

Numerical integrators for non-autonomous closed-loop systems with Riccati-feedback

Björn Baran* Peter Benner** Jens Saak† Tony Stillfjord‡

*Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg.

Email: baran@mpi-magdeburg.mpg.de, ORCID: 0000-0001-6570-3653

**Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg.

Email: benner@mpi-magdeburg.mpg.de, ORCID: 0000-0003-3362-4103

†Max Planck Institute for Dynamics of Complex Technical Systems, Sandtorstr. 1, 39106 Magdeburg.

Email: saak@mpi-magdeburg.mpg.de, ORCID: 0000-0001-5567-9637

‡Lund University, Box 117, SE-221 00, Lund, Sweden.

Email: tony.stillfjord@math.lth.se, ORCID: 0000-0001-6123-4271

Abstract: By computing a feedback control via the linear quadratic regulator (LQR) approach and simulating a non-linear non-autonomous closed-loop system using this feedback, we combine two numerically challenging tasks. For the first task, the computation of the feedback control, we use the non-autonomous generalized differential Riccati equation (DRE), whose solution determines the time-varying feedback gain matrix. Regarding the second task, we want to be able to simulate non-linear closed-loop systems for which it is known that the regulator is only valid for sufficiently small perturbations. Thus, one easily runs into numerical issues in the integrators when the closed-loop control varies greatly. For these systems, e.g., the A-stable implicit Euler methods fails.

On the one hand, we implement non-autonomous versions of splitting schemes and BDF methods for the solution of our non-autonomous DREs. These are well-established DRE solvers in the autonomous case. On the other hand, to tackle the numerical issues in the simulation of the non-linear closed-loop system, we apply a fractional-step-theta scheme with time-adaptivity tuned specifically to this kind of challenge. That is, we additionally base the time-adaptivity on the activity of the control. We compare this approach to the more classical error-based time-adaptivity.

We describe techniques to make these two tasks computable in a reasonable amount of time and are able to simulate closed-loop systems with strongly varying controls, while avoiding numerical issues. Our time-adaptivity approach requires fewer time steps than the error-based alternative and is more reliable.

Keywords: differential Riccati equation, non-autonomous, closed-loop simulation, feedback control, step size control

Novelty statement: The novelty of this work is twofold:

- Until now, numerical methods for non-autonomous differential Riccati equations have been restricted to very specific assumptions on the time-dependent matrices. We propose new low-rank splitting schemes and backward differentiation formulas to numerically solve general non-autonomous differential Riccati equations.
- We overcome numerical issues that occur during the simulation of the corresponding closed-loop system by utilizing the fractional-step-theta algorithm combined with a time-adaptive strategy that is specifically adapted to quickly varying controls.

1 Introduction

The combination of the two tasks, to compute a feedback control for a (possibly) non-linear non-autonomous system and to apply this control in a simulation of the closed-loop system, consists of three steps. First, we simulate the (open-loop) non-linear problem, (possibly) with a feedforward-control, to assemble the reference solution trajectory and data for the feedback computation. Then, we linearize the equations around the reference trajectory and apply the linear-quadratic regulator (LQR) approach, to the linearized system, to receive the feedback gain. In the third step, we apply the feedback gain to the non-linear problem during a closed-loop simulation. The novelty of this approach lies in the application of this framework to large-scale non-autonomous systems. This can include problems with free boundaries and moving interfaces.

The second step, the computation of a feedback control with the LQR approach on a finite time-horizon, results in the necessity to solve a differential Riccati equation (DRE) [15, 41]. In fact, DREs arise for other applications as well and have been studied in great detail [1]. We focus in particular on the generalized DREs [29] and DREs resulting from differential-algebraic equations (DAEs) [30]. The algebraic constraint equations can arise, e.g., from divergence-free conditions in convection-diffusion equations or from coupling constraints with free boundaries and moving interfaces.

In order to numerically solve the corresponding DRE as well as the non-linear problem in step one and three, we first semi-discretize in space. Thereby, the DRE becomes matrix-valued with large-scale, but sparse, coefficients and a solution that, in most cases, can be approximated sufficiently accurate by a low-rank factorization [49]. There is a range of methods available to numerically solve this kind of DREs. To enumerate several, splitting schemes can be applied to DREs [39, 40, 46–48]. Further, there are the backward differentiation formulas (BDF) [12, 13, 32, 38] as well as Rosenbrock and Peer methods [10, 31, 32, 38]. Recently, also Krylov subspace methods have been applied to DREs [8, 21, 25, 28] as well as exponential integrators [33] and an all-at-once space-time approach [16] (computing a low-rank tensor format solution).

The just mentioned methods are mainly focused on autonomous DREs. In contrast to this, DREs are especially challenging when the coefficient matrices are time-dependent. In detail, these non-autonomous DREs are particularly challenging for problems with free boundaries [4, 6] and moving interfaces like fluid structure interaction (FSI) [19] and the Stefan problem [6, 7]. Addressing such problems, the numerical solution of large-scale non-autonomous DREs has not been studied in a lot of detail so far. For instance, in [13, 38], the authors formulate non-autonomous BDF and Rosenbrock methods without going into detail about the implementation or performance. Moreover, non-autonomous BDF and Rosenbrock methods are used in [31, 32] as well as non-autonomous Peer methods in [10], but are restricted to the case of constant matrices scaled by time-dependent scalar functions. The applications we consider have no restriction on the non-autonomy of the matrices. Consequently, in this manuscript, we extend the existing BDF methods (and to a limited extent also splitting methods) for solving large-scale fully non-autonomous DREs. Further, we give details on how to correctly incorporate the weight parameter from the cost functional into the DRE and provide a numerical implementation in the software package M-M.E.S.S. [43].

In addition to these contributions, this manuscript focuses on solving closed-loop problems with free boundaries and moving interfaces. This is a very broad and active field of research, see e.g., [18, 23, 26, 27, 35], and there are numerous numerical methods available. During the simulation of problems with moving interfaces, we can observe numerical issues, which are similar to the ones in [20]. This is especially true for closed-loop problems with strongly varying controls. In this setting, common methods like the implicit Euler method and the trapezoidal rule run the risk to break down as in [20]. To overcome this, our method of choice is the fractional-step-theta algorithm [17] combined with adaptive time stepping. It has successfully been applied for Navier-Stokes equations [36, 37] as well as for FSI problems [20, 42, 52]. Part of this method can be classical error estimates for the time-adaptivity, e.g., heuristics, based on models of the actual error [22, 24, 50], like the one that is used in [20]. Others are residual-based like the dual weighted residual, which yields the best results in [20]. Unfortunately, this method has high computational costs and, thus, is not feasible for our purpose. Therefore, our approach is to modify an error-based heuristic and adapt it to the case of closed-loop systems with quickly varying controls.

For specific problems that have moving inner boundaries as well, there are time-adaptive strate-

gies available that differ from ours. One example is an adaptive time stepping applied for a two-phase flow with a proportional–integral–derivative (PID) control in [2]. Another example is [44], where temporal and spatial discretizations are used that are adaptive with respect to Courant–Friedrichs–Lewy and non-self-intersection conditions for a dendritic Stefan problem.

Structure of the Manuscript. In Section 1.1, we state a general definition of the (possibly) non-linear non-autonomous closed-loop system we consider. To design a feedback controller for this system, we choose the LQR approach and formulate the non-autonomous DRE in Section 1.2. To solve this DRE, the non-autonomous versions of the BDF methods and the splitting schemes are discussed in Section 2. With this solution, we can compute a feedback control, which is then used in the closed-loop system. For a robust numerical solver of this system, we describe several time stepping methods for closed-loop integrators, e.g., a time-adaptive fractional-step-theta algorithm with several indicators in Section 3. We demonstrate and compare the performance of the presented methods and their behavior with respect to different control-parameter choices in Section 4. Additionally, we specify how to access the codes and data for our experiments in [5] and summarize our conclusions in Section 5.

Notation. In most equations, we omit the time-dependence (t), the spatial dependence (x), or the combination of both (t, x). This is intended to improve the readability of the equations.

1.1 Problem Definition

Our goal is to derive and apply a feedback control $u_{\mathcal{K}}$ for possibly non-linear non-autonomous systems. This section describes the corresponding control problem in an abstract form such that it can be applied to different dynamical systems.

We define this abstract, controlled, dynamical system

$$\dot{x} = f(x, u) \quad (1)$$

on a finite time horizon $t \in \mathcal{I} = [t_0, t_{\text{end}}]$ and a domain $\Omega \subset \mathbb{R}^d$ of dimension $d > 0$. The domain has a fixed outer shape and can contain moving interfaces or free inner boundaries. However, we expect that our methods work as well with a time-dependent domain that is finite and with a volume that is uniformly bounded away from zero. Additionally, the function f is possibly non-linear in $x(t)$. This state $x \in \mathcal{V}_x$ as well as the control $u \in \mathcal{V}_u$ are unknowns, with the function spaces \mathcal{V}_x and \mathcal{V}_u being suitable spaces chosen with respect to the underlying dynamical problem.

Further, we assume that a reference pair $(x_{\text{ref}}, u_{\text{ref}}) \in \mathcal{V}_x \times \mathcal{V}_u$ is given, which is a solution to Equation (1). In particular, this represents the desired state trajectory of the system, which we want to stabilize with the feedback control $u_{\mathcal{K}}$. That means, in this scenario, that the resulting input is given as $u = u_{\text{ref}} + u_{\mathcal{K}}$. To clarify, the additional feedback control is necessary because we assume that the system (1) is influenced by model inaccuracies or perturbations, which cause the state to deviate from the desired trajectory. This deviation is expressed as

$$x_{\Delta} = x - x_{\text{ref}}, \quad u_{\mathcal{K}}(t, x_{\Delta}(t)) = u - u_{\text{ref}}.$$

Consequently, we reformulate system (1) in closed-loop form

$$\dot{x}_{\Delta} = f(x_{\Delta}, u_{\mathcal{K}}). \quad (2)$$

With this setup, the desired state of system (2) is the zero state.

In order to measure the deviation of the state in a performance index, we denote the function that maps the state to the output $y_{\Delta}(t)$ of the system by $\mathcal{C}: \mathcal{V}_x \rightarrow \mathcal{V}_{\text{out}}$, where \mathcal{V}_{out} is a suitable space:

$$y_{\Delta} = \mathcal{C}x_{\Delta}.$$

These output and control costs are measured in a performance index $\mathcal{J}: \mathcal{V}_x \times \mathcal{V}_u \rightarrow \mathbb{R}_{\geq 0}$, defined by

$$\mathcal{J}(x_{\Delta}, u_{\mathcal{K}}) = \int_{t_0}^{t_{\text{end}}} \|y_{\Delta}\|^2 + \lambda \|u_{\mathcal{K}}\|^2 dt + x_{\Delta}^{\text{T}}(t_{\text{end}}) S x_{\Delta}(t_{\text{end}}), \quad (3)$$

where S is positive semidefinite and $\lambda > 0$ is a regularization parameter. A smaller λ makes the problem numerically more challenging and allows the feedback control greater activity. As a result, minimizing the performance index corresponds to the objective to stabilize the system (2) to zero.

To perform this task computationally, we use numerical methods to solve the system (2) and compute a feedback control $t \mapsto u_{\mathcal{K}}(t, x_{\Delta}(t))$, thus approximating an optimal pair $(x^*, u^*) \in \mathcal{V}_x \times \mathcal{V}_u$ that minimizes the performance index $\mathcal{J}(x^*, u^*)$ with the LQR approach.

1.2 Linear Quadratic Regulator

The optimal feedback control for a linear control system in state-space formulation with a quadratic cost functional is given by the linear quadratic regulator, see e.g. [34]. Moreover, non-linear systems can be stabilized with this method if the deviation from the desired trajectory is sufficiently small [45, Section 8.5]. We use this approach because it is well studied for related types of problems, e.g., convection diffusion equations [51], and demonstrates promising performance for these. For the non-linear case, first, we linearize system (2) around the reference pair $(x_{\text{ref}}, u_{\text{ref}})$. Next, we apply the LQR approach to the linearization. Then, the derived feedback control can be applied to the non-linear system.

To begin with, the linearization of system (2) can always be written in the form

$$\begin{aligned} \mathcal{M}\dot{x}^h &= \mathcal{A}x^h + \hat{\mathcal{B}}u^h, \\ y^h &= \mathcal{C}x^h. \end{aligned} \quad (4)$$

In detail, the system is considered to be semi-discretized in space with the sparse square matrices $\mathcal{A}(t), \mathcal{M}(t) \in \mathbb{R}^{n \times n}$, the input matrix $\hat{\mathcal{B}}(t) \in \mathbb{R}^{n \times m}$, and the output matrix $\mathcal{C}(t) \in \mathbb{R}^{p \times n}$. These matrices are possibly time-dependent and we assume $\mathcal{M}(t)$ to be differentiable and uniformly non-singular. Further, $x^h \in \mathcal{V}_x^h \subset \mathcal{S} \times \mathbb{R}^n$, $u^h \in \mathcal{V}_u^h \subset \mathcal{S} \times \mathbb{R}^m$, and $y^h \in \mathcal{V}_{\text{out}}^h \subset \mathcal{S} \times \mathbb{R}^p$ with $\mathcal{V}_x^h \subset \mathcal{V}_x$, $\mathcal{V}_u^h \subset \mathcal{V}_u$, and $\mathcal{V}_{\text{out}}^h \subset \mathcal{V}_{\text{out}}$ being semi-discrete subspaces. Additionally, we assume that $m, p \ll n$, such that low-rank methods are applicable in the next step.

The main task for computing an optimal feedback control is to solve a matrix-valued non-autonomous generalized differential Riccati equation (DRE), given by

$$\begin{aligned} -\frac{d}{dt}(\mathcal{M}^T \mathbf{X} \mathcal{M}) &= \mathcal{C}^T \mathcal{C} + \mathcal{A}^T \mathbf{X} \mathcal{M} + \mathcal{M}^T \mathbf{X} \mathcal{A} - \mathcal{M}^T \mathbf{X} \mathcal{B} \mathcal{B}^T \mathbf{X} \mathcal{M}, \\ \mathcal{M}(t_{\text{end}})^T \mathbf{X}(t_{\text{end}}) \mathcal{M}(t_{\text{end}}) &= S. \end{aligned} \quad (5)$$

In contrast to this general non-autonomous problem, the DRE with constant mass matrix, i.e., $\mathcal{M}(t) \equiv \hat{\mathcal{M}}$, is more common in literature [10, 13, 31, 32, 38] and requires no special treatment of the time-derivative at the left-hand side of Equation (5). Here, this time-derivative can be reformulated using the chain rule:

$$-\frac{d}{dt}(\mathcal{M}^T \mathbf{X} \mathcal{M}) = -\dot{\mathcal{M}}^T \mathbf{X} \mathcal{M} - \mathcal{M}^T \dot{\mathbf{X}} \mathcal{M} - \mathcal{M}^T \mathbf{X} \dot{\mathcal{M}}.$$

With a constant mass matrix, the terms on the left and right vanish and Equation (5) simplifies to

$$\begin{aligned} -\hat{\mathcal{M}}^T \dot{\mathbf{X}} \hat{\mathcal{M}} &= \mathcal{C}^T \mathcal{C} + \mathcal{A}^T \mathbf{X} \hat{\mathcal{M}} + \hat{\mathcal{M}}^T \mathbf{X} \mathcal{A} - \hat{\mathcal{M}}^T \mathbf{X} \mathcal{B} \mathcal{B}^T \mathbf{X} \hat{\mathcal{M}}, \\ \hat{\mathcal{M}}(t_{\text{end}})^T \mathbf{X}(t_{\text{end}}) \hat{\mathcal{M}}(t_{\text{end}}) &= S. \end{aligned} \quad (6)$$

Otherwise, by subtracting the two terms containing $\dot{\mathcal{M}}(t)$ from Equation (5) we get a DRE that contains the time-derivative of $\mathcal{M}(t)$:

$$\begin{aligned} -\mathcal{M}^T \dot{\mathbf{X}} \mathcal{M} &= \mathcal{C}^T \mathcal{C} + (\dot{\mathcal{M}} + \mathcal{A})^T \mathbf{X} \mathcal{M} + \mathcal{M}^T \mathbf{X} (\dot{\mathcal{M}} + \mathcal{A}) - \mathcal{M}^T \mathbf{X} \mathcal{B} \mathcal{B}^T \mathbf{X} \mathcal{M}, \\ \mathcal{M}(t_{\text{end}})^T \mathbf{X}(t_{\text{end}}) \mathcal{M}(t_{\text{end}}) &= S. \end{aligned} \quad (7)$$

Here, the coefficient matrices are time-dependent and so is the $\mathbb{R}^{n \times n}$ -valued solution $\mathbf{X}(t)$. Note that the input matrix $\mathcal{B}(t) = \frac{1}{\sqrt{\lambda}} \hat{\mathcal{B}}(t)$ is scaled with the weight factor from the performance index (3).

With the solution of Equation (7), we can compute the feedback gain matrix

$$\mathcal{K} = \frac{1}{\lambda} \hat{\mathcal{B}}^\top \mathbf{X} \mathcal{M} = \frac{1}{\sqrt{\lambda}} \mathcal{B}^\top \mathbf{X} \mathcal{M}.$$

Then, we obtain the feedback control $u_{\mathcal{K}}^h(t, x_{\Delta}^h(t))$ that minimizes the performance index (3) by applying $-\mathcal{K}(t)$ to the semi-discrete state deviation $x_{\Delta}^h(t)$ as in [11, 34, 45]:

$$u_{\mathcal{K}}^h = -\mathcal{K} x_{\Delta}^h. \quad (8)$$

This semi-discrete feedback control can then be applied to the non-linear system (2).

One main contribution of this manuscript are the numerical methods to efficiently solve the DRE (7) and, thus, compute the feedback gain $\mathcal{K}(t)$. Such methods, which can also treat the time-dependent coefficients, are presented in the next section.

2 DRE Solver

In this section, we present methods to numerically compute a low-rank approximation of the solution $\mathbf{X}(t) \in \mathbb{R}^{n \times n}$ of the DRE (7). For the autonomous DRE (6), i.e., when all matrix coefficients are constant in time, many efficient low-rank methods exist, as we have discussed in the introduction. These methods are mostly tailored for the autonomous DRE and not yet suited to handle time-dependent coefficients. Here, we introduce the BDF methods and splitting schemes adapted to the non-autonomous DRE (7).

To this end, we introduce an equidistant temporal discretization consisting of the $n_t + 1$ grid points t_k defined by

$$\tau = \frac{t_{\text{end}} - t_0}{n_t}, \quad \mathcal{T} = \{t_k = t_0 + k \cdot \tau, k = 0, \dots, n_t\}.$$

The DRE is solved backwards in time starting at t_{end} . For reasons of better readability and for consistency with previously published material on the autonomous case, we perform a change of variables $t \rightarrow t_{\text{end}} + t_0 - t$ such that we can formulate the methods forward in time. The DRE to be solved is therefore

$$\begin{aligned} \mathcal{M}^\top \dot{\mathbf{X}} \mathcal{M} &= \mathcal{C}^\top \mathcal{C} + (\mathcal{A} + \dot{\mathcal{M}})^\top \mathbf{X} \mathcal{M} + \mathcal{M}^\top \mathbf{X} (\mathcal{A} + \dot{\mathcal{M}}) - \mathcal{M}^\top \mathbf{X} \mathcal{B} \mathcal{B}^\top \mathbf{X} \mathcal{M}, \\ \mathcal{M}(t_0)^\top \mathbf{X}(t_0) \mathcal{M}(t_0) &= \mathbf{X}_0 = S. \end{aligned} \quad (9)$$

The DRE solvers in this section then efficiently approximate $\mathbf{X}_k \approx \mathbf{X}(t_k)$, representing \mathbf{X}_k by a low-rank factorization.

2.1 Low-rank Methods

Motivated by [49], with $m, p \ll n$, we assume the solution of the DRE, $\mathbf{X}(t)$, to have a small (numerical) rank $s(t) \ll n$. Even if the matrices $\mathcal{A}(t), \mathcal{M}(t) \in \mathbb{R}^{n \times n}$ are sparse, at each time (t), the solution $\mathbf{X}(t) \in \mathbb{R}^{n \times n}$ is a dense matrix. However, since it has a low numerical rank, it can be well approximated by

$$\mathbf{X} \approx \mathbf{L} \mathbf{D} \mathbf{L}^\top, \quad (10)$$

where the low-rank factors $\mathbf{L}(t) \in \mathbb{R}^{n \times s(t)}$ and $\mathbf{D}(t) \in \mathbb{R}^{s(t) \times s(t)}$ have the rank $s(t) \ll n$ [32]. With this low-rank factorization, just the low-rank factors $\mathbf{L}(t)$ and $\mathbf{D}(t)$ need to be stored. Consequently, the memory requirement for the storage of the solution reduces to $\mathcal{O}(s(t)n + s(t)^2)$ instead of $\mathcal{O}(n^2)$, per time step. This factorization technique is used by the methods we introduce in this section.

Non-autonomous BDF Method We choose the BDF methods because the existing implementations can be adapted to the non-autonomous case relatively straight forward since the original method is, in theory, not restricted to autonomous data. However, the existing BDF methods for matrix-valued DREs like [32, Algorithm 3.2] are tuned for constant data. Nevertheless, we can extend these BDF methods to solve the non-autonomous DRE (7) with a feasible amount of extra

Algorithm 1 Non-autonomous low-rank factor BDF method of order φ

Input $\mathcal{A}(t), \mathcal{M}(t), \dot{\mathcal{M}}(t), \mathcal{B}(t), \mathcal{C}(t), \lambda, \mathcal{T}, \varphi, L_0, \dots, L_{\varphi-1}, D_0, \dots, D_{\varphi-1}$
Output $\mathcal{K}_k, k = 0, \dots, n_t - 1$

- 1: Invert the direction of time in all matrix functions, e.g. $\mathcal{A}(t) \rightarrow \mathcal{A}(t_{\text{end}} + t_0 - t)$
- 2: $\mathcal{K}_{n_t-k} = \frac{1}{\sqrt{\lambda}} \mathcal{B}(t_k)^\top L_k D_k L_k^\top \mathcal{M}(t_k), \quad k = 1, \dots, \varphi - 1$
- 3: **for** $k = \varphi, \dots, n_t$ **do**
- 4: $\mathcal{A}_k = \tau \beta (\dot{\mathcal{M}}(t_k) + \mathcal{A}(t_k)) - \frac{1}{2} \mathcal{M}(t_k)$
- 5: $\mathcal{M}_k = \mathcal{M}(t_k)$
- 6: $\mathcal{B}_k = \sqrt{\tau \beta} \mathcal{B}(t_k)$
- 7: $\mathcal{C}_k^\top = [\mathcal{C}(t_k)^\top, \mathcal{M}_k^\top L_{k-1}, \dots, \mathcal{M}_k^\top L_{k-\varphi}]$
- 8:
$$\mathcal{S}_k = \begin{bmatrix} \tau \beta I_p & & & & \\ & -\alpha_1 D_{k-1} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & -\alpha_\varphi D_{k-\varphi} \end{bmatrix}$$
- 9: solve ARE (11) for L_k and D_k
- 10: $\mathcal{K}_{n_t-k} = \frac{1}{\sqrt{\lambda}} \mathcal{B}(t_k)^\top L_k D_k L_k^\top \mathcal{M}_k$

Algorithm 2 BDF Start-up Time-steps

Input $\mathcal{T}, \varphi, n_\varphi$
Output \mathcal{T}

- 1: **if** $\varphi > 2$ **then**
- 2: $\tau = t_1 - t_0, \quad \tilde{\tau} = \frac{\tau}{2^{n_\varphi}}$
- 3: $\tilde{t}_0 = t_0, \quad \tilde{t}_1 = t_0 + \tilde{\tau}$
- 4: **if** $\varphi = 3$ **then**
- 5: **for** $k = 2, \dots, n_\varphi + 2$ **do**
- 6: $\tilde{t}_k = t_0 + 2^{k-1} \tilde{\tau}$
- 7: $\mathcal{T} = \{\tilde{t}_0, \dots, \tilde{t}_{n_\varphi+2}, t_2, \dots, t_{n_t}\}$
- 8: **else if** $\varphi = 4$ **then**
- 9: **for** $k = 1, \dots, n_\varphi + 2$ **do**
- 10: $\tilde{t}_{2k} = t_0 + 2^k \tilde{\tau}$
- 11: $\tilde{t}_{2k+1} = t_0 + 3 \cdot 2^{k-1} \tilde{\tau}$
- 12: $\mathcal{T} = \{\tilde{t}_0, \dots, \tilde{t}_{2n_\varphi+5}, t_4, \dots, t_{n_t}\}$

computational cost. This results in Algorithm 1 formulated forward in time. Note that the actual numerical implementation is backward in time. An equivalent formulation of the non-autonomous BDF methods backward in time can be found in [7].

The inputs of Algorithm 1 are the time-dependent matrices $\mathcal{A}(t), \mathcal{M}(t), \dot{\mathcal{M}}(t), \mathcal{B}(t), \mathcal{C}(t)$, the time grid \mathcal{T} , and the order $\varphi \in \{1, 2, 3, 4\}$ of the BDF method. Further, the low-rank factors $L_0, \dots, L_{\varphi-1}, D_0, \dots, D_{\varphi-1}$ of the initial values $\mathbf{X}_0, \dots, \mathbf{X}_{\varphi-1}$ are required as inputs with sufficient accuracy to obtain the desired order of convergence φ .

These initial values could be computed with a different DRE solver of at least order φ . Instead, we compute the initial solution snapshots with sufficiently small time steps of lower, but successively increasing, order BDF in a wind-up procedure. This way, the method becomes entirely self-contained. No extra time steps are required for $\varphi = 1$, where the initial solution \mathbf{X}_0 , which is available, is sufficient and for $\varphi = 2$, where the additional initial solution \mathbf{X}_1 can be computed with one BDF 1 step. But, we generate extra time steps for $\varphi \in \{3, 4\}$ with a specific wind-up procedure.

We summarize this wind-up procedure in Algorithm 2, which has the inputs \mathcal{T}, φ , and n_φ . With these, the smallest additional time step is of length $\tilde{\tau} = \frac{t_1 - t_0}{2^{n_\varphi}}$. This is used to compute small BDF steps with step size $\tilde{\tau}$ and increasing order from 1 up to $\varphi - 1$. Then, BDF steps are computed with time step lengths that are doubled n_φ times until the time steps coincide with the first $t_0, \dots, t_{\varphi-1}$. Usually, we choose $n_\varphi = 10$. Independently of n_t , this results in 10 extra time steps for BDF 3

and 22 for BDF 4. For implementation details of [Algorithms 1](#) and [2](#), see [\[43, mess_bdf_dre.m\]](#).

In each step of the BDF method, we solve an algebraic Riccati equation (ARE)

$$0 = \mathcal{C}_k^\top \mathcal{S}_k \mathcal{C}_k + (\dot{\mathcal{M}}_k + \mathcal{A}_k)^\top \mathbf{X}_k \mathcal{M}_k + \mathcal{M}_k^\top \mathbf{X}_k (\dot{\mathcal{M}}_k + \mathcal{A}_k) - \mathcal{M}_k^\top \mathbf{X}_k \mathcal{B}_k \mathcal{B}_k^\top \mathbf{X}_k \mathcal{M}_k \quad (11)$$

for the low-rank factors $L_k D_k L_k^\top \approx \mathbf{X}_k$. The matrices in [Equation \(11\)](#) are constructed in [Lines 4 to 8](#) of [Algorithm 1](#) and the coefficients α_j and β can be found in, e.g., [\[3, Table 5.2\]](#). Instead of storing the solutions for each time step, [Algorithm 1](#) computes the feedback gain matrices $\mathcal{K}_k \approx \mathcal{K}(t_k), k = 1, \dots, n_t$ directly, which further decreases the storage requirements to $O(mn)$ per time step. Note that we look at the on-disk storage after completion of the computation. Intermediately, we may need more memory. However, we use the Newton alternating-direction implicit (Newton-ADI) method to solve the ARE [\(11\)](#), which can accumulate the feedback gain matrices directly avoiding assembling the low-rank solution factors. This is a significant performance advantage in terms of the memory requirements during the computations, especially for large-scale DREs.

There are several new developments in [Algorithm 1](#) compared to [\[32, Algorithm 3.2\]](#). First, the method has the ability to handle non-autonomous DREs with time-dependent matrices and, especially, non-constant $\mathcal{M}(t)$. Also, the start-up with $\mathbf{X}_0, \dots, \mathbf{X}_\varphi$ for order φ with [Algorithm 2](#) is independent of any other DRE solver. Altogether, the extra computational cost of the non-autonomous method are the memory requirements for the time-dependent coefficient matrices and the matrix-matrix multiplications $\mathcal{M}_k^\top L_j$ in [Line 7](#). In the autonomous case, only $\hat{\mathcal{M}}^\top L_{k-1}$ is computed and $\hat{\mathcal{M}}^\top L_{k-2}, \dots, \hat{\mathcal{M}}^\top L_{k-\varphi}$ are reused from the previous time steps since $\hat{\mathcal{M}}$ is constant in time. Further, [Algorithm 1](#) is embedded in the freely available software package M-M.E.S.S. ([\[43\]](#)), where it benefits from efficient solvers for the ARE [\(11\)](#) and can handle various forms of structures in the coefficient matrices.

Non-autonomous Splitting Schemes. The splitting schemes for autonomous DREs utilize the fact that there are closed-form solutions for the two subproblems that arise when treating the affine and the quadratic part of the equation separately. This idea can be partially extended also to the non-autonomous case. The relevant subproblems are

$$\mathcal{M}^\top \dot{\mathbf{X}}_F \mathcal{M} = \mathcal{C}^\top \mathcal{C} + (\mathcal{A} + \dot{\mathcal{M}})^\top \mathbf{X}_F \mathcal{M} + \mathcal{M}^\top \mathbf{X}_F (\mathcal{A} + \dot{\mathcal{M}}), \quad (12)$$

$$\mathcal{M}^\top \dot{\mathbf{X}}_G \mathcal{M} = -\mathcal{M}^\top \mathbf{X}_G \mathcal{B} \mathcal{B}^\top \mathbf{X}_G \mathcal{M}. \quad (13)$$

The solution to [Equation \(13\)](#) is given implicitly by

$$\mathbf{X}_G(t) = \left(I + \mathbf{X}_G(s) \int_s^t \mathcal{B}(\tau) \mathcal{B}(\tau)^\top d\tau \right)^{-1} \mathbf{X}_G(s) \quad (14)$$

for $s \leq t$. This is easily seen by multiplying by the inverted term from the left and then differentiating. In the autonomous case (see [\[48\]](#)), the corresponding formula is

$$\mathbf{X}_G(t) = \left(I + (t-s) \mathbf{X}_G(s) \mathcal{B} \mathcal{B}^\top \right)^{-1} \mathbf{X}_G(s),$$

and the only difference when solving this subproblem in the non-autonomous case is, thus, that the integral $\int_s^t \mathcal{B}(\tau) \mathcal{B}(\tau)^\top d\tau$ needs to be evaluated rather than simply $\mathcal{B} \mathcal{B}^\top$.

For the affine subproblem, we introduce the two-parameter semigroup

$$\mathcal{T}(r, t) = \exp \left(\int_r^t ((\mathcal{A}(\tau) + \dot{\mathcal{M}}(\tau)) \mathcal{M}(\tau)^{-1})^\top d\tau \right).$$

Denote the integrand by $\mathcal{Q}(\tau)$. If $\mathcal{Q}(t)\mathcal{Q}(s) = \mathcal{Q}(s)\mathcal{Q}(t)$ for any t and s , then also the integrals commute:

$$\int_r^t \mathcal{Q}(\tau) d\tau \int_r^s \mathcal{Q}(\tau) d\tau = \int_r^s \mathcal{Q}(\tau) d\tau \int_r^t \mathcal{Q}(\tau) d\tau.$$

As a consequence, $((\mathcal{A}(t) + \dot{\mathcal{M}}(t)) \mathcal{M}(t)^{-1})^\top$ commutes with $\mathcal{T}(r, t)$ and

$$\frac{d}{dt} \mathcal{T}(r, t) = ((\mathcal{A}(\tau) + \dot{\mathcal{M}}(\tau)) \mathcal{M}(\tau)^{-1})^\top \mathcal{T}(r, t).$$

Then, the solution to Equation (12) is given implicitly by

$$\mathbf{X}_F(t) = \mathcal{T}(r, t)\mathbf{X}_F(r)\mathcal{T}(r, t)^\top + \int_r^t \mathcal{T}(s, t)\mathcal{M}(s)^{-\top}\mathcal{C}(s)^\top\mathcal{C}(s)\mathcal{M}(s)^{-1}\mathcal{T}(s, t)^\top ds, \quad (15)$$

for $t_0 \leq r \leq t$. This follows quickly by differentiating each of the two terms separately using the above identity and noting that the first term satisfies Equation (12) without the $\mathcal{C}(t)^\top\mathcal{C}(t)$ term, and the second satisfies Equation (12) but with $X_F(r) = 0$.

The condition that

$$\mathcal{Q}(t)\mathcal{Q}(s) = \mathcal{Q}(s)\mathcal{Q}(t) \quad (16)$$

for any t and s does not hold in general, and needs to be verified. Note that it does hold, e.g., when $\mathcal{A}(t) = \alpha(t)\bar{\mathcal{A}}$ and $\mathcal{M}(t) = \mu(t)\bar{\mathcal{M}}$, with constant matrices $\bar{\mathcal{A}}$ and $\bar{\mathcal{M}}$ and scalar functions $\alpha(t)$ and $\mu(t)$. This is a common application, e.g., when considering heat flow with a variable thermal conductivity. In case a different approach can be used to solve Equation (12) efficiently, the condition in Equation (16) is not required.

In the low-rank setting, we can utilize the Sherman-Morrison-Woodbury matrix inversion lemma to express the solution X_G to the nonlinear subproblem (13) in low-rank form. Given that $\mathbf{X}_G(s) = L_G D_G L_G^\top$, from Equation (14) we get

$$\mathbf{X}_G(t) = L_G \left(I + D_G L_G^\top \int_s^t \mathcal{B}(\tau)\mathcal{B}(\tau)^\top d\tau L_G \right)^{-1} D_G L_G^\top. \quad (17)$$

Since the integral does not depend on $\mathbf{X}_G(t)$ and the integrand has a low-rank factorization for each τ , it can easily and efficiently be approximated by applying a quadrature formula followed by a column compression step. This can potentially even be done in an offline phase.

For the affine problem, the constant term in Equation (15) is in low-rank form if the initial condition $\mathbf{X}_F(r)$ is, and so is the integrand in the integral term. We can, therefore, apply a quadrature rule to approximate the integral and compress the columns of the resulting sum to acquire an approximation to $\mathbf{X}_F(t)$. For each of the terms of this sum, we need to evaluate $\mathcal{T}(c, t)L$ for different values of c and L . For this, we utilize the fact that, at time $s = t$, it is the solution to the system

$$\mathcal{M}(s)^\top \dot{Y}(s) = (\mathcal{A}(s) + \dot{\mathcal{M}}(s))^\top Y(s), \quad Y(r) = L. \quad (18)$$

In the autonomous case, a single integral can be approximated once and then reused in every splitting step, but in the non-autonomous case we need to approximate an integral term in every step. In the autonomous case, we may additionally use the semigroup property $T(r, t)L = T(r, s)T(s, t)L$ to avoid recomputing the same values several times in the quadrature formula. This is no longer possible in the non-autonomous case, because the semigroup will be applied to different L at different quadrature nodes. As a consequence, most of the benefits arising from the splitting, that make these methods very competitive in the autonomous case, are lost in the non-autonomous case.

We summarize the procedure in Algorithm 3. For clarity, we only consider the Lie splitting case, where we first solve the nonlinear subproblem over a full time step, then the affine subproblem over a full time step. Other splitting schemes will use different combinations of these subproblems, as can be found in [48]. However, it is straightforward to adapt Algorithm 3 to those cases. The output of Algorithm 3 is the same as in Algorithm 1, and so are the basic inputs. The only difference is in the parameters $\{c_j\}_{j=1}^s$ and $\{w_j\}_{j=1}^s$, which denote the quadrature nodes and weights of a quadrature formula.

3 Time stepping Schemes for Closed-loop integrators

Once we have computed a feedback gain $\mathcal{K}(t)$ with a method from the previous section, we can apply it to the solution of the non-linear closed-loop system (2), as denoted in Equation (8), and get a closed-loop input $u_{\mathcal{K}}^h(t, x_{\Delta}^h(t))$. In that case, the numerical simulation of this closed-loop system can be challenging if $u_{\mathcal{K}}^h(t, x_{\Delta}^h(t))$ has very large variation as we demonstrate in Section 4, where numerical issues can cause blow-ups. Similar numerical issues occur in [20] for FSI problems

Algorithm 3 Non-autonomous (Lie) splitting scheme

Input $\mathcal{A}(t), \mathcal{M}(t), \dot{\mathcal{M}}(t), \mathcal{B}(t), \mathcal{C}(t), \lambda, \mathcal{T}, L_0, D_0, \{c_j\}_{j=1}^s, \{w_j\}_{j=1}^s$
Output $\mathcal{K}_k, k = 0, \dots, n_t - 1$

- 1: Invert the direction of time in all matrix functions, e.g., $\mathcal{A}(t) \rightarrow \mathcal{A}(t_{\text{end}} + t_0 - t)$
- 2: **for** $k = 0, \dots, n_t - 1$ **do**
- 3: Set $\hat{L} = L_k$ and compute $\hat{D} = \left(I + D_k L_k^\top \int_{t_k}^{t_{k+1}} \mathcal{B}(\tau) \mathcal{B}(\tau)^\top d\tau L_k \right)^{-1} D_k$ as in [Equation \(17\)](#)
- 4: Compute $\tilde{L} = \mathcal{T}(t_k, t_{k+1}) \hat{L}$ by solving [Equation \(18\)](#) and set $\tilde{D} = \hat{D}$
- 5: Approximate the integral term in [Equation \(15\)](#) by quadrature $\sum_{j=1}^s \tilde{L}_j \tilde{D}_j \tilde{L}_j^\top$, where $\tilde{L}_j = T(t_k + c_j, t_{k+1}) \mathcal{M}(t_k + c_j)^{-\top} \mathcal{C}(t_k + c_j)^\top$, $\tilde{D}_j = \tau_k w_j I$
- 6: Column compress $L_{k+1} D_{k+1} L_{k+1}^\top$ with $L_{k+1} = [\tilde{L} \quad \tilde{L}_1 \quad \dots \quad \tilde{L}_s]$ and $D_{k+1} = \text{blkdiag}(\tilde{D}, \tilde{D}_1, \dots, \tilde{D}_s)$
- 7: $\mathcal{K}_{n_t-k-1} = \frac{1}{\sqrt{\lambda}} \mathcal{B}(t_{k+1})^\top L_{k+1} D_{k+1} L_{k+1}^\top \mathcal{M}(t_{k+1})$

with, e.g., the implicit Euler method and the trapezoidal rule. There, the numerical issues can be overcome with time-adaptive fractional-step-theta schemes. Since these FSI problems model moving interfaces and use time-dependent meshes, similar to some of the problems we discuss in [Section 4](#), we follow [\[20\]](#) and the references therein and overcome the blow-up behavior in our examples with time-adaptive fractional-step-theta schemes, as well.

To this end, we consider the system to be spatially semi-discretized with finite elements on a mesh of triangular cells $\mathcal{Q}^h(t) = \{Q(t)\}$. The cells $Q(t)$ cover the domain Ω and are time-dependent in order to treat moving interfaces with time-dependent meshes. The spatial discretization is a problem-related choice, which we specify in [Section 4](#). Part of this is the finite element space \mathcal{V}^h , which we use to define $x_\Delta^h \in \mathcal{V}^h$. With this, the closed-loop system [\(2\)](#) in semi-discretized form reads

$$\dot{x}_\Delta^h(t) = f^h(t; x_\Delta^h(t), u_{\mathcal{K}}^h(t, x_\Delta^h(t))), \quad (19)$$

where the function $f^h: [t_0, t_{\text{end}}] \times \mathbb{R}^n \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ is the semi-discretized version of the function f .

For the time discretization, we use n_t sub-intervals $[t_{k-1}, t_k]$ to partition the time-interval $[t_0, t_{\text{end}}]$ with the time grid \mathcal{T}_{fwd} :

$$\mathcal{T}_{\text{fwd}} = \{t_0, \dots, t_{n_t} = t_{\text{end}}\}, \quad \tau_k = t_k - t_{k-1}.$$

In contrast to the time discretization \mathcal{T} for the DREs, the time steps in \mathcal{T}_{fwd} are not necessarily equidistant to allow time-adaptivity, but $\mathcal{T} \in \mathcal{T}_{\text{fwd}}$. With these time steps, we define discrete approximations (x_k, u_k) by

$$x_k \approx x_\Delta^h(t_k), \quad u_k \approx u_{\mathcal{K}}^h(t_k, x_k),$$

and the discrete approximation $(x_{k,\text{ref}}, u_{k,\text{ref}})$ of the reference trajectory $(x_{\text{ref}}, u_{\text{ref}})$ on the corresponding time grid:

$$\mathcal{T}_{\text{fwd}}^{\text{ref}} = \{t_0 = t_{0,\text{ref}}, \dots, t_{n_t,\text{ref}} = t_{\text{end}}\} \in \mathcal{T}_{\text{fwd}}.$$

For the computation of these discrete approximations, we formulate different time stepping schemes in the same framework by denoting the parameter $\Theta \in [0, 1]$. With this, the closed-loop system [\(19\)](#) in discrete form can be approximated by

$$\frac{x_k - x_{k-1}}{\tau_k} = \Theta f^h(x_k, u_k) + (1 - \Theta) f^h(x_{k-1}, u_{k-1}). \quad (20)$$

In case the function f^h is nonlinear, [Equation \(20\)](#) can be solved with, e.g., a Newton method, otherwise with a direct linear solver. As a result, we can obtain various time stepping schemes by choosing different parameters Θ . For $\Theta = 0$, [Equation \(20\)](#) is the explicit Euler scheme and for $\Theta = 1$, the implicit Euler scheme. With $\Theta = 0.5$, the scheme corresponds to a trapezoidal rule¹. These are some straight-forward choices for Θ . Though, the described framework can be used to formulate more time stepping schemes as well, where Θ can vary in between time steps, as for the fractional-step-theta scheme, which is described next.

¹In different PDE contexts, like in [\[20\]](#), the scheme with $\Theta = 0.5$ is also called Crank-Nicolson

Algorithm 4 Time-adaptive Step-size Control

Input $t_k, t_{k+1,\text{ref}}, x_k, u_k, \tau_k, \text{TOL}, \gamma, r, \underline{\delta}, \bar{\delta}$
Output $\tau_{k+1}, \text{RetryStep}$

- 1: compute $\mathcal{I}_k \in \{\mathcal{I}_k^e, \mathcal{I}_k^{\Delta u}, \mathcal{I}_k^u\}$ ▷ choose one variant
- 2: $\delta = \left(\gamma \frac{\text{TOL}}{\mathcal{I}_k}\right)^r$
- 3: $\text{RetryStep} = \delta < \underline{\delta}$
- 4: **if** $\underline{\delta} < \delta < \bar{\delta}$ **then**
- 5: $\tau_{k+1} = \tau_k$
- 6: **else**
- 7: $\tau_{k+1} = \delta \tau_k$
- 8: **if** $t_k + \tau_{k+1} > t_{k+1,\text{ref}}$ **then** ▷ ensure to meet the reference time steps
- 9: $\tau_{k+1} = t_{k+1,\text{ref}} - t_k$

3.1 Fractional-step-theta Scheme

To have a more reliable and robust method regarding the numerical issues that we demonstrate in [Section 4](#), we combine time-adaptivity with the fractional-step-theta scheme. It is known to be of second-order accuracy and to be A-stable [17]. Another advantage is that it has little numerical dissipation, which is beneficial for many problems. To formulate this method in our framework, we follow [19] and define the two parameters

$$\Theta = 2 - \sqrt{2} \quad \text{and} \quad \beta = \frac{\Theta - 1}{\Theta - 2} = 1 - \sqrt{\frac{1}{2}}.$$

In each time step, we perform three sub-steps, in which we solve [Equation \(20\)](#) with the parameters

$$(\Theta_{k,0}, \Theta_{k,1}, \Theta_{k,2}) = (\Theta, 1 - \Theta, \Theta), \quad (\tau_{k,0}, \tau_{k,1}, \tau_{k,2}) = (\beta\tau_k, (1 - 2\beta)\tau_k, \beta\tau_k).$$

For each sub-step, the initial condition is the solution of the previous sub-step.

Next, we describe the time-adaptive strategies in detail, which we combine with the fractional-step-theta scheme.

3.2 Time-adaptive Strategy

The numerical behavior of the methods for solving the closed-loop system significantly depends on the discretization in space and time. Namely, an equidistant time-discretization with too large time steps (n_t too small) can lead to several numerical issues. For instance, for nonlinear FSI problems, numerical blow-ups with the implicit Euler scheme and the trapezoidal rule have been observed in [20] even though these two methods are A-stable. Related to this, similar issues can appear when solving the closed-loop system (20). This is potentially be caused, e.g., by small weight parameters λ in the cost functional (3), which can cause the feedback control $u_{\mathcal{K}}^h(t_k, x_k)$ to have very large variation. Such numerical blow-ups can be prevented by fine time-discretizations. However, since $u_{\mathcal{K}}^h(t_k, x_k)$ is unknown a priori, also the required time step size is unknown. Especially, choosing a fine, equidistant time-discretization with possibly smaller time steps than necessary can become very computationally costly.

To overcome this difficulty, we refine and coarsen the time steps adaptively in the time-intervals that might suffer from numerical blow-ups. For this, we use an indicator in order to determine whether the time steps are supposed to be larger, smaller, or stay the same. We denote the indicator as $\mathcal{I}_k = \mathcal{I}(t_k, \zeta_k, u_k)$.

This indicator is used in [Algorithm 4](#), the time-adaptive step size control, which is similar to [20, Algorithm 3]. From the indicator \mathcal{I}_k , the threshold δ is computed with the two parameters $0 < \gamma \leq 1$ and $r > 0$. The bounds, $\underline{\delta} \leq 1 \leq \bar{\delta}$, for this threshold set the interval in which the current solution is accepted and the step size stays the same. Otherwise, the algorithm decides that, either, the current solution is considered reasonable ($\text{RetryStep} = \text{False}$), if $\bar{\delta} < \delta$, and the step size for the next time step is increased since we assume that we solved with more accuracy

than necessary. Or, the solution will be recomputed (`RetryStep = True`), if $\delta < \underline{\delta}$, with the smaller step size $\delta\tau_k$. Further, [Algorithm 4](#) returns the possibly updated step size τ_{k+1} .

Additionally, we ensure that the time steps from $\mathcal{T}_{\text{fwd}}^{\text{ref}}$ are respected in [Lines 8](#) and [9](#). For these time steps, we have computed \mathcal{K}_k as described in [Section 2](#). Otherwise, we use linear interpolation to compute \mathcal{K}_k between time steps from $\mathcal{T}_{\text{fwd}}^{\text{ref}}$.

A crucial part of this strategy are the heuristic indicators \mathcal{I}_k computed in [Line 1](#). Of these, we define three different variants, which indicate whether the time step size should be changed adaptively. One is an error-based indicator, which comes with significantly higher computational costs, while the other two monitor the feedback control and have no extra computational costs.

Error-based Indicator. This indicator is based on a heuristic error estimation. First, we compute the solution $(\tilde{x}_k, \tilde{u}_k)$ with one step of the fractional-step-theta scheme with step size τ_k . Additionally, we compute (x_k, u_k) with three steps of step size $\frac{\tau_k}{3}$. Then, the indicator

$$\mathcal{I}_k^e = \|\chi(\tilde{x}_k) - \chi(x_k)\|$$

estimates the error between \tilde{x}_k and x_k . The function $\chi: \mathbb{R}^n \rightarrow \mathbb{R}^{n_e}$, with $n_e \leq n$, elects certain entries from the solution vector. It can be the identity as well ($n_e = n$).

This indicator is supposed to refine the time steps especially in time-intervals that might suffer from numerical issues, which can lead to a blow-up behavior. Since this behavior strongly depends on the time-discretization, \mathcal{I}_k^e is expected to be large in these time-intervals. However, the computation of \mathcal{I}_k^e requires extra computational effort. Three additional time steps are calculated, which quadruples the cost per time step. We use the approximate solution with the finer discretization, (x_k, u_k) , to continue the time stepping.

The two control-based indicators, that we describe next, come without the extra computational effort.

Absolute Control-based Indicator. We assume that the reference time-discretization $\mathcal{T}_{\text{fwd}}^{\text{ref}}$ is chosen such that the reference trajectory snapshots $(x_{k,\text{ref}}, u_{k,\text{ref}})$ are computed with sufficient accuracy. Thus, no further refinement is required as long as the feedback control u_k is inactive which is the case as long as the state x_k does not deviate from $x_{k,\text{ref}}$. Therefore, no numerical issues occur during the solution of the closed-loop system as long as the feedback control is inactive. On the other hand, refinement of the time steps might be necessary, especially, when u_k changes quickly.

The indicator

$$\mathcal{I}_k^{\Delta u} = \|u_k - u_{k-1}\|$$

monitors the change of the feedback control. This indicator is computationally cheap to evaluate and tailored to the solution of closed-loop systems, i.e., the feedback control, which is the focus of this manuscript.

Scaled Control-based Indicator. Since $\mathcal{I}_k^{\Delta u}$ monitors an absolute value, this indicator might be sensitive when the time step size τ_k is large and less sensitive with a small time step size. Consequently, it can be strongly problem dependent. An alternative is to monitor the scaled change of the feedback control:

$$\mathcal{I}_k^{\dot{u}} = \left\| \frac{u_k - u_{k-1}}{\tau_k} \right\|.$$

The indicator $\mathcal{I}_k^{\dot{u}}$ is an approximation to the norm of the gradient of the feedback control. Accordingly, it indicates to refine the time steps if u_k has very large variation as well. At the same time, it is more robust with respected to the time step size.

We demonstrate the behavior of the time-adaptive strategies with several numerical examples in [Section 4.2](#).

4 Numerical Experiments

In this section, we use three numerical examples to put the methods from [Sections 2](#) and [3](#) to the test and demonstrate the interplay between the closed-loop system [\(2\)](#) and the DREs. In contrast to the theory in [Section 2](#), we do not perform the change of variables in time here, such that the DREs are solved backwards in time.

With the first example in [Section 4.1](#), we compare the DRE solvers from [Section 2](#). This example is a small-scale approximation of the Laplacian with time-dependent coefficients generated artificially for this purpose. Due to the size of 25×25 , it is small enough to compute a reference solution and accurate numerical errors.

The next example models the optimal control of heat flow in a 2D cross-section of a steel rail with a time-varying thermal conductivity and thus a time-dependent \mathcal{A} -matrix. Also for this example, we can compute a reference solution, since the matrices for this example are available in different sizes, starting from dimension 109×109 and going up to $n = 79\,841$. Consequently, we are able to compare the performance of the DRE solvers in terms of error, runtime, and rank of the computed low-rank factorizations.

The third example is the two-dimensional two-phase Stefan problem, which naturally results in large-scale time-dependent matrices due to a moving interface. This type of time-dependency in the matrices causes that condition [\(16\)](#) is not fulfilled. Consequently, the splitting schemes in their current state are not applicable to this example. However, the BDF methods can handle these time-dependent matrices, and we confirm that they still behave similarly to the previous examples. While we used the previous examples only to investigate the performance of the DRE solvers, for this example we additionally apply several of the computed feedback gain matrices in the closed-loop system [\(19\)](#). Like this, we compute feedback controls for the Stefan problem. With these feedback controls, we solve the non-linear closed-loop system [\(2\)](#) using the methods from [Section 3](#). In doing so, we put special emphasis on the time-adaptive techniques, since those can prevent the solution of the closed-loop system from blowing up.

The computations are run on a computer with 2 Intel Xeon Silver 4110 (Skylake) CPUs with 8 cores each and 192 GB DDR4 RAM on CentOS Linux release 7.9.2009. The used software packages are FEniCS 2018.1.0, NumPy 1.20.3, SciPy 1.3.3, and MATLAB[®] R2017b. We provide the codes and data for our experiments in [\[5\]](#), including execution logs of the experiments and the results.

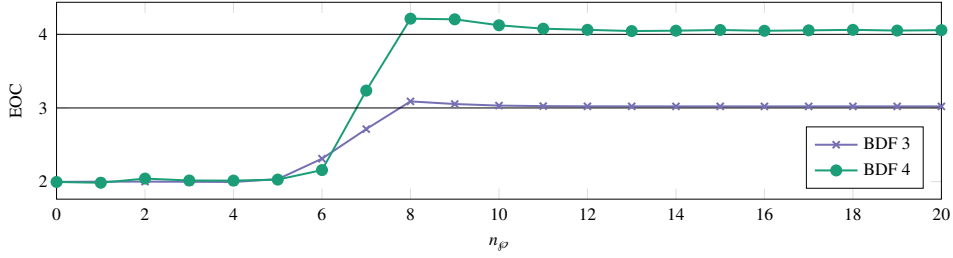
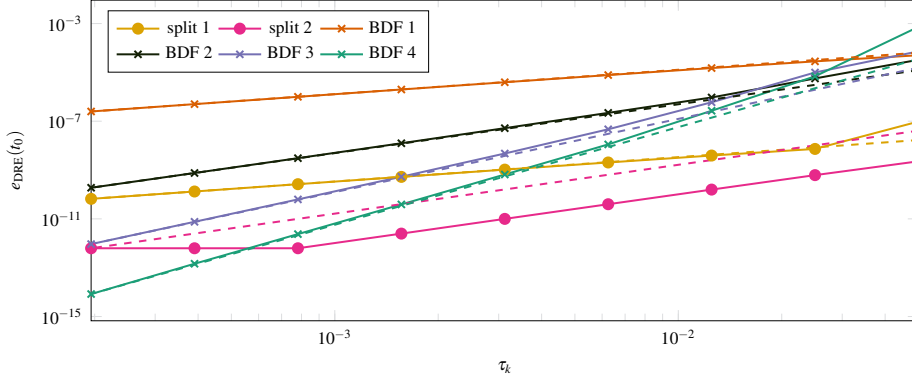
4.1 DRE Solver

A naive method to solve DREs, is to vectorize them. This means that the DRE is transformed into an equivalent vector-valued non-linear ordinary differential equation (ODE). Existing numerical solvers for ODEs can then be applied. However, this approach does not exploit the low-rank structure of the solution and is unfeasible for large-scale problems. Still, for the sufficiently small examples, we use this method and apply MATLAB's `ode*`-functions to generate a high order reference DRE solution for error comparisons.

Small-scale Academic Example. With this small example, we intend to analyze the convergence behavior of the DRE solvers. Therefore, for this example, we compute a reference solution and compare the numerical errors of the BDF methods and splitting schemes as well as the order of convergence. In particular, we show how the BDF method ([Algorithm 1](#)) of order 3 and 4 depends on the parameters for the BDF startup time steps from [Algorithm 2](#). Since they would have little meaning in this small-scale example, we do not compare the runtimes of these large-scale methods.

To generate this example, we use a finite-difference approximation of the Laplacian on the unit square $[0, 1]^2$ and assemble the matrices $\hat{\mathcal{A}} \in \mathbb{R}^{n \times n}$, $\hat{\mathcal{B}} \in \mathbb{R}^{n \times 3}$, $\hat{\mathcal{C}} \in \mathbb{R}^{1 \times n}$. For details about the implementation, see [\[5\]](#). Then, these matrices are multiplied with time-dependent scalar functions to get the time-dependent matrices

$$\begin{aligned} \mathcal{A}(t) &= \left(1 + \frac{1}{2} \sin(2\pi t)\right) \hat{\mathcal{A}}, & \mathcal{M}(t) &= \left(2 + \frac{1}{2} \sin(2\pi t)\right) \mathcal{I}_n, \\ \mathcal{B}(t) &= (3 + \cos(t)) \hat{\mathcal{B}}, & \mathcal{C}(t) &= (1 - \min(t, 1)) \hat{\mathcal{C}}. \end{aligned}$$


 Figure 1: Convergence orders of BDF with different n_φ , small scale example.

 Figure 2: Convergence of splitting and BDF and the theoretical convergence orders φ (dashed), small scale example.

Additionally, we choose $n = 25$, $[t_0, t_{\text{end}}] = [0, 0.1]$ and compute a reference approximation $\mathbf{X}_{\text{ref}}(t)$ applying `ode15s` from MATLAB after vectorization, as mentioned above. With this, in order to assess the quality of our approximate solutions, we compare the pointwise relative error:

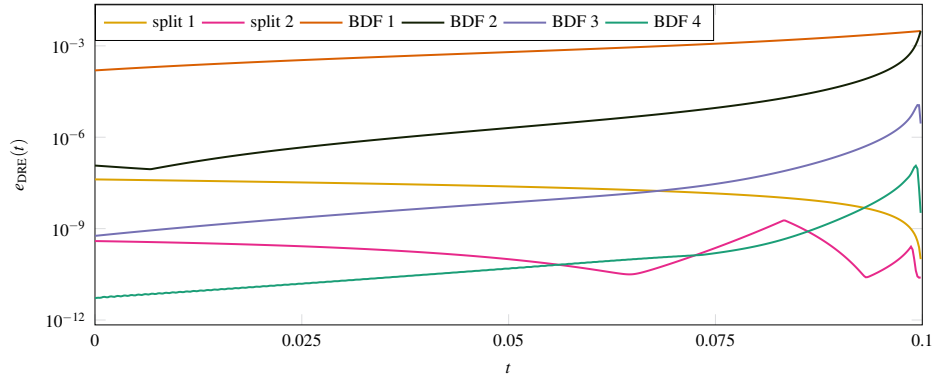
$$e_{\text{DRE}}(t) = \frac{\|\mathbf{X}(t) - \mathbf{X}_{\text{ref}}(t)\|_2}{\|\mathbf{X}_{\text{ref}}(t)\|_2}.$$

Following the theory, a BDF method with $\varphi \leq 6$ stages should converge with order φ , as well [3, Section 5.2.3, Example 5.9]. As described in Section 2.1, to achieve these convergence orders, the initialization must be sufficiently accurate. For $\varphi = 3$ (BDF 3) and $\varphi = 4$ (BDF 4), we compute the required initial values with the extra time steps from Algorithm 2. Here, we demonstrate the effect of this on the convergence order. However, for $\varphi = 1$ (BDF 1) and $\varphi = 2$ (BDF 2) no extra time steps are required.

Correspondingly, Figure 1 shows the experimental order of convergence (EOC) for BDF 3 and 4. In detail, with no extra time steps or n_φ too small both methods converge with order 2 only. At the same time, for $n_\varphi \geq 8$ both methods achieve the expected convergence order φ in this example. Moreover, larger values of n_φ have no (especially also no negative) further effect on the convergence. Note, that the choice of n_φ is problem-dependent, and that for the further computations in this section we use $n_\varphi = 10$.

With this, BDF 1, 2, 3, and 4 as well as the splitting methods of order $\varphi = 1$ (split 1) and $\varphi = 2$ (split 2) can meet their theoretical convergence orders. This can be seen in Figure 2, which displays the error of the DRE solvers for different time step sizes. The observation that the splitting methods yield significantly smaller errors than the BDF methods of the same order is also clearly visible here.

In addition, Figure 3 shows the error of the different DRE solvers over the whole time interval for a fixed time step size. The BDF methods start with larger transient errors at t_{end} . Then, the errors decrease during the backward-in-time solve of the BDF methods. Especially BDF 4 is more accurate than the splitting methods in the first half of the time-interval. The root cause is that


 Figure 3: Errors of splitting and BDF for $n_t = 512$ and $n_\varphi = 10$, small scale example.

the BDF schemes preserve fixed points, while the splitting schemes do not. At the same time, it is important to note that the errors of the splitting methods are varying less over the time-interval and stay on roughly the same level. Particularly for order $\varphi = 1$ and $\varphi = 2$, the splitting schemes outperform the BDF methods of the same order. In essence, the errors in Figure 3 are a typical behavior for these DRE solvers, which is already known for the autonomous variants of BDF and splitting methods [31, 48].

For the next example, the BDF and splitting methods benefit from the low-rank factorization of the solution since the spatial discretizations can result in large-scale DREs.

Partially Non-autonomous Steel Profile. The semi-discretized heat transfer model for optimal cooling of steel profiles [9] is an autonomous model. Hence, we modify the example to be non-autonomous and study the performance of the BDF and splitting methods for different matrix sizes and weight parameters in Equation (3). We choose this example, because the matrices \mathcal{M} , \mathcal{S} , \mathcal{M}_γ , \mathcal{B} and \mathcal{C} are available for different spatial discretizations $n \in \{109, 371, 1357, 5177, 20209, 79841\}$ (see [43, `mess_get_linear_rail.m`]) and we can easily modify the model to be non-autonomous such that it fulfills the condition, which is necessary for the splitting methods.

In detail, the model has 7 inputs and 6 outputs and the given matrices form a system equivalent to Equation (4):

$$\begin{aligned} \mathcal{M}\dot{x}^h &= \left(\frac{\kappa}{c \cdot \rho} \mathcal{S} + \frac{\gamma}{c \cdot \rho} \mathcal{M}_\gamma\right)x^h + \frac{\gamma}{c \cdot \rho} \mathcal{B}u^h, \\ y^h &= \mathcal{C}x^h. \end{aligned}$$

Here, the material parameters are the thermal conductivity $\kappa = 26.4 \frac{kg m}{s^3 \circ C}$, the heat capacity $c = 7620 \frac{m^2}{s^2 \circ C}$, the coefficient of thermal conductivity at the input boundary regions $\gamma = 7.0164 \frac{kg}{s^3 \circ C}$, and the density $\rho = 654 \frac{kg}{m^3}$. Following [14], we use the real time-interval $[0, 45s]$, which corresponds to model-time $[t_0, t_{end}] = [0, 4500]$. Further, to make this system partially non-autonomous, we augment the thermal conductivity with a non-linear, smooth, and scalar function:

$$\kappa(t) = 26.4 + 0.1 \cdot \left(2 + \cos\left(\frac{2 \cdot \pi \cdot t}{t_{end}}\right)\right) \frac{kg m}{s^3 \circ C}.$$

With this modification the example is similar to, e.g., [31]. The difference lies in the scalar function $\kappa(t)$, where we specifically picked a non-linear function. Also, due to the constant mass matrix, Algorithm 1 is equivalent to the BDF method therein. In addition, the default matrix sizes, number of time steps, and weight parameter for this example are given in Table 1. We vary those one by one to examine the behavior of the BDF and splitting methods.

Starting with different matrix dimensions n , Figure 4a displays the corresponding runtimes. The first thing to recognize is that the runtime of the BDF methods grows linearly with respect to the dimension and there is no significant difference between the orders $\varphi \in \{1, \dots, 4\}$. In

n	n_t	m	p	λ
109	128	7	6	1

Table 1: Partially non-autonomous steel profile parameters (smallest example).

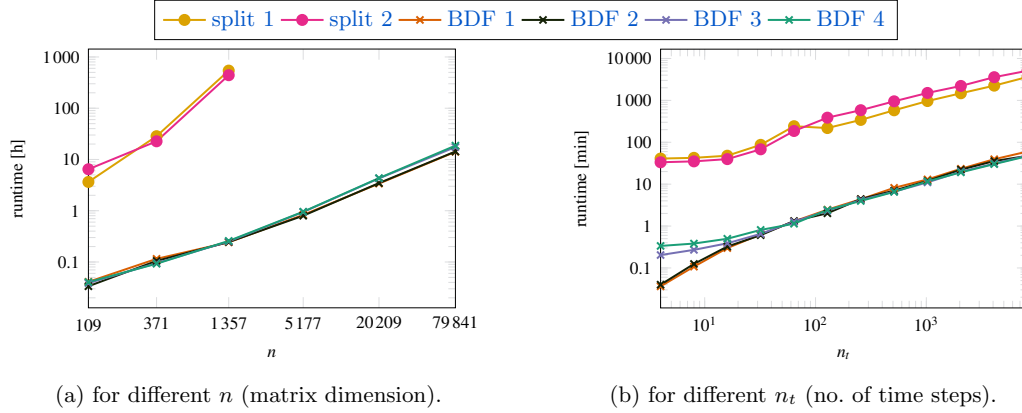


Figure 4: Runtime of splitting and BDF, steel profile example. All the methods use the same temporal grid.

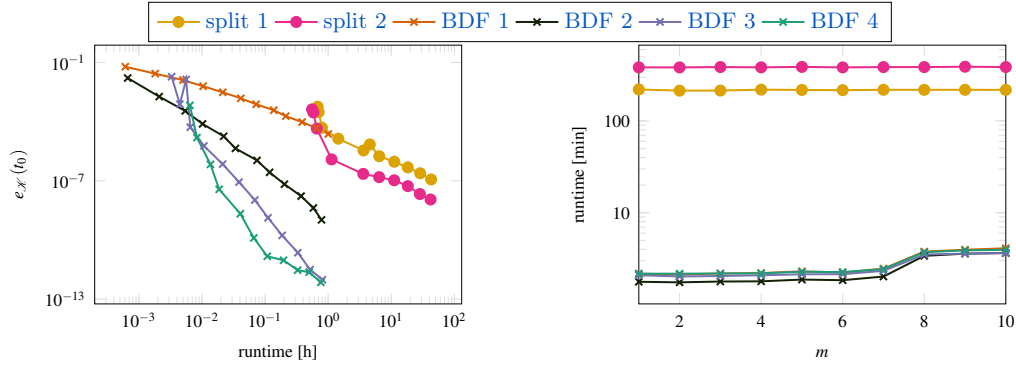
contrast, the splitting methods require considerably more runtime. As an illustration, while the BDF methods take around 2.3 minutes for the smallest dimension $n = 109$, the splitting methods require 3.6 and 6.4 hours. For $n = 371$ it is 6.1 minutes versus 28.5 and 22.6 hours for BDF and splitting, respectively. Forthwith, for dimensions $n > 1357$, we stopped the computations since the splitting methods did not finish in reasonable time. Accordingly, we continue the remaining tests with $n = 109$ to have reasonable runtimes for the splitting methods even though, like this, the problem is not large-scale.

Next, in [Figure 4b](#), for $n_t \in \{2^2, \dots, 2^{13}\}$, we compare the runtime, which grows with more time steps, as expected, for all methods. Notably, the time per time step is getting slightly smaller since the underlying subproblems converge faster. A possible explanation is, that in [Line 6](#) of [Algorithm 1](#) the quadratic term of the ARE is scaled with $\sqrt{\tau}$. Moreover, the additional overhead to compute the initial values for BDF 3 and 4 with [Algorithm 2](#) is visible for $n_t \in \{2^2, \dots, 2^6\}$.

To better evaluate the efficiency of the methods, [Figure 5a](#) shows the runtime in relation to the relative error. Similar to the small scale-example, we compute a reference solution through vectorizing the DRE. Since not all matrices are time dependent but only $\mathcal{A}(t)$ in the way that it is a scalar function times a constant matrix no stiff ODE solver is required and it is sufficient to use `ode45`. Since this example involves large-scale matrices, we do not store the DRE solution anymore but only the feedback gain matrices \mathcal{K}_k . Thus, for this example, the errors are computed by comparing \mathcal{K}_k at t_0 :

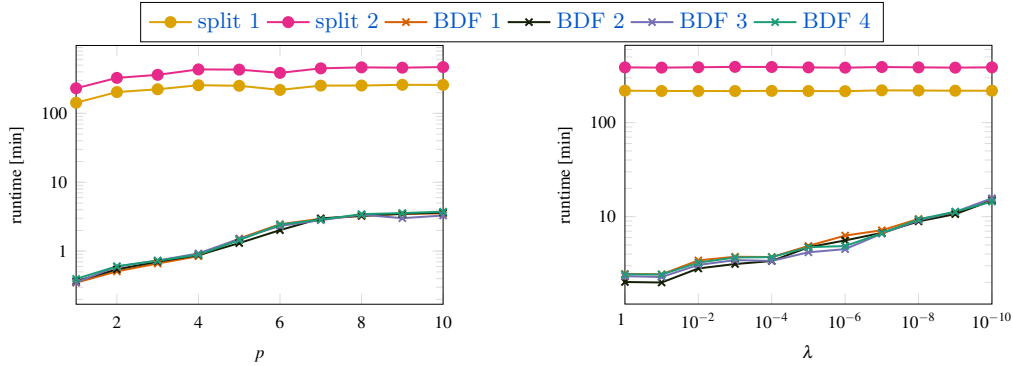
$$e_{\mathcal{K}}(t_0) = \frac{\|\mathcal{K}_0 - \mathcal{K}_{0,\text{ref}}\|_2}{\|\mathcal{K}_{0,\text{ref}}\|_2}$$

The splitting methods achieve better accuracy than the BDF methods of the same order. However, the main computational benefit of the splitting schemes is lost in the non-autonomous case. For an autonomous DRE, only one integral term needs to be approximated for the whole time interval, but now one such integral approximation is needed in each time step. The computational cost is therefore increased by a factor roughly as large as the number of time steps. This is not the case for the BDF schemes, which, therefore, in direct comparison, need significantly less computational effort. The effect can be alleviated by choosing larger tolerances for the splitting subproblem approximations, but as seen in [Figure 5a](#), where this was tried for the 2nd-order splitting scheme, this runs the risk of destroying the overall convergence behavior. We also note that the BDF methods converge faster for higher orders, while requiring almost the same runtime for the same n_t . Thus, BDF 4 is the most efficient method for an accuracy below 10^{-5} .



(a) Runtime vs. relative error (compared with ode45), $n_t = 16384$. (b) Runtime for different numbers of columns m in $\mathcal{B}(t)$.

Figure 5: Runtime of splitting and BDF, steel profile example. All the methods use the same temporal grid.



(a) for different p (rows in $\mathcal{C}(t)$). (b) for different λ (weight in cost functional).

Figure 6: Runtime of splitting and BDF, steel profile example. All the methods use the same temporal grid.

We further study the influence of the number of inputs m , the number of outputs p , and the weighting factor λ from the performance index (3) on the runtime.

By default, this model has 7 inputs and thus 7 columns in $\mathcal{B}(t)$. On the one hand, to test the methods with $m \in \{1, \dots, 7\}$, we truncate $\mathcal{B}(t)$ after m columns. On the other hand, for $m \in \{8, \dots, 10\}$, we generate additional columns that are filled with random values. As a result, Figure 5b shows that the BDF methods need slightly more time with those random columns, which are not physically motivated in the model. Besides this, the runtimes of the BDF and splitting methods are independent of m .

In contrast, the methods runtimes are affected by different numbers of rows in $\mathcal{C}(t)$. Figure 6a displays the runtimes for $p \in \{1, \dots, 10\}$. Again, we used the first 6 rows in $\mathcal{C}(t)$ that are defined in the model. This time, for $p \in \{7, \dots, 10\}$, we add additional rows for physically motivated quantities of interest, which are constructed similar to the first 6 outputs. The resulting runtimes grow with more rows in $\mathcal{C}(t)$, but the BDF methods are affected more strongly in this respect. This is, because in each BDF step, the underlying iterative ARE solver generates solutions, which have p times the number of iterations of columns. Thus, the ARE solution has significantly more columns for larger p and the column compression that we perform on these solutions is significantly more expensive.

Lastly, the weight λ in the cost functional is an important parameter. For instance, the performance of the BDF methods is strongly impacted by λ , which influences the balance between the

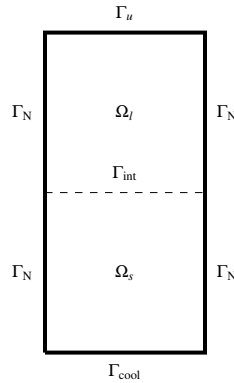


Figure 7: One instance of the domain $\Omega \subset \mathbb{R}^2$ of the Stefan problem.

quadratic and the constant term in the ARE (11). In detail, a smaller λ makes the quadratic part more dominant and consequently the inner ARE solver in the BDF methods requires more iterations to converge. This results in larger runtimes of the BDF methods for smaller λ , as illustrated in Figure 6b. In contrast, the runtimes of the splitting methods are not affected by λ since it only affects Equation (13), which is solved exactly. However, we note that the temporal discretization errors for both the BDF and splitting schemes depend on the problem being solved and thus on λ . To acquire an error of a specific size we might therefore need to use differently sized time steps for different λ . The actual computational effort required for a certain accuracy thus varies for both classes of methods. A thorough investigation of these error structures is out of the scope of this paper, but will be the subject of future research.

Further, the solutions computed by the splitting methods have a smaller numerical rank of $s(t_0) = 68$ compared to $s(t_0) = 84$ for the BDF methods. This results in smaller memory requirements for the low-rank solution factors as well as lower computational effort to compute the feedback gain matrices \mathcal{K}_k . Ultimately, we apply \mathcal{K}_k to generate a feedback control and do not need the actual DRE solution for this once we have computed \mathcal{K}_k . Thus, we can reduce the on-disk memory requirement by storing $\mathcal{K}_k \in \mathbb{R}^{n \times m}$ only, which is of fixed size and usually much smaller than the low-rank solution factors.

Two-dimensional Two-phase Stefan Problem. The Stefan problem, which is treated in [6, 7] as well, provides large-scale DREs with time-dependent matrices. This time-dependency goes beyond condition (16). For this example, we demonstrate the interplay between the feedback gain computation and the application of the feedback control during the simulation in Section 4.2.

The two-dimensional two-phase Stefan problem can model solidification and melting of pure materials. In our test setup, we consider the time-interval $[t_0, t_{\text{end}}] = [0, 1]$ and the domain Ω is a rectangle $[0, 0.5] \times [0, 1] \subset \mathbb{R}^2$ with the initial interface position at height 0.5. This interface $\Gamma_{\text{int}}(t)$ splits the domain Ω into two regions that correspond to the solid phase $\Omega_s(t)$ and the liquid phase $\Omega_l(t)$, as illustrated in Figure 7. Because the inner phase-boundary can move, its position is time-dependent. Thus, also the two regions $\Omega_s(t)$ and $\Omega_l(t)$ are time-dependent, as well as their corresponding boundary regions. They are $\Gamma_u(t)$, $\Gamma_{\text{cool}}(t)$ and $\Gamma_N(t)$ as denoted in Figure 7. The boundary region for the input $\Gamma_u(t)$ is constant in this representation of the domain. However, we do not restrict ourselves to this case and, therefore keep the time-dependence in our notation.

We follow the definition of the Stefan problem from [6], but use it in a more compact form omitting the couplings with the Navier-Stokes equations and the interface graph formulation. Both introduce additional nonlinearities to the problem and the Navier-Stokes equations add additional DAE structure. Before treating these technical challenges, we study the feedback control problem in this compact form, without the couplings, to develop the general numerical strategy.

n_t	m	p	λ	n_φ
401	1	2	10^{-4}	10

Table 2: Two-dimensional two-phase Stefan problem parameters.

The temperature is denoted as $\Theta(t)$ and modeled by [Equation \(21\)](#):

$$\frac{d}{dt}\Theta - \Upsilon \cdot \nabla\Theta - \alpha\Delta\Theta = 0, \quad \text{on } (0, t_{\text{end}}] \times \Omega, \quad (21a)$$

$$\partial_{\mathbf{n}}\Theta = u, \quad \text{on } (0, t_{\text{end}}] \times \Gamma_u, \quad (21b)$$

$$\Theta = \Theta_{\text{cool}}, \quad \text{on } (0, t_{\text{end}}] \times \Gamma_{\text{cool}}, \quad (21c)$$

$$\Delta\Upsilon = 0, \quad \text{on } (0, t_{\text{end}}] \times \Omega, \quad (21d)$$

$$\Upsilon = \left(\frac{1}{\ell}[k(\nabla\Theta)]_l^s\right) \cdot \mathbf{n}_{\text{int}}, \quad \text{on } (0, t_{\text{end}}] \times \Gamma_{\text{int}}. \quad (21e)$$

In [Equations \(21a\)](#) and [\(21e\)](#), the temperature is coupled with the extended interface movement $\Upsilon(t)$, which we model with the system of algebraic equations [\(21d\)](#) and [\(21e\)](#). On the interface, $\Upsilon|_{\Gamma_{\text{int}}}(t) = \Upsilon_{\text{int}}(t)$ is the interface movement in the normal direction where $\mathbf{n}_{\text{int}}(t)$ is the unit normal vector pointing from $\Omega_s(t)$ to $\Omega_l(t)$. The extended interface movement $\Upsilon(t)$ is the smooth extension of $\Upsilon_{\text{int}}(t)$ to the whole domain. We use $\Upsilon(t)$ to implement a spatial discretization that respects the position of $\Gamma_{\text{int}}(t)$ and follows the interface movement, while preventing mesh tangling and too strong deformations.

[Equation \(21e\)](#) is the Stefan condition, with the latent heat constant ℓ and the jump of the temperature gradient

$$[k(\nabla\Theta)]_l^s = k_s \partial_{\mathbf{n}_{\text{int}}}\Theta|_{\Omega_s} - k_l \partial_{-\mathbf{n}_{\text{int}}}\Theta|_{\Omega_l}, \quad (22)$$

across $\Gamma_{\text{int}}(t)$. Thus, $\Upsilon(t)$ depends on $\Theta(t)$ as well as the heat conductivities in the solid phase k_s and in the liquid phase k_l collected in $\alpha(t)$:

$$\alpha = \begin{cases} k_s, & \text{on } \Omega_s, \\ k_l, & \text{on } \Omega_l. \end{cases}$$

Due to this temperature-dependence, the terms $\Upsilon \cdot \nabla\Theta$ and $\alpha\Delta\Theta$ in [Equation \(21a\)](#) are non-linear. Further, in [Equation \(21b\)](#), we apply the control $u(t)$ as a Neumann condition on the control boundary $\Gamma_u(t)$. Additionally, [Equation \(21c\)](#) is a Dirichlet condition on the cooling boundary $\Gamma_{\text{cool}}(t)$ with the constant Θ_{cool} . A more detailed description including all boundary conditions, material parameters, details on the linearization and discretization, and the matrix assembly for the DRE can be found in [\[7\]](#).

In summary, the Stefan problem is a non-linear system of DAEs with a moving interface $\Gamma_{\text{int}}(t)$, which is represented by a moving mesh. Our aim is to control the position of $\Gamma_{\text{int}}(t)$.

This example results in a large-scale problem with dimension $n = 11\,429$ (max. cell width: 0.0268, min. cell width: 0.0061, initial mesh). For the Stefan problem example, we evaluate the runtime performance of the BDF methods in relation to different parameters and matrices $\mathcal{B}(t)$ and $\mathcal{C}(t)$. However, the splitting methods are not applicable to this example since the time-dependent matrices do not fulfill the necessary assumptions in [Equation \(16\)](#). To examine their influence on the BDF methods, we individually change the default parameters from [Table 2](#).

First, we vary the number of time steps n_t from 101 to 8001. As a result, the runtime of the BDF methods grows linearly with n_t as can be seen in [Figure 8](#). Another key point is that, the runtimes for $\varphi = 1, \dots, 4$ are similar. For $n_t = 8001$, BDF 3 and 4 are even slightly faster than BDF 1 and 2. This results from the fact that the underlying ARE solver (Newton-ADI) requires fewer iterations to converge in these cases. A possible explanation is that the solutions computed by BDF 3 and 4 are more accurate resulting in faster convergence.

Then, as for the example of the steel profile, we vary the number of inputs m , the number of outputs p , and the weight λ in the cost functional [\(3\)](#). The results are consistent with the findings from the previous example and the resulting figures can be found in the supplementary material. Moreover, the average numerical ranks of the approximated solutions are $\text{avg}(s(t)) = 28$

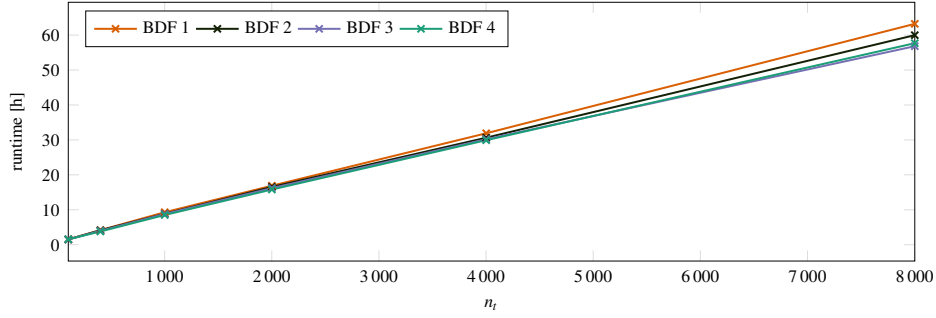


Figure 8: Runtime of BDF for different n_t (no. of time steps), Stefan problem example.

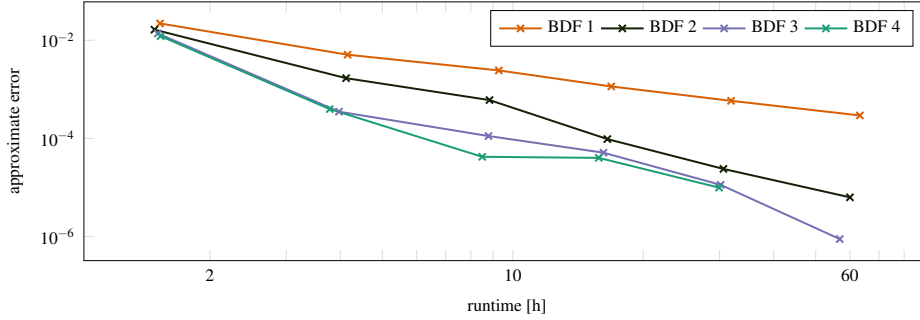


Figure 9: Runtime vs. approximate error (compared with BDF 4, $n_t = 8001$) of BDF, Stefan problem example. All the methods use the same temporal grid.

for BDF 1 and 2, $\text{avg}(s(t)) = 27.9$ for BDF 3, and $\text{avg}(s(t)) = 27.8$ for BDF 4 (for details see supplementary material).

In contrast to the previous example, we evaluate only the low-rank methods since the transformation to an ODE is infeasible to compute. Thus, we do not compare any accurate errors in the sense that we compare to a reference solution for the DRE. Instead, we assume the BDF 4 method with the finest time-discretization ($n_t = 8001$) to be most accurate and use that as the reference solution for the comparison in Figure 9. Similar to the results in Figure 5a, the BDF methods are more efficient with higher orders. However, the difference is smaller especially between BDF 3 and 4, which might not fully meet their order of convergence.

In the next section, we will use several of the feedback gain matrices that we computed in this paragraph and apply them in the closed-loop system.

4.2 Closed-loop System

In this section, we use the time stepping methods from Section 3 to numerically solve the non-linear Equation (2) for the two-dimensional two-phase Stefan problem from the previous paragraph. In this case, we introduce a perturbation to Equation (21) by augmenting the Dirichlet boundary condition (21c) at the bottom of the domain with the function $\varphi(t)$:

$$\Theta = \Theta_{\text{cool}} + \varphi, \quad \text{on } (0, t_{\text{end}}] \times \Gamma_{\text{cool}}.$$

Here, the control objective is to stabilize the interface position from the discrete reference approximation $(x_{k,\text{ref}}, u_{k,\text{ref}})$ and the perturbation $\varphi(t)$ drives the interface away from the desired trajectory in the experiments in this section. For this, we use a single input ($m = 1$) at the top of the domain as depicted in Equation (21b). Note, that with the present formulation of the Stefan problem in Equation (21) the interface position can not be measured explicitly. Thus, the interface position does not enter the performance index (3). With this in mind, we use two temperature measurements at the walls Γ_N of the domain ($p = 2$) as output. These two outputs are intended to indicate the deviation of the interface from the reference position.

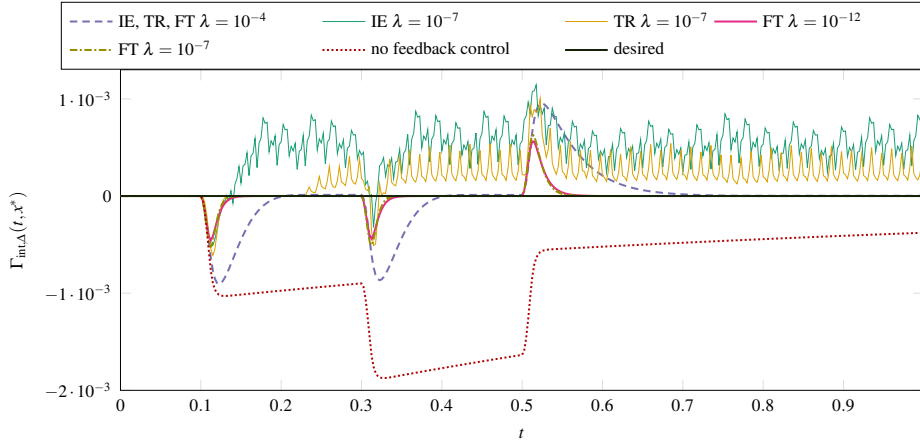


Figure 10: Perturbed interface position relative to the reference trajectory with different feedback controls.

As a result, we obtain several feedback controls by applying the feedback gain matrices computed in Section 4.1 with different weight factors λ .

With these feedback controls, we test the time stepping methods implicit Euler (IE), trapezoidal rule (TR) and our fractional-step-theta scheme (FT) from Section 3.1 to numerically solve Equation (2). For FT, we further compare the different time-adaptivity strategies. Those are the error-based indicator (FT-err), the absolute control-based indicator (FT-u) and the scaled control-based indicator (FT-dt-u). Additional examples for the Stefan problem with more emphasis on different numbers of inputs and outputs as well as different types of perturbations can be found in [7]. From there, we use the simpler example since the numerical effects that are highlighted in this manuscript are more clearly visible. Further, for the more complicated example from [7], IE as well as TR are not able to compute a feedback control that successfully stabilizes the interface position due to the issues that are discussed next.

Figure 10 displays the relative interface position $\Gamma_{\text{int},\Delta}(t, x^*) = \Gamma_{\text{int}}(t, x^*) - \Gamma_{\text{int,ref}}(t, x^*)$ at the point x^* on the interface which has the largest deviation from the reference trajectory. With a weight factor larger than $\lambda = 10^{-4}$, the computed feedback control shows very low activity and is not able to steer the interface back towards the reference trajectory in the computed time-frame. Decreasing λ means that the performance index (3) is more dominated by the output deviation and the control cost term has less impact. Thus, a smaller λ leads to a more active feedback control, which prevents the interface from deviating. In Figure 10 this is demonstrated with $\lambda = 10^{-4}$ for IE, TR, and FT. Note that all three lead to very similar results and are summarized in one line for this value of λ .

For smaller λ , the interface can be steered back even faster but IE and TR show a numerical blow-up behavior similar to the experiments in [20]. This is shown in Figure 10 for $\lambda = 10^{-7}$. In our experiments, these blow-ups occur when the feedback control has very large variation and the time step size is too large to properly resolve this.

To overcome this issue, we use time-adaptive FT-dt-u. In Figure 10, this is demonstrated with $\lambda = 10^{-12}$ and FT-dt-u (TOL = 10^{-2}). Here, the computed feedback control is able to steer the interface position back to the reference trajectory shortly after the perturbation.

The time-adaptive FT-dt-u comes with extra computational cost compared to IE and TR, which both require 15.6 minutes to numerically solve the closed-loop system with $n_t = 401$ time steps. On the other hand, for the same closed-loop system ($\lambda = 10^{-4}$), FT-dt-u computes 5761 time steps and requires 546.3 minutes. This is partly due to the long time that is needed to steer the interface back to the desired trajectory. Hence, the control is active in this time period and the time-adaptivity is very expensive. In contrast, with $\lambda = 10^{-12}$, the interface is back to the reference position in shorter time and the feedback control becomes inactive earlier as well as the time-adaptivity. Thus, FT-dt-u computes 2405 time steps and requires 232.3 minutes. However, IE and TR fail with $n_t = 401$ and $\lambda \leq 10^{-7}$. The number of time steps $n_t = 401$ is chosen as

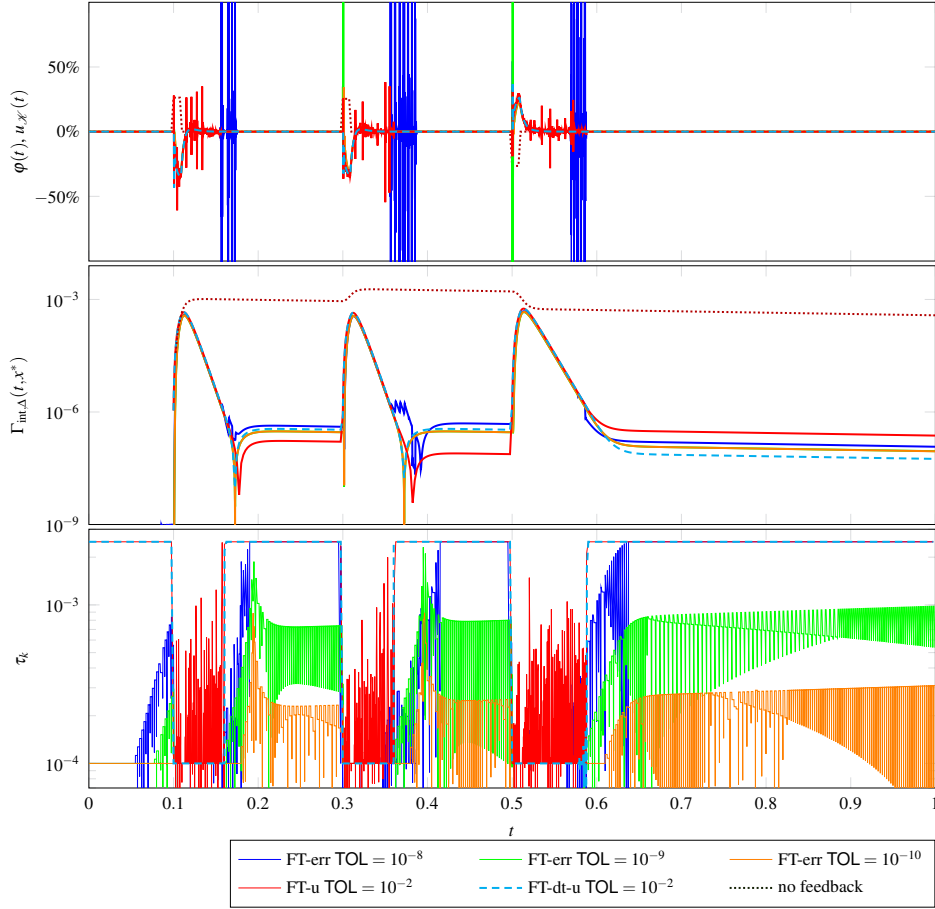


Figure 11: Perturbation and different time-adaptive feedback controls (clipped, max. value: 8.4, min. value: -39.3) (top), Perturbed interface position (center), and time step sizes (bottom) for $\lambda = 10^{-12}$.

default since we observed that this is the smallest n_t for which the error of the DRE solution and the simulation of the reference trajectory stabilize.

The corresponding feedback control that is computed with FT-dt-u for $\lambda = 10^{-12}$ is shown in comparison with the feedback controls computed with the different time-adaptive strategies FT-u and FT-err for the same closed-loop system in Figure 11 (top). Further, the figure displays the perturbation $\varphi(t)$ as a dotted line. To clarify the y-axis labeling, it is important to note that it is relative to the original boundary value Θ_{cool} .

An important parameter for the adaptive step size computation in Algorithm 4 is TOL. With smaller TOL, Algorithm 4 computes smaller time step sizes when the indicator grows. On the one hand, a larger TOL can save some time steps by increasing the time step size earlier with dropping indicator values and reduce the computational effort. On the other hand, Algorithm 4 can fail to prevent a numerical blow-up if TOL is chosen too large and the time step sizes grow too early. This effect is presented in Figure 11 for FT-u with TOL = 10^{-2} and FT-err with TOL = 10^{-8} . Both increase the time step size when the feedback control is nearly inactive but still active enough to lead to a blow-up. This blow-up then leads to an increase of the indicator, and consequently a decrease of the time step size such that the feedback controls become inactive again. The result is that the feedback control is quickly bouncing back and forth between large and small values.

In contrast to this, for FT-dt-u, the indicator monitors the relative change of the feedback control. Thus, as long as the feedback control is active the algorithm sets τ_k to the minimum value ($\tau_k = 0.0001$). However, when the control is inactive the algorithm sets τ_k back to the maximum

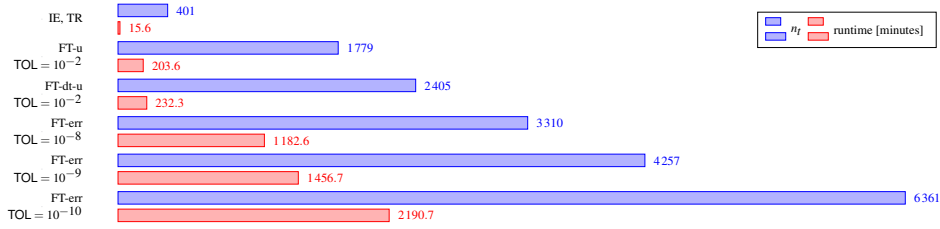


Figure 12: Number of time steps n_t and runtimes of IE, TR, and FT with different time-adaptive strategies, $\lambda = 10^{-12}$.

value ($\tau_k = 0.0025$) as can be seen in Figure 11 (bottom, dashed line). We expect this to be the general behavior of FT-dt-u if TOL is chosen sufficiently small.

To emphasize a meaningful comparison, for FT-u, the indicator monitors the absolute change of the feedback control. This means, with a smaller TOL, it behaves the same as FT-dt-u ($\text{TOL} < 10^{-5}$). In particular, it depends on the magnitude of the feedback control values, and the right choice of TOL is therefore strongly problem dependent. From our numerical experiments, we expect FT-dt-u to be more robust with respect to the magnitude of the feedback control values and, thus, the problem dependent choice of TOL. Further, we can choose a smaller TOL to be on the safe side and get the same n_t since it jumps back and forth from the maximum to the minimum τ_k .

Different from the other two, the indicator FT-err monitors the difference of the computed solutions for two different time step sizes. This results in a larger computational effort for a single time step compared to the alternative indicators. For instance, if the feedback control has very large variation and consequently the solution changes more, then τ_k is reduced. However, FT-err can not distinguish between high activity of the reference solution and of the perturbed solution. In addition, it enlarges τ_k gradually. Altogether, it requires significantly smaller TOL, larger n_t , and, consequently, increases the runtime to prevent a blow-up. To illustrate this, the computed time steps for three different choices of TOL are presented in Figure 11 (bottom). To clarify the behavior of FT-err, Algorithm 4 is restricted to meet the reference time steps $\mathcal{T}_{\text{fwd}}^{\text{ref}}$ for all the time-adaptive strategies. This can result in time step sizes that are smaller than the minimum τ_k if the adaptively computed time steps would otherwise jump past the next time step from $\mathcal{T}_{\text{fwd}}^{\text{ref}}$. Some quickly alternating time step size adjustments computed by FT-err result from this restriction.

To illustrate the just described effects and differences between the time-adaptive strategies, Figure 12 displays the different n_t and runtimes. Notably, IE and TR require 2.3 seconds per time step. With the latter two methods, it would take 390 minutes to solve the closed-loop system with 10 000 equidistant time steps. This number of time steps corresponds to the step size that FT-dt-u uses while the control is active. Compared to that, FT-u and FT-dt-u require shorter runtime and significantly fewer time steps while using the same small time step size locally. Further, they simulate the closed-loop system more reliably in presence of strongly varying controls due to their adaptivity, which is tuned to this problem. In detail, the runtime per time step is 6.9 seconds for FT-u and 5.8 seconds for FT-dt-u on average. FT-u takes longer on average since it discards time steps more often than FT-dt-u since it tries to take larger time steps too early. A discarded time step results in a re-computation of this time step with a smaller step size, which makes this time step more expensive. Finally, with around 21 seconds per time step and a larger number of time steps, FT-err is significantly more expensive than FT-dt-u.

5 Conclusions

The two numerical challenges, to derive a feedback control for problems with moving interfaces or free boundaries and to reliably simulate the resulting non-linear closed-loop system, are tackled successfully in this work. In detail, we apply the LQR approach and implement non-autonomous versions of the BDF methods and splitting schemes to solve the arising DREs with time-dependent coefficients. Overall, the splitting methods, in their current state, are not competitive for the non-autonomous case. Other advantages of the splitting methods carry over to the non-autonomous

case, like better predictability of the accuracy due to a more uniform error distribution and less sensitivity to several problem-parameters. Besides the shorter runtimes, the BDF methods can handle more general time-dependencies in the coefficient matrices. They are not restricted to the case that, e.g., $\mathcal{A}(t) = \alpha(t)\bar{\mathcal{A}}$ and $\mathcal{M}(t) = \mu(t)\bar{\mathcal{M}}$ with constant matrices $\bar{\mathcal{A}}$ and $\bar{\mathcal{M}}$. This condition does, e.g., not hold for problems with a time-dependent domain resulting from moving boundaries or interfaces. Thus, we exclusively use the BDF methods for the next example, which incorporates a moving interface.

In addition, we have established that the fractional-step-theta scheme with our time-adaptive strategy can simulate the closed-loop systems with different feedback controls reliably and efficiently. Namely, our time-adaptivity is tuned to strongly varying controls and prevents blow-ups that can occur with other time stepping schemes. Furthermore, it requires significantly less time steps and runtime compared to classical error-based time-adaptivity.

Forthwith, we observed that the non-autonomous BDF methods and time-adaptive fractional-step-theta scheme, in interaction with each other, are well suited for non-linear large-scale control-problems with moving interfaces. They perform efficiently and reliably with respect to various problem parameters and problem dimensions. After all, despite several stages of approximation and a linearization for the LQR approach, the computed feedback control stabilizes the non-linear problem successfully.

Besides these results, the runtime performance of the BDF methods, as well as the splitting schemes, can potentially be improved significantly by exploring alternative ways to solve the underlying subproblems. For example, regarding the ARE also available are RADI and projection based methods like EKSM and RKSM. However, for large problems, a clear advantage of the Newton ADI used in this work in terms of memory requirements is that only the feedback gain matrices can be accumulated, thus avoiding the formation of the low-rank solution factors. In this case, the restrictions on time-dependent coefficients for splitting schemes could also be removed.

References

- [1] H. ABOU-KANDIL, G. FREILING, V. IONESCU, AND G. JANK, *Matrix Riccati Equations in Control and Systems Theory*, Systems & Control: Foundations & Applications, Birkhäuser, Basel, Switzerland, 2003, <https://doi.org/10.1007/978-3-0348-8081-7>.
- [2] D. AKAKPO AND E. GILDIN, *A control perspective to adaptive time-stepping in reservoir simulation*, in Society of Petroleum Engineers - SPE Reservoir Simulation Conference 2017, 2017, pp. 213–232, <https://doi.org/10.2118/182601-ms>.
- [3] U. M. ASCHER AND L. R. PETZOLD, *Computer Methods for Ordinary Differential Equations and Differential-Algebraic Equations*, SIAM, Philadelphia, 1998, <https://doi.org/10.1137/1.9781611971392>.
- [4] E. BÄNSCH, J. PAUL, AND A. SCHMIDT, *An ALE finite element method for a coupled Stefan problem and Navier–Stokes equations with free capillary surface*, *Internat. J. Numer. Methods Fluids*, 71 (2013), pp. 1282–1296, <https://doi.org/10.1002/flid.3711>.
- [5] B. BARAN, *Linear Quadratic Regulator Computation for a Stefan Problem using M.-M.E.S.S. and FEniCS*, April 2021, <https://doi.org/10.5281/zenodo.4671886>.
- [6] B. BARAN, P. BENNER, J. HEILAND, AND J. SAAK, *Optimal control of a Stefan problem fully coupled with incompressible Navier–Stokes equations and mesh movement*, *Analele Stiintifice ale Universitatii Ovidius Constanta: Seria Matematica*, XXVI (2018), pp. 11–40, <https://doi.org/10.2478/auom-2018-0016>.
- [7] B. BARAN, P. BENNER, AND J. SAAK, *Riccati feedback control of a two-dimensional two-phase Stefan problem*, e-print arXiv:2209.05476v2, arXiv, September 2022, <https://arxiv.org/abs/2209.05476>. math.NA.
- [8] M. BEHR, P. BENNER, AND J. HEILAND, *Invariant Galerkin trial spaces and Davison-Maki methods for the numerical solution of differential Riccati equations*, *Appl. Math. Comput.*, 410 (2021), p. 126401, <https://doi.org/10.1016/j.amc.2021.126401>.

- [9] M. BEHR AND J. SAAK, *Oberwolfach Rail Reimplementation in FEniCS*, <https://gitlab.mpi-magdeburg.mpg.de/models/fenicssrail>.
- [10] P. BENNER AND N. LANG, *Peer methods for the solution of large-scale differential matrix equations*, e-print arXiv:1804.08524, arXiv, July 2018, <https://arxiv.org/abs/1804.08524>. math.NA.
- [11] P. BENNER, J.-R. LI, AND T. PENZL, *Numerical solution of large-scale Lyapunov equations, Riccati equations, and linear-quadratic optimal control problems*, Numer. Lin. Alg. Appl., 15 (2008), pp. 755–777, <https://doi.org/10.1002/nla.622>.
- [12] P. BENNER AND H. MENA, *BDF methods for large-scale differential Riccati equations*, in Proc. 16th Intl. Symp. Mathematical Theory of Network and Systems, MTNS 2004, B. De Moor, B. Motmans, J. Willems, P. Van Dooren, and V. Blondel, eds., 2004.
- [13] P. BENNER AND H. MENA, *Numerical solution of the infinite-dimensional LQR-problem and the associated differential Riccati equations*, J. Numer. Math., 26 (2018), pp. 1–20, <https://doi.org/10.1515/jnma-2016-1039>. published online May 2016.
- [14] P. BENNER AND J. SAAK, *Linear-quadratic regulator design for optimal cooling of steel profiles*, Tech. Report SFB393/05-05, Sonderforschungsbereich 393 *Parallele Numerische Simulation für Physik und Kontinuumsmechanik*, TU Chemnitz, D-09107 Chemnitz (Germany), 2005, <http://nbn-resolving.de/urn:nbn:de:swb:ch1-200601597>.
- [15] R. R. BITMEAD AND M. GEVERS, *Riccati Difference and Differential Equations: Convergence, Monotonicity and Stability*, Springer, Berlin, Heidelberg, 1991, pp. 263–291, https://doi.org/10.1007/978-3-642-58223-3_10.
- [16] T. BREITEN, S. DOLGOV, AND M. STOLL, *Solving differential Riccati equations: a nonlinear space-time method using tensor trains*, Numerical Algebra, Control and Optimization, 11 (2021), pp. 407–429, <https://doi.org/10.3934/naco.2020034>.
- [17] M. O. BRISTEAU, R. GLOWINSKI, AND J. PERIAUX, *Numerical methods for the Navier-Stokes equations. applications to the simulation of compressible and incompressible viscous flows*, Comput. phys. rep., 6 (1987), pp. 73 – 187, [https://doi.org/10.1016/0167-7977\(87\)90011-6](https://doi.org/10.1016/0167-7977(87)90011-6).
- [18] I. CVOK, I. RATKOVIC, AND J. DEUR, *Optimization of control parameters of vehicle air-conditioning system for maximum efficiency*, SAE Technical Papers, 2020-April (2020), <https://doi.org/10.4271/2020-01-1242>.
- [19] L. FAILER, *Optimal Control of Time-Dependent Nonlinear Fluid-Structure Interaction*, Dissertation, Technische Universität München, München, 2017, <https://mediatum.ub.tum.de/doc/1355578/1355578.pdf>.
- [20] L. FAILER AND T. WICK, *Adaptive time-step control for nonlinear fluid-structure interaction*, J. Comput. Phys., 366 (2018), pp. 448–477, <https://doi.org/10.1016/j.jcp.2018.04.021>.
- [21] Y. GÜLDOĞAN, M. HACHED, K. JBILOU, AND M. KURULAY, *Low rank approximate solutions to large-scale differential matrix Riccati equations*, Appl. Math. (Warsaw), 45 (2018), pp. 233–254, <https://doi.org/10.4064/am2355-1-2018>. math.NA.
- [22] K. GUSTAFSSON, M. LUNDH, AND G. SÖDERLIND, *A PI stepsize control for the numerical solution of ordinary differential equations*, BIT, 28 (1988), pp. 270–287, <https://doi.org/10.1007/BF01934091>.
- [23] Q. JIN, J. T. WEN, AND S. NARAYANAN, *Moving boundary model for dynamic control of multi-evaporator cooling systems facing variable heat loads [modèle à frontières mobiles pour la régulation dynamique des systèmes de refroidissement à évaporateurs multiples confrontés à des charges thermiques variables]*, Int. J. Refrig., 120 (2020), pp. 481–492, <https://doi.org/10.1016/j.ijrefrig.2020.09.014>.

- [24] V. JOHN AND J. RANG, *Adaptive time step control for the incompressible Navier–Stokes equations*, *Comp. Meth. Appl. Mech. Eng.*, 199 (2010), pp. 514 – 524, <https://doi.org/10.1016/j.cma.2009.10.005>.
- [25] G. KIRSTEN AND V. SIMONCINI, *Order reduction methods for solving large-scale differential matrix Riccati equations*, *SIAM J. Sci. Comput.*, 42 (2020), pp. A2182–A2205, <https://doi.org/10.1137/19M1264217>.
- [26] S. KOGA, M. DIAGNE, AND M. KRSTIC, *Control and state estimation of the one-phase Stefan problem via backstepping design*, *IEEE Trans. Autom. Control*, 64 (2019), pp. 510–525, <https://doi.org/10.1109/TAC.2018.2836018>.
- [27] S. KOGA, M. MAKIHATA, R. CHEN, M. KRSTIC, AND A. PISANO, *Energy storage in paraffin: A PDE backstepping experiment*, *IEEE Trans. Control Syst. Technol.*, (2020), pp. 1–13, <https://doi.org/10.1109/TCST.2020.3014295>.
- [28] A. KOSKELA AND H. MENA, *Analysis of Krylov subspace approximation to large-scale differential Riccati equations*, *Electron. Trans. Numer. Anal.*, 52 (2020), pp. 431–454, https://doi.org/10.1553/etna_vol52s431.
- [29] P. KUNKEL AND V. MEHRMANN, *Numerical solution of differential algebraic Riccati equations*, *Linear Algebra Appl.*, 137/138 (1990), pp. 39–66, [https://doi.org/10.1016/0024-3795\(90\)90126-W](https://doi.org/10.1016/0024-3795(90)90126-W).
- [30] P. KUNKEL AND V. MEHRMANN, *Numerical solution of Riccati differential algebraic equations*, in *Proceedings of the International Symposium on the Mathematical Theory of Networks and Systems*, Amsterdam, Netherlands, June 1989, M. A. Kaashoek, J. H. Schuppen, and A. S. M. Ran, eds., vol. III, Basel, 1990, Birkhäuser, pp. 479–487.
- [31] N. LANG, *Numerical Methods for Large-Scale Linear Time-Varying Control Systems and related Differential Matrix Equations*, Dissertation, Technische Universität Chemnitz, Germany, June 2017, <https://www.logos-verlag.de/cgi-bin/buch/isbn/4700>. Logos-Verlag, Berlin, ISBN 978-3-8325-4700-4.
- [32] N. LANG, H. MENA, AND J. SAAK, *On the benefits of the LDL^T factorization for large-scale differential matrix equation solvers*, *Linear Algebra Appl.*, 480 (2015), pp. 44–71, <https://doi.org/10.1016/j.laa.2015.04.006>.
- [33] D. LI, X. ZHANG, AND R. LIU, *Exponential integrators for large-scale stiff Riccati differential equations*, *J. Comput. Appl. Math.*, 389 (2020), p. 113360, <https://doi.org/10.1016/j.cam.2020.113360>.
- [34] A. LOCATELLI, *Optimal Control: An Introduction*, Birkhäuser, Basel, Switzerland, 2001, <https://doi.org/10.1007/978-3-0348-8328-3>.
- [35] L. LU AND Y. FENG, *Observability and stabilization of 1-D wave equations with moving boundary feedback*, *Acta Appl. Math.*, 170 (2020), pp. 731–753, <https://doi.org/10.1007/s10440-020-00356-4>.
- [36] D. MEIDNER AND T. RICHTER, *Goal-oriented error estimation for the fractional step theta scheme*, *Comput. Methods Appl. Math.*, 14 (2014), pp. 203 – 230, <https://doi.org/10.1515/cmam-2014-0002>.
- [37] D. MEIDNER AND T. RICHTER, *A posteriori error estimation for the fractional step theta discretization of the incompressible Navier–Stokes equations*, *Comp. Meth. Appl. Mech. Eng.*, 288 (2015), pp. 45 – 59, <https://doi.org/10.1016/j.cma.2014.11.031>.
- [38] H. MENA, *Numerical Solution of Differential Riccati Equations Arising in Optimal Control Problems for Parabolic Partial Differential Equations*, Unidad de Publicaciones de la Facultad de Ciencias, Quito-Ecuador, first ed., May 2012. Available as ISBN: 978-9978-383-09-4.

- [39] H. MENA, A. OSTERMANN, L.-M. PFURTSCHELLER, AND C. PIAZZOLA, *Numerical low-rank approximation of matrix differential equations*, J. Comput. Appl. Math., 340 (2018), pp. 602–614, <https://doi.org/10.1016/j.cam.2018.01.035>.
- [40] A. OSTERMANN, C. PIAZZOLA, AND H. WALACH, *Convergence of a low-rank Lie–Trotter splitting for stiff matrix differential equations*, SIAM J. Numer. Anal., 57 (2019), pp. 1947–1966, <https://doi.org/10.1137/18M1177901>.
- [41] W. T. REID, *Riccati Differential Equations*, vol. 86 of Mathematics in Science and Engineering, Academic Press, New York, 1972, <https://doi.org/10.1115/1.3426914>.
- [42] T. RICHTER AND T. WICK, *On time discretizations of fluid-structure interactions*, in Multiple Shooting and Time Domain Decomposition Methods, Cham, 2015, Springer International Publishing, pp. 377–400, https://doi.org/10.1007/978-3-319-23321-5_15.
- [43] J. SAAK, M. KÖHLER, AND P. BENNER, *M-M.E.S.S. – the Matrix Equations Sparse Solvers library*, <https://doi.org/10.5281/zenodo.632897>. see also: <https://www.mpi-magdeburg.mpg.de/projects/mess>.
- [44] T. SONG, K. UPRETI, AND G. SUBBARAYAN, *A sharp interface isogeometric solution to the Stefan problem*, Comp. Meth. Appl. Mech. Eng., 284 (2015), pp. 556–582, <https://doi.org/10.1016/j.cma.2014.10.013>.
- [45] E. D. SONTAG, *Mathematical Control Theory*, Texts in Applied Mathematics, Springer-Verlag, New York, NY, 2nd ed., 1998, <https://doi.org/10.1007/978-1-4612-0577-7>.
- [46] T. STILLFJORD, *Low-rank second-order splitting of large-scale differential Riccati equations*, IEEE Trans. Autom. Control, 60 (2015), pp. 2791–2796, <https://doi.org/10.1109/TAC.2015.2398889>.
- [47] T. STILLFJORD, *Splitting schemes for nonlinear parabolic problems*, PhD thesis, Lund University, 2015, <https://lup.lub.lu.se/search/ws/files/3905802/5277358.pdf>.
- [48] T. STILLFJORD, *Adaptive high-order splitting schemes for large-scale differential Riccati equations*, Numer. Algorithms, 78 (2018), pp. 1129–1151, <https://doi.org/10.1007/s11075-017-0416-8>.
- [49] T. STILLFJORD, *Singular value decay of operator-valued differential Lyapunov and Riccati equations*, SIAM J. Control Optim., 56 (2018), pp. 3598–3618, <https://doi.org/10.1137/18M1178815>.
- [50] S. TUREK, *Efficient solvers for incompressible flow problems*, vol. 6 of Lect. Notes Comput. Sci. Eng., Springer-Verlag, Berlin, 1999, <https://doi.org/10.1007/978-3-642-58393-3>.
- [51] H. K. WEICHEL, *Numerical Aspects of Flow Stabilization by Riccati Feedback*, Dissertation, Otto-von-Guericke-Universität, Magdeburg, Germany, January 2016, <https://doi.org/10.25673/4493>.
- [52] T. WICK, *Adaptive Finite Element Simulation of Fluid-Structure Interaction with Application to Heart-Valve Dynamics*, Dissertation, Heidelberg University, Heidelberg, Germany, December 2011, <https://doi.org/10.11588/heidok.00012992>.

1 Supplementary Material

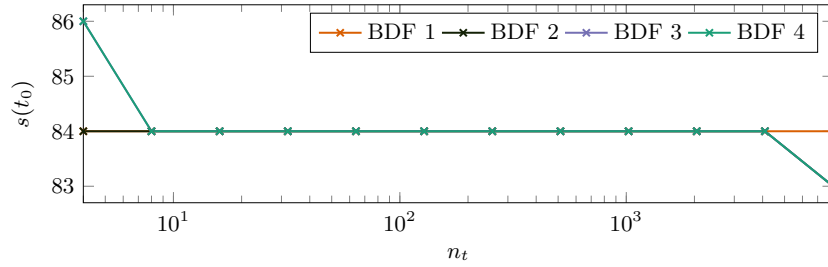


Figure 1: $s(t_0)$ (numerical rank of the solution) for BDF for different n_t (no. of time steps), steel profile example.

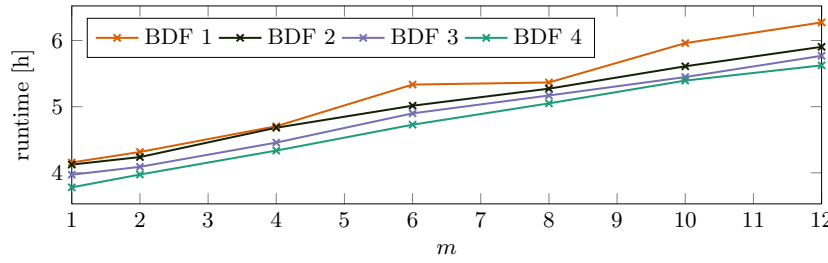


Figure 2: Runtime of BDF for different m (columns in $\mathcal{B}(t)$), Stefan problem example.

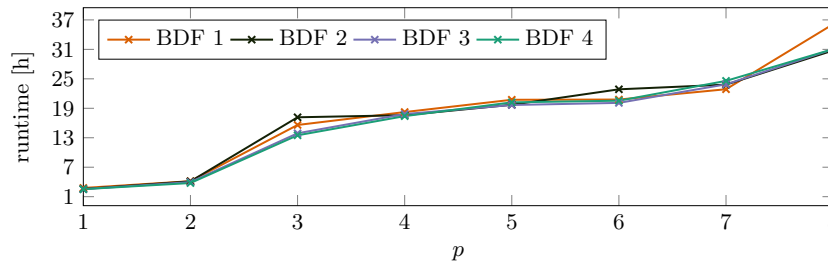


Figure 3: Runtime of BDF for different p (rows in $\mathcal{C}(t)$), Stefan problem example.

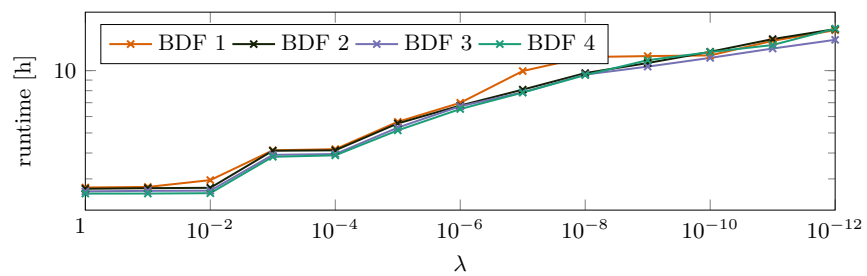


Figure 4: Runtime of BDF for different λ (weight in cost functional), Stefan problem example.

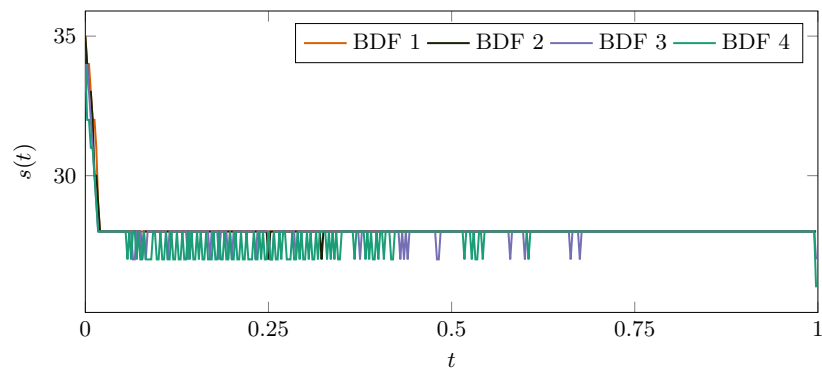


Figure 5: Numerical rank of the solution $s(t)$ for BDF 1 to 4, Stefan problem example. (used truncation tolerance: machine precision times square of largest singular value)