A framework for randomized time-splitting in linear-quadratic optimal control

D.W.M. Veldman[1] and E. Zuazua[1,2,3]

[1]Chair in Dynamics, Control, and Numerics (Alexander-von-Humboldt Professorship), Friedrich-Alexander Universität (FAU) Erlangen-Neuremberg, Cauerstrasse 11, 91052, Erlangen, Germany.

[2]Departamento do Matemáticas, Universidad Autonoma de Madrid, Ciudad Universitaria de Cantoblanco, 28049, Madrid, Spain.

[3]Chair of Computational Mathematics, Fundación Deusto, Av. de las Universidades 24, 48007, Bilbao, Basque-Country, Spain.

Contributing authors: daniel.veldman@math.fau.de; enrique.zuazua@fau.de;

Abstract

Inspired by the successes of stochastic algorithms in the training of deep neural networks and the simulation of interacting particle systems, we propose and analyze a framework for randomized time-splitting in linear-quadratic optimal control. In our proposed framework, the linear dynamics of the original problem is replaced by a randomized dynamics. To obtain the randomized dynamics, the system matrix is split into simpler submatrices and the time interval of interest is split into subintervals. The randomized dynamics is then found by selecting randomly one or more submatrices in each subinterval. We show that the dynamics, the minimal values of the cost functional, and the optimal control obtained with the proposed randomized time-splitting method converge in expectation to their analogues in the original problem when the time grid is refined. The derived convergence rates are validated in several numerical experiments. Our numerical results also indicate that the proposed method can lead to a reduction in computational cost for the simulation and optimal control of large-scale linear dynamical systems.

Keywords: Random Batch Method, Operator Splitting, Optimal Control, Model Predictive Control

MSC Classification: 65C99 , 49M99 , 65L20 , 37M05

1 Introduction

Solving an optimal control problem for a large-scale dynamical system can be computationally demanding. This problem appears in numerous applications. One example is Model Predictive Control (MPC), which requires the solution of several optimal control problems on a receding time horizon [12, 19]. Another example is the training of Deep Neural Networks (DNNs), which can be approached as an optimal control problem for a large-scale nonlinear dynamical system, see, e.g., [9, 4, 11, 10, 28]. Because the computational cost for gradient-based deterministic optimization algorithms explodes on large training data sets, Neural Networks (NNs) are typically trained using stochastic optimization algorithms such as stochastic gradient descent or stochastic (mini-)batch methods, see, e.g., [6]. In such methods, the update direction for the parameters of the NN is not computed based on the complete training data set, but on a subset of the available training data that is chosen randomly in each iteration. It can be shown that such methods converge in expectation to a (local) minimum of the considered cost functional, see, e.g., [6].

These successes inspired the development of Random Batch Methods (RBMs) for the simulation of interacting particle systems [15, 22, 16]. Because the number of interactions between N particles is of order N^2, the forward simulation of a system with a large number of particles is computationally demanding. A RBM reduces the required computational cost by reducing the number of considered interactions as follows. First, the considered time interval is divided into a number of subintervals of length $\leq h$. In each subinterval, particles are grouped in randomly chosen batches (of at least two particles) and only the interactions between particles in the same batch are considered. The number of considered interactions now grows as PN, where P is the size of the considered batches, and a significant reduction in computational time can be achieved when $P \ll N$. It can be shown that the expected error introduced by this process is proportional to $\sqrt{h}$, where h denotes (an upper bound on) the length of the considered time intervals, see [15].

The computation of optimal controls for interacting particle systems is even more computationally demanding than the forward simulation because it requires several simulations of the forward dynamics and the associated adjoint problem, see, e.g., [21]. Because the optimal control for the RBM-approximated dynamics can be computed significantly faster than the control for the original dynamics, it has been proposed in [19] to control the original system with the controls optimized for the RBM dynamics. The numerical experiments in [19]

indeed indicate that this approach can lead to a reasonably good approximation of the control for the original system. In [19], the control of the original dynamics with the RBM-optimal controls is combined with an MPC strategy, which creates additional robustness against the errors introduced by the RBM-approximation. However, even for the simplest case that does not consider the combination with MPC, a formal proof that the optimal control computed for the RBM-approximated dynamics indeed converges to the optimal control for the original system for $h \to 0$ was not given.

In this paper, we study, motivated by the ideas from [19], the classical linear-quadratic (LQ) optimal control problem constrained by randomized dynamics. Extensions of these results to a nonlinear setting are not only of interest for the control of interacting particle systems as considered in [19], but have also applications in the training of certain DNNs which can be viewed as (the time discretization) of an optimal control problem, see, e.g., [9, 4, 11, 10, 28]. The results for the LQ problem in this paper form a starting point for the study of these more involved problem settings.

In this paper, we propose a framework for the simulation and optimal control of large-scale linear dynamical systems. In our proposed framework, the system matrix is split into submatrices and the time interval of interest is split into subintervals of length $\leq h$. The randomized dynamics is then found based on the randomly selected submatrices in each subinterval. Similarly as in [15, 22, 16], we show that the randomized dynamics converges to the dynamics of the original system at a rate $\sqrt{h}$. The main contributions of this paper concern the LQ optimal control problem in which the original dynamics is replaced by these randomized dynamics. In particular, we show that the minimal values of the cost functional and the corresponding optimal controls for the RBM-dynamics converge (in L^2 and in expectation) to their analogues for the original dynamics when $h \to 0$. The found convergence rates are validated by several numerical examples. Numerical results also indicate that the proposed method can lead to a reduction in computational cost.

The remainder of this paper is structured as follows. Section 2 contains a precise description of our proposed stochastic simulation method and a summary of the main results of the paper. Section 3 contains the detailed proofs of the convergence of the proposed method. The proposed method and the obtained convergence results are illustrated by several numerical examples in Section 4. The conclusions and discussions are presented in Section 5.

2 Proposed method and main results

2.1 Proposed method

We consider the evolution of a large-scale Linear Time Invariant (LTI) dynamical system of the form

$$\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0, \tag{1}$$

where the state $x(t)$ evolves in $\mathbb{R}^N$, the control $u(t)$ evolves in $\mathbb{R}^q$, $A \in \mathbb{R}^{N\times N}$ is the system matrix, $B \in \mathbb{R}^{N\times q}$ is the input matrix, and $x_0 \in \mathbb{R}^N$ is the initial condition.

A typical problem associated to the dynamics (1) is to find the optimal control $u^*(t)$ that minimizes the quadratic cost functional

$$J(u) = \frac{1}{2}\int_0^T \left((x(t) - x_d(t))^\top Q(x(t) - x_d(t)) + u(t)^\top R u(t)\right)\, \mathrm{d}t, \tag{2}$$

where the given target trajectory $x_d(t)$ evolves in $\mathbb{R}^N$, the weighting matrix $Q \in \mathbb{R}^{N\times N}$ is symmetric and positive semi-definite, and the weighting matrix $R \in \mathbb{R}^{q\times q}$ is symmetric and positive definite. It is well known that the optimal control $u^*(t)$ exists and that it is unique, see, e.g., [23, 18].

Remark 1 When the state-dimension N is large, the optimal control $u^*(t)$ is typically computed using a gradient-based algorithm in which the gradient of $J(u)$ is computed from the adjoint state $\varphi(t)$ that satisfies (see, e.g., [18])

$$-\dot{\varphi}(t) = A^\top \varphi(t) + Q(x(t) - x_d(t)), \qquad \varphi(T) = 0, \tag{3}$$

where $x(t)$ is the solution of (1). Note that the adjoint state $\varphi(t)$ is computed by integrating (3) backward in time starting from the final condition $\varphi(T) = 0$. The gradient of the cost functional $J(u)$ is then obtained as

$$(\nabla J(u))(t) = B^\top \varphi(t) + Ru(t). \tag{4}$$

In our proposed randomized time-splitting method, the matrix A is written as the sum of M submatrices A_m

$$A = \sum_{m=1}^{M} A_m. \tag{5}$$

Typically, the submatrices A_m will be more sparse than the original matrix A. For ease of presentation, the results in this paper are presented under the following assumption.

Assumption 1 The submatrices A_m in (5) are dissipative, i.e. $\langle x, A_m x\rangle \le 0$ for all $x \in \mathbb{R}^N$ and all $m \in \{1, 2, \ldots, M\}$.

Remark 2 Note that there always exists a constant $a > 0$ such that the matrices $A_m - aI$ are dissipative for $m \in \{1, 2, \ldots, M\}$. Assumption 1 is therefore not essential for the convergence of the proposed method, but without Assumption 1 the error estimates are less clean and grow exponentially in time. This idea is made more precise in Remark 9 in Section 3.2.

We then choose a temporal grid in the time interval $[0, T]$

$$0 = t_0 < t_1 < t_2 < \ldots < t_{K-1} < t_K = T, \tag{6}$$

and denote

$$h_k = t_k - t_{k-1}, \qquad h = \max_{k \in \{1,2,\ldots,K\}} h_k. \tag{7}$$

In each of the K subintervals $[t_{k-1}, t_k)$, we randomly select a subset of indices in $\{1, 2, \ldots, M\}$. The idea of the proposed method is to consider a linear combination of the submatrices A_m with the indices that have been selected for each time interval. This can lead to a significant reduction in computational time when the submatrices A_m are well-chosen and only a small number of submatrices A_m are selected in each time interval.

To make this idea more precise, we enumerate all of the 2^M subsets of $\{1, 2, \ldots, M\}$ as $S_1, S_2, \ldots S_{2^M}$. Note that one of the subsets S_ω will be the empty set. To every subset S_ω ($\omega \in \Omega := \{1, 2, \ldots, 2^M\}$) we then assign a probability p_ω with which this subset is selected. This probability is the same in each of the time intervals $[t_{k-1}, t_k)$. Because we select only one subset S_ω in each time interval, the probabilities p_ω should satisfy

$$\sum_{\omega=1}^{2^M} p_\omega = 1. \tag{8}$$

From the chosen probabilities p_ω, we then compute the probability π_m that an index $m \in \{1, 2, \ldots, M\}$ is an element of the selected subset

$$\pi_m = \sum_{\omega \in \Omega_m} p_\omega, \qquad \Omega_m = \{\omega \in \{1, 2, \ldots, 2^M\} \mid m \in S_\omega\}. \tag{9}$$

Observe that Ω_m is the set of the indices ω of the sets S_ω that contain the index m. We need the following (weak) assumption on the selected probabilities p_ω.

Assumption 2 The probabilities p_ω ($\omega \in \{1, 2, \ldots, 2^M\}$) are assigned such that

- (8) is satisfied and
- the probabilities π_m defined in (9) are positive for all $m \in \{1, , 2, \ldots, M\}$.

In each of the K time intervals $[t_{k-1}, t_k)$, we then randomly select an index $\omega_k \in \{1, 2, \ldots, 2^M\}$ according to the chosen probabilities p_ω (and independently of the other indices $\omega_1, \omega_2, \ldots \omega_{k-1}, \omega_{k+1}, \omega_{k+1}, \ldots, \omega_K$). The selected indices form a vector

$$\boldsymbol{\omega} := (\omega_1, \omega_2, \ldots, \omega_K) \in \{1, 2, \ldots, 2^M\}^K =: \Omega^K. \tag{10}$$

For the selected $\boldsymbol{\omega} \in \Omega^K$, we then define a piece-wise constant matrix $t \mapsto \mathcal{A}_h(\boldsymbol{\omega}, t)$

$$\mathcal{A}_h(\boldsymbol{\omega}, t) = \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m}, \qquad t \in [t_{k-1}, t_k). \tag{11}$$

The scaling by $1/\pi_m$ assures that the expected value of $\mathcal{A}_h$ is A because

$$\sum_{\omega=1}^{2^M} \sum_{m \in S_\omega} \frac{A_m}{\pi_m} p_\omega = \sum_{m=1}^{M} \sum_{\omega \in \Omega_m} \frac{A_m}{\pi_m} p_\omega = \sum_{m=1}^{M} \frac{A_m}{\pi_m} \pi_m = \sum_{m=1}^{M} A_m = A, \tag{12}$$

where the first identity follows after interchanging the two summations using the definition of Ω_m in (9), the second from the definition of π_m in (9), and the last identity from the decomposition of A in (5).

Example 1 In the simplest situation, we decompose the original matrix A into $M = 2$ matrices as $A = A_1 + A_2$. We then need to assign $2^M = 4$ probabilities p_ℓ to the subsets $S_1 = \{1\}$, $S_2 = \{2\}$, $S_3 = \{1, 2\}$, and $S_4 = \emptyset$. In this example, we choose $p_1 = p_2 = \frac{1}{2}$ and $p_3 = p_4 = 0$. This choice indeed satisfies Assumption 2 because $\pi_1 = p_1 + p_3 = \frac{1}{2} > 0$ and $\pi_2 = p_2 + p_3 = \frac{1}{2} > 0$. The matrix $\mathcal{A}_h(\boldsymbol{\omega}, t)$ is thus either equal to $2A_1$ with probability $p_1 = \frac{1}{2}$ or equal to $2A_2$ with probability $p_2 = \frac{1}{2}$. The expected value of $\mathcal{A}_h$ is then indeed $\frac{1}{2}2A_1 + \frac{1}{2}2A_2 = A_1 + A_2 = A$.

To reduce the computational cost for solving (1), the matrix A is replaced by a $\mathcal{A}_h(\boldsymbol{\omega}, t)$ in the RBM. For the selected vector of indices $\boldsymbol{\omega} \in \Omega^K$, we thus obtain a solution $t \mapsto x_h(\boldsymbol{\omega}, t)$

$$\dot{x}_h(\boldsymbol{\omega}, t) = \mathcal{A}_h(\boldsymbol{\omega}, t) x_h(\boldsymbol{\omega}, t) + Bu(t), \qquad x_h(\boldsymbol{\omega}, 0) = x_0. \tag{13}$$

The main contribution of this paper concerns the optimal controls computed based on the RBM-dynamics (13). In particular, we consider the minimization of the functional

$$J_h(\boldsymbol{\omega}, u) = \frac{1}{2} \int_0^T \left((x_h(\boldsymbol{\omega}, t) - x_d(t))^\top Q (x_h(\boldsymbol{\omega}, t) - x_d(t)) + u(t)^\top R u(t) \right) \, \mathrm{d}t, \tag{14}$$

over all $u \in L^2(0, T;\ \mathbb{R}^q)$ subject to the dynamics (13). The minimizer of $J_h(\boldsymbol{\omega}, \cdot)$ depends on the selected indices $\boldsymbol{\omega} \in \Omega^K$ and is denoted by $u_h^*(\boldsymbol{\omega}, t)$. Because R is positive definite, the minimizer $u_h^*(\boldsymbol{\omega}, t)$ exists and is unique. As we will show in (48)–(50) in Section 3.1, the minimizers $u_h^*(\boldsymbol{\omega}, t)$ are uniformly bounded because R is positive definite.

Remark 3 Similarly as for the original cost functional $J(u)$ in (2), we can compute the optimal control $u_h(\boldsymbol{\omega}, t)$ that minimizes $J_h(\boldsymbol{\omega}, u)$ by a gradient-based algorithm.

We can again compute the gradient of $J_h(\boldsymbol{\omega}, u)$ from the adjoint state $\varphi_h(\boldsymbol{\omega}, t)$ which satisfies

$$-\dot{\varphi}_h(\boldsymbol{\omega}, t) = (\mathcal{A}_h(\boldsymbol{\omega}, t))^\top \varphi_h(\boldsymbol{\omega}, t) + Q(x_h(\boldsymbol{\omega}, t) - x_d(t)), \qquad \varphi_h(\boldsymbol{\omega}, T) = 0. \quad (15)$$

The gradient of $J_h(\boldsymbol{\omega}, u)$ is then obtained as

$$\nabla J_h(\boldsymbol{\omega}, u) = B^\top \varphi_h(\boldsymbol{\omega}, t) + Ru(t). \quad (16)$$

Note that when the randomized dynamics for $x_h(\boldsymbol{\omega}, t)$ in (13) can be solved faster than the original dynamics for $x(t)$ in (1), the same reduction in computational cost is typically also obtained for the randomized adjoint equation (15) compared to the original adjoint equation (3). Because the computation of the optimal control $u^*(t)$ (resp. $u_h^*(\boldsymbol{\omega}, t)$) requires several evaluations of the forward dynamics (1) (resp. (13)) and the adjoint equation (3) (resp. (15)), it is natural to expect the same relative speed-up for $u_h^*(\boldsymbol{\omega}, t)$ (compared to $u_h^*(t)$) as for $x_h(\boldsymbol{\omega}, t)$ (compared to $x(t)$). This idea is confirmed by the numerical experiments in Section 4.

To conclude this subsection, we summarize the proposed approach to approximate the solution $x(t)$ of (1) for a given control $u(t)$ and/or the optimal control $u^*(t)$ that minimizes $J(\cdot)$ in (2) subject to (1) in Algorithm 1. The accuracy of the obtained solutions $x_h(\boldsymbol{\omega}, t)$ and/or $u_h^*(\boldsymbol{\omega}, t)$ depends on the chosen submatrices A_m in Step 1, the chosen probabilities p_ω in Step 2, and the chosen time grid $t_0, t_1, \ldots, t_K$ in Step 3. This dependence is captured by the error estimates in the next subsection.

It should be emphasized that we do not have that $\mathbb{E}[x_h(t)] = x(t)$ (for a fixed control $u(t)$) or that $\mathbb{E}[u_h^*(t)] = u^*(t)$. Repeating Step 4 in Algorithm 1 for different realizations of $\boldsymbol{\omega}$ and averaging the obtained results leads to better approximations of $\mathbb{E}[x_h(t)]$ and/or $\mathbb{E}[u_h^*(t)]$ and can therefore only improve the approximation of $x(t)$ and $u^*(t)$ to some extend. A better way to increase the accuracy of the proposed method is to repeat Algorithm 1 for a choice of submatrices A_m, probabilities p_ω, and a time grid $t_0, t_1, \ldots, t_K$ that reduce the error estimates presented in the next subsection.

Step 1 Decompose the matrix A into M submatrices A_m as in (5), preferably such that Assumption 1 is satisfied.
Step 2 Enumerate the 2^M subsets of $\{1, 2, \ldots, M\}$ and assign probabilities $p_1, p_2, \ldots, p_{2^M}$ such that Assumption 2 is satisfied.
Step 3 Divide the considered time interval $[0, T]$ into K subintervals $[t_{k-1}, t_k)$ and choose an index ω_k according to the selected probabilities in Step 2 for each subinterval. Store the selected indices in a vector $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots \omega_K)$.
Step 4 Compute the solution $x_h(\boldsymbol{\omega}, t)$ of the dynamics (13) for a certain given control $u(t)$ and/or compute the control $u^*(\boldsymbol{\omega}, t)$ that minimizes $J_h(\boldsymbol{\omega}, \cdot)$ in (14) subject to the dynamics (13).

Algorithm 1: The proposed randomized time-splitting method

Remark 4 The presented framework is somewhat different from the problem setting considered in previous publications on RBMs for interacting particle systems, see, e.g., [15, 22, 16, 19]. Appendix A shows how these RBMs can be accommodated in our proposed framework.

2.2 Main results

The main results of this paper concern the effect of replacing the system matrix A in the original LQ optimal control problem (1)–(2) by the randomized matrix $\mathcal{A}_h(\boldsymbol{\omega}, t)$ defined in (11). Clearly, the deviation of the randomized matrix $\mathcal{A}_h(\boldsymbol{\omega}, t)$ from the original matrix A will influence the accuracy of the obtained results. The deviation of $\mathcal{A}_h(\boldsymbol{\omega}, t)$ from A is measured by

$$\mathrm{Var}[\mathcal{A}] := \sum_{\omega=1}^{2^M} \left\| \sum_{m \in S_\omega} \frac{A_m}{\pi_m} - A \right\|^2 p_\omega, \tag{17}$$

where $\|\cdot\|$ denotes the operator norm. The quantity $\mathrm{Var}[\mathcal{A}]$ is thus the average squared distance of $\mathcal{A}_h(\boldsymbol{\omega}, t)$ from A, weighted with the probabilities p_1, p_2, $\ldots$, p_{2^M} with which different values of $\mathcal{A}_h(\boldsymbol{\omega}, t)$ occur. Naturally, the error estimates below show that reducing $\mathrm{Var}[\mathcal{A}]$ will also reduce the errors introduced by the proposed randomized time-splitting method.

Example 1 (continued) We again consider the situation from Example 1 in which A is decomposed into $M = 2$ submatrices as $A = A_1 + A_2$ and $\mathcal{A}_h(\boldsymbol{\omega}, t)$ is either $2A_1$ or $2A_2$, both with probability $\frac{1}{2}$. We now compute the variance $\mathrm{Var}[\mathcal{A}]$ according to (17) and find

$$\mathrm{Var}[\mathcal{A}] = \|2A_1 - A\|^2 p_1 + \|2A_2 - A\|^2 p_2 = \|A_1 - A_2\|^2. \tag{18}$$

Examples 2 and 3 at the end of this subsection further illustrate how $\mathrm{Var}[\mathcal{A}]$ depends on the decomposition of A into submatrices A_m and the selected probabilities p_ω.

Remark 5 When A in an approximation of an unbounded operator as in the examples in Section 4, it is natural to introduce an additional (invertible) weighting matrix W and compute

$$\mathrm{Var}_W[\mathcal{A}] := \sum_{\ell=1}^{2^M} \left\| \left(\sum_{m \in S_\ell} \frac{A_m}{\pi_m} - A \right) W \right\|^2 p_\ell. \tag{19}$$

Clearly, we want to choose W such that AW and the matrices $A_m W$ can be considered as approximations of bounded operators. In that case, $\mathrm{Var}_W[\mathcal{A}]$ is also an approximation of a finite quantity. A natural choice is $W = (A - \lambda I)^{-1}$ for some λ in the resolvent of A.

The first main result of this paper is an estimate for the difference

$$e_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x(t) \tag{20}$$

between the RBM-dynamics (13) and the original dynamics (1).

Main result 1 *Assume that Assumptions 1 and 2 hold and that the input $u(t)$ in (1) is the same as in the input $u(t)$ in (13), then*

$$\mathbb{E}[|e_h(t)|^2] \leq C_{[A,B,x_0,T,u]} h \mathrm{Var}[\mathcal{A}], \tag{21}$$

for all $t \in [0, T]$.

The first main result follows directly from Theorem 1 in Subsection 3.2.

The expectation operator $\mathbb{E}$ is taken w.r.t. all possible outcomes $\boldsymbol{\omega} \in \Omega^K$. A precise definition will be given in Section 3.1. The constant $C_{[A,B,x_0,T,u]}$ can be taken as $(\|A\|T^2+2T)(|x_0|+|Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2$. The estimate thus only depends on the used submatrices A_m, the probabilities p_ω, and the used temporal grid $t_0, t_1, \ldots, t_K$ through $h\mathrm{Var}[\mathcal{A}]$ defined in (17). The proof of Main result 1 is inspired by the proofs of convergence of the RBM in [15, 16].

The estimate (21) shows that the expected squared error is proportional to the temporal grid spacing h. We can thus make the expected squared error in the forward dynamics arbitrary small by reducing the grid spacing. Note that Markov's inequality, see, e.g., [27], shows that

$$\mathbb{P}[|e_h(\boldsymbol{\omega}, t)| > \varepsilon] = \mathbb{P}[|e_h(\boldsymbol{\omega}, t)|^2 > \varepsilon^2] < \frac{\mathbb{E}[|e_h(t)|^2]}{\varepsilon^2}. \tag{22}$$

The probability that we select an $\boldsymbol{\omega} \in \Omega^K$ for which $|e_h(\boldsymbol{\omega}, t)|$ exceeds any given treshold $\varepsilon > 0$ is thus controlled by $\mathbb{E}[|e_h(t)|^2]$. According to (21), we can make this probability as small as desired by choosing the temporal grid spacing h small enough. However, one should also keep in mind that decreasing h will increase the computational cost for the RBM-dynamics (13) and that the computational advantage of the RBM is lost when the required grid spacing is too small.

Example 1 (continued) To illustrate why Main result 1 could be true, we again consider the situation from Example 1 in which A is decomposed as $A = A_1 + A_2$ and $\mathcal{A}_h(\boldsymbol{\omega}, t)$ is equal to $2A_1$ or $2A_2$, both with probability $\frac{1}{2}$. We additionally assume that $u(t) \equiv 0$, that the time grid $t_k = kT/K$ ($k \in \{0, 1, 2, \ldots, K\}$) is uniform with grid spacing $h = T/K$, and that A_1 and A_2 commute. Because $u(t) = 0$, the solution of (1) is $x(t) = e^{At}x_0$ and the solution of (13) is

$$x_h(\boldsymbol{\omega}, T) = e^{2A_{\omega_K}h} \cdots e^{2A_{\omega_2}h} e^{2A_{\omega_1}h} x_0 = e^{2A_1 T_1(\boldsymbol{\omega}) + 2A_2 T_2(\boldsymbol{\omega})} x_0. \tag{23}$$

Here, $T_1(\boldsymbol{\omega})$ and $T_2(\boldsymbol{\omega})$ denote the times during which A_1 and A_2 are used, i.e.

$$T_1(\boldsymbol{\omega}) = \frac{T}{K} \sum_{\ell=1}^{K} \chi_1(\omega_\ell), \qquad T_2(\boldsymbol{\omega}) = \frac{T}{K} \sum_{\ell=1}^{K} \chi_2(\omega_\ell), \tag{24}$$

where the characteristic functions $\chi_1(\omega)$ and $\chi_2(\omega)$ are defined by the property that $\chi_i(\omega) = 1$ when $\omega = i$ and $\chi_i(\omega) = 0$ otherwise ($i \in \{1, 2\}$). Note that the second identity in (23) uses that A_1 and A_2 commute. Because $\mathbb{E}[\chi_1] = \mathbb{E}[\chi_2] = \frac{1}{2}$, it follows that $\mathbb{E}[T_1] = \mathbb{E}[T_2] = T/2$. When we now consider the limit $K \to \infty$ (so $h \to 0$), the law of large numbers states that T_1 and T_2 converge to $T/2$ (in probability). The RHS of (23) thus converges (in probability) to $e^{AT}x_0 = x(T)$ for $K \to \infty$. Note that the convergence in Main result 1 is in expectation, which is stronger than convergence in probability.

We now present the main results aimed at the LQ optimal control problem constrained by randomized dynamics. Because the optimal control $u_h^*(\boldsymbol{\omega}, t)$ depends on the selected indices $\boldsymbol{\omega}$, we need the following result. The key difference with the first main result is that the input $u_h(\boldsymbol{\omega}, t)$ may now depend on the randomly selected indices $\boldsymbol{\omega}$. As will be explained at the start of Section 3, this makes the arguments for the convergence of the RBM in [15, 16] break down.

Note that replacing $u(t)$ in (1) and (13) by $u_h(\boldsymbol{\omega}, t)$ results in solutions $x(\boldsymbol{\omega}, t)$ and $x_h(\boldsymbol{\omega}, t)$ that now both depend on the selected indices $\boldsymbol{\omega}$. The second main result now gives a bound for the expected value of the difference

$$e_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x(\boldsymbol{\omega}, t). \tag{25}$$

Main result 2 *Consider any control* $u_h : \Omega^K \to L^2(0, T;\ \mathbb{R}^q)$. *Assume that Assumptions 1 and 2 are satisfied and let* U *be such that* $|u_h(\boldsymbol{\omega})|_{L^2(0,T;\,\mathbb{R}^q)} \leq U$ *for all* $\boldsymbol{\omega} \in \Omega^K$, *then*

$$\mathbb{E}[|e_h(t)|^2] \leq C_{[A,B,x_0,T,U]} h\mathrm{Var}[\mathcal{A}]. \tag{26}$$

The second result follows directly from Theorem 2 in Subsection 3.3.

Just as in the first main result, the expectation is taken over all possible values of $\boldsymbol{\omega} \in \Omega^K$ and the constant $C_{[A,B,x_0,T,U]}$ does not depend on the chosen submatrices A_m in (5), the chosen probabilities p_ω, and the used temporal grid.

Using this result, we can now obtain a no-gap result which shows that the minimal value of the cost functional $J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega}))$ is (in expectation) close to the minimal value $J(u^*)$ in the original problem when $h\mathrm{Var}[\mathcal{A}]$ is small enough.

Main result 3 *Let* $u^*(t)$ *be the control that minimizes the cost functional* $J(u)$ *in* (2) *and let* $u_h^*(\boldsymbol{\omega}, t)$ *be the control that minimizes the cost functional* $J_h(\boldsymbol{\omega}, u)$ *in* (14). *Then*

$$\mathbb{E}[|J_h(u_h^*) - J(u^*)|] \leq C_{[A,B,x_0,Q,R,x_d,T]} \left(\sqrt{h\mathrm{Var}[\mathcal{A}]} + h\mathrm{Var}[\mathcal{A}] \right). \tag{27}$$

The third main result is identical to Theorem 3 in Subsection 3.4.

For $h\mathrm{Var}[\mathcal{A}]$ small enough, Main result 3 clearly implies that $\mathbb{E}[|J_h(u_h^*) - J(u^*)|] \leq C_{[A,B,x_0,Q,R,x_d,T]} \sqrt{h\mathrm{Var}[\mathcal{A}]}$, which is also the rate that is observed in numerical experiments. We keep the second term on the RHS of (27) to

assure that the estimate is valid for all values of $h\text{Var}[\mathcal{A}]$, and not just for sufficiently small values of $h\text{Var}[\mathcal{A}]$.

By Markov's inequality, this result thus implies that, for any $\varepsilon > 0$, the probability that $|J(u_h^*(\boldsymbol{\omega})) - J(u^*)| > \varepsilon$ can be made arbitrarily small by reducing the temporal grid spacing h.

The next main result shows that the optimal control for the RBM-problem $u_h^*(\boldsymbol{\omega})$ also converges (in expectation) to the optimal control of the original problem u^* when $h \to 0$.

Main result 4 *Let $u_h^*(\boldsymbol{\omega}, t)$ be the minimizer of $J_h(\boldsymbol{\omega}, \cdot)$ in (14) and $u^*(t)$ be the minimizer of J in (2), then*

$$\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\,\mathbb{R}^q)}] \leq C_{[A,B,x_0,Q,R,x_d,T]} h\text{Var}[\mathcal{A}]. \tag{28}$$

The fourth main result follows directly from Theorem 4 in Subsection 3.5.

The fourth main result justifies the use of the optimal control $u_h^*(\boldsymbol{\omega})$, that is optimized for the RBM-dynamics, to control the original dynamics, as proposed in [19]. An almost immediate corollary of Main result 4 is that the trajectories of the original dynamics (1) resulting from the controls $u_h^*(\boldsymbol{\omega}, t)$ and $u^*(t)$ will also be close to each other, see Corollary 2 in Subsection 3.5. This further justifies the strategy in [19].

When the control $u_h^*(\boldsymbol{\omega})$ is close to the control u^* that is optimal for the original dynamics, the performance $J(u_h^*(\boldsymbol{\omega}))$ should also be close to the optimal performance $J(u^*)$. This idea is formalized by the fifth and last main result.

Main result 5 *Let $u^*(t)$ be the control that minimizes the cost functional $J(u)$ in (2) and let $u_h^*(\boldsymbol{\omega}, t)$ be the control that minimizes the cost functional $J_h(\boldsymbol{\omega}, u)$ in (14). Then*

$$\mathbb{E}[|J(u_h^*) - J(u^*)|] \leq C_{[A,B,x_0,Q,R,x_d,T]} h\text{Var}[\mathcal{A}]. \tag{29}$$

The fifth main result is identical to Corollary 3 in Subsection 3.5. Main result 5 is proven as a corollary of Main result 4/Theorem 4.

The fifth main result is particularly important because it shows that the performance $J(u_h^*(\boldsymbol{\omega}))$ obtained with control $u_h^*(\boldsymbol{\omega})$ optimized for the randomized dynamics is close to the optimal performance $J(u^*)$ when $h\text{Var}[\mathcal{A}]$ is sufficiently small. This further motivates strategies in which the original system is controlled by a control $u_h^*(\boldsymbol{\omega})$ that is optimized for the randomized dynamics, as was proposed in [19].

2.3 Further examples for Var[$\mathcal{A}$]

The quantity $\text{Var}[\mathcal{A}]$ describes how the derived estimates depend on the decomposition of A into submatrices and the selected probabilities $p_1, p_2, \ldots, p_{2^M}$. We therefore conclude this section with two other examples that illustrate

how $\mathrm{Var}[\mathcal{A}]$ depends on the decomposition of A into submatrices A_m and the selected probabilities p_ω.

Example 2 We decompose the matrix A into $M = 3$ parts $A = A_1 + A_2 + A_3$ and consider two choices for the probabilities p_ω. In the first case, we only use one of the submatrices A_m simultaneously. We thus assign probabilities $p_1 = p_2 = p_3 = \frac{1}{3}$ to the subsets $S_1 = \{1\}$, $S_2 = \{2\}$, and $S_3 = \{3\}$ and zero probability to the other 5 subsets of $\{1, 2, 3\}$. We then have that $\pi_1 = \pi_2 = \pi_3 = \frac{1}{3}$ and the variance $\mathrm{Var}[\mathcal{A}]$ in (17) becomes

$$\begin{aligned}\mathrm{Var}[\mathcal{A}] &= \|3A_1 - A\|^2 p_2 + \|3A_2 - A\|^2 p_3 + \|3A_3 - A\|^2 p_4 \\ &= \tfrac{1}{3}\left(\|2A_1 - A_2 - A_3\|^2 + \|2A_2 - A_1 - A_3\|^2 + \|2A_3 - A_1 - A_2\|^2\right). \end{aligned} \tag{30}$$

In the second case, we always use two of the three submatrices A_m simultaneously. We thus assign probabilities $p_4 = p_5 = p_6 = \frac{1}{3}$ to the subsets $S_4 = \{1, 2\}$, $S_5 = \{2, 3\}$, and $S_6 = \{1, 3\}$ and zero probability to the other 5 subsets of $\{1, 2, 3\}$. We then have that $\pi_1 = p_4 + p_6$, $\pi_2 = p_4 + p_5$, and $\pi_3 = p_5 + p_6$, so that $\pi_1 = \pi_2 = \pi_3 = \frac{2}{3}$. The variance $\mathrm{Var}[\mathcal{A}]$ in (17) becomes

$$\begin{aligned}\mathrm{Var}[\mathcal{A}] &= \|\tfrac{3}{2}(A_1 + A_2) - A\|^2 p_5 + \|\tfrac{3}{2}(A_2 + A_3) - A\|^2 p_6 + \|\tfrac{3}{2}(A_1 + A_3) - A\|^2 p_7 \\ &= \tfrac{1}{3}\left(\|\tfrac{1}{2}(A_1 + A_2) - A_3\|^2 + \|\tfrac{1}{2}(A_2 + A_3) - A_1\|^2 + \|\tfrac{1}{2}(A_1 + A_3) - A_2\|^2\right). \end{aligned} \tag{31}$$

Observe that $\|\frac{1}{2}(A_1 + A_2) - A_3\|^2 = \frac{1}{4}\|2A_3 - A_1 - A_2\|^2$ and that similar expressions relate the other terms in (30) and (31). The variance for the first case in (30) is thus four times larger than the variance for the second case in (31). Increasing the overlap between the possible values of $\mathcal{A}_h(\boldsymbol{\omega}, t)$ thus reduces $\mathrm{Var}[\mathcal{A}]$ and will improve the accuracy of the proposed method. It is worth noting that similar observations have been made for domain decomposition methods, for which it is well-known that increasing the overlap between subdomains increases the convergence rate (see, e.g., [8, Section 1.5]). Note however that increasing the overlap will also reduce the sparsity of $\mathcal{A}_h(t)$ and thus also increase the computational cost. This will be illustrated further by the numerical examples in Section 4.

Example 3 It is not always optimal to choose the probabilities uniform. To illustrate this, we assume $A = A_1 + A_2$ has a block-diagonal decomposition

$$A = \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}, \qquad A_1 = \begin{bmatrix} A_{11} & 0 \\ 0 & 0 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0 & 0 \\ 0 & A_{22} \end{bmatrix}. \tag{32}$$

It easy to verify that $\|\alpha A_1 + \beta A_2\| = \max\{|\alpha|\|A_1\|, |\beta|\|A_2\|\}$ for any $\alpha, \beta \in \mathbb{R}$. We assign the (at this point undetermined) probability $p_1 = p$ to the subset $S_1 = \{1\}$, the probability $p_2 = 1 - p$ to the subset $S_2 = \{2\}$, and probabilities $p_3 = p_4 = 0$ to the subsets $S_3 = \emptyset$ and $S_4 = \{1, 2\}$. It follows that $\pi_1 = p$ and $\pi_2 = 1 - p$ and that

$$\begin{aligned}\mathrm{Var}[\mathcal{A}] &= \|\tfrac{1}{p}A_1 - A\|^2 p + \|\tfrac{1}{1-p}A_2 - A\|^2 (1-p) \\ &= \|\tfrac{1}{p}((1-p)A_1 - pA_2)\|^2 p + \|\tfrac{1}{1-p}(pA_2 - (1-p)A_1)\|^2 (1-p) \\ &= \|(1-p)A_1 - pA_2\|^2 \left(\tfrac{1}{p} + \tfrac{1}{1-p}\right) = \left\|\sqrt{\tfrac{1-p}{p}}A_1 + \sqrt{\tfrac{p}{1-p}}A_2\right\|^2 \\ &= \left(\max\left\{\sqrt{\tfrac{1-p}{p}}\|A_1\|, \sqrt{\tfrac{p}{1-p}}\|A_2\|\right\}\right)^2. \end{aligned} \tag{33}$$

It is now easy to see that $\mathrm{Var}[\mathcal{A}]$ is minimal when $\sqrt{\frac{1-p}{p}}\|A_1\| = \sqrt{\frac{p}{1-p}}\|A_2\|$. Solving this equation for p, we find optimal probability

$$p^* = \frac{\|A_1\|}{\|A_1\| + \|A_2\|}. \tag{34}$$

We observe that the larger the submatrix A_1 is compared to A_2, the larger the probability p with which the submatrix A_1 is selected should be. Inserting the optimal probability p^* in (34) into the expression for $\mathrm{Var}[\mathcal{A}]$, we find that

$$\mathrm{Var}[\mathcal{A}]^* = \|A_1\|\|A_2\|. \tag{35}$$

With uniform probabilities, i.e., with $p = 1/2$, $\mathrm{Var}[\mathcal{A}] = \max\{\|A_1\|^2, \|A_2\|^2\}$, see (33). When $\|A_1\|/\|A_2\| \gg 1$ or $\|A_1\|/\|A_2\| \ll 1$, using the optimal probability p^* in (34) can thus reduce $\mathrm{Var}[\mathcal{A}]$ significantly.

3 Convergence analysis

The proof of convergence for the RBM optimal control problem is divided into several stages.

In the first stage, we consider a control $u \in L^2(0,T;\ \mathbb{R}^q)$ that does not depend on the selected indices $\boldsymbol{\omega}$. We then show that the expected difference between the RBM-dynamics (13) and the original dynamics (1) can be bounded in terms of $h\mathrm{Var}[\mathcal{A}]$ as in Main result 1. The proof of this statement is inspired by the results for interacting particles systems in [15, 16].

Because we will also need to deal with the optimal control $u_h^*(\boldsymbol{\omega}, t)$ that minimizes $J_h(\boldsymbol{\omega}, \cdot)$, we consider a general family of controls $u_h(\boldsymbol{\omega}, t)$ (with $\boldsymbol{\omega} \in \Omega^K$) in the second stage. This is a nontrivial extension of the results in the previous stage because the crucial idea in the proof for the first stage and in [15, 16] is that the solutions $x(t_{k-1})$ and $x_h(\boldsymbol{\omega}, t_{k-1})$ do not depend on ω_k (the index that is used in the time interval $[t_{k-1}, t_k)$). This is clearly no longer the case when we insert an input $u_h(\boldsymbol{\omega}, t)$ that depends on $\boldsymbol{\omega}$, so also on ω_k, into the dynamics (1) and (13). This problem is particularly clear when we consider the family of optimal controls $u_h^*(\boldsymbol{\omega})$ for which $u_h^*(\boldsymbol{\omega}, t_{k-1})$ will depend on the choices for the 'future' indices $\omega_k, \omega_{k+1}, \ldots \omega_K$.

In the third stage, we prove the no-gap condition presented in Main result 3. A crucial result for the proof is an auxiliary lemma (Lemma 1) that bounds the differences $J_h(\boldsymbol{\omega}, u) - J(u)$ and $J_h(\boldsymbol{\omega}, u_h(\boldsymbol{\omega})) - J(u_h(\boldsymbol{\omega}))$ (in expectation). For controls u that do not depend on ω, a bound on $J_h(\boldsymbol{\omega}, u) - J(u)$ can be obtained directly from Main result 1. For controls $u_h(\boldsymbol{\omega})$ that do depend on $\boldsymbol{\omega}$, we need to use Main result 2 to find the bound on the expected difference $J_h(\boldsymbol{\omega}, u_h(\boldsymbol{\omega})) - J(u_h(\boldsymbol{\omega}))$. For brevity, Lemma 1 considers controls $u_h(\boldsymbol{\omega})$ that depend on $\boldsymbol{\omega}$ (which of course also covers the case in which the control does not depend on $\boldsymbol{\omega}$). The no-gap condition (i.e., a bound on $J_h(u_h^*(\boldsymbol{\omega})) - J(u^*)$) can then be obtained using classical arguments from the calculus of variations and Lemma 1 applied to the optimal controls u^* and $u_h^*(\boldsymbol{\omega})$.

In the fourth stage, we bound the difference between the RBM-optimal control $u_h^*(\boldsymbol{\omega})$ and the control u^* optimized for the original dynamics. To this

end, we first bound the expected difference between the gradients of $J_h(\boldsymbol{\omega}, \cdot)$ and J. The bound on the difference in the optimal controls then follows from classical arguments based on the α-convexity of the functional $J_h(\boldsymbol{\omega}, \cdot)$. Finally, the bound for the difference $J(u_h^*(\boldsymbol{\omega})) - J(u^*)$ follows easily from the previously derived bound on the difference between the optimal controls $u_h^*(\boldsymbol{\omega})$ and u^*.

The four stages discussed above will be proved in detail in Subsections 3.2–3.5. We first present some preliminaries in Subsection 3.1.

3.1 Preliminaries

We will use the following notation. The transpose of a real column vector x is denoted by $x^\top$. Similarly, the transpose of a real matrix A is denoted by $A^\top$. The entry in the i-th row and j-th column of A is denoted by $[A]_{ij}$. The standard Euclidean innerproduct of two vectors $x, y \in \mathbb{R}^N$ is denoted by $\langle x, y \rangle := x^\top y$. The corresponding norm is denoted by $|x| := \sqrt{x^\top x}$. The (operator) norm of a matrix $A \in \mathbb{R}^{N \times N}$ is denoted by

$$\|A\| := \max_{|x|=1} |Ax|. \tag{36}$$

We use $C_{[a,b,\ldots,d]}$ to denote a constant that only depends on the parameters $a, b, \ldots, d$. The value of $C_{[a,b,\ldots,d]}$ may vary from line to line. The L^p-norm of a function in $u \in L^p(0,T;\, \mathbb{R}^q)$ (for $1 \leq p < \infty$ and $p = \infty$) is defined as

$$|u|_{L^p(0,T;\, \mathbb{R}^q)} := \sqrt[p]{\int_0^T |u(t)|^p \, \mathrm{d}t}, \qquad |u|_{L^\infty(0,T;\, \mathbb{R}^q)} := \operatorname*{ess\,sup}_{t \in [0,T]} |u(t)|. \tag{37}$$

We now set up the precise probabilistic setting for our problem. The set Ω^K defined in (10) is the natural sample space for the considered problem. To turn Ω^K into a probability space, we assign a probability $p(\boldsymbol{\omega})$ to each $\boldsymbol{\omega} \in \Omega^K$ according to

$$p(\boldsymbol{\omega}) = p_{\omega_1} p_{\omega_2} \ldots p_{\omega_K}. \tag{38}$$

Note that we use here that each index ω_k is chosen independently from the other indices $\omega_1, \omega_2, \ldots, \omega_{k-1}, \omega_{k+1}, \omega_{k+1}, \ldots, \omega_K$.

A random element on the sample space Ω^K is a function $X : \Omega^K \to V$ from the sample space Ω^K to a vector space V. When $V = \mathbb{R}$, $X : \Omega^K \to \mathbb{R}$ is also called a random variable. Note that we can embed V into V^{Ω^K} by associating to each element $x \in V$ the constant function $X(\boldsymbol{\omega}) = x$ for all $\boldsymbol{\omega} \in \Omega^K$. Constant functions $X(\boldsymbol{\omega}) = x$ are called deterministic. Functions $X(\boldsymbol{\omega})$ that are not deterministic are called stochastic. The expectation operator $\mathbb{E}$ assigns to a random variable $X \in V^{\Omega^K}$ an element of the vector space V

$$\mathbb{E}[X] = \sum_{\boldsymbol{\omega} \in \Omega^K} X(\boldsymbol{\omega}) p(\boldsymbol{\omega})$$

$$= \sum_{\omega_1=1}^{2^M} \sum_{\omega_2=1}^{2^M} \cdots \sum_{\omega_K=1}^{2^M} X(\omega_1, \omega_2, \ldots, \omega_K) p_{\omega_1} p_{\omega_2} \cdots p_{\omega_K}. \tag{39}$$

It is immediate from this definition that $\mathbb{E}$ is linear. When $V = \mathbb{R}$, we also see that $\mathbb{E}[X] \geq 0$ when $X(\boldsymbol{\omega}) \geq 0$ for all $\boldsymbol{\omega} \in \Omega^K$.

Several random elements appear in the proposed randomized splitting method. One example is the matrix $\mathcal{A}_h(\boldsymbol{\omega}, t)$ defined in (11). When $t \in [t_{k-1}, t_k)$, $\mathcal{A}_h(\boldsymbol{\omega} t)$ only depends on ω_k. Therefore, the definitions in (39) and (11) show that (for $t \in [t_{k-1}, t_k)$)

$$\begin{aligned}\mathbb{E}[\mathcal{A}_h(t)] &= \sum_{\boldsymbol{\omega} \in \Omega^K} \mathcal{A}_h(\boldsymbol{\omega}, t) p(\boldsymbol{\omega}) = \sum_{\omega_1=1}^{2^M} \sum_{\omega_2=1}^{2^M} \cdots \sum_{\omega_K=1}^{2^M} \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m} p_{\omega_1} p_{\omega_2} \cdots p_{\omega_K} \\ &= \sum_{\omega_k=1}^{2^M} \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m} p_{\omega_k} = A, \end{aligned} \tag{40}$$

where the second to last identity follows from (8) and the last identity from (12). Again using that $\mathcal{A}_h(\boldsymbol{\omega}, t)$ only depends on ω_k for $t \in [t_{k-1}, t_k)$, we also find that

$$\begin{aligned}\mathbb{E}[\|\mathcal{A}_h(t) - A\|^2] &= \sum_{\boldsymbol{\omega} \in \Omega^K} \|\mathcal{A}_h(\boldsymbol{\omega}, t) - A\|^2 p(\boldsymbol{\omega}) \\ &= \sum_{\omega_k=1}^{2^M} \left\| \sum_{m \in S_{\omega_k}} \frac{A_m}{\pi_m} - A \right\|^2 p_{\omega_k} = \mathrm{Var}[\mathcal{A}], \end{aligned} \tag{41}$$

where the last identity follows from the definition of $\mathrm{Var}[\mathcal{A}]$ in (17). Note that (41) holds for every time instant t and that $\mathbb{E}[\|\mathcal{A}_h(t) - A\|^2]$ therefore does not depend on the considered time instant t.

Another random element is the solution $x_h : \Omega^K \to L^2(0, T; \mathbb{R}^N)$ in (13). We will frequently use that $|x_h(\boldsymbol{\omega}, t)|$ can be bounded as follows. First of all, observe that

$$\frac{d}{dt}|x_h(\boldsymbol{\omega}, t)|^2 = 2\langle x_h(\boldsymbol{\omega}, t), \mathcal{A}_h(\boldsymbol{\omega}, t) x_h(\boldsymbol{\omega}, t) + Bu(t) \rangle \leq 2|x_h(\boldsymbol{\omega}, t)||Bu(t)|, \tag{42}$$

where is was used that $\langle x, \mathcal{A}_h(\boldsymbol{\omega}, t) x \rangle \leq 0$ for all $x \in \mathbb{R}^N$ and $\boldsymbol{\omega} \in \Omega^K$ because of Assumption 1. Now observe that

$$\frac{d}{dt}|x_h(\boldsymbol{\omega}, t)| = \frac{1}{2|x_h(\boldsymbol{\omega}, t)|} \frac{d}{dt}|x_h(\boldsymbol{\omega}, t)|^2 \leq |Bu(t)|, \tag{43}$$

from which we conclude that

$$|x_h(\boldsymbol{\omega})|_{L^\infty(0,T;\,\mathbb{R}^N)} \leq |x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)}. \tag{44}$$

For $x(t)$, a similar derivation shows that

$$|x|_{L^\infty(0,T;\,\mathbb{R})} \leq |x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)}. \tag{45}$$

We will also consider situations in which we apply an input $u_h(\boldsymbol{\omega}, t)$ to the dynamics (1) and (13) that depends on $\boldsymbol{\omega}$. The resulting solutions are then both random elements $x(\boldsymbol{\omega}, t)$ and $x_h(\boldsymbol{\omega}, t)$ which satisfy

$$\dot{x}(\boldsymbol{\omega}, t) = Ax(\boldsymbol{\omega}, t) + Bu_h(\boldsymbol{\omega}, t), \qquad x(\boldsymbol{\omega}, 0) = x_0, \tag{46}$$

$$\dot{x}_h(\boldsymbol{\omega}, t) = \mathcal{A}_h(\boldsymbol{\omega}, t)x_h(\boldsymbol{\omega}, t) + Bu_h(\boldsymbol{\omega}, t), \qquad x_h(\boldsymbol{\omega}, 0) = x_0, \tag{47}$$

In this case we can obtain estimates similar to (44) and (45) with u and x replaced by $u_h(\boldsymbol{\omega})$ and $x(\boldsymbol{\omega})$, respectively.

The third important random element in this paper is the optimal control $u_h^*(\boldsymbol{\omega}, \cdot)$ that minimizes $J_h(\boldsymbol{\omega}, \cdot)$ in (14). The coercivity of the functional $J_h(\boldsymbol{\omega}, \cdot)$ allows us to bound $|u_h^*(\boldsymbol{\omega})|_{L^2(0,T;\,\mathbb{R}^q)}$ as follows. Denote the smallest eigenvalue of the matrix R by $\alpha > 0$, then

$$\frac{\alpha}{2}|u_h^*(\boldsymbol{\omega})|^2_{L^2(0,T;\,\mathbb{R}^q)} \leq \frac{1}{2}\int_0^T u_h^*(t)^\top R u_h^*(t)\,\mathrm{d}t \leq J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega})) \leq J_h(\boldsymbol{\omega}, 0), \tag{48}$$

where the last inequality follows because $u_h^*(\boldsymbol{\omega})$ is the minimizer of $J_h(\boldsymbol{\omega}, \cdot)$. Next, observe that

$$\begin{aligned} J_h(\boldsymbol{\omega}, 0) &\leq \frac{1}{2}\int_0^T (x_h(\boldsymbol{\omega}, t) - x_d(t))^\top Q(x_h(\boldsymbol{\omega}, t) - x_d(t))\,\mathrm{d}t \\ &\leq \tfrac{1}{2}\|Q\|\left(|x_h(\boldsymbol{\omega})|_{L^2(0,T;\,\mathbb{R}^N)} + |x_d|_{L^2(0,T;\,\mathbb{R}^N)}\right)^2 \\ &\leq \tfrac{1}{2}\|Q\|\left(T|x_0| + |x_d|_{L^2(0,T;\,\mathbb{R}^N)}\right)^2 = C_{[x_0,Q,x_d,T]}, \end{aligned} \tag{49}$$

where $x_h(\boldsymbol{\omega}, t)$ denotes the solution of (13) with $u(t) = 0$ and the last inequality follows from (44). Looking back at (48), we find

$$|u_h^*(\boldsymbol{\omega})|^2_{L^2(0,T;\,\mathbb{R}^N)} \leq C_{[x_0,Q,R,x_d,T]}. \tag{50}$$

Finally, we repeat some standard definitions from the theory of the convex optimization, see, e.g., [23]. A functional $J : V \to \mathbb{R}$ on a normed vector space V is α-convex if there exists an $\alpha \geq 0$ such that for all $u, v \in V$ and $\theta \in [0, 1]$

$$J((1-\theta)u + \theta v) \leq (1-\theta)J(u) + \theta J(v) - \tfrac{\alpha}{2}\theta(1-\theta)|u - v|_V^2. \tag{51}$$

One can easily verify that the functional $J_h(\boldsymbol{\omega}, \cdot)$ is α-convex (for all $\boldsymbol{\omega} \in \Omega^K$) when we take α as the smallest eigenvalue of the positive definite matrix R. The Gâteaux-derivative of J at the point u in the direction v is denoted by $\delta J(u;\ v)$, i.e.

$$\delta J(u;\ v) := \lim_{h \to 0} \frac{J(u + hv) - J(u)}{h}. \tag{52}$$

By subtracting $J(u)$ from both sides of (51), dividing the resulting inequality by θ, and then taking the limit $\theta \to 0$, we find the well-known inequality

$$J(v) \geq J(u) + \delta J(u;\ v - u) + \tfrac{\alpha}{2}|v - u|_V^2. \tag{53}$$

3.2 The forward dynamics with a deterministic input

In this subsection, we consider a deterministic $u(t)$ and derive a bound for the error

$$e_h(\boldsymbol{\omega}, t) := x_h(\boldsymbol{\omega}, t) - x(t), \tag{54}$$

where $x_h(\boldsymbol{\omega}, t)$ and $x(t)$ are the solutions of (13) and (1) resulting from the same input $u(t)$, respectively.

Remark 6 It is important to stress that $x_h(t)$ is not an unbiased estimator for $x(t)$, i.e., we do *not* have $\mathbb{E}[e(t)] = \mathbb{E}[x_h(t)] - x(t) = 0$. This can for example be observed when we write the error dynamics as

$$\begin{aligned} \dot{e}_h(\boldsymbol{\omega}, t) &= \mathcal{A}_h(\boldsymbol{\omega}, t) x_h(\boldsymbol{\omega}, t) + Bu(t) - Ax(\boldsymbol{\omega}, t) - Bu(t) \\ &= Ae_h(\boldsymbol{\omega}, t) + (\mathcal{A}_h(\boldsymbol{\omega}, t) - A) x_h(\boldsymbol{\omega}, t) \qquad e_h(\boldsymbol{\omega}, 0) = 0, \end{aligned} \tag{55}$$

where we have substituted $x(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - e_h(\boldsymbol{\omega}, t)$. Taking the expected value in (55) we find

$$\frac{d}{dt}\mathbb{E}[e_h(t)] = A\mathbb{E}[e_h(t)] + \mathbb{E}[(\mathcal{A}_h(t) - A)x_h(t)], \qquad \mathbb{E}[e_h(0)] = 0. \tag{56}$$

However, (56) does not imply that $\mathbb{E}[e_h(t)] = 0$ for all t because generally

$$\mathbb{E}[(\mathcal{A}_h(t) - A)x_h(t)] \neq \mathbb{E}[\mathcal{A}_h(t) - A]\mathbb{E}[x_h(t)] = 0, \tag{57}$$

where the equality follows because $\mathbb{E}[\mathcal{A}_h(t)] = A$, see (40). This would be the case when $\mathcal{A}_h(\boldsymbol{\omega}, t)$ and $x_h(\boldsymbol{\omega}, t)$ are independent, but they are correlated by the dynamics (13). Note, however, that at the beginning of each time interval $[t_{k-1}, t_k)$, the value of $\mathcal{A}_h(\boldsymbol{\omega}, t)$ changes and that $\mathcal{A}_h(\boldsymbol{\omega}, t_{k-1})$ is independent of the values of $\mathcal{A}_h(\boldsymbol{\omega}, t)$ for $t < t_{k-1}$ so that

$$\mathbb{E}[(\mathcal{A}_h(t_{k-1}) - A)x_h(t_{k-1})] = \mathbb{E}[\mathcal{A}_h(t_{k-1}) - A]\mathbb{E}[x_h(t_{k-1})] = 0, \tag{58}$$

where the second identity again follows because $\mathbb{E}[\mathcal{A}_h(t)] = A$, see (40). This observation is crucial to obtain the main result of this subsection.

The main result in this subsection is the following.

Theorem 1 *Assume that the input $u(t)$ in (13) is deterministic and equal to the input $u(t)$ in (1) and that Assumptions 1 and 2 hold, then*

$$\mathbb{E}[|e_h(t)|^2] \leq h\mathrm{Var}[\mathcal{A}](\|A\|t^2 + 2t)(|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2. \tag{59}$$

Proof Observe that

$$\begin{aligned}\dot{e}_h(\boldsymbol{\omega},t) &= \mathcal{A}_h(\boldsymbol{\omega},t)x_h(\boldsymbol{\omega},t) + Bu(t) - Ax(\boldsymbol{\omega},t) - Bu(t)\\ &= \mathcal{A}_h(\boldsymbol{\omega},t)e_h(\boldsymbol{\omega},t) + (\mathcal{A}_h(\boldsymbol{\omega},t) - A)x(t) \qquad e_h(\boldsymbol{\omega},0) = 0, \end{aligned} \tag{60}$$

where the last equation follows after substituting $x_h(\boldsymbol{\omega},t) = x(\boldsymbol{\omega},t) + e_h(\boldsymbol{\omega},t)$.

Fix $t \in [0,T]$ and let $k \le K$ be such that $t \in [t_{k-1}, t_k)$.

Consider an arbitrary time instant $s \in [0,t)$ and let $\ell \in \{1,2,\dots,k\}$ be such that $s \in [t_{\ell-1}, t_\ell)$. Then (60) shows that

$$\begin{aligned}\frac{d}{ds}|e_h(\boldsymbol{\omega},s)|^2 &= 2\langle e_h(\boldsymbol{\omega},s), \mathcal{A}_h(\boldsymbol{\omega},s)e_h(\boldsymbol{\omega},s)\rangle + 2\langle e_h(\boldsymbol{\omega},s), (\mathcal{A}_h(\boldsymbol{\omega},s) - A)x(s)\rangle\\ &= 2\langle e_h(\boldsymbol{\omega},s), \mathcal{A}_h(\boldsymbol{\omega},s)e_h(\boldsymbol{\omega},s)\rangle + 2\langle e_h(\boldsymbol{\omega},t_{\ell-1}), (\mathcal{A}_h(\boldsymbol{\omega},s) - A)x(s)\rangle\\ &\quad + 2\langle \Delta e_h(\boldsymbol{\omega},s), (\mathcal{A}_h(\boldsymbol{\omega},s) - A)x(s)\rangle, \end{aligned} \tag{61}$$

where, in the second equality, we have introduced

$$\Delta e_h(\boldsymbol{\omega},s) := e_h(\boldsymbol{\omega},s) - e_h(\boldsymbol{\omega},t_{\ell-1}). \tag{62}$$

The first term on the RHS of (61) is nonpositive due to Assumption 1. We thus find after taking the expected value in (61) that

$$\begin{aligned}\frac{d}{ds}\mathbb{E}[|e_h(s)|^2] &\le 2\mathbb{E}[\langle e_h(t_{\ell-1}), (\mathcal{A}_h(s) - A)x(s)\rangle]\\ &\quad + 2\mathbb{E}[\langle \Delta e_h(s), (\mathcal{A}_h(s) - A)x(s)\rangle]. \end{aligned} \tag{63}$$

For the first term on the RHS of (63), observe that $e_h(\boldsymbol{\omega},t_{\ell-1}) = x_h(\boldsymbol{\omega},t_{\ell-1}) - x(t_{\ell-1})$ only depends on $\omega_1,\dots\omega_{\ell-1}$, so that

$$\begin{aligned}\mathbb{E}[\langle e_h(t_{\ell-1}), (\mathcal{A}_h(s) - A)x(s)\rangle] &= \sum_{\boldsymbol{\omega}\in\Omega^K} \langle e_h(\boldsymbol{\omega},t_{\ell-1}), (\mathcal{A}_h(\boldsymbol{\omega},s) - A)x(s)\rangle p(\boldsymbol{\omega})\\ &= \sum_{\omega_1=1}^{2^M}\cdots\sum_{\omega_{\ell-1}=1}^{2^M}\sum_{\omega_\ell=1}^{2^M}\left\langle e_h(\boldsymbol{\omega},t_{\ell-1}), \left(\sum_{m\in S_{\omega_\ell}}\frac{A_m}{\pi_m} - A\right)x(s)\right\rangle p_{\omega_1}\cdots p_{\omega_{\ell-1}}p_{\omega_\ell}\\ &= \sum_{\omega_1=1}^{2^M}\cdots\sum_{\omega_{\ell-1}=1}^{2^M}\left\langle e_h(\boldsymbol{\omega},t_{\ell-1}), \left(\sum_{\omega_\ell=1}^{2^M}\sum_{m\in S_{\omega_\ell}}\frac{A_m}{\pi_m}p_{\omega_\ell} - A\right)x(s)\right\rangle p_{\omega_1}\cdots p_{\omega_{\ell-1}}\\ &= 0, \end{aligned} \tag{64}$$

where the second identity uses (8), the third identity follows from (8) and the fact that $e_h(\boldsymbol{\omega},t)$ does not depend on ω_ℓ, and the last identity follows because (12) shows that the factor between round brackets vanishes.

For the second term on the RHS of (63), we use that

$$\begin{aligned}\mathbb{E}[\langle \Delta e_h(s), (\mathcal{A}_h(s) - A)x(s)\rangle] &\le \mathbb{E}[|\Delta e_h(s)|\|\mathcal{A}_h(s) - A\||x(s)|]\\ &\le \sqrt{\mathbb{E}[|\Delta e_h(s)|^2]\mathbb{E}[\|\mathcal{A}_h(s) - A\|^2|x(s)|^2]} = \sqrt{\mathbb{E}[|\Delta e_h(s)|^2]}\sqrt{\mathrm{Var}[\mathcal{A}]}|x(s)|\\ &\le \sqrt{\mathbb{E}[|\Delta e_h(s)|^2]}\sqrt{\mathrm{Var}[\mathcal{A}]}(|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)}), \end{aligned} \tag{65}$$

where the first identity follows from the Cauchy-Schwartz inequality in $\mathbb{R}^N$, the second inequality from Cauchy-Schwartz inequality in the probability space, and the last inequality follows from (45).

We now claim that

$$\mathbb{E}[|\Delta e_h(s)|^2] \le h^2\mathrm{Var}[\mathcal{A}](\|A\|s+1)^2(|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2. \tag{66}$$

We will prove (66) at the end of the proof. Inserting the claim (66) into (65), we find

$$\mathbb{E}[\langle \Delta e_h(s), (\mathcal{A}_h(s) - A)x(s)\rangle] \leq h\mathrm{Var}[\mathcal{A}](\|A\|s+1)(|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2. \quad (67)$$

Inserting (64) and (67) into (63) shows that

$$\frac{d}{ds}\mathbb{E}[|e_h(s)|^2] \leq 2h\mathrm{Var}[\mathcal{A}](\|A\|s+1)(|x_0| + |Bu_d|_{L^1(0,T;\,\mathbb{R}^N)})^2. \quad (68)$$

Integrating (68) from $s = 0$ to $s = t$ using that $e_h(\omega, 0) = 0$ now shows that

$$\mathbb{E}[|e_h(t)|^2] \leq h\mathrm{Var}[\mathcal{A}](\|A\|t^2 + 2t)(|x_0| + |Bu_d|_{L^1(0,T;\,\mathbb{R}^N)})^2, \quad (69)$$

which is the desired estimate (59).

It thus remains to show that (66) holds. Recall that, for $\tau \in [t_{\ell-1}, s)$, (62) shows that $\Delta e_h(\boldsymbol{\omega}, \tau) = e_h(\boldsymbol{\omega}, \tau) - e_h(\boldsymbol{\omega}, t_{\ell-1})$. Using (55), we thus see that $\Delta e_h(\boldsymbol{\omega}, \tau)$ is the solution of the ODE

$$\tfrac{d}{d\tau}\Delta e_h(\boldsymbol{\omega}, \tau) = \dot{e}_h(\boldsymbol{\omega}, t) = Ae_h(\boldsymbol{\omega}, t) + (\mathcal{A}_h(\boldsymbol{\omega}, t) - A)x_h(\boldsymbol{\omega}, t), \quad (70)$$

with initial condition $\Delta e_h(\boldsymbol{\omega}, t_{\ell-1}) = 0$. We therefore also have that

$$\frac{d}{d\tau}|\Delta e_h(\boldsymbol{\omega}, \tau)| = \frac{\langle \Delta e_h(\boldsymbol{\omega}, \tau), \dot{e}_h(\boldsymbol{\omega}, \tau)\rangle}{|\Delta e_h(\boldsymbol{\omega}, \tau)|} \leq |Ae_h(\boldsymbol{\omega}, \tau)| + |(\mathcal{A}_h(\boldsymbol{\omega}, \tau) - A)x_h(\boldsymbol{\omega}, \tau)|. \quad (71)$$

Using that $\Delta e_h(\boldsymbol{\omega}, t_{\ell-1}) = 0$, integrating (71) from $\tau = t_{\ell-1}$ to $\tau = s$ yields

$$|\Delta e_h(\boldsymbol{\omega}, s)| \leq \int_{t_{\ell-1}}^{s} (\|A\||e_h(\boldsymbol{\omega}, \tau)| + |(\mathcal{A}_h(\boldsymbol{\omega}, \tau) - A)x_h(\boldsymbol{\omega}, \tau)|)\ \mathrm{d}\tau. \quad (72)$$

To bound $e_h(\boldsymbol{\omega}, \tau)$, we apply the variation of constants formula to the error dynamics in (55) and obtain

$$\begin{aligned} |e_h(\boldsymbol{\omega}, \tau)| &= \left| \int_0^\tau e^{A(\tau-\sigma)}(\mathcal{A}_h(\boldsymbol{\omega}, \sigma) - A)x_h(\boldsymbol{\omega}, \sigma)\ \mathrm{d}\sigma \right| \\ &\leq \int_0^\tau \|\mathcal{A}_h(\boldsymbol{\omega}, \sigma) - A\|\ \mathrm{d}\sigma\ (|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)}), \end{aligned} \quad (73)$$

where we have used the bound for $x_h(\boldsymbol{\omega}, \sigma)$ in (44) and that $\|e^{A(\tau-\sigma)}\| \leq 1$ because Assumption 1 implies that A is dissipative. Using this result in (72), we find

$$|\Delta e_h(\boldsymbol{\omega}, s)| \leq \int_{t_{\ell-1}}^{s} g(\boldsymbol{\omega}, \tau)\ \mathrm{d}\tau\ (|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)}), \quad (74)$$

where we have again used the bound on $x_h(\boldsymbol{\omega}, t)$ in (44) for the second term in (72) and introduced

$$g(\boldsymbol{\omega}, \tau) := \left(\|A\| \int_0^\tau \|\mathcal{A}_h(\boldsymbol{\omega}, \sigma) - A\|\ \mathrm{d}\sigma + \|\mathcal{A}_h(\boldsymbol{\omega}, \tau) - A\| \right). \quad (75)$$

Squaring both sides in (74) and taking the expectation, we find

$$\begin{aligned} \mathbb{E}[|\Delta e_h(s)|^2] &\leq \mathbb{E}\left[\left(\int_{t_{\ell-1}}^{s} g(\tau)\ \mathrm{d}\tau \right)^2 \right] (|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2 \\ &\leq (s - t_{\ell-1}) \int_{t_{\ell-1}}^{s} \mathbb{E}[(g(\tau))^2]\ \mathrm{d}\tau\ (|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2, \end{aligned} \quad (76)$$

where the second inequality follows from the Cauchy-Schwartz inequality in $L^2(t_{\ell-1}, s)$. Now observe that (75) shows that

$$\mathbb{E}[(g(\tau))^2] = \|A\|^2 \int_0^\tau \int_0^\tau \mathbb{E}[\|\mathcal{A}_h(\sigma) - A\|\|\mathcal{A}_h(\sigma') - A\|]\ \mathrm{d}\sigma\ \mathrm{d}\sigma'$$

$$+ 2\|A\| \int_0^\tau \mathbb{E}[\|\mathcal{A}_h(\sigma) - A\|\|\mathcal{A}_h(\tau) - A\|]\, \mathrm{d}\sigma + \mathbb{E}[\|\mathcal{A}_h(\tau) - A\|^2]. \quad (77)$$

Because $\mathbb{E}[\|\mathcal{A}_h(t) - A\|^2] = \mathrm{Var}[\mathcal{A}]$ for all t, we also have that

$$\mathbb{E}[\|\mathcal{A}_h(\sigma) - A\|\|\mathcal{A}_h(\tau) - A\|] \leq \sqrt{\mathbb{E}[\|\mathcal{A}_h(\sigma) - A\|^2]\mathbb{E}[\|\mathcal{A}_h(\tau) - A\|^2]} = \mathrm{Var}[\mathcal{A}]. \quad (78)$$

Using this result in (77), we obtain

$$\mathbb{E}[(g(\tau))^2] \leq \mathrm{Var}[\mathcal{A}](\|A\|\tau + 1)^2. \quad (79)$$

Using this result again in (76), also using that $s - t_{\ell-1} \leq h$ and $\tau \leq s$, we find the claimed inequality (66). □

Some remarks regarding Theorem 1 are in order.

Remark 7 The error estimate in Theorem 1 involves the operator norm of the matrix A. This suggests that the expected error $\mathbb{E}[|e_h(t)|^2]$ grows when we are considering better approximations A of an unbounded operator, which for example happens when we consider a discretization of a PDE and refine the spatial grid. However, Figure 4a in Section 4 indicates that $\mathbb{E}[|e_h(t)|] \leq C\sqrt{h\mathrm{Var}[\mathcal{A}]}$ for a constant C that does not increase (but even seems to decrease) when the spatial grid is refined.

A first step in understanding the infinite-dimensional case better is taken in Appendix B, where we prove that

$$\mathbb{E}[|e_h(t)|^2] \leq 2ht\mathrm{Var}_W[\mathcal{A}]|W^{-1}x_0|. \quad (80)$$

under the additional assumptions that $u(t) \equiv 0$ and that all matrices A_m commute pairwise. Here, W is any invertible matrix and $\mathrm{Var}_W[\mathcal{A}]$ is the weighted variance introduced in Remark 5. Observe that the operator norm $\|A\|$ does not appear in this estimate. The result from Appendix B extends naturally to an infinite dimensional setting in which all operators A_m have the same domain $D(A_m) = D(A)$.

Recall from Remark 5 that a typical choice for W is $W = (A - \lambda I)^{-1}$ for some λ in the resolvent of A. For $|W^{-1}x_0|$ to be bounded, we thus require that $x_0 \in D(A)$, where $D(A)$ denotes the domain of the operator A. In an infinite dimensional setting we thus need an additional smoothness assumption on the initial condition x_0. Such conditions are typical for (deterministic) splitting algorithms, see e.g. [13, 14]. Further details can be found in Appendix B.

Remark 8 The error estimate in Theorem 1 is derived based on the error dynamics (60). Considering the error dynamics (55) leads to a less clean proof because instead of the 3 terms on the RHS of (61), we then get 4 terms

$$\begin{aligned}\frac{d}{ds}|e_h(\boldsymbol{\omega}, s)|^2 &= 2\langle e_h(\boldsymbol{\omega}, s), Ae_h(\boldsymbol{\omega}, s)\rangle + 2\langle e_h(\boldsymbol{\omega}, s), (\mathcal{A}_h(\boldsymbol{\omega}, s) - A)x_h(\boldsymbol{\omega}, s)\rangle \\ &= 2\langle e_h(\boldsymbol{\omega}, s), Ae_h(\boldsymbol{\omega}, s)\rangle + 2\langle e_h(\boldsymbol{\omega}, t_{\ell-1}), (\mathcal{A}_h(\boldsymbol{\omega}, s) - A)x_h(\boldsymbol{\omega}, t_{\ell-1})\rangle \\ &\quad + 2\langle \Delta e_h(\boldsymbol{\omega}, s), (\mathcal{A}_h(\boldsymbol{\omega}, s) - A)x_h(\boldsymbol{\omega}, s)\rangle \\ &\quad + 2\langle e_h(\boldsymbol{\omega}, s), (\mathcal{A}_h(\boldsymbol{\omega}, s) - A)\Delta x_h(\boldsymbol{\omega}, s)\rangle, \end{aligned} \quad (81)$$

where $\Delta e_h(\boldsymbol{\omega}, s) := e_h(\boldsymbol{\omega}, s) - e_h(\boldsymbol{\omega}, t_{\ell-1})$ and $\Delta x_h(\boldsymbol{\omega}, s) := x_h(\boldsymbol{\omega}, s) - x_h(\boldsymbol{\omega}, t_{\ell-1})$. This approach is closer to proofs for interacting particle systems in [15].

Note that the fourth term in (81) is needed because $x_h(\boldsymbol{\omega}, s)$ is correlated to $\mathcal{A}_h(\boldsymbol{\omega}, s)$ for $s \in (t_{\ell-1}, t_\ell)$. Because $x(s)$ is not correlated to $\mathcal{A}_h(\boldsymbol{\omega}, s)$, it was not necessary to introduce such a term in (61). The proof of Theorem 1 based on the error dynamics (60) presented above is thus simpler than a proof based on (55).

Remark 9 When we look back at the proof of Theorem 1, we see that Assumption 1 is only used to assure that the matrices A and $\mathcal{A}_h(\boldsymbol{\omega}, t)$ are dissipative (for all $\boldsymbol{\omega}$ with $p(\boldsymbol{\omega}) > 0$ and all $t \in [0, T]$). When Assumption 1 is not satisfied, there must exist a constant $a > 0$ such that $\hat{A} = A - aI$ and $\hat{\mathcal{A}}_h(\boldsymbol{\omega}, t) = \mathcal{A}_h(\boldsymbol{\omega}, t) - aI$ are dissipative (for all $\boldsymbol{\omega}$ with $p(\boldsymbol{\omega}) > 0$ and all $t \in [0, T]$). Because $\mathbb{E}[\mathcal{A}_h(t)] = A$, it follows that $\mathbb{E}[\hat{\mathcal{A}}_h(t)] = \mathbb{E}[\mathcal{A}_h(t)] - aI = A - aI = \hat{A}$ and $\mathrm{Var}[\|\hat{\mathcal{A}}_h(t) - \hat{A}\|^2] = \mathrm{Var}[\mathcal{A}]$. When we let $\hat{x}(t)$ and $\hat{x}_h(\boldsymbol{\omega}, t)$ denote the solutions generated by $\hat{A}$ and $\hat{\mathcal{A}}_h(\boldsymbol{\omega}, t)$, respectively, we can now proof in a similar way as in Theorem 1 that the error $\hat{e}_h(\boldsymbol{\omega}, t) = \hat{x}_h(\boldsymbol{\omega}, t) - \hat{x}(t)$ can be bounded as

$$\mathbb{E}[|\hat{e}_h(t)|^2] \leq h\mathrm{Var}[\mathcal{A}](\|\hat{A}\|t^2 + 2t)(|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2. \tag{82}$$

Because $x(t) = e^{at}\hat{x}(t)$ and $x_h(\boldsymbol{\omega}, t) = e^{at}\hat{x}_h(\boldsymbol{\omega}, t)$, also

$$e_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x(t) = e^{at}\hat{x}_h(\boldsymbol{\omega}, t) - e^{at}\hat{x}(t) = e^{at}\hat{e}_h(\boldsymbol{\omega}, t). \tag{83}$$

Taking the expectation and using (82), we find

$$\mathbb{E}[|e_h(t)|^2] \leq he^{at}\mathrm{Var}[\mathcal{A}](\|\hat{A}\|t^2 + 2t)(|x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)})^2. \tag{84}$$

The error estimate now grows exponentially in time.

3.3 The forward dynamics with a stochastic input

In this subsection, we prove a result similar to Theorem 1 for inputs $u_h(\boldsymbol{\omega}, t)$ that are stochastic, i.e., which depend on $\boldsymbol{\omega}$. We thus want to bound the error

$$e_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x(\boldsymbol{\omega}, t), \tag{85}$$

where $x_h(\boldsymbol{\omega}, t)$ and $x(\boldsymbol{\omega}, t)$ are the solutions of (47) and (46), respectively.

To this end, we consider the semi-group e^{At} generated by the matrix A and the evolution operator $S_h(\boldsymbol{\omega}, t, s)$ associated to $\mathcal{A}_h(\boldsymbol{\omega}, t)$. The evolution operator $S_h(\boldsymbol{\omega}, t, s)$ is defined by property that for all vectors $x_s \in \mathbb{R}^N$ (and all $t \geq s$), $S_h(\boldsymbol{\omega}, t, s)x_s$ is equal to the solution $y_h(\boldsymbol{\omega}, t)$ of

$$\dot{y}_h(\boldsymbol{\omega}, t) = \mathcal{A}_h(\boldsymbol{\omega}, t)y_h(\omega, t), \qquad y_h(\boldsymbol{\omega}, s) = x_s. \tag{86}$$

Remark 10 An explicit formula for the evolution operator $S_h(\boldsymbol{\omega}, t, s)$ can be obtained as follows. Let $0 \leq s \leq t \leq T$ and let $\ell, k \in \{1, 2, \ldots, K\}$ be selected such that

$$s \in [t_{\ell-1}, t_\ell), \qquad t \in [t_{k-1}, t_k). \tag{87}$$

By restricting the given time grid $0 = t_0 < t_1 < t_2 < \ldots < t_{K-1} < t_K = T$ to the interval $[s, t]$, we obtain a grid with $\tilde{K} = k - \ell + 1$ grid points

$$\tilde{t}_0 := s < \tilde{t}_1 := t_\ell < \tilde{t}_2 := t_{\ell+1} < \ldots < \tilde{t}_{\tilde{K}-1} := t_{k-1} < \tilde{t}_{\tilde{K}} := t. \tag{88}$$

The construction of the time grid $\tilde{t}_0, \tilde{t}_1, \ldots \tilde{t}_{\tilde{K}}$ is illustrated in Figure 1. We also denote $\tilde{h}_p := \tilde{t}_p - \tilde{t}_{p-1}$ (for $p \in \{1, 2, \ldots, \tilde{K}\}$) and introduce (for each $\omega \in \{1, 2, \ldots, 2^M\}$)

$$\mathcal{A}_\omega := \sum_{m \in S_\omega} \frac{A_m}{\pi_m}. \tag{89}$$

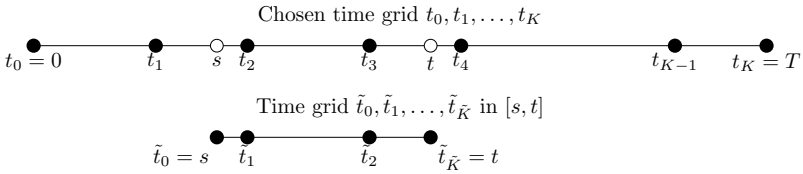


Fig. 1: The relation between the chosen time grid $t_0, t_1, \ldots, t_K$ and the time grid $\tilde{t}_0, \tilde{t}_1, \ldots, \tilde{t}_{\tilde{K}}$ used in Remark 10. In the displayed example, $\ell = 2$, $k = 4$, and $\tilde{K} = 3$.

Because $\mathcal{A}_h(\boldsymbol{\omega}, \tau) = \mathcal{A}_{\omega_p}$ is constant for $\tau \in [\tilde{t}_{p-1}, \tilde{t}_p)$, it is now easy to see that

$$S_h(\boldsymbol{\omega}, t, s) = e^{\mathcal{A}_{\omega_k}\tilde{h}_{\tilde{K}}} \cdots e^{\mathcal{A}_{\omega_{\ell+1}}\tilde{h}_2} e^{\mathcal{A}_{\omega_\ell}\tilde{h}_1} = \prod_{p=1}^{\tilde{K}} e^{\mathcal{A}_{\omega_{p+\ell-1}}\tilde{h}_p}. \tag{90}$$

Under Assumption 1, all matrices $\mathcal{A}_{\omega_p}$ are dissipative and (90) shows that

$$\|S_h(\boldsymbol{\omega}, t, s)\| \leq 1. \tag{91}$$

Using the variation of constants formula, the solutions of $x_h(\boldsymbol{\omega}, t)$ and $x(\boldsymbol{\omega}, t)$ can expressed as

$$x_h(\boldsymbol{\omega}, t) = S_h(\boldsymbol{\omega}, t, 0)x_0 + \int_0^t S_h(\boldsymbol{\omega}, t, s)Bu_h(\boldsymbol{\omega}, s)\, \mathrm{d}s, \tag{92}$$

$$x(\boldsymbol{\omega}, t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu_h(\boldsymbol{\omega}, s)\, \mathrm{d}s. \tag{93}$$

Subtracting (93) from (92) we find the following expression for the error $e_h(\boldsymbol{\omega}, t)$

$$e_h(\boldsymbol{\omega}, t) = E_h(\boldsymbol{\omega}, t, 0)x_0 + \int_0^t E_h(\boldsymbol{\omega}, t, s)Bu_h(\boldsymbol{\omega}, s)\, \mathrm{d}s, \tag{94}$$

where $E_h(\boldsymbol{\omega}, t, s) = S_h(\boldsymbol{\omega}, t, s) - e^{A(t-s)}$. The following corollary of Theorem 1 shows that we can bound $E_h(\boldsymbol{\omega}, t, s) = S_h(\boldsymbol{\omega}, t, s) - e^{A(t-s)}$.

Corollary 1 *Under Assumptions 1 and 2, we have that*

$$\mathbb{E}[\|S_h(t, s) - e^{A(t-s)}\|^2] \leq (\|A\|T^2 + 2T)h\mathrm{Var}[\mathcal{A}], \tag{95}$$

for all $0 \leq s \leq t \leq T$.

Proof Fix $s \in [0, T]$ and an initial condition $x_s \in \mathbb{R}^N$.

Define $y(t) = e^{A(t-s)}x_s$ and let $y_h(\boldsymbol{\omega}, t)$ be the solution of (86), both for $t \in [s, T]$. We then apply Theorem 1 with $u(t) \equiv 0$ to the time-shifted solutions $\tilde{y}(\tilde{t}) = y(\tilde{t} + s)$

and $\tilde{y}_h(\boldsymbol{\omega}, \tilde{t}) = y_h(\boldsymbol{\omega}, \tilde{t} + s)$ and the time-shifted matrix $\tilde{\mathcal{A}}_h(\boldsymbol{\omega}, \tilde{t}) = \mathcal{A}_h(\boldsymbol{\omega}, \tilde{t} + s)$ defined on $\tilde{t} \in [0, T - s]$. We thus conclude that (writing $\tilde{t} = t - s$)

$$\mathbb{E}[|y_h(t) - y(t)|^2] = \mathbb{E}[|\tilde{y}_h(\tilde{t}) - \tilde{y}(\tilde{t})|^2] \leq h\mathrm{Var}[\mathcal{A}](\|A\|\tilde{t}^2 + 2\tilde{t})|x_s|^2. \tag{96}$$

Noting that, by definition, $y(t) = e^{A(t-s)}x_s$ and $y_h(\boldsymbol{\omega}, t) = S_h(\boldsymbol{\omega}, t, s)x_s$, we find that (for $x_s \neq 0$)

$$\mathbb{E}\left[\frac{|(S_h(\boldsymbol{\omega}, t, s) - e^{A(t-s)})x_s|^2}{|x_s|^2}\right] \leq h\mathrm{Var}[\mathcal{A}](\|A\|T^2 + 2T), \tag{97}$$

where it was used that $\tilde{t} = t - s \leq T$. The result now follows from the definition of the operator-norm. □

Remark 11 In Appendix B, we prove a result similar to Corollary 1 under the additional assumption that all matrices A_m commute pairwise. The result in Appendix B extends naturally to an infinite dimensional setting under the additional assumption that the domains of the operators A_m are the same. This is not the case for Corollary 1 because the operator norm $\|A\|$ appears in (95).

We are now ready for the main result of this subsection.

Theorem 2 *Consider any control* $u_h : \Omega^K \to L^2(0, T;\ \mathbb{R}^q)$. *Assume that Assumptions 1 and 2 are satisfied and let* U *be such that*

$$|Bu_h(\boldsymbol{\omega})|_{L^2(0,T;\,\mathbb{R}^q)} \leq U, \tag{98}$$

for all $\boldsymbol{\omega} \in \Omega^K$, *then*

$$\mathbb{E}[|e_h(t)|^2] \leq (\|A\|T^2 + 2T)h\mathrm{Var}[\mathcal{A}]\left(|x_0| + U\sqrt{T}\right)^2. \tag{99}$$

Proof Using the triangle inequality in (94), we find

$$\begin{aligned}|e_h(\boldsymbol{\omega}, t)| &\leq \|E_h(\boldsymbol{\omega}, t, 0)\||x_0| + \int_0^t \|E_h(\boldsymbol{\omega}, t, s)\||Bu_h(\boldsymbol{\omega}, s)|\ \mathrm{d}s \\ &\leq \|E_h(\boldsymbol{\omega}, t, 0)\||x_0| + \sqrt{\int_0^t \|E_h(\boldsymbol{\omega}, t, s)\|^2\ \mathrm{d}s}|Bu_h(\boldsymbol{\omega})|_{L^2(0,T;\,\mathbb{R}^q)},\end{aligned} \tag{100}$$

where the second inequality follows from the Cauchy-Schwarz inequality in $L^2(0, t)$. Squaring both sides and using the bound (98), we find

$$\begin{aligned}|e_h(\boldsymbol{\omega}, t)|^2 \leq \|E_h(\boldsymbol{\omega}, t, 0)\|^2|x_0|^2 + U^2\int_0^t \|E_h(\boldsymbol{\omega}, t, s)\|^2\ \mathrm{d}s \\ + 2U|x_0|\|E_h(\boldsymbol{\omega}, t, 0)\|\sqrt{\int_0^t \|E_h(\boldsymbol{\omega}, t, s)\|^2\ \mathrm{d}s}.\end{aligned} \tag{101}$$

In order to use the bound from Corollary 1 to estimate the last term, note that we can use the Cauchy-Schwartz inequality in the probability space to find

$$\mathbb{E}\left[\|E_h(t, 0)\|\sqrt{\int_0^t \|E_h(t, s)\|^2\ \mathrm{d}s}\right] \leq \sqrt{\mathbb{E}[\|E_h(t, 0)\|^2]\int_0^t \mathbb{E}[\|E_h(t, s)\|^2]\ \mathrm{d}s} \tag{102}$$

Taking the expected value in (101) and using that the bound on $\mathbb{E}[\|E_h(t,s)\|^2]$ from Corollary 1 does not depend on t and s, we find

$$\mathbb{E}[|e_h(t)|^2] \leq (|x_0| + U\sqrt{t})^2(\|A\|T^2 + 2T)h\mathrm{Var}[\mathcal{A}], \tag{103}$$

which gives the desired estimate. □

Remark 12 Because Ω^K is finite, we can always find a constant U such that (98) is satisfied for a given $u_h : \Omega^K \to L^2(0,T;\ \mathbb{R}^q)$. However, when we consider a family of temporal grids for which $h \to 0$, the constant U may depend on h (depending on the considered family of controls $u_h(\boldsymbol{\omega}, t)$). Fortunately, we only need to apply Theorem 2 with $u_h(\boldsymbol{\omega}, t) = u_h^*(\boldsymbol{\omega}, t)$, where $u_h^*(\boldsymbol{\omega}, t)$ is the control that minimizes the cost functional $J_h(\boldsymbol{\omega}, \cdot)$ in (14). For this control, the coercivity of the cost functional $J_h(\boldsymbol{\omega}, \cdot)$ implies that the constant U can be chosen independent of the considered temporal grid, see (50).

Remark 13 Note that the estimate in Theorem 1 depends on the L^1-norm of the control but that estimate in Theorem 2 depends through (98) on the L^2-norm. Setting $u_h(\boldsymbol{\omega}, t) = u(t)$ in Theorem 2 therefore does not give the estimate in Theorem 1. This underlines the additional difficulty posed by stochastic controls.

3.4 A no-gap condition

With the results regarding forward dynamics from the previous two subsections, we are now ready to address the optimal control problem. The main result of this subsection is the no-gap condition in Theorem 3. To prove this result, we need the following technical lemma.

Lemma 1 *Consider any control $u_h : \Omega^K \to L^2(0,T;\ \mathbb{R}^q)$. Assume that Assumptions 1 and 2 hold and let $U > 0$ be such that (98) is satisfied. Then*

$$\mathbb{E}[|J_h(u_h) - J(u_h)|] \leq C_{[A,x_0,Q,x_d,T,U]} \left(\sqrt{h\mathrm{Var}[\mathcal{A}]} + h\mathrm{Var}[\mathcal{A}]\right). \tag{104}$$

Proof Let $x(\boldsymbol{\omega}, t)$ and $x_h(\boldsymbol{\omega}, t)$ be the solutions of (46) and (47) for the considered control $u_h(\boldsymbol{\omega}, t)$. For brevity, we write $\tilde{x}(\boldsymbol{\omega}, t) = x(\boldsymbol{\omega}, t) - x_d(t)$ and $\tilde{x}_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x_d(t)$. By definition of the cost functionals $J(\cdot)$ and $J_h(\boldsymbol{\omega}, \cdot)$ in (2) and (14), we have

$$\begin{aligned}
J_h(\omega, u_h(\boldsymbol{\omega})) - J(u_h(\boldsymbol{\omega})) &= \tfrac{1}{2}\int_0^T \left(\tilde{x}_h(\boldsymbol{\omega}, t)^\top Q\tilde{x}_h(\boldsymbol{\omega}, t) - \tilde{x}(\boldsymbol{\omega}, t)^\top Q\tilde{x}(\boldsymbol{\omega}, t)\right)\,\mathrm{d}t \\
&= \int_0^T \tilde{x}(\boldsymbol{\omega}, t)^\top Q(\tilde{x}_h(\boldsymbol{\omega}, t) - \tilde{x}(\boldsymbol{\omega}, t))\,\mathrm{d}t \\
&\qquad + \tfrac{1}{2}\int_0^T (\tilde{x}_h(\boldsymbol{\omega}, t) - \tilde{x}(\boldsymbol{\omega}, t))^\top Q(\tilde{x}_h(\boldsymbol{\omega}, t) - \tilde{x}(\boldsymbol{\omega}, t))\,\mathrm{d}t \\
&= \int_0^T \left(\tilde{x}(\boldsymbol{\omega}, t)^\top Qe_h(\boldsymbol{\omega}, t) + \tfrac{1}{2}e_h(\boldsymbol{\omega}, t)^\top Qe_h(\boldsymbol{\omega}, t)\right)\,\mathrm{d}t,
\end{aligned} \tag{105}$$

where the last identity follows because $e_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x(t) = \tilde{x}_h(\boldsymbol{\omega}, t) - \tilde{x}(t)$. Taking the absolute value and estimating the RHS, we find

$$|J_h(\boldsymbol{\omega}, u_h) - J(u_h(\boldsymbol{\omega}))| \leq \|Q\| \int_0^T \left(|\tilde{x}(\boldsymbol{\omega}, t)||e_h(\boldsymbol{\omega}, t)| + \tfrac{1}{2}|e_h(\boldsymbol{\omega}, t)|^2 \right) \, \mathrm{d}t$$
$$\leq \|Q\| \left(|\tilde{x}(\boldsymbol{\omega})|_{L^2(0,T;\, \mathbb{R}^N)} |e_h(\boldsymbol{\omega})|_{L^2(0,T;\, \mathbb{R}^N)} + \tfrac{1}{2}|e_h(\boldsymbol{\omega})|^2_{L^2(0,T;\, \mathbb{R}^N)} \right). \quad (106)$$

Taking the expectation and using the Cauchy-Schwartz inequality, we find that

$$\mathbb{E}[|J_h(u_h) - J(u_h)|] \leq$$
$$\|Q\| \left(\sqrt{\mathbb{E}[|\tilde{x}|^2_{L^2(0,T;\, \mathbb{R}^N)}]} \sqrt{\mathbb{E}[|e_h|^2_{L^2(0,T;\, \mathbb{R}^N)}]} + \tfrac{1}{2}\mathbb{E}[|e_h|^2_{L^2(0,T;\, \mathbb{R}^N)}] \right). \quad (107)$$

Using the estimate from Theorem 2, we find

$$\mathbb{E}[|e_h|^2_{L^2(0,T;\, \mathbb{R}^N)}] = \int_0^T \mathbb{E}[|e_h(t)|^2] \, \mathrm{d}t \leq h\mathrm{Var}[\mathcal{A}] C_{[A,x_0,T,U]}. \quad (108)$$

Because $\tilde{x}(\boldsymbol{\omega}, t) = x(\boldsymbol{\omega}, t) - x_d(t)$, (45) shows that

$$|\tilde{x}(\boldsymbol{\omega})|^2_{L^2(0,T;\, \mathbb{R}^N)} \leq (\sqrt{T}(|x_0| + |Bu_h(\boldsymbol{\omega})|_{L^1(0,T;\, \mathbb{R}^N)}) + |x_d|_{L^2(0,T;\, \mathbb{R}^N)})^2. \quad (109)$$

Because $|Bu_h(\boldsymbol{\omega})|_{L^1(0,T;\, \mathbb{R}^N)} \leq \sqrt{T}|Bu_h(\boldsymbol{\omega})|_{L^2(0,T;\, \mathbb{R}^N)} \leq \sqrt{T}U$, we see from (109) that $\mathbb{E}[|\tilde{x}|^2_{L^2(0,T;\, \mathbb{R}^N)}] \leq C_{[x_0,x_d,T,U]}$. The result now follows by inserting this estimate and (108) into (107). □

We are now ready to proof the main result of this section which can be considered as a no-gap condition for the RBM optimal control problem.

Theorem 3 *Let $u^*(t)$ be the (deterministic) control that minimizes the cost functional $J(u)$ in (2) and let $u_h^*(\boldsymbol{\omega}, t)$ be the control that minimizes the cost functional $J_h(\boldsymbol{\omega}, u)$ in (14). Then*

$$\mathbb{E}[|J_h(u_h^*) - J(u^*)|] \leq C_{[A,B,x_0,Q,R,x_d,T]} \left(\sqrt{h\mathrm{Var}[\mathcal{A}]} + h\mathrm{Var}[\mathcal{A}] \right). \quad (110)$$

Proof We have that

$$J(u^*) \leq J(u_h^*(\boldsymbol{\omega})) = J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega})) + \delta(\boldsymbol{\omega})$$
$$\leq J_h(\boldsymbol{\omega}, u^*) + \delta(\boldsymbol{\omega}) = J(u^*) + \delta(\boldsymbol{\omega}) + \varepsilon(\boldsymbol{\omega}), \quad (111)$$

where $\delta(\boldsymbol{\omega}) = J(u_h^*(\boldsymbol{\omega})) - J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega}))$ and $\varepsilon(\boldsymbol{\omega}) = J_h(\boldsymbol{\omega}, u^*) - J(u^*)$. Note that the first inequality follows because u^* is the minimizer of J and the second inequality because $u_h^*(\boldsymbol{\omega})$ is the minimizer of $J_h(\boldsymbol{\omega}, \cdot)$. Subtracting $J(u^*) + \delta(\boldsymbol{\omega})$ from the first, third, and fifth expressions in (111), shows that

$$-\delta(\boldsymbol{\omega}) \leq J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega})) - J(u^*) \leq \varepsilon(\boldsymbol{\omega}). \quad (112)$$

Taking the absolute value, we find

$$|J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega})) - J(u^*)| \leq \max\{|\delta(\boldsymbol{\omega})|, |\varepsilon(\boldsymbol{\omega})|\} \leq |\delta(\boldsymbol{\omega})| + |\varepsilon(\boldsymbol{\omega})|. \quad (113)$$

Therefore also

$$\mathbb{E}[|J_h(u_h^*) - J(u^*)|] \leq \mathbb{E}[|\delta|] + \mathbb{E}[|\varepsilon|]. \quad (114)$$

Lemma 1 can now be used to find bounds for $\mathbb{E}[|\delta|] = \mathbb{E}[|J_h(u_h^*) - J(u_h^*)|]$ and $\mathbb{E}[|\varepsilon|] = \mathbb{E}[|J_h(u^*) - J(u^*)|]$.

For the bound on $\mathbb{E}[|\delta|]$, we use that (50) shows that there exists a constant such that $|Bu_h^*(\boldsymbol{\omega})|_{L^2(0,T;\,\mathbb{R}^N)} \leq C_{[B,x_0,Q,R,x_d,T]}$ so that (98) is satisfied with a constant U that does not depend on the used temporal grid $t_0, t_1, \ldots, t_K$. Lemma 1 thus implies that

$$\mathbb{E}[|\delta|] \leq C_{[A,B,x_0,Q,R,x_d,T]} \left(\sqrt{h\mathrm{Var}[\mathcal{A}]} + h\mathrm{Var}[\mathcal{A}] \right). \tag{115}$$

For the bound on $\mathbb{E}[|\varepsilon|]$, we can simply take $U = |Bu^*(t)|_{L^2(0,T;\,\mathbb{R}^N)}$, which is a constant that only depends on the parameters A, B, x_0, Q, R, x_d, T that define the deterministic problem (1)–(2). Lemma 1 thus also shows that

$$\mathbb{E}[|\varepsilon|] \leq C_{[A,B,x_0,Q,R,x_d,T]} \left(\sqrt{h\mathrm{Var}[\mathcal{A}]} + h\mathrm{Var}[\mathcal{A}] \right). \tag{116}$$

Inserting (115) and (116) into (114) we find (110). □

3.5 Convergence in the controls

In the last stage of our analysis of the RBM-optimal control problem, we bound the expected difference between the optimal control u_h^* that minimizes J_h in (14) and the optimal control u^* for the original problem. The proof is based on the strong convexity of the functional J_h in (14).

To prove the main result, we need the following lemma which bounds the difference between the Gâteaux derivative of J_h and the Gâteaux derivative of J in expectation.

Lemma 2 *For any deterministic control $u \in L^2(0,T;\ \mathbb{R}^q)$ and any stochastic perturbation $v_h : \Omega^K \to L^2(0,T;\ \mathbb{R}^q)$,*

$$\mathbb{E}[|\delta J_h(u;\ v_h) - \delta J(u;\ v_h)|] \leq C_{[A,B,x_0,Q,x_d,T,u]} \sqrt{h\mathrm{Var}[\mathcal{A}]} \sqrt{\mathbb{E}[|v_h|^2_{L^2(0,T;\,\mathbb{R}^q)}]}. \tag{117}$$

Proof Let $x(t)$ and $x_h(\boldsymbol{\omega}, t)$ be the solutions of (1) and (13), respectively. Furthermore, denote

$$y(\boldsymbol{\omega}, t) = \int_0^t e^{A(t-s)} B v_h(\boldsymbol{\omega}, s)\ \mathrm{d}s, \qquad y_h(\boldsymbol{\omega}, t) = \int_0^t S_h(\boldsymbol{\omega}, t, s) B v_h(\boldsymbol{\omega}, s)\ \mathrm{d}s. \tag{118}$$

Directly from the definition of the Gâteaux derivative, we find that

$$\delta J(u, v_h(\boldsymbol{\omega})) = \int_0^T \left(\tilde{x}(t)^\top Q y(\boldsymbol{\omega}, t) + u(t)^\top R v_h(\boldsymbol{\omega}, t) \right)\ \mathrm{d}t, \tag{119}$$

$$\delta J_h(\boldsymbol{\omega}, u, v_h(\boldsymbol{\omega})) = \int_0^T \left(\tilde{x}_h(\boldsymbol{\omega}, t)^\top Q y_h(\boldsymbol{\omega}, t) + u(t)^\top R v_h(\boldsymbol{\omega}, t) \right)\ \mathrm{d}t, \tag{120}$$

where we write $\tilde{x}(t) = x(t) - x_d(t)$ and $\tilde{x}_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x_d(t)$.

Subtracting (119) from (120), we find

$$\delta J_h(\boldsymbol{\omega}, u, v_h(\boldsymbol{\omega})) - \delta J(u, v_h(\boldsymbol{\omega}))$$

$$= \int_0^T \left(\tilde{x}_h(\boldsymbol{\omega}, t)^\top Q y_h(\boldsymbol{\omega}, t) - \tilde{x}(t)^\top Q y(\boldsymbol{\omega}, t) \right) \, \mathrm{d}t$$
$$= \int_0^T \left(\tilde{x}_h(\boldsymbol{\omega}, t)^\top Q (y_h(\boldsymbol{\omega}, t) - y(\boldsymbol{\omega}, t)) + (\tilde{x}_h(\boldsymbol{\omega}, t) - \tilde{x}(t))^\top Q y(\boldsymbol{\omega}, t) \right) \, \mathrm{d}t$$
$$= \int_0^T \left(\tilde{x}_h(\boldsymbol{\omega}, t)^\top Q f_h(\boldsymbol{\omega}, t) + e_h(\boldsymbol{\omega}, t)^\top Q y(\boldsymbol{\omega}, t) \right) \, \mathrm{d}t, \tag{121}$$

where $e_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x(t) = \tilde{x}_h(\boldsymbol{\omega}, t) - \tilde{x}(t)$ and $f_h(\boldsymbol{\omega}, t) = y_h(\boldsymbol{\omega}, t) - y(\boldsymbol{\omega}, t)$. Taking the absolute value, we find

$$|\delta J_h(\boldsymbol{\omega}, u, v_h(\boldsymbol{\omega})) - \delta J(u, v_h(\boldsymbol{\omega}))|$$
$$\leq \|Q\| \int_0^T (|\tilde{x}_h(\boldsymbol{\omega}, t)||f_h(\boldsymbol{\omega}, t)| + |e_h(\boldsymbol{\omega}, t)||y(\boldsymbol{\omega}, t)|) \, \mathrm{d}t. \tag{122}$$

Using (44), we find the following bound for $\tilde{x}_h(\boldsymbol{\omega}, t) = x_h(\boldsymbol{\omega}, t) - x_d(t)$

$$|\tilde{x}_h(\boldsymbol{\omega}, t)| \leq |x_h(\boldsymbol{\omega}, t)| + |x_d(t)| \leq |x_0| + |Bu|_{L^1(0,T;\,\mathbb{R}^N)} + |x_d(t)|. \tag{123}$$

We thus have $|\tilde{x}_h(\boldsymbol{\omega}, t)| \leq C_{[B,x_0,x_d,T,u]}$ for all $\boldsymbol{\omega} \in \Omega^K$.

Taking the expectation in (122) using this result shows that

$$\mathbb{E}[|\delta J_h(u, v_h) - \delta J(u, v_h)|]$$
$$\leq \|Q\| \int_0^T \left(C_{[B,x_0,x_d,T,u]} \mathbb{E}[|f_h(t)|] - \sqrt{\mathbb{E}[|e_h(t)|^2]} \sqrt{\mathbb{E}[|y(t)|^2]} \right) \, \mathrm{d}t, \tag{124}$$

where the second term on the RHS follows from the Cauchy-Schwartz inequality.

Again using the notation $E_h(\boldsymbol{\omega}, t, s) := S_h(\boldsymbol{\omega}, t, s) - e^{A(t-s)}$, (118) shows that

$$f_h(\boldsymbol{\omega}, t) = y_h(\boldsymbol{\omega}, t) - y(\boldsymbol{\omega}, t) = \int_0^t E_h(\boldsymbol{\omega}, t, s) B v_h(\boldsymbol{\omega}, s) \, \mathrm{d}s. \tag{125}$$

Therefore,

$$\mathbb{E}[|f_h(t)|] \leq \int_0^t \mathbb{E}[\|E_h(t, s)\||Bv_h(s)|] \, \mathrm{d}s$$
$$\leq \int_0^t \sqrt{\mathbb{E}[\|E_h(t, s)\|^2]} \sqrt{\mathbb{E}[|Bv_h(s)|^2]} \, \mathrm{d}s$$
$$\leq C_{[A,T]} \sqrt{h \mathrm{Var}[\mathcal{A}]} \int_0^t \sqrt{\mathbb{E}[|Bv_h(s)|^2]} \, \mathrm{d}s$$
$$\leq C_{[A,T]} \sqrt{h \mathrm{Var}[\mathcal{A}]} \sqrt{t} \sqrt{\int_0^t \mathbb{E}[|Bv_h(s)|^2] \, \mathrm{d}s}$$
$$\leq C_{[A,T]} \sqrt{h \mathrm{Var}[\mathcal{A}]} \sqrt{\mathbb{E}[|Bv_h|^2_{L^2(0,T;\,\mathbb{R}^N)}]}, \tag{126}$$

where the second inequality follows from the Cauchy-Schwartz inequality in the probability space, the third inequality from Corollary 1, and the third inequality from the Cauchy-Schwartz inequality in $L^2(0, t)$.

Because the control $u(t)$ is deterministic, Theorem 1 shows that

$$\mathbb{E}[|e_h(t)|^2] \leq h \mathrm{Var}[\mathcal{A}] C_{[A,B,x_0,T,u]}. \tag{127}$$

Finally, note

$$|y(\boldsymbol{\omega}, t)|^2 = \left(\int_0^t \|e^{A(t-s)}\||Bv_h(\boldsymbol{\omega}, s)| \, \mathrm{d}s \right)^2$$

$$\leq \int_0^t \|e^{A(t-s)}\|^2 \, \mathrm{d}s \int_0^t |Bv_h(\boldsymbol{\omega}, s)r|^2 \, \mathrm{d}s \leq t|Bv_h(\boldsymbol{\omega})|^2_{L^2(0,T;\, \mathbb{R}^N)}. \quad (128)$$

Therefore, also

$$\mathbb{E}[|y(t)|^2] \leq C_{[B,T]}\mathbb{E}[|v_h|^2_{L^2(0,T;\, \mathbb{R}^N)}]. \quad (129)$$

Inserting (126), (127), and (129) into (124) completes the proof. □

We are now ready to prove the convergence result for the optimal controls.

Theorem 4 *Suppose that the functional $J_h(\boldsymbol{\omega}, \cdot)$ in (14) is α-convex for all $\boldsymbol{\omega} \in \Omega^K$. Let $u_h^*(\boldsymbol{\omega}, t)$ be the minimizer of $J_h(\boldsymbol{\omega}, \cdot)$ in (14) and $u^*(t)$ be the minimizer of J in (2), then*

$$\alpha^2\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}] \leq C_{[A,B,x_0,Q,R,x_d,T]}h\mathrm{Var}[\mathcal{A}]. \quad (130)$$

Proof We apply (53) with $J(\cdot) = J_h(\boldsymbol{\omega}, \cdot)$, $v = u_h^*(\boldsymbol{\omega})$, and $u = u^*$

$$J_h(\omega, u_h^*(\boldsymbol{\omega})) \geq J_h(\boldsymbol{\omega}, u^*) + \delta J_h(\boldsymbol{\omega}, u^*;\, u_h^*(\boldsymbol{\omega}) - u^*) + \tfrac{\alpha}{2}|u_h^*(\boldsymbol{\omega}) - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}. \quad (131)$$

Because $u_h^*(\boldsymbol{\omega})$ is the minimizer of $J_h(\boldsymbol{\omega}, \cdot)$, $J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega})) \leq J_h(\boldsymbol{\omega}, u^*)$ and

$$0 \geq \delta J_h(\boldsymbol{\omega}, u^*;\, u_h^*(\boldsymbol{\omega}) - u^*) + \tfrac{\alpha}{2}|u_h^*(\boldsymbol{\omega}) - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}. \quad (132)$$

Bringing δJ_h to the other side, taking the absolute value and then the expectation, yields

$$\frac{\alpha}{2}\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}] \leq \mathbb{E}[|\delta J_h(u^*;\, u_h^* - u^*)|]. \quad (133)$$

Since u^* is the minimizer of J, $\delta J(u^*, v) = 0$ for all perturbation $v \in L^2(0, T;\, \mathbb{R}^q)$. In particular, we have that $\delta J(u^*, u_h^*(\boldsymbol{\omega}) - u^*) = 0$ for all $\boldsymbol{\omega} \in \Omega^K$ so that also

$$\frac{\alpha}{2}\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}] \leq \mathbb{E}[|\delta J_h(u^*;\, u_h^* - u^*) - \delta J(u^*;\, u_h^* - u^*)|]. \quad (134)$$

We now apply Lemma 2 to the RHS with $u = u^*$ and $v_h(\boldsymbol{\omega}) = u_h^*(\boldsymbol{\omega}) - u^*$, which shows that

$$\frac{\alpha}{2}\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}] \leq C_{[B,x_0,Q,x_d,T,u^*]}\sqrt{h\mathrm{Var}[\mathcal{A}]}\sqrt{\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}]}. \quad (135)$$

Next, we divide (135) by $\frac{1}{2}\sqrt{\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}]}$ to find

$$\alpha\sqrt{\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}]} \leq C_{[A,B,x_0,Q,x_d,T,u^*]}\sqrt{h\mathrm{Var}[\mathcal{A}]}. \quad (136)$$

Squaring both sides we arrive at

$$\alpha^2\mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\, \mathbb{R}^q)}] \leq C_{[A,B,x_0,Q,x_d,T,u^*]}h\mathrm{Var}[\mathcal{A}]. \quad (137)$$

The result follows because the optimal control $u^*(t)$ only depends on the parameters A, B, x_0, Q, R, x_d, and T that define the original problem (1)–(2). □

We now point out two corollaries of Theorem 4 that are important when we use the control $u_h^*(\boldsymbol{\omega}, t)$ (optimized for the RBM-dynamics) to control the original dynamics. For the first corollary, we introduce the notation

$$x_h^*(\boldsymbol{\omega}, t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu_h^*(\boldsymbol{\omega}, s) \, \mathrm{d}s, \quad (138)$$

$$x^*(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu^*(s)\ \mathrm{d}s, \tag{139}$$

i.e., $x_h^*(\boldsymbol{\omega}, t)$ is the solution of the original dynamics (1) resulting from the control $u_h^*(\boldsymbol{\omega}, t)$ optimized for the RBM-dynamics and $x^*(t)$ is the solution of the original dynamics (1) resulting from the optimal control $u^*(t)$.

Corollary 2 *Suppose that the functional $J_h(\boldsymbol{\omega}, \cdot)$ in (14) is α-convex for all $\boldsymbol{\omega} \in \Omega^K$ and let $x_h^*(\boldsymbol{\omega}, t)$ and $x^*(t)$ be as in (138) and (139), respectively. Then*

$$\alpha^2 \mathbb{E}[|x_h^*(t) - x^*(t)|^2] \leq C_{[A,B,x_0,Q,R,x_d,T]} h \mathrm{Var}[\mathcal{A}], \tag{140}$$

for all $t \in [0, T]$.

Proof Note that

$$x_h^*(\boldsymbol{\omega}, t) - x^*(t) = \int_0^t e^{A(t-s)}B(u_h^*(\boldsymbol{\omega}, s) - u^*(s))\ \mathrm{d}s. \tag{141}$$

Therefore also

$$\begin{aligned} |x_h^*(\boldsymbol{\omega}, t) - x^*(t)| &\leq \int_0^t \|e^{A(t-s)}\|\|B\||u_h^*(\boldsymbol{\omega}, s) - u^*(s)|\ \mathrm{d}s \\ &\leq \|B\||u_h^*(\boldsymbol{\omega}) - u^*|_{L^1(0,T;\,\mathbb{R}^q)} \leq \|B\|\sqrt{T}\sqrt{|u_h^*(\boldsymbol{\omega}) - u^*|_{L^2(0,T;\,\mathbb{R}^q)}}, \end{aligned} \tag{142}$$

where the second inequality uses that $\|e^{At}\| \leq 1$ in view of Assumption 1. The result now follows after squaring this inequality, taking the expectation, and using (130). □

Corollary 3 *Suppose that the cost functional $J_h(\boldsymbol{\omega}, \cdot)$ is α-convex for all $\boldsymbol{\omega} \in \Omega^K$. Let $u^*(t)$ be the (deterministic) control that minimizes the cost functional $J(u)$ in (2) and let $u_h^*(\boldsymbol{\omega}, t)$ be the control that minimizes the cost functional $J_h(\boldsymbol{\omega}, u)$ in (14). Then*

$$\alpha^2 \mathbb{E}[|J(u_h^*) - J(u^*)|] \leq C_{[A,B,x_0,Q,R,x_d,T]} h \mathrm{Var}[\mathcal{A}]. \tag{143}$$

Proof Denote $v_h(\boldsymbol{\omega}, t) := u_h^*(\boldsymbol{\omega}, t) - u^*(t)$ and $y(\boldsymbol{\omega}, t) := \int_0^t e^{A(t-s)}Bv_h(\boldsymbol{\omega}, s)\ \mathrm{d}s$. Because the considered functional is quadratic,

$$\begin{aligned} J(u_h^*(\boldsymbol{\omega})) - J(u^*) &= J(u^* + v_h(\boldsymbol{\omega})) - J(u^*) \\ &= \delta J(u^*, v_h(\boldsymbol{\omega})) + \delta^2 J(v_h(\boldsymbol{\omega}), v_h(\boldsymbol{\omega})), \end{aligned} \tag{144}$$

where the Hessian $\delta^2 J(v_h(\boldsymbol{\omega}), v_h(\boldsymbol{\omega}))$ is given by

$$\delta^2 J(v_h(\boldsymbol{\omega}), v_h(\boldsymbol{\omega})) = \frac{1}{2}\int_0^T \left(y(\boldsymbol{\omega}, t)^\top Q y(\boldsymbol{\omega}, t) + v_h(\boldsymbol{\omega}, t)^\top R v_h(\boldsymbol{\omega}, t)\right)\ \mathrm{d}t. \tag{145}$$

Because u^* is the minimizer of $J(\cdot)$, $\delta J(u^*, v) = 0$ for all $v \in L^2(0, T;\, \mathbb{R}^q)$. The first term on the RHS of (144) thus vanishes. Also observe that

$$\delta^2 J(v_h(\boldsymbol{\omega}), v_h(\boldsymbol{\omega})) \leq \tfrac{1}{2}\|Q\||y(\boldsymbol{\omega})|^2_{L^2(0,T;\,\mathbb{R}^N)} + \tfrac{1}{2}\|R\||v_h(\boldsymbol{\omega})|^2_{L^2(0,T;\,\mathbb{R}^q)}. \tag{146}$$

A similar estimate as (128) shows that $|y(\boldsymbol{\omega})|^2_{L^2(0,T;\,\mathbb{R}^N)} \leq C_{[B,T]}|v_h(\boldsymbol{\omega})|^2_{L^2(0,T;\,\mathbb{R}^q)}$. Combining these results in (144), we conclude

$$\begin{aligned}|J(u_h^*(\boldsymbol{\omega})) - J(u^*)| &\leq J(u_h^*(\boldsymbol{\omega})) - J(u^*)\\ &\leq \delta^2 J(v_h(\boldsymbol{\omega}), v_h(\boldsymbol{\omega})) \leq C_{[B,Q,R,T]}|v_h(\boldsymbol{\omega})|^2_{L^2(0,T;\,\mathbb{R}^q)}. \end{aligned} \tag{147}$$

The result now follows after taking the expectation and using the result from Theorem 4 to bound $\mathbb{E}[|v_h|_{L^2(0,T;\,\mathbb{R}^q)}{}^2] = \mathbb{E}[|u_h^* - u^*|^2_{L^2(0,T;\,\mathbb{R}^q)}]$. □

4 Numerical results

In this section, we apply our proposed method to three medium to large scale linear dynamical systems that are obtained after spatial discretization of a linear PDE.

4.1 A discretized 1D heat equation

We consider a controlled heat equation on the 1-D spatial domain $[-L, L]$,

$$y_t(t,\xi) = y_{\xi\xi}(t,\xi) + \chi_{[-L/3,0]}(\xi)u(t), \qquad \xi \in [-L, L], \tag{148}$$

$$y_\xi(t,-L) = y_\xi(t,L) = 0, \qquad y(0,\xi) = e^{-\xi^2} + \xi^2 e^{-L^2}, \tag{149}$$

where $\chi_{[-L/3,0]}(\xi)$ denotes the characteristic function for the interval $[-L/3, 0]$. We want to compute the optimal control $u^*(t)$ that minimizes

$$\mathcal{J}(u) = \frac{100}{2}\int_0^T \int_{-L}^0 y(t,\xi)^2 \,\mathrm{d}\xi \,\mathrm{d}t + \frac{1}{2}\int_0^T u(t)^2 \,\mathrm{d}t. \tag{150}$$

The spatial discretization of the dynamics (148)–(149) is made by finite differences and the cost functional in (150) is discretized by the trapezoid rule. We choose a uniform spatial grid with $N = 61$ grid points $\xi_i = (i-1)\Delta\xi - L$ ($i \in \{1, 2, \ldots, N\}$), where $\Delta\xi = 2L/(N-1)$ is the grid spacing, and obtain a system of the form (1).

The resulting A-matrix is of the form

$$A = \frac{1}{\Delta\xi^2}\begin{bmatrix} -2 & 2 & 0 & \cdots & 0 & 0 & 0\\ 1 & -2 & 1 & & 0 & 0 & 0\\ 0 & 1 & -2 & & 0 & 0 & 0\\ \vdots & & & \ddots & & & \vdots\\ 0 & 0 & 0 & & -2 & 1 & 0\\ 0 & 0 & 0 & & 1 & -2 & 1\\ 0 & 0 & 0 & \cdots & 0 & 2 & -2 \end{bmatrix}. \tag{151}$$

Observe that A can be written as

$$A = \sum_{i=1}^{n} \tilde{A}_i, \tag{152}$$

where the $n := N-1 = 60$ matrices $\tilde{A}_i \in \mathbb{R}^{N\times N}$ are zero except for the entries

$$\begin{aligned}
\begin{bmatrix} [\tilde{A}_1]_{11} & [\tilde{A}_1]_{12} \\ [\tilde{A}_1]_{21} & [\tilde{A}_1]_{22} \end{bmatrix} &= \begin{bmatrix} -2 & 2 \\ 1 & -1 \end{bmatrix}, \\
\begin{bmatrix} [\tilde{A}_i]_{ii} & [\tilde{A}_i]_{i,i+1} \\ [\tilde{A}_i]_{i+1,i} & [\tilde{A}_i]_{i+1,i+1} \end{bmatrix} &= \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad 2 \le i \le n-1, \\
\begin{bmatrix} [\tilde{A}_n]_{nn} & [\tilde{A}_n]_{n,n+1} \\ [\tilde{A}_n]_{n+1,n} & [\tilde{A}_n]_{n+1,n+1} \end{bmatrix} &= \begin{bmatrix} -1 & 1 \\ 2 & -2 \end{bmatrix}.
\end{aligned}$$

One can easily verify that the matrices $\tilde{A}_i$ are dissipative. We now define the M submatrices A_m (for $M = 1, 2, 3, 4$) as

$$A_m = \sum_{i=i_{m-1}+1}^{i_m} \tilde{A}_i, \tag{153}$$

where $i_m = nm/M$. Because of (152), it is easy to see that the submatrices A_m satisfy (5). Because the submatrices $\tilde{A}_i$ are dissipative, the submatrices A_m in (153) are dissipative and Assumption 1 is satisfied.

Example 4 For $M = 2$ and $N = 61$, we obtain the splitting of the A-matrix in (151) as $A = A_1 + A_2$, with

$$A_1 = \begin{bmatrix} A_{11} & 0_{31\times 30} \\ 0_{30\times 31} & 0_{30\times 30} \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 0_{30\times 30} & 0_{30\times 31} \\ 0_{31\times 30} & A_{22} \end{bmatrix}, \tag{154}$$

where A_{11} and A_{22} are the 31×31-matrices

$$A_{11} = \frac{1}{\Delta\xi^2} \begin{bmatrix} -2 & 2 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & & 0 & 0 & 0 \\ 0 & 1 & -2 & & 0 & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & & -2 & 1 & 0 \\ 0 & 0 & 0 & & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{bmatrix}, \tag{155}$$

$$A_{22} = \frac{1}{\Delta\xi^2} \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & & 0 & 0 & 0 \\ 0 & 1 & -2 & & 0 & 0 & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & 0 & 0 & & -2 & 1 & 0 \\ 0 & 0 & 0 & & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 2 & -2 \end{bmatrix}. \tag{156}$$

We will present numerical results for four cases:

Case i We decompose A into $M = 2$ submatrices and assign a probability $\frac{1}{2}$ to the subsets $\{1\}$ and $\{2\}$ and a probability 0 to the subsets $\emptyset$ and $\{1, 2\}$.

Case ii We decompose A into $M = 3$ submatrices and assign a probability $\frac{1}{3}$ to the subsets $\{1\}$, $\{2\}$, and $\{3\}$ and a probability 0 to the other subsets of $\{1, 2, 3\}$.
Case iii We decompose A into $M = 4$ submatrices and assign a probability $\frac{1}{4}$ to the subsets $\{1\}$, $\{2\}$, $\{3\}$, and $\{4\}$ and a probability 0 to the other subsets of $\{1, 2, 3, 4\}$.
Case iv We decompose A into $M = 4$ submatrices and assign a probability $\frac{1}{2}$ to the subsets $\{1, 3\}$ and $\{2, 4\}$ and a probability 0 to the other subsets of $\{1, 2, 3, 4\}$.

In all 4 cases, we fix $N = 61$, $L = \frac{3}{2}$, and $T = \frac{1}{2}$.

We use a uniform grid $0 = t_0 < t_1 < \ldots < t_{K-1} < t_K = T$ with a uniform grid spacing h. We will present results for $h = 2^{-5}$, 2^{-7}, 2^{-9}, 2^{-11}, 2^{-13}, and 2^{-15}. For each of the $K = T/h$ time intervals $[t_{k-1}, t_k)$, we select an index ω_k according to the probabilities specified in Cases i–iv above. The state $x_h(\boldsymbol{\omega}, t)$ that satisfies (13) is computed using a single Crank-Nicholson step in each time interval $[t_{k-1}, t_k)$. We use precomputed LU-factorizations of the matrices $I - \frac{h}{2} \sum_{m \in S_\omega} \frac{A_m}{\pi_m}$ (for subsets S_ω with a nonzero probability p_ω) that need to be inverted frequently.

The optimal control $u_h^*(\boldsymbol{\omega}, t)$ that minimizes $J_h(\boldsymbol{\omega}, u)$ in (14) is computed with a gradient-descent algorithm. The gradient is computed using the adjoint state $\varphi_h(\boldsymbol{\omega}, t)$, see Remark 3. The time discretization for the adjoint state equation (15) is done using the scheme proposed in [1] that leads to discretely consistent gradients. The iterates u^k are computed as $u^{k+1} = u^k - \beta \nabla J_h(\boldsymbol{\omega}, u^k)$. The step size β is chosen such that $J_h(\boldsymbol{\omega}, u^k - \beta \nabla J_h(\boldsymbol{\omega}, u^k))$ is minimal. The algorithm is terminated when the relative change in $J_h(\boldsymbol{\omega}, u)$ is below 10^{-6}.

The results for the four considered cases are displayed in Figure 2. Because the obtained results depend on the randomly selected indices stored in $\boldsymbol{\omega}$, each marker in the subfigures in Figure 2 represents the average error or duration over 25 random realizations of $\boldsymbol{\omega}$. The errorbars represent the 2σ-confidence interval estimated from these 25 realizations. The errors are computed w.r.t. the solutions $x(t)$ and $u^*(t)$ that are computed on the same time grid as the corresponding solutions $x_h(\boldsymbol{\omega}, t)$ and $u_h^*(\boldsymbol{\omega}, t)$. The displayed errors therefore do not reflect the errors due to the temporal (or spatial) discretization but capture only the error introduced by the proposed randomized splitting method.

Because the matrices A and A_m represent approximations of unbounded operators, the variance $\mathrm{Var}[\mathcal{A}]$ defined in (17) will grow unbounded when the mesh is refined. This is also reflected by the large values of $\mathrm{Var}[\mathcal{A}]$ given in Table 1. It is therefore more natural to consider the variance $\mathrm{Var}_W[\mathcal{A}]$ in (19) weighted by a matrix of the form $W = (A - \lambda I)^{-1}$. The values of $\mathrm{Var}_W[\mathcal{A}]$ are indeed much smaller than the values of $\mathrm{Var}[\mathcal{A}]$ in Table 1. The results at the end of this subsection (in Figure 4) also indicate that the weighted variance $\mathrm{Var}_W[\mathcal{A}]$ reflects the behavior of the error better when the mesh is refined.

The error estimates in Theorems 1, 3, and 4 and in Corollary 3 are proportional to $h\mathrm{Var}[\mathcal{A}]$. We therefore plot the errors in Figures 2a–2d against

Table 1: Values of $\mathrm{Var}[\mathcal{A}]$ and $\mathrm{Var}_W[\mathcal{A}]$ for $W = (A - \lambda I)^{-1}$ with $\lambda = 0.1$

	Case i	Case ii	Case iii	Case iv
$\mathrm{Var}[\mathcal{A}]$	$4.16 \cdot 10^7$	$1.65 \cdot 10^8$	$3.68 \cdot 10^8$	$4.16 \cdot 10^7$
$\mathrm{Var}_W[\mathcal{A}]$	57.32	133.91	246.54	96.68

$\sqrt{h\mathrm{Var}_W[\mathcal{A}]}$ (with $W = (A - 0.1I)^{-1}$) and expect that the errors for the different cases will be (approximately) on one line.

Figure 2a shows the difference $|x_h(\boldsymbol{\omega}, t) - x(t)|$ between the solutions $x(t)$ and $x_h(\omega, t)$ of (1) and (13) with $u(t) = 0$. Recall that the markers in this figure indicate the average error observed over 25 realizations of $\boldsymbol{\omega}$, and are thus estimates for $\mathbb{E}[\max_{t\in[0,T]}|x_h(t) - x(t)|]$. Because $\mathbb{E}[|x_h(t) - x(t)|] \leq \sqrt{\mathbb{E}[|x_h(t) - x(t)|^2]}$, we expect (based on the bound in Theorem 1) that the errors in Figure 2a are proportional to $\sqrt{h\mathrm{Var}_W[\mathcal{A}]}$. This is indeed confirmed by Figure 2a.

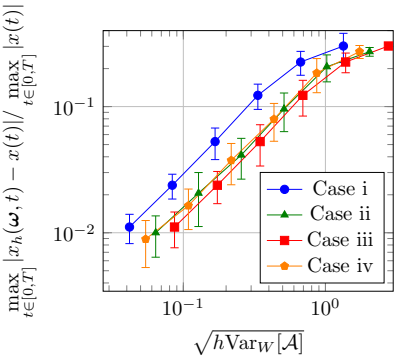
Figure 2b shows the difference $|u_h^* - u^*|_{L^2(0,T)}$ between the optimal controls $u^*(t)$ and $u_h^*(\boldsymbol{\omega}, t)$ that minimize (2) and (14), respectively. Based on the estimate in Theorem 4, we again expect that the observed errors are proportional to $\sqrt{h\mathrm{Var}_W[\mathcal{A}]}$. This is indeed the case and the proportionality constants for the different cases are again (approximately) equal, which is also expected based on the error estimate in Theorem 4.

The convergence in the optimal controls in Figure 2b is also illustrated in Figure 3. This figure shows the optimal controls $u_h^*(\boldsymbol{\omega}, t)$ obtained for 25 randomly selected realizations of $\boldsymbol{\omega} \in \Omega^K$ (light red) for the six considered grid spacings h of the temporal grid. The figure also shows the average of the 25 optimal controls $u_h^*(\boldsymbol{\omega}, t)$ (dark red) and the optimal control $u^*(t)$ for the original system (black). Figure 3 indeed shows that the optimal controls $u_h^*\boldsymbol{\omega}, t)$ get closer to the optimal control $u^*(t)$ when the spacing of the temporal grid h is reduced. Especially in Figures 3a and 3b, it is also clear that the average of the 25 optimal controls $u_h^*(\boldsymbol{\omega}, t)$ (dark red) is not equal to the optimal control $u^*(t)$ for the original system (black). This indicates that $\mathbb{E}[u_h^*] \neq u^*$, see also Remark 6. This means that u_h^* is a biased estimator for u^* and averaging several realizations of $u^*(\boldsymbol{\omega}, t)$ can only improve the approximation of $u^*(t)$ to a limited extend. Note, however, that

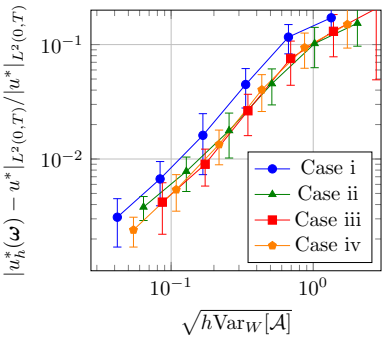
$$|\mathbb{E}[u_h^*] - u^*| = |\mathbb{E}[u_h^* - u^*]| \leq \mathbb{E}[|u_h^* - u^*|] \leq \sqrt{\mathbb{E}[|u_h^* - u^*|^2]}, \tag{157}$$

so that Theorem 4 shows that $\mathbb{E}[u_h^*] \to u^*$ at a rate of $\sqrt{h\mathrm{Var}[\mathcal{A}]}$. An analysis of the numerical results (that is not presented in Figure 2) also indicates that the average of the 25 realizations of $u_h^*(\boldsymbol{\omega}, t)$ converges to $u^*(t)$ at this rate.

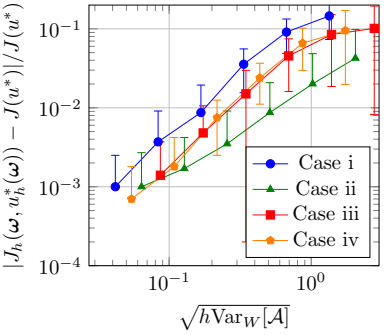
Figures 2c and 2d illustrate the convergence of $J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega}))$ and $J(u_h^*(\boldsymbol{\omega}))$ to $J(u^*)$. Figure 2c illustrates the error estimate in Theorem 3 and shows that the optimality gap $|J_h(\omega, u_h^*(\boldsymbol{\omega})) - J(u^*)|$ is indeed proportional to $\sqrt{h\mathrm{Var}_W[\mathcal{A}]}$. The difference between the different cases is more visible than in Figures 2a and 2b. Figure 2d illustrates the error estimate in Corollary 3,



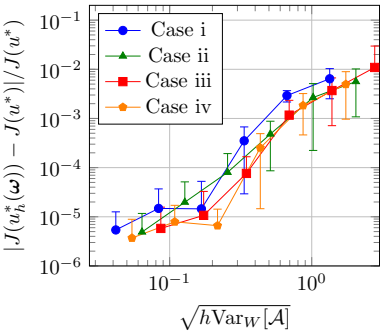
(a) Error in $x_h(t)$



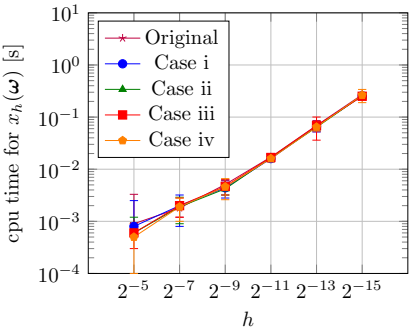
(b) Error in $u_h^*(t)$



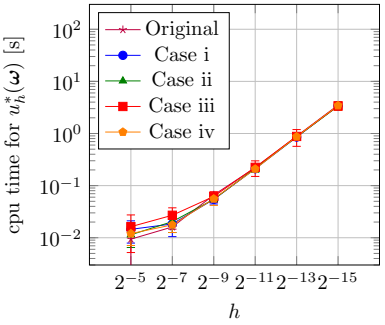
(c) Error in $J_h(u_h^*)$



(d) Error in $J(u_h^*)$



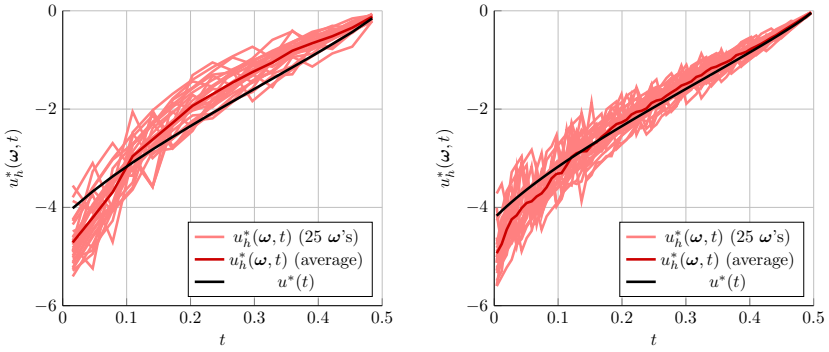
(e) Computational time for $x_h(t)$



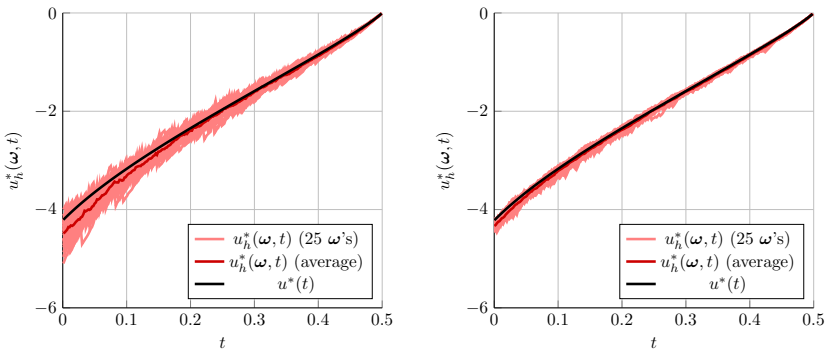
(f) Computational time for $u_h^*(t)$

Fig. 2: Simulation results for the discretized 1D heat equation

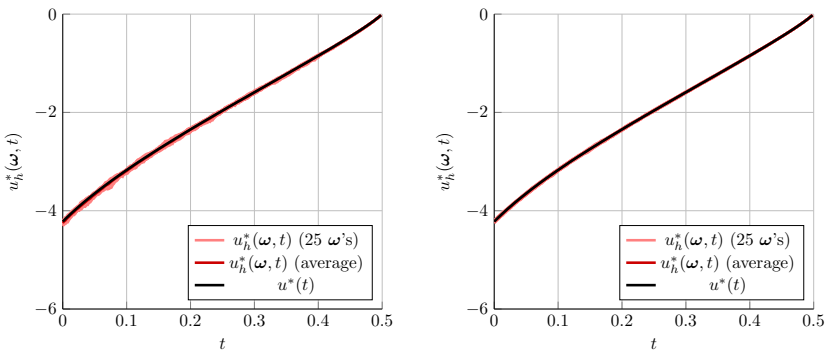
which shows that the suboptimality of the RBM-control $|J(u_h^*(\boldsymbol{\omega})) - J(u^*)|$ is proportional to $h\mathrm{Var}_W[\mathcal{A}]$. The convergence rate is now twice as high as in the previous cases and the relative error stabilizes around 10^{-5}, which seems



(a) $h = 2^{-5}$ (b) $h = 2^{-7}$



(c) $h = 2^{-9}$ (d) $h = 2^{-11}$



(e) $h = 2^{-13}$ (f) $h = 2^{-15}$

Fig. 3: The optimal controls computed for the 1D heat equation for different time steps h. The controls $u_h^*(\boldsymbol{\omega}, t)$ computed with the proposed randomized time-splitting method are shown for 25 realizations of $\boldsymbol{\omega}$ and compared to the optimal control $u^*(t)$ for the original system.

to be related to the tolerance of 10^{-6} used in the computation of the optimal controls.

Figures 2e and 2f show the computational times for (one realization of) $x_h(\boldsymbol{\omega}, t)$ and $u_h^*(\boldsymbol{\omega}, t)$ in Cases i–iv and the computational time for the original problem (labeled 'Original'). Note that the results have been generated on temporal grids with different grid spacings h and that the computational time generally increases when the more time steps are used, i.e. when h is smaller. The figures indicate that $x_h(\boldsymbol{\omega}, t)$ and $u_h^*(\boldsymbol{\omega}, t)$ are not computed faster than the solutions $x(t)$ and $u^*(t)$ of the original problem. The proposed method does thus not lead to any reduction in computational time in this example. It seems that we cannot observe any reduction in computational time for this example because the original A-matrix is quite small ($N = 61$) and sparse (A is tridiagonal). The examples in the following two subsections indicate that a reduction in computational cost is obtained when the state dimension N is significantly higher or when A has significantly more nonzero off-diagonal elements.

To conclude this example, we study the dependence of our results on the number of grid points N. This gives us some indication whether the RBM can also be applied to infinite dimensional problems. In particular, the results give us some indication whether the proposed randomized splitting also works for the underlying PDE problem (148)–(150). As we also noted in Remarks 5 and 7, the main concerns are related to operator norm of A, that appears in $\mathrm{Var}[\mathcal{A}]$ and in the estimate in Theorem 1, which grows unbounded when the mesh is refined. These concerns also motivated the introduction of the weighted variance $\mathrm{Var}_W[\mathcal{A}]$, see Remark 5.

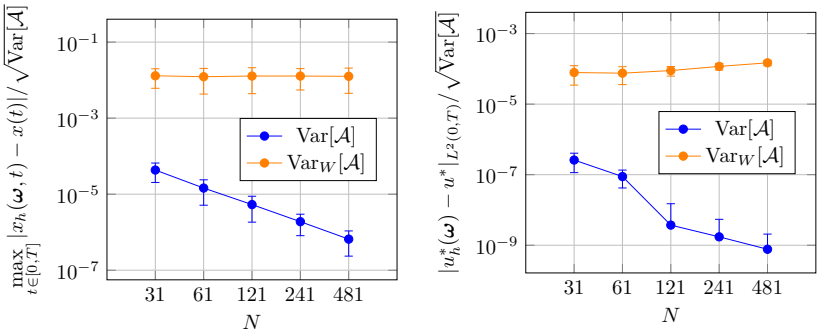
When the estimate in Theorem 1 indeed depends on $\|A\|$, the error $|x_h(\boldsymbol{\omega}, t) - x(t)|$ divided by $\mathrm{Var}[\mathcal{A}]$ should grow when N is increased. Figure 4a shows that this is not the case, but that this ratio actually decreases when N is increased. However, when we divided the errors by $\mathrm{Var}_W[\mathcal{A}]$, the result seems to be independent of the mesh size. Figure 4b shows that the same trend is observed for the errors in the optimal control.

The numerical results in Figure 4 match well with the result from Appendix B, where we prove an error estimate proportional to $\mathrm{Var}_W[\mathcal{A}]$ under the additional assumption that all matrices A_m commute. This result also extends to an infinite-dimensional setting when the domains the operators A_m coincide. However, in the setting considered here, the matrices A_m do not commute and are not approximations of operators with the same domains. Proving the convergence of the proposed randomized time splitting method for the underlying PDE problem (148)–(150) with the proposed randomized time splitting method is a challenging topic for future research.

4.2 A discretized 3D heat equation

We now consider a heat equation on the a 3-D spatial domain $V = [-L, L]^3$,

$$y_t(t, \boldsymbol{\xi}) = \Delta y(t, \boldsymbol{\xi}), \qquad \boldsymbol{\xi} \in [-L, L]^3, \tag{158}$$



(a) Error in $x_h(t)$ (b) Error in $u_h^*(t)$

Fig. 4: The errors in the forward dynamics $x_h(\omega, t)$ and the optimal control $u_h^*(\omega, t)$ divided by $\mathrm{Var}[\mathcal{A}]$ and $\mathrm{Var}_W[\mathcal{A}]$ (with $W = (A - 0.1I)^{-1}$) as a function of the number of nodes N. The results are presented for case i, so A is decomposed in $M = 2$ parts.

$$\nabla y(t, \boldsymbol{\xi}) \cdot \mathbf{n} = u(t), \qquad \boldsymbol{\xi} \in S_{\text{top}}, \tag{159}$$

$$\nabla y(t, \boldsymbol{\xi}) \cdot \mathbf{n} = 0, \qquad \boldsymbol{\xi} \in \partial V \backslash S_{\text{top}}, \tag{160}$$

$$y(0, \boldsymbol{\xi}) = e^{-|\boldsymbol{\xi}|^2/(8L^2)}, \tag{161}$$

where ∇ and Δ are the gradient and Laplacian operators w.r.t. $\boldsymbol{\xi}$, $\mathbf{n}$ is the outward pointing normal, and S_{top} denotes the top surface $S_{\text{top}} = \{(\xi_1, \xi_2, \xi_3) \in [-L, L]^3 \mid \xi_3 = L\}$. The control $u(t)$ can be considered as a uniform heat load on the top surface. We want to compute the control $u^*(t)$ that minimizes

$$J = 1000 \int_0^T \iint_{S_{\text{side}}} (y(t, \boldsymbol{\xi}))^2 \, \mathrm{d}\boldsymbol{\xi} \, \mathrm{d}t + \int_0^T (u(t))^2 \, \mathrm{d}t, \tag{162}$$

where $S_{\text{side}} = \{(\xi_1, \xi_2, \xi_3) \in [-L, L]^3 \mid \xi_1 = -L\}$. We fix $L = 0.75$ and $T = 2$.

The spatial discretization of (158)–(162) is made by finite differences using $16 \times 16 \times 16$ grid points the ξ_1-, ξ_2-, and ξ_3-directions. This leads to a model of the form (1)–(2) with $N = 16^3 = 4096$ states. The resulting A-matrix is again dissipative. We create the decomposition of A into submatrices A_m by observing that A is diagonally dominant. In particular, we have that

$$[A]_{ii} = -\sum_{\substack{j=1 \\ j \neq i}}^{N} [A]_{ij}, \tag{163}$$

where the off-diagonal elements $[A]_{ij}$ $(j \neq i)$ are positive and the diagonal elements $[A]_{ii}$ are negative. By associating a matrix $\tilde{A}_{ij} \in \mathbb{R}^{N \times N}$ to each pair

(i, j) with $j > i$, we obtain a decomposition of A as

$$A = \sum_{\substack{j=1 \\ j>i}}^{N} \tilde{A}_{ij}, \tag{164}$$

where the matrices $\tilde{A}_{ij}$ $(j > i)$ are zero except for the entries

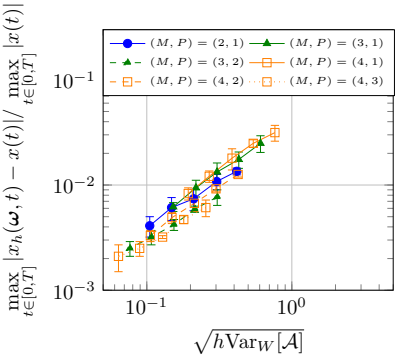
$$\begin{bmatrix} [\tilde{A}_{ij}]_{ii} & [\tilde{A}_{ij}]_{ij} \\ [\tilde{A}_{ij}]_{ji} & [\tilde{A}_{ij}]_{jj} \end{bmatrix} = [A]_{ij} \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \tag{165}$$

Because the off-diagonal elements $[A]_{ij} \geq 0$ $(j \neq i)$, it is easy to verify that all the matrices $\tilde{A}_{ij}$ are dissipative. Also note that the matrix A contains many zero off-diagonal elements, so that many of the matrices $\tilde{A}_{ij}$ are zero. There are only $3(16-1)16^2 = 11,520$ nonzero off-diagonal elements and thus only $11,520$ nonzero matrices $\tilde{A}_{ij}$. The $11,520$ nonzero matrices $\tilde{A}_{ij}$ are randomly divided into M groups of (approximately) equal size. The matrices A_m in (5) are formed by summing the matrices $\tilde{A}_{ij}$ in each group.

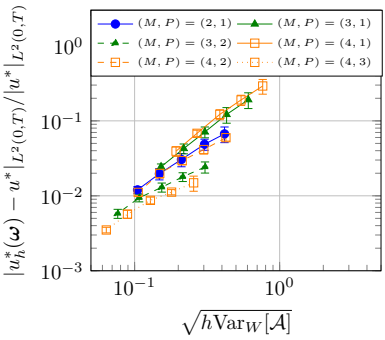
We again consider uniform time grids with a grid spacing h. In each time interval $[t_{k-1}, t_k)$, we randomly use P of the M submatrices simultaneously. In our formalism, we thus assign a probability $1/\binom{M}{P}$ to each of the $\binom{M}{P}$ subsets of $\{1, 2, \ldots, M\}$ of size P. The states $x_h(\boldsymbol{\omega}, t)$ and the optimal controls $u_h^*(\boldsymbol{\omega}, t)$ are computed in the same way as for the example in the previous subsection.

The obtained results are presented in Figure 5. The average errors (indicated by the markers) and the 2σ-confidence intervals (indicated by the error bars) are now estimated based on 10 realizations of $\boldsymbol{\omega}$. Figures 5a–5d again show the convergence rates expected based on our theoretical results, just as in Figures 2a–2d for the example in the previous subsection. We also observe that the errors are smaller when larger parts of A are used simultaneously, i.e., when P/M is larger.

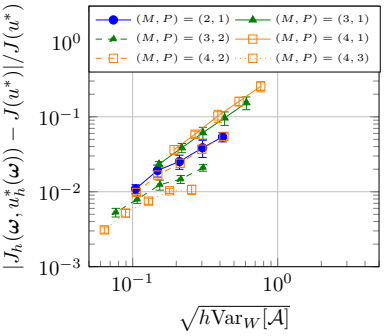
Figures 5e and 5f also show a computational advantage of the proposed method. Naturally, the computational advantage increases when the matrix $\mathcal{A}_h(\boldsymbol{\omega}, t)$ is more sparse, i.e., when P/M is smaller. This situation is significantly different from the 1D heat equation considered in the previous subsection. For that example, the proposed method did not lead to any computational advantage. Apart from the larger state dimension N in the 3D example, this difference seems to be related to the more 'dense interconnection structure' of the 3D problem (in which every node is typically connected to 6 neighboring nodes) compared to the 1D problem (in which every node is connected to two neighboring nodes). This idea will be explored further in the next subsection in which we consider a model with an even denser interconnection structure.



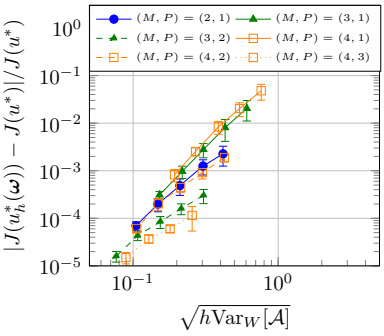
(a) Error in $x_h(t)$



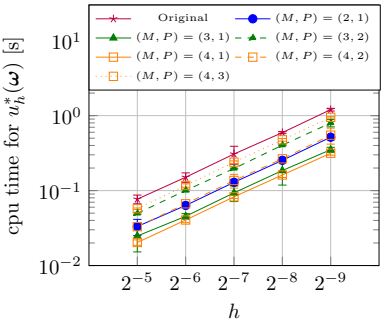
(b) Error in $u_h^*(t)$



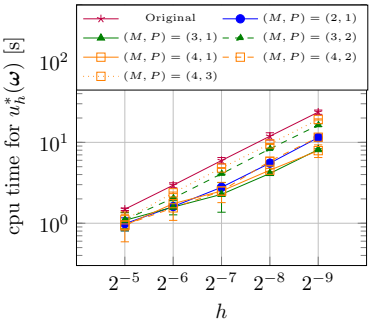
(c) Error in $J_h(u_h^*)$



(d) Error in $J(u_h^*)$



(e) Computational time for $x_h(t)$



(f) Computational time for $u_h^*(t)$

Fig. 5: Results for the discretized 3D heat equation

4.3 A FE discretization of the fractional Laplacian

We consider a controlled fractional heat equation on the a 1-D spatial domain $\xi \in [-L, L]$,

$$y_t(t,\xi) = -(-d_\xi^2)^s y(t,\xi) + \chi_{[-L/3,0]}(\xi)u_1(t) + \chi_{[L/3,2L/3]}(\xi)u_2(t), \tag{166}$$

$$y(t,-L) = y(t,L) = 0, \qquad y(0,\xi) = e^{-\beta^2\xi^2} - e^{-\beta^2 L^2}, \tag{167}$$

with the fractional power $s \in (0,1)$. We fix $s = 0.7$, $L = 5$, and $\beta = 0.4$. Note that the control $u(t) = [u_1(t), u_2(t)]^\top$ now has two components. Our aim is to compute the optimal control $u^*(t) = [u_1^*(t), u_2^*(t)]^\top$ that minimizes

$$\mathcal{J}(u) = \frac{100}{2}\int_0^T \int_{-L}^L y(t,\xi)^2 \,\mathrm{d}\xi \,\mathrm{d}t + \frac{1}{2}\int_0^T \left(u_1(t)^2 + u_2^2(t)\right) \,\mathrm{d}t. \tag{168}$$

A Finite Element (FE) discretization of (166)–(167) with $N+1$ linear elements of equal length takes the form

$$E\dot{x}(t) = Ax(t) + Bu(t), \qquad x(0) = x_0, \tag{169}$$

where the state $x(t)$ evolves in $\mathbb{R}^N$. Note that (169) now also contains the symmetric and positive definite mass matrix E and is thus not exactly of the form (1), but that the proposed method also applies to systems of this form. An explicit expression for the stiffness matrix A can be found in [5]. Because the fractional Laplacian is a nonlocal operator, all elements of A are nonzero. From the expressions for the coefficients of A in [5] we can verify that A is symmetric and diagonally dominant, i.e.

$$-[A]_{ii} > \sum_{\substack{j=1\\ j\neq i}}^N |[A]_{ij}|. \tag{170}$$

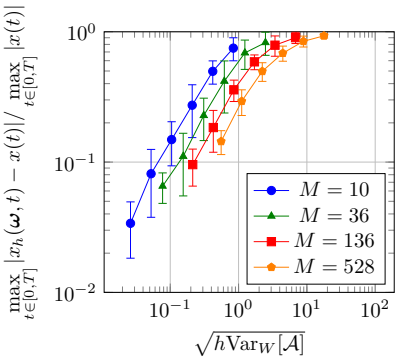
We can now write

$$A = \sum_{\substack{j=1\\ j\geq i}}^N \tilde{A}_{ij} = \sum_{\substack{j=1\\ j>i}}^N \tilde{A}_{ij} + \sum_{i=1}^N \tilde{A}_{ii}, \tag{171}$$

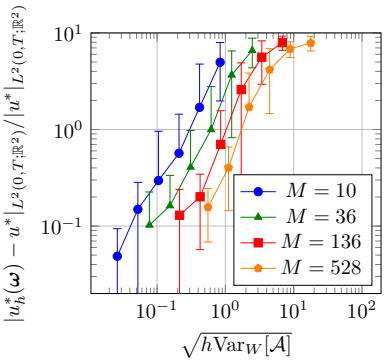
where the matrices $A_{ij} \in \mathbb{R}^{N\times N}$ $(j \geq i)$ are zero except for the coefficients

$$\begin{bmatrix} [\tilde{A}_{ij}]_{ii} & [\tilde{A}_{ij}]_{ij} \\ [\tilde{A}_{ij}]_{ji} & [\tilde{A}_{ij}]_{jj} \end{bmatrix} = \begin{bmatrix} -|[A]_{ij}| & [A]_{ij} \\ [A]_{ij} & -|[A]_{ij}| \end{bmatrix}, \quad [A_{ii}]_{ii} = [A]_{ii} + \sum_{\substack{j=1\\ j\neq i}}^N |[A]_{ij}|. \tag{172}$$

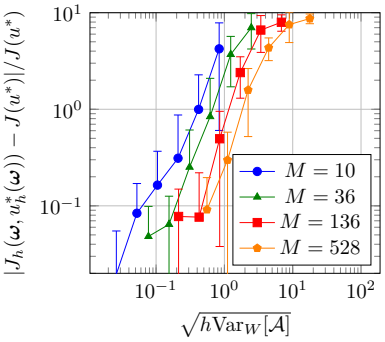
Again, it is easy to verify that the matrices A_{ij} $(j \geq i)$ are dissipative.



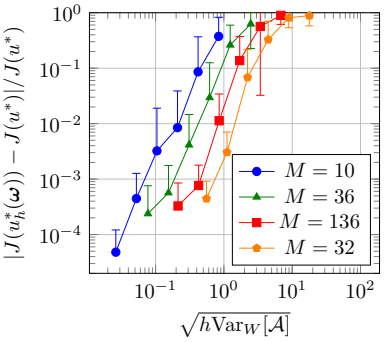
(a) Error in $x_h(t)$



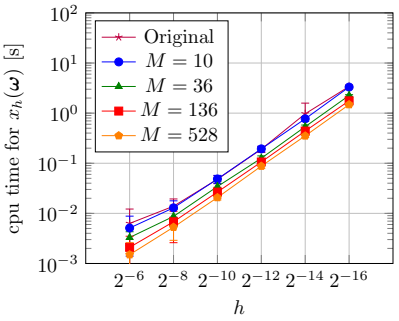
(b) Error in $u_h^*(t)$



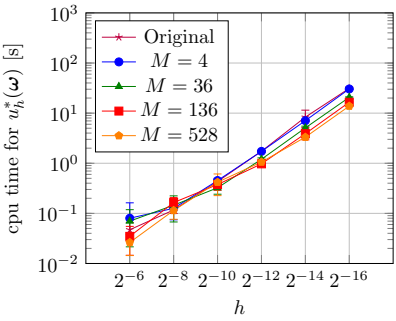
(c) Error in $J_h(u_h^*)$



(d) Error in $J(u_h^*)$



(e) Computational time for $x_h(t)$



(f) Computational time for $u_h^*(t)$

Fig. 6: Results for the discretized 1D fractional heat equation with $s = 0.7$

Now assume that N is divisable by some number P. We then decompose A into $M = P(P+1)/2$ submatrices A_m as in (5) by setting

$$A_{m(p,q)} = \sum_{i=i_{p-1}+1}^{i_p} \sum_{j=i_{q-1}+1}^{i_q} \tilde{A}_{ij}, \qquad q \geq p \in \{1, 2, \ldots, P\}, \tag{173}$$

where $i_p = pN/P$ and $m(p,q)$ is a bijection

$$m : \{(p,q) \in \{1,2,\ldots,P\}^2 \mid q \geq p\} \to \{1,2,\ldots,P(P+1)/2\}. \tag{174}$$

We thus effectively decompose A into $N/P \times N/P$ blocks, but we treat the diagonal in such a way that all submatrices A_m are dissipative.

We only use one of the matrices A_m in each time interval $[t_{k-1}, t_k)$ and thus assign uniform probabilities $2/(P(P+1))$ to each of the $M = P(P+1)/2$ subsets of $\{1,2,\ldots M\}$ of size 1.

The results obtained for $N = 96$ are shown in Figure 6. The markers and the error bars in this figure again indicate the average and 2σ-confidence interval estimated from 10 realizations of $\boldsymbol{\omega}$. Results are presented for for $P = 4$, 8, 16, and 32, which correspond to values of $M = 10$, 36, 136, and 528, respectively. Note that the number of submatrices M is now much larger than in the previous two examples, and that also $h\mathrm{Var}[\mathcal{A}]$ and the relative errors are larger than in the previous examples. Figures 6b and 6c even show relative errors that exceed 100%. However, we still observe the convergence rates predicted by the theoretical results in Section 3 in Figures 6a–6d. In particular, the convergence rate in Figure 6d is again twice as high as in the other figures.

When we inspect the computational times in Figures 6e and 6f, we see that increasing M decreases the computational time. In particular, solutions for $M = 528$ are typically computed 2-3 times faster than the solutions for the original dynamics. We expect that the computational advantage of the proposed method increases further when we increase the state dimension N.

5 Conclusions and discussions

5.1 Conclusions

We have proposed a general framework for randomized time-splitting in LQ optimal control problems. It has been shown that the dynamics, the minimal values of the cost functional, and the optimal control obtained with the proposed randomized time-splitting method converge in expectation to their analogues in the original problem when the grid spacing of the time grid goes to zero. The convergence rates in our theoretical results are also observed in three numerical examples.

In two of the three considered examples, the proposed method leads to a typical reduction in computational cost of a factor 2-3. Only in the first example of a heat equation on a 1-D spatial domain, no reduction in computational cost could be observed. This seems to be the case because the matrix A is not very large and already very sparse in this example.

5.2 Extension to unbounded operators

We have considered finite-dimensional systems in this paper, but the numerical examples in Section 4 are all obtained after spatial discretization of an infinite-dimensional system. A natural question is therefore whether our results can be extended to an infinite-dimensional setting. We already touched on this question in Remarks 5 and 7 and in Appendix B. In particular, at the end of Appendix B we indicate how results can be extended to an infinite dimensional setting under the (strong) additional assumptions that all operators A_m commute and have the same domain $D(A_m)$.

It should be noted that the assumption that $D(A_m) = D(A)$ is very strong and will not be satisfied in many applications. A prototypical example is the splitting of an advection diffusion problem with zero Dirichlet boundary conditions (represented by A) in an advective part (represented by A_1) and a diffusive part (represented by A_2). Functions in $D(A_2)$ can then satisfy the zero Dirichlet boundary conditions on the whole boundary, but the functions in $D(A_1)$ only satisfy the zero Dirichlet boundary conditions on the parts of the boundary where the velocity field is pointing inward. The analysis of the RBM becomes much more subtle in these kind of situations. The numerical results in Figure 4 also seem to indicate that the proposed randomized time splitting method converges under weaker assumptions than the ones in Appendix B.

The technical difficulties encountered when weakening these assumptions are related to the difficulties in deterministic operator splitting with unbounded operators. These date back to the paper [29] by Trotter, and have been an active field of research since then, see, e.g., [17, 20, 24, 13, 25]. As the large literature on this topic indicates, determining the necessary conditions for the convergence of the proposed stochastic operator splitting method with unbounded operators is an interesting but challenging topic for future research.

5.3 Extension to nonlinear dynamics

Another important topic for future research is the extension of our results for the linear quadratic optimal control problem to problems with non-quadratic cost functions constrained by nonlinear dynamics. This extension is particularly interesting because of the connections between the training of certain types of Deep Neural Networks (DNNs) and optimal control, see, e.g., [9, 4, 11, 10, 28], and is also important for the control of interacting particles systems, see [19].

In the most general setting, we would replace the linear dynamics (1) by the nonlinear dynamics

$$\dot{x}(t) = f(x(t), u(t)), \qquad x(0) = x_0, \tag{175}$$

where $f : \mathbb{R}^N \times \mathbb{R}^q \to \mathbb{R}^N$ is Lipschitz in the first variable x. As an analogue of (5), we then write (for $x \in \mathbb{R}^N$ and $u \in \mathbb{R}^q$)

$$f(x,u) = \sum_{m=1}^{M} f_m(x,u), \tag{176}$$

for certain Lipschitz continuous functions $f_m : \mathbb{R}^N \times \mathbb{R}^q \to \mathbb{R}^N$. Similarly as in this paper, we choose a time grid $0 = t_0 < t_1 < t_2 < \ldots < t_K = T$, enumerate the subsets $S_1, S_2, \ldots, S_{2^M}$ of $\{1, 2, \ldots, M\}$ and assign probabilities $p_1, p_2, \ldots, p_{2^M}$ to them, and randomly select a K-tuple $\boldsymbol{\omega} = (\omega_1, \omega_2, \ldots, \omega_K)$ of indices $\omega_k \in \{1, 2, \ldots 2^M\}$ according to the selected probabilities. We then consider the (typically simpler) dynamics

$$\dot{x}_h(\boldsymbol{\omega}, t) = \sum_{m \in S_{\omega_k}} \frac{f_m(x_h(\boldsymbol{\omega}, t), u_h(\boldsymbol{\omega}, t))}{\pi_m}, \qquad t \in [t_{k-1}, t_k). \tag{177}$$

Extending Theorem 1 (which considers the forward dynamics with a deterministic control $u_h(\boldsymbol{\omega}, t) = u(t)$) to such a nonlinear setting seems possible along the lines of the results for interacting-particle systems in [15]. The main difficulty is in Theorem 2 where we use the variation of constants formula to obtain an estimate for a stochastic control $u_h(\boldsymbol{\omega}, t)$ (which depends on the randomly selected indices in $\boldsymbol{\omega}$). The variation of constants formula can be extended to a nonlinear setting, see, e.g., [7], but this leads to several additional complications which we aim to address in a future work.

When an analogue of Theorem 2 for nonlinear dynamics can be been obtained, a bound on $\mathbb{E}[|J_h(u_h) - J(u_h)|]$ as in Lemma 1 should follow relatively easily from a Lipschitz condition on the integrand in the considered cost function. An analogue of the no-gap condition, i.e., a bound on $\mathbb{E}[|J(u_h^*) - J(u^*)|]$, can then be obtained using classical arguments from the calculus of variations and the bound on $\mathbb{E}[|J_h(u_h) - J(u_h)|]$, similarly as for the linear-quadratic case in Theorem 3.

With these results, the suboptimality gap $\mathbb{E}[|J_h(u_h^*) - J(u^*)|]$ be bounded using the analogues of Lemma 1 and Theorem 3 as follows. We start by noting that the triangle inequality shows that

$$|J(u_h^*(\boldsymbol{\omega})) - J(u^*)| \leq |J(u_h^*(\boldsymbol{\omega})) - J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega}))| + |J_h(\boldsymbol{\omega}, u_h^*(\boldsymbol{\omega})) - J(u^*)|. \tag{178}$$

Taking the expectation in this inequality, we see that the first term on the RHS can be bounded using (the analogue of) Lemma 1 and the second term on the RHS can be bounded using (the analogue of) Theorem 3. We thus obtain a bound on $\mathbb{E}[|J_h(u_h^*) - J(u^*)|]$ that is of order $\sqrt{h}$. It is interesting to observe that this rate is slower than the rate of order h found for the linear-quadratic case in Corollary 3. This difference seems to occur because Corollary 3 relies on the strict convexity of the functional, which is lost in a setting in which the dynamics are nonlinear.

5.4 Combination with model predictive control

As suggested in [19], it is natural to combine the proposed randomized time-splitting method with an MPC strategy. The resulting algorithm is essentially a receding horizon strategy, see, e.g., [26, 2, 3], but we now use the proposed stochastic time-splitting method to approximate the optimal controls that need to be computed in each step. An important element of such a receding horizon strategy is that the optimal control is computed based on the current state of the original dynamics (1). This creates a feedback mechanism that provides additional robustness against the errors introduced by the proposed stochastic time-splitting method.

The receding horizon strategy introduces two additional parameters in the control algorithm: the prediction horizon $\hat{T}$ and the control horizon τ. When the prediction horizon $\hat{T}$ is too short, the difference between the controls computed on the prediction horizon $[0, \hat{T}]$ and the desired optimal control on $[0, \infty)$ will be large. Decreasing the control horizon τ strengthens the feedback mechanism of the MPC strategy, which will likely allow for larger errors in the proposed stochastic time-splitting method. This idea could be formalized further by deriving an explicit error estimate that demonstrates the interaction of the control horizon τ and $h\mathrm{Var}[\mathcal{A}]$ (which characterizes the accuracy of the proposed random time-splitting method).

Acknowledgments. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement NO: 694126-DyCon), the Alexander von Humboldt-Professorship program, the European Unions Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No.765579-ConFlex and the Transregio 154 Project "Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks", project C08, of the German DFG, the Grant MTM2017-92996-C2-1-R COSNET of MINECO (Spain) and the Elkartek grant KK-2020/00091 CONVADP of the Basque government.

Appendix A Interacting particle systems in the proposed framework

In this appendix, we explain the connection of our framework to the previously proposed RBMs for interacting particle systems in [15, 16, 22, 19]. We consider a (linearized first-order) system of N interacting particles

$$\dot{x}_i(t) = \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} a_{ij}(x_j(t) - x_i(t)), \quad x_i(0) = x_{0,i}, \quad i \in \{1, 2, \dots N\}, \quad \text{(A1)}$$

where the $a_{ij} \in \mathbb{R}$ $(j \neq i)$ are constants. To simplify the following exposition, we assume that the number of particles N is divisible by some number $P > 1$.

We discuss here one particular RBM called RBM-1 in [15], but other variants can be treated similarly. We first choose a time grid $0 = t_0 < t_1 < t_2 < \ldots < t_{K_1} < t_K = T$ in the time interval $[0, T]$. In each time interval $[t_{k-1}, t_k)$, we then choose a random partition of the index set $\{1, 2, \ldots, n\}$ into disjoint subsets $\mathcal{B}_r^k$ (also called batches) of size P ($r \in \{1, 2, \ldots, N/P\}$). We consider only the interactions between particles that are in the same batch. To formalize this idea, note that, in each every time interval $[t_{k-1}, t_k)$, every particle i is contained in precisely one batch $\mathcal{B}_{r(i,k)}^k$. We thus consider the dynamics

$$\dot{x}_{\mathrm{RBM},i}(t) = \frac{1}{P-1} \sum_{\substack{j \in \mathcal{B}_{r(i,k)}^k \\ j \neq i}} a_{ij}(x_{\mathrm{RBM},j}(t) - x_{\mathrm{RBM},i}(t)), \qquad x_i(0) = x_{0,i}. \tag{A2}$$

To connect this idea to our framework, we write (A1) in matrix form

$$\dot{x}(t) = Ax(t), \qquad x(0) = x_0, \qquad A = \frac{1}{N-1} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \tilde{A}_{ij}, \tag{A3}$$

where $x(t) = [x_1(t), x_2(t), \ldots x_N(t)]^\top$ and $x_0 = [x_{0,1}, x_{0,2}, \ldots, x_{0,N}]$ and the entries of the matrices $\tilde{A}_{ij}$ ($j \neq i$) are zero except for the entries

$$\left[[\tilde{A}_{ij}]_{ij} \ [\tilde{A}_{ij}]_{ii} \right] = a_{ij} \begin{bmatrix} 1 & -1 \end{bmatrix}. \tag{A4}$$

Also the RBM-dynamics (A2) can be written in matrix form as

$$\dot{x}_{\mathrm{RBM}}(t) = \mathcal{A}_{\mathrm{RBM}}(t) x_{\mathrm{RBM}}(t), \qquad x_{\mathrm{RBM}}(0) = x_0, \tag{A5}$$

where

$$\mathcal{A}_{\mathrm{RBM}}(t) = \frac{1}{P-1} \sum_{r=1}^{N/P} \sum_{\{i,j\} \subseteq \mathcal{B}_r^k} \tilde{A}_{ij}, \qquad t \in [t_{k-1}, t_k). \tag{A6}$$

Note that the probability that two distinct indices i and j are in the same batch (i.e., the probability that $j \neq i$ is in the batch $\mathcal{B}_{r(i,k)}^k$) is $(P-1)/(N-1)$ because there are $P-1$ of the $N-1$ places in $\mathcal{B}_{r(i,k)}^k$ remaining after the index i has been fixed. This factor is also visible in the definitions of A and $\mathcal{A}_{\mathrm{RBM}}(t)$.

To make the connection to our proposed framework, we enumerate the $M = N(N-1)$ interaction matrices A_{ij}, i.e., we choose a bijection

$$\mathfrak{m} : \{(i,j) \in \{1, 2, \ldots, N\}^2 \mid i \neq j\} \to \{1, 2, \ldots, N(N-1)\}, \tag{A7}$$

and set

$$A_{\mathfrak{m}(i,j)} := \frac{1}{N-1} \tilde{A}_{ij}. \tag{A8}$$

We then need to assign probabilities p_ω to the 2^M subsets S_ω of $\{1, 2, \ldots, M\}$. Naturally, we only assign nonzero probabilities to subsets S_ω that correspond to a partition $\dot{\cup}_r \mathcal{B}_r = \{1, 2, \ldots, N\}$, i.e. sets of the form

$$S_\omega = \{\mathfrak{m}(i,j) \mid \exists_{i,j,r} \text{ such that } i \neq j \text{ and } \{i,j\} \subseteq \mathcal{B}_r\}. \tag{A9}$$

Standard combinatorics shows that there are

$$\mathcal{N} = \frac{N!}{(P!)^{N/P}\,(N/P)!}, \tag{A10}$$

distinct partitions of N indices into N/P subsets of size P. We assign a probability $p_\omega = 1/\mathcal{N}$ to each of the subsets of the form (A9).

It remains to compute the probabilities $\pi_m = \pi_{\mathfrak{m}(i,j)}$ defined in (9), i.e. to determine how many of the subsets S_ω of the form (A9) contain $m = \mathfrak{m}(i,j)$. When a certain batch $\mathcal{B}_{r^*}$ contains i and j ($j \neq i$) there are $\binom{N-2}{P-2}$ ways to fill the remaining positions in $\mathcal{B}_{r^*}$ with $P-2$ of the $N-2$ remaining indices. Once the indices in $\mathcal{B}_{r^*}$ are fixed, there are

$$\mathcal{M} = \frac{(N-P)!}{(P!)^{N/P-1}\,(N/P-1)!}, \tag{A11}$$

ways to distribute the remaining $N-P$ indices into $N/P-1$ subsets of size P. We thus conclude that

$$\pi_m = \frac{\binom{N-2}{P-2}\mathcal{M}}{\mathcal{N}} \tag{A12}$$

Using the formulas for $\mathcal{N}$ and $\mathcal{M}$, it can be verified that

$$\pi_m = \frac{P-1}{N-1}. \tag{A13}$$

It is now easy to verify that the definition of $\mathcal{A}_h(\boldsymbol{\omega}, t)$ in (11) is equivalent to the definition of $\mathcal{A}_{\mathrm{RBM}}(t)$ in (A6).

Appendix B An alternative for Corollary 1

In this appendix, we will prove a result similar to Corollary 1 under the additional assumption that all matrices commute. The proof is quite intuitive and gives an idea about how the results in this paper can be generalized to an infinite dimensional setting.

The analysis in this appendix uses the following additional assumption.

Assumption 3 Suppose that the matrices $A_1, A_2, \ldots, A_M$ all commute pairwise, i.e.

$$A_m A_{m'} = A_{m'} A_m, \tag{B14}$$

for all $m, m' \in \{1, 2, \ldots, M\}$.

Also observe that for any two dissipative matrices $X, Y \in \mathbb{R}^{N\times N}$ and vector $x_0 \in \mathbb{R}^N$ we have that

$$\begin{aligned}|e^X x_0 - e^Y x_0| &= \left|\int_0^1 \frac{d}{d\tau} e^{X\tau + Y(1-\tau)} x_0 \, \mathrm{d}\tau\right| \\ &\leq \int_0^1 \|e^{X\tau+Y(1-\tau)}\| |(X-Y)x_0| \mathrm{d}\tau \leq |(X-Y)x_0|, \end{aligned} \tag{B15}$$

where it was used that $X\tau + Y(1-\tau)$ is dissipative for $\tau \in [0,1]$ because X and Y are dissipative by assumption.

Theorem 5 *Under Assumptions 1, 2, and 3, we have that*

$$\mathbb{E}[\|S_h(t,s)x_0 - e^{A(t-s)}x_0\|^2] \leq 2h(t-s)\mathrm{Var}_W[\mathcal{A}]|W^{-1}x_0|^2, \tag{B16}$$

for all $0 \leq s \leq t \leq T$, all $x_0 \in \mathbb{R}^N$, and all invertible matrices W.

Proof We use the notation from Remark 10, so ℓ and k are such that $s \in [t_{\ell-1}, t_\ell)$ and $t \in [t_{k-1}, t_k)$, $\tilde{K} = k - \ell + 1$, and

$$\tilde{t}_0 := s < \tilde{t}_1 := t_\ell < \tilde{t}_2 := t_{\ell+1} < \ldots < \tilde{t}_{\tilde{K}-1} := t_{k-1} < \tilde{t}_{\tilde{K}} := t, \tag{B17}$$

see also Figure 1 on page 22. Furthermore, we denote $\tilde{h}_p := \tilde{t}_p - \tilde{t}_{p-1}$ for $p \in \{1, 2, \ldots, \tilde{K}\}$ and denote $\mathcal{A}_\omega := \sum_{m\in S_\omega} A_m/\pi_m$ for $\omega \in \{1, 2, \ldots, 2^M\}$. Note that $\mathcal{A}_h(\boldsymbol{\omega}, \tau) = \mathcal{A}_{\omega_p}$ for $\tau \in [\tilde{t}_{p-1}, \tilde{t}_p)$ and that $\mathcal{A}_\omega$ is dissipative for all $\omega \in \{1, 2, \ldots, 2^M\}$ because of Assumption 1.

Because the matrices $\mathcal{A}_\omega$ (with $\omega \in \{1, 2, \ldots, 2^M\}$) all commute pairwise due to Assumption 3, the formula for $S_h(\boldsymbol{\omega}, t, s)$ in (90) in Remark 10 reduces to

$$S_h(\boldsymbol{\omega}, t, s)x_0 = \exp\left(\sum_{p=1}^{\tilde{K}} \mathcal{A}_{\omega_{p+\ell-1}} \tilde{h}_p\right) x_0. \tag{B18}$$

Because Assumption 1 implies that the matrix in the exponent in the formula above and A are both dissipative, (B15) can be applied to find that

$$|S_h(\boldsymbol{\omega}, t, s)x_0 - e^{A(t-s)}x_0| \leq \left|\sum_{p=1}^{\tilde{K}} \left(\mathcal{A}_{\omega_{p+\ell-1}} - A\right) \tilde{h}_p x_0\right|, \tag{B19}$$

where it was used that $\sum_{p=1}^{\tilde{K}} \tilde{h}_p = t - s$. Squaring this expression yields

$$\begin{aligned}|S_h(\boldsymbol{\omega}, t, s)x_0 - e^{A(t-s)}x_0|^2 \leq \\ \sum_{p,p'=1}^{\tilde{K}} \tilde{h}_p \tilde{h}_{p'} \langle (\mathcal{A}_{\omega_{p+\ell-1}} - A)x_0, (\mathcal{A}_{\omega_{p'+\ell-1}} - A)x_0\rangle. \end{aligned} \tag{B20}$$

When we take the expected value, the terms with $p \neq p'$ disappear because

$$\mathbb{E}[\langle (\mathcal{A}_{\omega_{p+\ell-1}} - A)x_0, (\mathcal{A}_{\omega_{p'+\ell-1}} - A)x_0\rangle]$$

$$= \sum_{\omega=1}^{2^M} \sum_{\omega'=1}^{2^M} \langle (\mathcal{A}_\omega - A)x_0, (\mathcal{A}_{\omega'} - A)x_0 \rangle p_\omega p_{\omega'}$$
$$= \left\langle \sum_{\omega=1}^{2^M} (\mathcal{A}_\omega - A)x_0, \sum_{\omega'=1}^{2^M} (\mathcal{A}_{\omega'} - A)x_0 \right\rangle = \langle 0, 0 \rangle = 0 \tag{B21}$$

where the first identity follows after writing $\omega = \omega_{p-\ell+1}$ and $\omega' = \omega_{p'-\ell+1}$, and the second to last identity from (12) and (8). Therefore, only the terms with $p = p'$ remain after taking the expected value of (B20) and

$$\begin{aligned}
&\mathbb{E}[|S_h(t,s)x_0 - e^{A(t-s)}x_0|^2] \\
&\leq \sum_{\omega_\ell=1}^{2^M} \sum_{\omega_{\ell+1}=1}^{2^M} \cdots \sum_{\omega_{\ell+\tilde{K}-1}=1}^{2^M} \sum_{p=1}^{\tilde{K}} \tilde{h}_p^2 |(\mathcal{A}_{\omega_{p+\ell-1}} - A)x_0|^2 p_{\omega_\ell} p_{\omega_{\ell+1}} \cdots p_{\omega_{\ell+\tilde{K}-1}} \\
&= \sum_{p=1}^{\tilde{K}} \tilde{h}_p^2 \sum_{\omega=1}^{2^M} |(\mathcal{A}_{\omega_p+\ell-1} - A)x_0|^2 p_\omega.
\end{aligned} \tag{B22}$$

The proof is completed with two straightforward observations. First of all, note that because $\tilde{h}_p \leq h$

$$\sum_{p=1}^{\tilde{K}} \tilde{h}_p^2 \leq \sum_{p=1}^{\tilde{K}} h\tilde{h}_p = h \sum_{p=1}^{\tilde{K}} \tilde{h}_p = h(t-s). \tag{B23}$$

Secondly, we have that

$$\begin{aligned}
\sum_{\omega=1}^{2^M} |(\mathcal{A}_{\omega_p+\ell-1} - A)x_0|^2 p_\omega &= \sum_{\omega=1}^{2^M} |(\mathcal{A}_{\omega_p+\ell-1} - A)WW^{-1}x_0|^2 p_\omega \\
&\leq \sum_{\omega=1}^{2^M} \|(\mathcal{A}_{\omega_p+\ell-1} - A)W\|^2 |W^{-1}x_0|^2 p_\omega.
\end{aligned} \tag{B24}$$

The result follows after inserting (B23) and (B24) into (B22). □

The proof of Theorem 5 extends naturally to an infinite dimensional setting as follows. Most of the definitions and notations from Section 2 remain unchanged, apart from the following.

- The state and the control no longer evolve in the finite-dimensional spaces $\mathbb{R}^N$ and $\mathbb{R}^q$, but in the (potentially) infinite-dimensional Hilbert spaces X and U, respectively.
- A and A_m (with $m \in \{1, 2, \ldots, M\}$) now represent the generators of C_0-semigroups e^{At} and $e^{A_m t}$ on the Hilbert space X with domains $D(A)$ and $D(A_m)$, respectively.
- B is now a bounded linear operator from U to X.

For simplicity we assume that the domains of the operators A_m are all the same and equal to the domain of A, i.e. $D(A_m) = D(A)$. For a value of λ in the resolvent set of A, the resolvent $W = (A - \lambda I)^{-1}$ is a bounded operator $X \to D(A) \subset X$ with (unbounded) inverse $A - \lambda I$ and one now easily verifies that AW and $A_m W$ represent bounded operators on X, meaning that

$\mathrm{Var}_W[\mathcal{A}]$ as introduced in Remark 5 is bounded. For $|W^{-1}x_0| = |(A - \lambda I)x_0|$ to be bounded, we require that $x_0 \in D(A)$. The proof of Theorem 5 can thus be applied in this setting with the additional assumption that $x_0 \in D(A)$. The proof remains effectively unchanged.

Note that when we want to use Theorem 5 to obtain a result similar to Theorem 2, we also need a smoothness assumption on the input operator B. In particular, similarly as (100) in Theorem 2, we would then like to bound

$$\int_0^t \left| (S_h(\boldsymbol{\omega}, t, s) - e^{A(t-s)}) B u_h(\boldsymbol{\omega}, s) \right| \, \mathrm{d}s, \tag{B25}$$

which is only possible with Theorem 5 when $|W^{-1}Bu_h(\boldsymbol{\omega}, s)|$ is finite. To this end one would typically require that the range of B is contained in $D(A)$.

References

[1] T. Apel and T. G. Flaig. Crank-Nicolson schemes for optimal control problems with evolution equations. *SIAM J. Numer. Anal.*, 50(3):1484–1512, 2012. ISSN 0036-1429. doi: 10.1137/100819333. URL https://doi.org/10.1137/100819333.

[2] B. Azmi and K. Kunisch. On the stabilizability of the Burgers equation by receding horizon control. *SIAM J. Control Optim.*, 54(3):1378–1405, 2016. ISSN 0363-0129. doi: 10.1137/15M1030352. URL https://doi.org/10.1137/15M1030352.

[3] B. Azmi and K. Kunisch. Receding horizon control for the stabilization of the wave equation. *Discrete Contin. Dyn. Syst.*, 38(2):449–484, 2018. ISSN 1078-0947. doi: 10.3934/dcds.2018021. URL https://doi.org/10.3934/dcds.2018021.

[4] M. Benning, E. Celledoni, M. J. Ehrhardt, B. Owren, and C.-B. Schönlieb. Deep learning as optimal control problems: models and numerical methods. *J. Comput. Dyn.*, 6(2):171–198, 2019. ISSN 2158-2491. doi: 10.3934/jcd.2019009. URL https://doi.org/10.3934/jcd.2019009.

[5] U. Biccari and V. Hernández-Santamaría. Controllability of a one-dimensional fractional heat equation: theoretical and numerical aspects. *IMA Journal of Mathematical Control and Information*, 36(4):1199–1235, 07 2018. ISSN 0265-0754. doi: 10.1093/imamci/dny025. URL https://doi.org/10.1093/imamci/dny025.

[6] L. Bottou, F. E. Curtis, and J. Nocedal. Optimization methods for large-scale machine learning. *SIAM Rev.*, 60(2):223–311, 2018. ISSN 0036-1445. doi: 10.1137/16M1080173. URL https://doi.org/10.1137/16M1080173.

[7] F. Brauer. Perturbations of nonlinear systems of differential equations. *J. Math. Anal. Appl.*, 14:198–206, 1966. ISSN 0022-247X. doi: 10.1016/0022-247X(66)90021-7. URL https://doi.org/10.1016/0022-247X(66)90021-7.

[8] V. Dolean, P. Jolivet, and F. Nataf. *An introduction to domain decomposition methods: Algorithms, theory, and parallel implementation.* Society

for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2015. ISBN 978-1-611974-05-8. doi: 10.1137/1.9781611974065.ch1. URL https://doi.org/10.1137/1.9781611974065.ch1.

[9] W. E. A proposal on machine learning via dynamical systems. *Commun. Math. Stat.*, 5(1):1–11, 2017. ISSN 2194-6701. doi: 10.1007/s40304-017-0103-z. URL https://doi.org/10.1007/s40304-017-0103-z.

[10] C. Esteve and B. Geshkovski. Sparse approximation in learning via neural ODEs, 2021.

[11] C. Esteve, B. Geshkovski, D. Pighin, and E. Zuazua. Large-time asymptotics in deep learning, 2021.

[12] L. Grüne and J. Pannek. *Nonlinear model predictive control*. Communications and Control Engineering Series. Springer, Cham, 2017. ISBN 978-3-319-46023-9; 978-3-319-46024-6. doi: 10.1007/978-3-319-46024-6. URL https://doi.org/10.1007/978-3-319-46024-6. Theory and algorithms, Second edition [of MR3155076].

[13] E. Hansen and A. Ostermann. Dimension splitting for evolution equations. *Numer. Math.*, 108(4):557–570, 2008. ISSN 0029-599X. doi: 10.1007/s00211-007-0129-3. URL https://doi.org/10.1007/s00211-007-0129-3.

[14] L. I. Ignat. A splitting method for the nonlinear Schrödinger equation. *J. Differential Equations*, 250(7):3022–3046, 2011. ISSN 0022-0396. doi: 10.1016/j.jde.2011.01.028. URL https://doi.org/10.1016/j.jde.2011.01.028.

[15] S. Jin, L. Li, and J.-G. Liu. Random batch methods (RBM) for interacting particle systems. *J. Comput. Phys.*, 400:108877, 30, 2020. ISSN 0021-9991. doi: 10.1016/j.jcp.2019.108877. URL https://doi.org/10.1016/j.jcp.2019.108877.

[16] S. Jin, L. Li, and J.-G. Liu. Convergence of random batch method for interacting particles with disparate species and weights, 2020.

[17] T. Kato. Trotter's product formula for an arbitrary pair of self-adjoint contraction semigroups. In *Topics in functional analysis (essays dedicated to M. G. Kreĭn on the occasion of his 70th birthday)*, volume 3 of *Adv. in Math. Suppl. Stud.*, pages 185–195. Academic Press, New York-London, 1978.

[18] D. E. Kirk. *Optimal control theory: an introduction*. Dover, 2004.

[19] D. Ko and E. Zuazua. Model predictive control with random batch methods for a guiding problem. *Math. Models Methods Appl. Sci.*, 31(8):1569–1592, 2021. ISSN 0218-2025. doi: 10.1142/S0218202521500329. URL https://doi.org/10.1142/S0218202521500329.

[20] M. L. Lapidus. Generalization of the Trotter-Lie formula. *Integral Equations Operator Theory*, 4(3):366–415, 1981. ISSN 0378-620X. doi: 10.1007/BF01697972. URL https://doi.org/10.1007/BF01697972.

[21] E. B. Lee and L. Markus. *Foundations of optimal control theory*. John Wiley & Sons, Inc., New York-London-Sydney, 1967.

[22] L. Li, Z. Xu, and Y. Zhao. A random-batch Monte Carlo method for many-body systems with singular kernels. *SIAM J. Sci. Comput.*, 42(3):A1486–A1509, 2020. ISSN 1064-8275. doi: 10.1137/19M1302077. URL

https://doi.org/10.1137/19M1302077.

[23] M. Minoux and S. Vajda. *Mathematical Programming: Theory and Algorithms*. A Wiley-Interscience publication. Wiley, 1986. ISBN 9780471901709. URL https://books.google.de/books?id=5kDvAAAAMAAJ.

[24] H. Neidhardt and V. A. Zagrebnov. On error estimates for the Trotter-Kato product formula. *Lett. Math. Phys.*, 44(3):169–186, 1998. ISSN 0377-9017. doi: 10.1023/A:1007494816401. URL https://doi.org/10.1023/A:1007494816401.

[25] A. Ostermann and K. Schratz. Stability of exponential operator splitting methods for noncontractive semigroups. *SIAM J. Numer. Anal.*, 51(1): 191–203, 2013. ISSN 0036-1429. doi: 10.1137/110846580. URL https://doi.org/10.1137/110846580.

[26] M. Reble and F. Allgöwer. Unconstrained model predictive control and suboptimality estimates for nonlinear continuous-time systems. *Automatica J. IFAC*, 48(8):1812–1817, 2012. ISSN 0005-1098. doi: 10.1016/j.automatica.2012.05.067. URL https://doi.org/10.1016/j.automatica.2012.05.067.

[27] V. K. Rohatgi and A. K. M. Ehsanes Saleh. *An introduction to probability and statistics*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, third edition, 2015. ISBN 978-1-118-79964-2. doi: 10.1002/9781118799635. URL https://doi.org/10.1002/9781118799635.

[28] D. Ruiz-Balet and E. Zuazua. Neural ODE control for classification, approximation and transport, 2021.

[29] H. F. Trotter. On the product of semi-groups of operators. *Proc. Amer. Math. Soc.*, 10:545–551, 1959. ISSN 0002-9939. doi: 10.2307/2033649. URL https://doi.org/10.2307/2033649.