note that the algebraic part of the solution is invariant against changes in $\mathrm{im}(E^DE)$, which is why the values are equal for $e_{11}=0.1,1$.

Radau-IIA, $s=2$	
$x_5\geq 0$	$\tau\in(0,0.19]$
$x_6\geq 0$	$\tau\in(0,0.08]$
$x_7\geq 0$	$\tau\in(0,5]$

Lobatto-IIIC, $s=2$	
$x_5\geq 0,$	$\tau\in(0,5]$
$x_6\geq 0,$	$\tau\in(0,5]$
$x_7\geq 0,$	$\tau\in(0,5]$

Radau-IIA, $s=3$	
$x_5\geq 0$	$\tau\in(0,0.05]$
$x_6\geq 0$	$\tau\in(0,5]$
$x_7\geq 0$	$\tau\in(0,5]$

Lobatto-IIIC, $s=3$	
$x_5\geq 0,$	$\tau\in(0,5]$
$x_6\geq 0,$	$\tau\in(0,0.10]$
$x_7\geq 0,$	$\tau\in(0,5]$

Table 11: Positivity preserving stepsizes for the algebraic part of the discretization with Radau-IIA and Lobatto-IIIC methods.

5 Conclusion

The positivity of Runge-Kutta and linear multistep methods for linear differential-algebraic equations with constant coefficients has been discussed. The crucial tool in the analysis is the decomposition of $E\dot{x} = Ax + f$ into its differential and algebraic part, which allows to discuss the positivity for each of these parts separately. For the differential part, the result of [6] for ODEs is generalized to DAEs by extending the corresponding definitions and notations from single matrices to matrix pairs. For the algebraic part, the exact solution is overestimated by its approximation to ensure that the discretization is nonnegative. Conveniently, this condition holds if the considered method not only satisfies the consistency condition up to a certain order, but overestimates them for higher orders. For polynomial input functions f, this condition was shown to be even necessary.

Checking the positivity conditions, it has been shown that for the differential part, the assumptions are as strict as for ODEs; positive discretizations require the stepsize to be smaller than the radius of absolute monotonicity of the considered method, cp. Table 1, 2. This restriction can be relaxed, if the structure of the problem is taken into consideration, like e. g. the existence of an eigenvector basis, see [21] or the choice of special initial values.

For the algebraic part, Tables 4, 3 and 5 show that for most of the common methods the positivity conditions are not satisfied except for DAEs of index at most one. Thus, for higher index problems the construction of positivity preserving index reduction methods is necessary.

References

[1] A.R.A. Anderson, M.A.J. Chaplain, E.L. Newmann, R.J.C. Steele, and A.M. Thompson. Mathematical modelling of tumour invasion and metastasis. *J. Theor. Med.*, 2:129–154, 2000.

[2] B. D. O. Anderson. New developments in the theory of positive systems. In C.I. Byrnes, editor, *Systems and Control in the Twenty-First Century*. 1997.

[3] L. Benvenuti, A. De Santis, and L. Farina, editors. *Positive systems*, volume 294 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, 2003.

[4] L. Benvenuti and L. Farina. Positive and compartmental systems. *IEEE Trans. Automat. Control*, 47:370–373, 2002.

[5] G. Birkhoff and R. S. Varga. Reactor criticality and non-negative matrices. *J. Soc. Indust. Appl. Math.*, 6:354–377, 1958.

[6] C. Bolley and M. Crouzeix. Conservation de la positivite lors de la discretisation des problemes d'evolution paraboliques. *R.A.I.R.O. Analyse numerique*, 12:81–88, 1978.

[7] J.C. Butcher. *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods*. Wiley, Chichester, UK, 1987.

[8] S.L. Campbell. *Singular Systems of Differential Equations I*. Pitman, San Francisco, CA, 1980.

[9] S.L. Campbell and C.D. Meyer. *Generalized Inverses of Linear Transformations*. Pitman, San Francisco, CA, 1979.

[10] V. Capasso. *Mathematical Structures of Epidemic Systems*. Springer Verlag, Berlin, 1993.

[11] C. Commault and N. Marchand, editors. *Positive systems*, volume 341 of *Lecture Notes in Control and Information Sciences*. Springer, Berlin, 2006.

[12] G. Dahlquist. Convergence and stability in the numerical integration of ordinary differential equations. *Math. Scand.*, 4:33–53, 1956.

[13] L. Farina and S. Rinaldi. *Positive Linear Systems: Theory and its Applications*. John Wiley and Sons Inc., New York, 2000.

[14] G. Gandolfo. *Economic Dynamics*. Springer Verlag, Heidelberg, 1997.

[15] F.R. Gantmacher. *The Theory of Matrices*, volume 1. Chelsea Publishing Company, New York, NY, 1959.

[16] J.A. van de Griend and J.F.B.M. Kraaijevanger. Absolute monotonicity of rational functions occuring in the numerical study of initial value problems. *Numer. Math.*, 49:413–424, 1986.

[17] E. Hairer, S.P. Noersett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer Verlag, Berlin, 2nd edition, 1993.

[18] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations. II. Stiff and Differential-Algebraic Problems*. Springer Verlag, Berlin, 2nd edition, 1996.

[19] R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[20] Z. Horvath. Positivity of Runge-Kutta and diagonally split Runge-Kutta methods. *App. Num. Math.*, 28:309–326, 1998.

[21] Z. Horvath. Positively invariant cones of dynamical systems under Runge-Kutta and Rosenbrock-type discretization. *Proc. Appl. Math. Mech. (PAMM)*, 4:688–689, 2004.

[22] Z. Horvath. On the positivity step size threshold of Runge-Kutta methods. *App. Numer. Math.*, 53:341–356, 2005.

[23] W. Hundsdorfer, B. Koren, M. van Loon, and J.G. Verwer. A positive finite-difference advection scheme. *J. Comp. Phys.*, 117:35–46, 1994.

[24] W. Hundsdorfer and J. G. Verwer. *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Verlag, Berlin, 2003.

[25] T. Kaczorek. *Positive 1D and 2D Systems*. Springer Verlag, London, 2002.

[26] J.F.B.M. Kraaijevanger. Absolute monotonicity of polynomials occuring in the numerical soution of initial value problems. *Numer. Math.*, 48:303–322, 1986.

[27] P. Kunkel and V. Mehrmann. *Differential-Algebraic Equations. Analysis and Numerical Solution*. EMS Publishing House, Zürich, CH, 2006.

[28] P. Lancaster and M. Tismenetsky. *The Theory of Matrices*. Academic Press, New York, NY, 1985.

[29] A.J. Laub. *Matrix analysis for scientists and engineers*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 2005.

[30] H.W.J. Lenferink. Contractivity preserving explicit linear multistep methods. *Numer. Math.*, 55:213–223, 1989.

[31] H.W.J. Lenferink. Contractivity preserving implicit linear multistep methods. *Math. Comp.*, 56:177–199, 1991.

[32] D. G. Luenberger. *Introduction to Dynamic Systems*. John Wiley and Sons Inc., New York, NY, 1979.

[33] J.D. Murray, S.R. Lubkin, and R. Tyson. Model and analysis of chemotactic bacterial patters in a liquid medium. *J. Math. Biol.*, 38:359–75, 1999.

[34] B. G. Olivier and J. L. Snoep. Web-based kinetic modelling using JWS Online. *Bioinformatics*, 20(13):2143–2144, 2004.

[35] M.B. Reddy, R.S.H. Yang, H.J.C. Clewell, and M.E. Andersen. *Physiologically Based Pharmacokinetic Modelling*. John Wiley & Sons, New York, 2005.

[36] M.N. Spijker. Stepsize conditions for general monotonicity in numerical initial value problems. *SIAMNumAnal*, 45:1226–1245, 2007.

[37] H. Tian. On the reverse order law $(AB)^D = B^D A^D$. *Numerical Mathematics, A Journal of Chinese Universities*, 9(1):355–358, 2000.

[38] R.S. Varga. *Matrix Iterative Analysis*. Springer Verlag, Berlin, 2nd edition, 2000.

[39] E. Virnik. Stability analysis of positive descriptor systems. *Linear Algebra Appl.*, 429:2640–2659, 2008.