



Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

**ScienceDirect**

Journal of Differential Equations 371 (2023) 83–115

---

**Journal of  
Differential  
Equations**

---

[www.elsevier.com/locate/jde](http://www.elsevier.com/locate/jde)

# The hypocoercivity index for the short time behavior of linear time-invariant ODE systems

Franz Achleitner <sup>a,\*</sup>, Anton Arnold <sup>a</sup>, Eric A. Carlen <sup>b</sup>

<sup>a</sup> *TU Wien, Institute of Analysis and Scientific Computing, Wiedner Hauptstr. 8-10, A-1040, Wien, Austria*

<sup>b</sup> *Rutgers University, Department of Mathematics, 110 Frelinghuysen Rd., NJ 08854, Piscataway, USA*

Received 5 August 2022; revised 3 March 2023; accepted 13 June 2023

Available online 28 June 2023

---

## Abstract

We consider the class of conservative-dissipative ODE systems, which is a subclass of Lyapunov stable, linear time-invariant ODE systems. We characterize asymptotically stable, conservative-dissipative ODE systems via the hypocoercivity (theory) of their system matrices. Our main result is a concise characterization of the hypocoercivity index (an algebraic structural property of matrices with positive semi-definite Hermitian part introduced in Achleitner, Arnold, and Carlen (2018)) in terms of the short time behavior of the norm of the matrix exponential for the associated conservative-dissipative ODE system.

© 2023 Elsevier Inc. All rights reserved.

*MSC:* 34D05; 34A30; 34Exx

*Keywords:* Semi-dissipative ODE systems; Hypocoercivity (index)

---

## 1. Introduction

In this paper we shall use hypocoercivity techniques to characterize the short time behavior of linear time-invariant ODE systems of the form

$$\mathbf{x}'(t) = -B\mathbf{x}(t), \quad (1.1)$$

---

\* Corresponding author.

E-mail addresses: [franz.achleitner@tuwien.ac.at](mailto:franz.achleitner@tuwien.ac.at) (F. Achleitner), [anton.arnold@tuwien.ac.at](mailto:anton.arnold@tuwien.ac.at) (A. Arnold), [carlen@math.rutgers.edu](mailto:carlen@math.rutgers.edu) (E.A. Carlen).

with matrices  $B \in \mathbb{C}^{n \times n}$  whose Hermitian part  $B_H := (B + B^*)/2$  is positive semi-definite,<sup>1</sup> such that  $-B$  is a conservative-dissipative (or semi-dissipative) matrix, see Definition 1.1 below. An extension to *stable systems* of the form (1.1), i.e. matrices  $B$  having their spectrum in the closed right half plane where, additionally, purely imaginary eigenvalues are non-defective,<sup>2</sup> are discussed in the parallel article [5].

Concerning the short time behavior of the propagator  $P(t)$  of (1.1), given by the fundamental matrix  $P(t) := e^{-Bt}$ , we shall be interested in estimates on the spectral norm of the matrix exponential  $e^{-Bt}$  of the form

$$\|P(t)\|_2 = 1 - c t^a + \mathcal{O}(t^{a+1}) \quad \text{for } t \rightarrow 0^+, \quad (1.2)$$

where  $c > 0$  and  $a \in \mathbb{N}$ . For practical reason, we shall refer to the spectral norm of the matrix exponential  $e^{-Bt}$  as the *propagator norm*. The large time behavior, including conditions guaranteeing exponential decay estimates of the form

$$\|P(t)\|_2 \leq c e^{-\mu t}, \quad t \geq 0, \quad (1.3)$$

with some  $c \geq 1$ , and  $\mu > 0$  will be discussed in the forthcoming article [3].

Following [18], the matrix  $B$  is called *hypocoercive* if (1.3) holds, and  $B$  is *coercive* if and only if  $\|P(t)\|_2 \leq e^{-\mu t}$  for some  $\mu > 0$  and all  $t \geq 0$ . For the intermediate time regime we shall be interested in a *waiting time*  $t_0$  such that all solutions will have decayed (at least) by a factor  $1/e$ , i.e.

$$\|P(t)\|_2 \leq \frac{1}{e}, \quad t \geq t_0. \quad (1.4)$$

Here we shall include numerical evidence on this intermediate time regime; a detailed analysis will be provided in [3].

Our interest in systems (1.1) is motivated by non-equilibrium statistical physics and, in particular, kinetic theory. There, a frequently encountered class of linear equations has the form

$$\frac{\partial}{\partial t} f(t, x, v) = -v \cdot \nabla_x f(t, x, v) + Qf(t, x, v),$$

where  $Q$  is a linear operator, the so-called *linearized collision operator*, describing changes in velocity resulting from binary collisions. Take the domain of the  $x$  variable to be a  $d$ -dimensional torus  $\mathcal{T}$  of side-length  $L$ . The operator  $v \cdot \nabla_x$ , called the streaming operator is anti-Hermitian on a weighted  $L^2$ -space on  $\mathcal{T} \times \mathbb{R}^d$ , whereas  $Q$  is negative semi-definite on the same weighted  $L^2$ -space. Concerning examples we refer to [10, §1.4], [1,2] where a modal decomposition (in  $x$ ) of kinetic BGK-type equations<sup>3</sup> led to ODE systems like (1.1); in the case of continuous velocities it actually led to “infinite matrices”  $B$ . In the follow-up paper [9] a spectral decomposition of Fokker–Planck equations allowed for a precise analysis of their short time behavior in the spirit of (1.2).

<sup>1</sup> We use the following notation: The conjugate transpose of a matrix  $B \in \mathbb{C}^{n \times n}$  is denoted by  $B^*$ . Positive definiteness (resp. semi-definiteness) of Hermitian matrices is denoted by  $B > 0$  (resp.  $B \geq 0$ ).

<sup>2</sup> An eigenvalue is non-defective if its algebraic and geometric multiplicities coincide.

<sup>3</sup> named after the physicists Bhatnagar–Gross–Krook.

Very often, one is interested in initial data  $f_0(x, v)$  for the equation that are nearly constant on spatial scales much smaller than  $L$ , where  $L$  is large, so that  $v \cdot \nabla_x f(t, x, v)$  will be of order  $1/L$ . Finite dimensional approximations of such systems result in systems of ODEs of the form (1.1), where its Hermitian part  $B_H$  is positive semi-definite. Hence it is often natural to write (1.1) in the form

$$\mathbf{x}'(t) = -\epsilon A\mathbf{x}(t) - C\mathbf{x}(t), \quad (1.5)$$

where  $\epsilon \in \mathbb{R}$ ,  $A \in \mathbb{C}^{n \times n}$  is anti-Hermitian, and  $C \in \mathbb{C}^{n \times n}$  is positive semi-definite Hermitian. The systems (1.1) and (1.5) are related via  $B_A = \epsilon A$  and  $B_H = C$ , where  $B_A := (B - B^*)/2$  denotes the anti-Hermitian part of  $B$ . One is then led to study the asymptotics of the solution as  $\epsilon \rightarrow 0$ . This is often referred to as *asymptotic limit*.

The main goal of this paper is to give a concise interpretation of the *hypocoercivity index* (an algebraic structural property of the matrix  $B$ , introduced in [2]) in terms of the exponent  $a$  in short time estimate (1.2).

The rest of this paper is organized as follows: In the remainder of §1 we define and characterize *conservative-dissipative ODE systems* and *hypocoercive matrices*. In §2.1 we give three equivalent definitions of the *hypocoercivity index* of a matrix and three corresponding *Kalman rank conditions*. Our main result, Theorem 2.7, gives a sharp characterization of the hypocoercivity index of a matrix  $B$  in terms of the short time decay of the spectral norm of the associated matrix exponential  $e^{-Bt}$ . It is presented in §2.2, while technical parts of the proof are deferred to Appendix A. Finally, §2.3 gives a numerical illustration of Theorem 2.7.

### 1.1. Conservative-dissipative systems of ODEs

It is well known that the null solution  $\mathbf{x}(t) \equiv 0$  of a linear system (1.1) is (*Lyapunov*) *stable* if all eigenvalues of  $-B$  have non-positive real part and the eigenvalues on the imaginary axis are non-defective, and the null solution  $\mathbf{x}(t) \equiv 0$  of (1.1) is *asymptotically stable* if all eigenvalues of  $-B$  have negative real part. For practical reasons, if the null solution  $\mathbf{x}(t) \equiv 0$  of a linear system (1.1) is (asymptotically) stable then we will call (1.1) an (asymptotically) stable system.

Consider a linear system of ODEs (1.1) with matrix  $B \in \mathbb{C}^{n \times n}$ . Then the derivative of the squared Euclidean norm of a solution  $\mathbf{x}(t)$  satisfies

$$\frac{d}{dt} \|\mathbf{x}(t)\|_2^2 = \langle -B\mathbf{x}(t), \mathbf{x}(t) \rangle + \langle \mathbf{x}(t), -B\mathbf{x}(t) \rangle = -2\langle \mathbf{x}(t), B_H\mathbf{x}(t) \rangle. \quad (1.6)$$

Therefore a sufficient condition for  $B$  to generate a stable system (1.1) is that its Hermitian part is positive semi-definite.<sup>4</sup> This fact and the importance of this subclass of stable systems in kinetic theory inspires the following definition:

**Definition 1.1.** A matrix  $-B \in \mathbb{C}^{n \times n}$  is called *dissipative* (resp. *conservative-dissipative* or *semi-dissipative*) if the Hermitian part of  $-B$  is negative definite (resp. negative semi-definite).

For a (conservative-)dissipative matrix  $-B \in \mathbb{C}^{n \times n}$ , the associated system of ODEs (1.1) is called a (*conservative-)*dissipative ODE system.

<sup>4</sup> However, the Hermitian part of a matrix  $B$  pertaining to a stable system does not have to be positive semi-definite; a different equivalent norm will generally yield a different sufficient condition.

For practical reasons, a matrix  $B \in \mathbb{C}^{n \times n}$  is called *positive conservative-dissipative* if the Hermitian part of  $B$  is positive semi-definite.

This conservative-dissipative property of a matrix  $B$  is invariant under unitary transformations, but it is not invariant under similarity transformations (see [5] for details). Let  $\lambda_{min}^{B_H}$  and  $\lambda_{max}^{B_H}$  denote the least and the greatest eigenvalues of  $B_H$ , respectively. By the Rayleigh-Ritz variational principle

$$\lambda_{min}^{B_H} = \min_{\|\mathbf{x}\|_2=1} \langle \mathbf{x}, B_H \mathbf{x} \rangle \quad \text{and} \quad \lambda_{max}^{B_H} = \max_{\|\mathbf{x}\|_2=1} \langle \mathbf{x}, B_H \mathbf{x} \rangle.$$

It follows immediately that

$$\lambda_{min}^{B_H} \leq \min\{\Re \lambda : \lambda \in \sigma(B)\} \leq \max\{\Re \lambda : \lambda \in \sigma(B)\} \leq \lambda_{max}^{B_H}. \quad (1.7)$$

## 1.2. Hypocoercive matrices

In the introduction, matrices  $B \in \mathbb{C}^{n \times n}$  are called *hypocoercive* if the associated propagator  $P(t) := e^{-tB}$  satisfies (1.3) for some  $c \geq 1$  and  $\mu > 0$ . In other words,  $B$  is hypocoercive if (the null-solution  $\mathbf{x}(t) \equiv 0$  of) the associated linear system (1.1) is exponentially stable. For linear time-invariant systems (1.1), exponential stability is equivalent to asymptotic stability (i.e. all solutions approach the origin in the large time limit), which is equivalent to the condition that all eigenvalues of  $-B$  have negative real part. Starting with the classical notion of a coercive operator/matrix, we characterize hypocoercive matrices as follows (see e.g. [2]):

**Definition 1.2.** A matrix  $B \in \mathbb{C}^{n \times n}$  is called *coercive* if its Hermitian part  $B_H$  is positive definite, and it is called *hypocoercive* if the spectrum of  $B$  lies in the *open* right half plane.

Hypocoercive matrices are often called *positive stable*. We use the notion of hypocoercivity to emphasize the link to the analogous situation in partial differential equations, see [2,8,18].

It is well-known, see (1.7), that for positive conservative-dissipative matrices  $B \in \mathbb{C}^{n \times n}$ , the spectrum of  $B$  lies in the closed right half plane, but there may be purely imaginary eigenvalues. In this case, the existence of purely imaginary eigenvalues can be characterized as follows:

**Proposition 1.3** ([16, Lemma 3.1], [2, Lemma 2.4 with Proposition 1(B2), (B4)]). *Let  $B \in \mathbb{C}^{n \times n}$  be (positive) conservative-dissipative. Then,  $B$  has an eigenvalue on the imaginary axis if and only if  $B_H v = 0$  for some eigenvector  $v$  of  $B_A$ .*

Hence, a positive conservative-dissipative matrix  $B$  is hypocoercive if and only if no eigenvector of the anti-Hermitian part lies in the kernel of the Hermitian part. The latter condition is well known in control theory, and there exists a range of equivalent characterizations, see e.g. [2, Proposition 1] and [16, Lemma 3.1]: For example, positive conservative-dissipative matrices  $B \in \mathbb{C}^{n \times n}$  are hypocoercive if and only if

$$\text{“No non-trivial subspace of } \ker(B_H) \text{ is invariant under } B_A.” \quad (1.8)$$

The characterization (1.8) implies the following result:

**Lemma 1.4.** *Let  $B \in \mathbb{C}^{n \times n}$  be positive conservative-dissipative. Then  $B = B_A + B_H$  is hypocoercive if and only if  $\epsilon B_A + B_H$  is hypocoercive for all  $\epsilon \neq 0$ .*

But even if hypocoercivity of  $B$  is known, it is not trivial to obtain an exponential decay estimate (1.3) with a quantitative (or even optimal) decay rate  $\mu$ . Indeed, a simple energy estimate (i.e. pre-multiplying (1.1) by  $\mathbf{x}^*$ ) or using Trotter's product formula only yields conservative-dissipativity of the system, but no decay: So, its propagator  $P(t) = e^{-Bt}$  satisfies at least  $\|P(t)\|_2 \leq 1$  for all  $t \geq 0$ .

Let us also comment on the short time behavior. If a decay estimate (1.2) holds with some exponent  $a > 1$ , then an estimate of the form  $\|P(t)\|_2 \leq e^{-\mu t}$  (as it is typical for coercive matrices  $B$ ) is impossible (consider the Taylor expansion of  $\|P(t)\|_2$  around  $t = 0$ ). In such cases the system can only be hypocoercive, along with an estimate (1.3) with  $c > 1$ .

Our conditions for hypocoercivity lead naturally to the notion of an *index of hypocoercivity*, which is a non-negative integer whenever  $B$  is hypocoercive, and is zero if and only if  $B$  is coercive (see §2.1).

## 2. Hypocoercivity index and the short time decay of conservative-dissipative ODE systems

In this section we shall present our main result, i.e. a concise characterization of the hypocoercivity index in terms of the short time behavior of conservative-dissipative ODE systems (1.1).

### 2.1. Hypocoercivity index

First we recall from [2, §2.2] the definition of the hypocoercivity index:

**Definition 2.1.** *Let  $B \in \mathbb{C}^{n \times n}$  be positive conservative-dissipative. Its *hypocoercivity index* (*HC-index*)  $m_{HC}$  is defined as the smallest integer  $m \in \mathbb{N}_0$  (if it exists) such that the matrix*

$$T_m := \sum_{j=0}^m B_A^j B_H (B_A^*)^j \quad (2.1)$$

is positive definite. The matrix  $B$  is coercive iff  $m_{HC} = 0$ , hypocoercive iff  $m_{HC} \in \mathbb{N}_0$  (due to [2, Lemma 2.4]), and for non-hypocoercive matrices  $B$  we set  $m_{HC} = \infty$ .

The Hermitian matrix  $T_m$  in Definition 2.1 readily shows that the hypocoercivity index of a positive conservative-dissipative matrix  $B$  is invariant under unitary congruence transformations but, in general, not under similarity transformations (see [4] for details).

In [7, Lemma 2.3] it was proven that this index equals the smallest integer  $m \in \mathbb{N}_0$  such that

$$\text{rank} \{ \sqrt{B_H}, B_A \sqrt{B_H}, \dots, B_A^m \sqrt{B_H} \} = n, \quad (2.2)$$

which is often called *Kalman rank condition* (see also [2, Proposition 1(B1)]).

**Remark 2.2.** *When considering rather  $\epsilon B_A + B_H$ , its hypocoercivity property and its index of hypocoercivity are independent of  $\epsilon \neq 0$ , which follows trivially from (2.2). Due to the above mentioned equality of the indices, the positive definiteness of  $\sum_{j=0}^{m_{HC}} (\epsilon B_A)^j B_H (\epsilon B_A^*)^j$  is hence also independent of  $\epsilon \neq 0$ .*

Next we shall present four (equivalent) variants of the Kalman rank condition:

**Lemma 2.3.** *Let  $B \in \mathbb{C}^{n \times n}$  be positive conservative-dissipative. Consider the following four Kalman rank conditions:*

$$\exists m \in \mathbb{N}_0 : \text{rank}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}\} = n, \quad (2.3a)$$

$$\exists m \in \mathbb{N}_0 : \text{rank}\{\sqrt{B_H}, B\sqrt{B_H}, \dots, B^m\sqrt{B_H}\} = n, \quad (2.3b)$$

$$\exists m \in \mathbb{N}_0 : \text{rank}\{\sqrt{B_H}, B^*\sqrt{B_H}, \dots, (B^*)^m\sqrt{B_H}\} = n, \quad (2.3c)$$

$$\exists m \in \mathbb{N}_0 : \text{rank}\{C_0, C_1, \dots, C_m\} = n$$

$$\text{with } C_0 := \sqrt{B_H}; C_{j+1} := [C_j, B_A] = C_j B_A - B_A C_j, j \in \mathbb{N}_0. \quad (2.3d)$$

The conditions (2.3a)–(2.3d) are equivalent in the sense that, if there exists  $m \in \mathbb{N}_0$  such that one condition holds, then the other three conditions hold as well for the same  $m$ .

**Proof.** In fact, we prove that for all  $m \in \mathbb{N}_0$  the ranges of all four matrices in (2.3) are equal,

$$\text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}\} = \text{range}\{\sqrt{B_H}, B\sqrt{B_H}, \dots, B^m\sqrt{B_H}\} \quad (2.4a)$$

$$= \text{range}\{\sqrt{B_H}, B^*\sqrt{B_H}, \dots, (B^*)^m\sqrt{B_H}\} \quad (2.4b)$$

$$= \text{range}\{C_0, C_1, \dots, C_m\}, \quad (2.4c)$$

which can be done inductively:

The base case  $m = 0$  is trivial, since all four matrices in (2.4) are equal to  $\sqrt{B_H}$ . We start with the induction step for the equivalence of the first identity in (2.4): Assume (2.4a) holds for some  $m \in \mathbb{N}_0$ . Then, the following representations hold:  $(B^{m+1} - B_A^{m+1})\sqrt{B_H} = \sum_{j=0}^m B_A^j \sqrt{B_H} X_j$ , with appropriate matrices  $X_j \in \mathbb{C}^{n \times n}$ , and e.g.,  $X_m = B_H$ . Therefore

$$\text{range}((B^{m+1} - B_A^{m+1})\sqrt{B_H}) \subset \text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}\}$$

which implies

$$\begin{aligned} & \text{range}\{\sqrt{B_H}, B\sqrt{B_H}, \dots, B^m\sqrt{B_H}, B^{m+1}\sqrt{B_H}\} \\ &= \text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}, B^{m+1}\sqrt{B_H}\} \\ &= \text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}, B_A^{m+1}\sqrt{B_H}\}. \end{aligned}$$

Hence,  $\text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^{m+1}\sqrt{B_H}\} = \text{range}\{\sqrt{B_H}, B\sqrt{B_H}, \dots, B^{m+1}\sqrt{B_H}\}$  follows.

In the same way, for  $m \in \mathbb{N}_0$ , the identity

$$\text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}\} = \text{range}\{\sqrt{B_H}, B^*\sqrt{B_H}, \dots, (B^*)^m\sqrt{B_H}\}$$

is proven.

Since the equivalence of  $\text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}\}$  and  $\text{range}\{C_0, C_1, \dots, C_m\}$  follows very similarly, we only give the key steps: Note that the term of  $C_{j+1}$  with the leading maximum exponent of  $B_A$  is of the form  $(-B_A)^{j+1}\sqrt{B_H}$ . This allows to compute  $C_{m+1} - (-B_A)^{m+1}\sqrt{B_H} = \sum_{j=0}^m B_A^j \sqrt{B_H} X_j$ , with appropriate matrices  $X_j \in \mathbb{C}^{n \times n}$  and, e.g.,  $X_m = (-1)^m(m+1)B_A$ . Therefore

$$\text{range}(C_{m+1} - (-B_A)^{m+1}\sqrt{B_H}) \subset \text{range}\{\sqrt{B_H}, B_A\sqrt{B_H}, \dots, B_A^m\sqrt{B_H}\}$$

and the rest of the proof is as before.  $\square$

Next we shall present two (equivalent) variants of Definition 2.1 for the hypocoercivity index that are related to similar concepts in the literature: In [8, §2], the definition of the hypocoercivity index of a degenerate Fokker–Planck equation with linear drift involves a positive stable matrix  $B$  and its Hermitian part  $B_H \geq 0$  (rather than the Hermitian part and anti-Hermitian part of  $B$  as in Definition 2.1 here). In order to connect these two situations we shall establish the equivalence of these two definitions in the subsequent lemma.

The condition  $T_m > 0$  from Definition 2.1 can also be related to Hörmander’s “rank  $r$ ” bracket condition for hypoellipticity, cf. [13]. In particular, in [18] iterated commutators were used to establish hypocoercivity of kinetic PDEs by constructing an appropriate Lyapunov functional. In Lemma 2.4, Equation (2.5d), below we shall mimic condition (3.5) of [18] for the ODE-system (1.1).

**Lemma 2.4.** *Let  $B \in \mathbb{C}^{n \times n}$  be positive conservative-dissipative. Consider the following four hypocoercivity conditions:*

$$\exists m \in \mathbb{N}_0 : T_m := \sum_{j=0}^m B_A^j B_H (B_A^*)^j > 0, \quad (2.5a)$$

$$\exists m \in \mathbb{N}_0 : \tilde{T}_m := \sum_{j=0}^m B^j B_H (B^*)^j > 0, \quad (2.5b)$$

$$\exists m \in \mathbb{N}_0 : \tilde{\tilde{T}}_m := \sum_{j=0}^m (B^*)^j B_H B^j > 0, \quad (2.5c)$$

$$\exists m \in \mathbb{N}_0 : \hat{T}_m := \sum_{j=0}^m C_j^* C_j > 0$$

$$\text{with } C_0 := \sqrt{B_H}; C_{j+1} := [C_j, B_A] = C_j B_A - B_A C_j, \quad j \in \mathbb{N}_0. \quad (2.5d)$$

The conditions (2.5a)–(2.5d) are equivalent in the sense that, if there exists  $m \in \mathbb{N}_0$  such that one condition holds, then the other three conditions hold as well for the same  $m$ .

**Proof.** According to [7, Lemma 2.3], each of these four matrix inequalities (2.5a)–(2.5d) is equivalent to the corresponding Kalman rank conditions (2.3a)–(2.3d) where we used in the last case that  $C_j^* = C_j$  (verifiable by a simple induction). Moreover, the Kalman rank conditions (2.3a)–(2.3d) are equivalent due to Lemma 2.3.  $\square$

This lemma shows that the hypocoercivity index can equally be defined as the smallest integer  $m \in \mathbb{N}_0$  such that  $\tilde{T}_m > 0$ ,  $\tilde{\tilde{T}}_m > 0$ , or  $\widehat{T}_m > 0$ . Hence, it also gives the smallest necessary number of iterated commutators of  $B_H \geq 0$  with the matrix  $B_A$  such that their ranges span all of  $\mathbb{C}^n$  — in the spirit of Hörmander’s hypoellipticity theorem.

As an example we consider two matrices with the same Hermitian part  $B_H = \text{diag}(0, 0, 1, 1)$  such that  $\text{rank}(\ker(B_H)) = 2$  but two different anti-Hermitian parts such that one and, respectively, two iterated commutators are needed:

**Example 2.5.** Consider

$$B^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \text{ such that } B_A^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

$$[\sqrt{B_H^{(1)}}, B_A^{(1)}] = \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix},$$

and

$$B^{(2)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \text{ such that } B_A^{(2)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$[\sqrt{B_H^{(2)}}, B_A^{(2)}] = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hence  $\text{rank}\{\sqrt{B_H^{(1)}}, [\sqrt{B_H^{(1)}}, B_A^{(1)}]\} = 4$ , but  $\text{rank}\{\sqrt{B_H^{(2)}}, [\sqrt{B_H^{(2)}}, B_A^{(2)}]\} = 3$ . Thus, by Lemma 2.3,  $m_{HC}(B^{(1)}) = 1$  and  $m_{HC}(B^{(2)}) = 2$ .

While finite dimensional positive conservative-dissipative matrices are hypocoercive if and only if they have a finite hypocoercivity index, this is not true in the infinite dimensional case:

**Example 2.6.** Consider a block-diagonal “ODE”, with each block of the form

$$E_k = \begin{pmatrix} 0 & 1 & & & \\ -1 & \ddots & \ddots & & \\ & \ddots & \ddots & 1 & \\ & & -1 & 0 & 1 \\ & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{k \times k},$$

$$(E_k)_H = \text{diag}(0, 0, \dots, 0, 1), \quad (E_k)_A = \text{toeplitz}(-1, 0, 1). \quad (2.6)$$

Then, the matrices  $E_k$ ,  $k \in \mathbb{N}$  are positive conservative-dissipative. Evidently,  $(E_k)_H \mathbf{v} = 0$  if and only if  $\mathbf{v} = (v_1, \dots, v_{k-1}, 0)$ . For such a vector,

$$(E_k)_A \mathbf{v} = (v_2, v_3 - v_1, \dots, -v_{k-2}, -v_{k-1}) .$$

A simple inductive argument shows that if  $(E_k)_A \mathbf{v} = \lambda \mathbf{v}$  for any  $\lambda \neq 0$ , then  $\mathbf{v} = 0$ . By Proposition 1.3,  $E_k$  has no eigenvalue on  $i\mathbb{R}$ . Hence, the matrices  $E_k$ ,  $k \in \mathbb{N}$  are hypocoercive, satisfying the estimate

$$\|e^{-E_k t}\|_2 \leq c_k e^{-\lambda_k t} , \quad t \geq 0 , \quad (2.7)$$

for constants  $c_k \geq 1$  and  $\mu_k > 0$ ,  $k \in \mathbb{N}$ . We now make a simple rescaling, defining  $\tilde{E}_k := r_k E_k$  for  $r_k > 0$  to be chosen below, and consider the propagator  $P(t)$  for  $\text{diag}(\tilde{E}_1, \tilde{E}_2, \dots)$ , and let  $\|\cdot\|_2$  be the spectral norm on  $\ell^2$ . By (2.7),

$$\|e^{-\tilde{E}_k}\|_2 = \|e^{-r_k E_k}\|_2 \leq c_k e^{-r_k \mu_k} \leq \frac{1}{e} \quad \text{for } r_k = \frac{1 + \log c_k}{\mu_k} .$$

Therefore, making this choice of  $r_k$ ,  $\|P(1)\|_2 \leq 1/e$ , so that (1.4) is satisfied for  $t_0 = 1$ . Thus this infinite dimensional system is hypocoercive.

We now show that the combined system has an infinite hypocoercivity index. For this we can ignore the scaling and work with the original matrices  $E_k$ . Let  $\mathbf{e}_j$  denote the  $j$ -th standard basis vector in  $\mathbb{C}^k$ . Then,  $(E_k)_H = \mathbf{e}_k \mathbf{e}_k^*$  and  $T_m$  as specified in (2.1) is given by

$$T_m = \sum_{j=0}^m (-1)^j (E_k)_A^j \mathbf{e}_k \mathbf{e}_k^* (E_k)_A^j .$$

For  $m = 1, \dots, k-1$ , it is evident that

$$|\mathbf{e}_m^* (E_k)_A^j \mathbf{e}_k|^2 = 0 \quad \text{for } j < k-m \quad \text{but} \quad |\mathbf{e}_m^* (E_k)_A^{k-m} \mathbf{e}_k|^2 = 1 .$$

Hence,  $m_{HC}(E_k) = k-1$ . Thus, the combined system  $\text{diag}(E_1, E_2, \dots)$  can not have a finite  $m_{HC}$ .

## 2.2. Short time decay of conservative-dissipative ODE systems

Here we shall prove that the hypocoercivity index of a conservative-dissipative ODE system characterizes the decay of its propagator norm for short time. We denote the solution semigroup pertaining to (1.1) by  $P(t) := e^{-Bt} \in \mathbb{C}^{n \times n}$ , and its spectral norm by  $\|P(t)\|_2 := \sup\{\|P(t)\mathbf{x}\|_2 : \|\mathbf{x}\|_2 = 1\}$ , which is also the largest singular value of  $P(t)$ . Its short time decay is related to the hypocoercivity index as follows:

**Theorem 2.7.** *Let the ODE system (1.1) be conservative-dissipative with (positive conservative-dissipative) matrix  $B \in \mathbb{C}^{n \times n}$ .*

(a) The (positive conservative-dissipative) matrix  $B$  is hypocoercive (with hypocoercivity index  $m_{HC} \in \mathbb{N}_0$ ) if and only if

$$\|e^{-Bt}\|_2 = 1 - ct^a + \mathcal{O}(t^{a+1}) \quad \text{for } t \in [0, \epsilon), \quad (2.8)$$

for some  $a, c, \epsilon > 0$ . In this case, necessarily  $a = 2m_{HC} + 1$ .

(b) Moreover, for  $m_{HC} \geq 1$ , the optimal multiplicative factor  $c$  in (2.8) is given by

$$\begin{aligned} c &= \frac{1}{(2m_{HC} + 1)! \binom{2m_{HC}}{m_{HC}}} \min_{\mathbf{x} \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}\|=1} \langle \mathbf{x}, (B^*)^{m_{HC}} B_H B^{m_{HC}} \mathbf{x} \rangle \\ &= \frac{1}{(2m_{HC} + 1)! \binom{2m_{HC}}{m_{HC}}} \min_{\mathbf{x} \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}\|=1} \langle \mathbf{x}, (B_A^*)^{m_{HC}} B_H B_A^{m_{HC}} \mathbf{x} \rangle, \end{aligned} \quad (2.9)$$

and for  $m_{HC} = 0$  we have  $c = \min_{\|\mathbf{x}\|=1} \langle \mathbf{x}, B_H \mathbf{x} \rangle$ .

**Proof.** Since the proof for the coercive case, i.e.  $m_{HC} = 0$ , is trivial, we shall confine ourselves now to  $m_{HC} \geq 1$ :

Part (a): For sufficiently small time  $t_0 > 0$ , there exists a real analytic function  $\Phi : [0, t_0] \rightarrow \mathbb{R}$  such that

$$\|P(t)\|_2 = \Phi(t) \quad \text{for all } t \in [0, t_0], \quad (2.10)$$

e.g. see [15, Lemma 1]. Alternatively, the statement can be derived from [14, Part II.§6] or [12, Theorem 4.3.17].

**For the forward direction** we assume that  $B$  has a finite HC-index  $m_{HC} \in \mathbb{N}$ . Hence we are left to prove that the Taylor expansion of  $\|P(t)\|_2$  has the form (2.8). This proof will be split into two separate parts, the lower and the (technically more subtle) upper bound.

Lower bound: First, we shall prove that there exists  $c_1 > 0$  such that

$$\|P(t)\|_2 \geq 1 - c_1 t^a + \mathcal{O}(t^{a+1}) \quad \text{as } t \rightarrow 0^+ \quad (2.11a)$$

or, equivalently,

$$\|P(t)\|_2^2 \geq 1 - 2c_1 t^a + \mathcal{O}(t^{a+1}) \quad \text{as } t \rightarrow 0^+ \quad (2.11b)$$

holds with  $a = 2m_{HC} + 1$ .<sup>5</sup> To this end suppose  $\mathbf{x}_0$  is any unit vector such that for some  $m \in \mathbb{N}$  (not necessarily  $m = m_{HC}$ ),

$$\mathbf{x}_0 \in \ker \left( \sum_{j=0}^{m-1} (B^*)^j B_H B^j \right) = \ker(\tilde{T}_{m-1}), \quad (2.12)$$

<sup>5</sup> In inequalities, such as (2.11), the Landau symbol  $\mathcal{O}$  is used in the sense that there exist constants  $\epsilon, M > 0$  such that  $\|P(t)\|_2 \geq 1 - c_1 t^a - M t^{a+1}$  for  $t \in [0, \epsilon)$ .

which is equivalent to

$$\|\sqrt{B_H} B^j \mathbf{x}_0\|_2 = 0 \quad \text{for } 0 \leq j \leq m-1. \quad (2.13)$$

If such an  $\mathbf{x}_0$  exists, we will show for the corresponding trajectory that

$$\|P(t)\mathbf{x}_0\|_2^2 = 1 - t^{2m+1} \frac{2}{(2m+1)!} \binom{2m}{m} \langle \mathbf{x}_0, (B^*)^m B_H B^m \mathbf{x}_0 \rangle + \mathcal{O}(t^{2m+2}). \quad (2.14)$$

Note that our hypotheses allow the possibility that  $\langle \mathbf{x}_0, (B^*)^m B_H B^m \mathbf{x}_0 \rangle = 0$  in which case we have simply  $\|P(t)\mathbf{x}_0\|_2^2 = 1 + \mathcal{O}(t^{2m+2})$ .

Now take  $\mathbf{x}_0$  as in Equation (2.12) with  $\|\mathbf{x}_0\| = 1$ , and define  $q(t) := \|P(t)\mathbf{x}_0\|^2$ . Standard theory gives that (see also Equation (1.6))

$$q'(t) = -2\|\sqrt{B_H} P(t)\mathbf{x}_0\|^2 = -2\langle \sqrt{B_H} P(t)\mathbf{x}_0, \sqrt{B_H} P(t)\mathbf{x}_0 \rangle. \quad (2.15)$$

Using the assumption on  $\mathbf{x}_0$ , we have that  $\sqrt{B_H} P(t)\mathbf{x}_0 = \sum_{k=m}^{\infty} \frac{t^k}{k!} \sqrt{B_H} (-B)^k \mathbf{x}_0$ . Substituting this in (2.15) gives

$$q'(t) = -2 \frac{t^{2m}}{m! \cdot m!} \langle \sqrt{B_H} B^m \mathbf{x}_0, \sqrt{B_H} B^m \mathbf{x}_0 \rangle + \sum_{k=2m+1}^{\infty} u_k t^k \quad (2.16)$$

for some scalars  $u_k$ . Since this sequence converges absolutely (The expression in (2.15) is real analytic), we can integrate this expression term-wise to obtain  $q(t)$ . Using that  $q(0) = \|\mathbf{x}_0\|^2 = 1$ , we find

$$q(t) = 1 - \frac{2}{2m+1} \frac{t^{2m+1}}{m! \cdot m!} \langle \sqrt{B_H} B^m \mathbf{x}_0, \sqrt{B_H} B^m \mathbf{x}_0 \rangle + \sum_{k=2m+1}^{\infty} \frac{u_k}{k+1} t^{k+1}, \quad (2.17)$$

which shows (2.14).

Now, since the positive conservative-dissipative matrix  $B$  is hypocoercive with HC-index  $m_{HC} \in \mathbb{N}_0$ , there exists a normalized vector  $\mathbf{x}_0$  such that

$$\|\sqrt{B_H} B^j \mathbf{x}_0\|_2 = 0 \quad \text{for } 0 \leq j \leq m_{HC} - 1, \quad \text{and } \sqrt{B_H} B^{m_{HC}} \mathbf{x}_0 \neq 0, \quad (2.18)$$

but none that would satisfy instead also  $\sqrt{B_H} B^{m_{HC}} \mathbf{x}_0 = 0$ . Hence, for such  $\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1})$  the identity (2.14) holds with  $m = m_{HC}$ . Finally, taking the supremum over all initial conditions  $\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1})$ ,  $\|\mathbf{x}_0\|_2 = 1$  (which is a closed set) yields the estimates (2.11a) and (2.11b) with  $a = 2m_{HC} + 1$  and

$$\begin{aligned} c_1 &:= \frac{1}{2} \min_{\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}_0\|_2=1} \frac{2}{(2m_{HC}+1)!} \binom{2m_{HC}}{m_{HC}} \langle \mathbf{x}_0, (B^*)^{m_{HC}} B_H B^{m_{HC}} \mathbf{x}_0 \rangle \\ &= \frac{1}{(2m_{HC}+1)!} \binom{2m_{HC}}{m_{HC}} \min_{\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}_0\|_2=1} \langle \mathbf{x}_0, (B_A^*)^{m_{HC}} B_H B_A^{m_{HC}} \mathbf{x}_0 \rangle. \end{aligned} \quad (2.19)$$

In the last step we used

$$B^{m_{HC}} \mathbf{x}_0 = B_A B^{m_{HC}-1} \mathbf{x}_0 = \dots = B_A^{m_{HC}} \mathbf{x}_0, \quad (2.20)$$

since all other terms vanish due to  $\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1})$ . Due to (2.18),  $c_1 > 0$ .

In the proof of Part (b) we shall actually improve this constant by the factor  $(\frac{2m_{HC}}{m_{HC}})^{-2}$ , see Lemma A.3.

Upper bound: Second, we prove that there exists  $c_2 > 0$  such that

$$\|P(t)\|_2 \leq 1 - c_2 t^a + \mathcal{O}(t^{a+1}) \quad \text{as } t \rightarrow 0^+ \quad (2.21a)$$

or, equivalently,

$$\|P(t)\|_2^2 \leq 1 - 2c_2 t^a + \mathcal{O}(t^{a+1}) \quad \text{as } t \rightarrow 0^+ \quad (2.21b)$$

holds with  $a = 2m_{HC} + 1$ .<sup>6</sup> Here, we consider the case  $m_{HC} = 1$  whereas the general case  $m_{HC} \in \mathbb{N}$  is treated in Appendix A.

If matrix  $B$  has hypocoercivity index  $m_{HC} = 1$ , there exists  $\kappa > 0$  such that

$$\tilde{\tilde{T}}_{m_{HC}} = \sum_{j=0}^{m_{HC}} (B^*)^j B_H B^j = B_H + B^* B_H B \geq \kappa I > 0. \quad (2.22)$$

Since  $m_{HC} = 1$ ,  $B^* B_H B$  is positive definite on  $\ker(B_H)$ . For  $\mathbf{x} \in \mathbb{C}^n$  with  $\|\mathbf{x}\| = 1$ , we define

$$\lambda_{\mathbf{x}} := \langle \mathbf{x}, B_H \mathbf{x} \rangle \geq 0, \quad \mu_{\mathbf{x}} := \langle \mathbf{x}, B^* B_H B \mathbf{x} \rangle \geq 0, \quad \text{such that } \lambda_{\mathbf{x}} + \mu_{\mathbf{x}} \geq \kappa > 0. \quad (2.23)$$

Note that  $\|P(t)\mathbf{x}_0\|_2^2 = \langle \mathbf{x}_0, Q(t)\mathbf{x}_0 \rangle$  where

$$Q(t) := P^*(t) P(t) = e^{-B^* t} e^{-B t} = \sum_{j=0}^{\infty} \frac{t^j}{j!} \sum_{k=0}^j \binom{j}{k} (-B^*)^k (-B)^{j-k}. \quad (2.24)$$

Since  $\left\| \sum_{k=0}^j \binom{j}{k} (-B^*)^k (-B)^{j-k} \right\|_2 \leq (2\|B\|_2)^j$ , the Taylor series for the matrix family  $Q(t)$  converges uniformly on bounded  $t$ -intervals. Hence we have

$$\|P(t)\|_2^2 = \|Q(t)\|_2 = \|Q_j(t)\|_2 + \mathcal{O}(t^{j+1}), \quad (2.25)$$

where  $Q_j(t)$  denotes the Taylor expansion (2.24), but truncated after the  $t^j$ -term. Consider

$$Q_3(t) = \sum_{j=0}^3 \frac{t^j}{j!} U_j = I + t U_1 + \frac{t^2}{2!} U_2 + \frac{t^3}{3!} U_3. \quad (2.26)$$

<sup>6</sup> In inequalities, such as (2.21), the Landau symbol  $\mathcal{O}$  is used in the sense that there exist constants  $\epsilon, M > 0$  such that  $\|P(t)\|_2 \leq 1 - c_2 t^a + M t^{a+1}$  for  $t \in [0, \epsilon]$ .

Then, its spectral norm satisfies

$$\|Q_3(t)\|_2 := \sup_{\|\mathbf{x}_0\|=1} \|Q_3(t)\mathbf{x}_0\|_2 = \sup_{\|\mathbf{x}_0\|=1} \langle \mathbf{x}_0, Q_3(t)\mathbf{x}_0 \rangle, \quad (2.27)$$

for sufficiently small  $t \geq 0$ . The latter identity holds, since the self-adjoint matrix  $Q_3(t)$  satisfies  $Q_3(0) = I$ .

Let  $\mathbf{x}$  be a unit vector, we estimate each expression  $\langle \mathbf{x}, U_j \mathbf{x} \rangle$ ,  $j = 1, 2, 3$  separately as

$$\langle \mathbf{x}, U_1 \mathbf{x} \rangle = -2\langle \mathbf{x}, B_H \mathbf{x} \rangle = -2\lambda_{\mathbf{x}} \quad \text{using} \quad \lambda_{\mathbf{x}} = \langle \mathbf{x}, B_H \mathbf{x} \rangle \geq 0.$$

Next, since  $U_2 = 2(B^* B_H + B_H B)$ , so that by the Cauchy–Schwarz inequality,

$$\langle \mathbf{x}, U_2 \mathbf{x} \rangle = 2\langle \mathbf{x}, (B^* B_H + B_H B) \mathbf{x} \rangle \leq 4\|\sqrt{B_H} \mathbf{x}\| \|\sqrt{B_H} B \mathbf{x}\| = 4\sqrt{\lambda_{\mathbf{x}}} \sqrt{\mu_{\mathbf{x}}}$$

using (2.23). In the same way, we derive for  $U_3 = -2(B_H B^2 + 2B^* B_H B + (B^*)^2 B_H)$  that

$$\begin{aligned} \langle \mathbf{x}, U_3 \mathbf{x} \rangle &\leq -4\langle \mathbf{x}, B^* B_H B \mathbf{x} \rangle - 2\langle \mathbf{x}, (B_H B^2 + (B^*)^2 B_H) \mathbf{x} \rangle \\ &\leq -4\mu_{\mathbf{x}} + 4\|\sqrt{B_H} \mathbf{x}\| \|\sqrt{B_H} B^2 \mathbf{x}\| = -4\mu_{\mathbf{x}} + 4\sqrt{\lambda_{\mathbf{x}}} \|\sqrt{B_H} B^2 \mathbf{x}\|. \end{aligned}$$

Altogether then,

$$\begin{aligned} \langle \mathbf{x}, Q_3(t) \mathbf{x} \rangle &\leq 1 - 2\lambda_{\mathbf{x}} t + 2\sqrt{\lambda_{\mathbf{x}}} \sqrt{\mu_{\mathbf{x}}} t^2 - \frac{2}{3}\mu_{\mathbf{x}} t^3 + \frac{2}{3}\sqrt{\lambda_{\mathbf{x}}} \|\sqrt{B_H} B^2\| t^3 \\ &= 1 - 2t \left( \sqrt{\lambda_{\mathbf{x}}} - \frac{\sqrt{\mu_{\mathbf{x}}}}{2} t \right)^2 - \frac{1}{6}\mu_{\mathbf{x}} t^3 + \frac{2}{3}\sqrt{\lambda_{\mathbf{x}}} \|\sqrt{B_H} B^2\| t^3. \end{aligned} \quad (2.28)$$

To estimate (2.28), we will distinguish two cases for  $\sqrt{\lambda_{\mathbf{x}}} \in [0, \|\sqrt{B_H}\|]$ : If  $\sqrt{\lambda_{\mathbf{x}}}$  is sufficiently small we will use the third term in (2.28) to compensate the non-negative fourth term in (2.28), while in the other case we will use the second term to do so.

Case a: If

$$\sqrt{\lambda_{\mathbf{x}}} \leq \alpha \quad \text{with } \alpha := \min \left( 1, \frac{\kappa}{2(1 + 4\|\sqrt{B_H} B^2\|)} \right) > 0, \quad (2.29)$$

then the following estimates hold: Due to (2.23) and  $\sqrt{\lambda_{\mathbf{x}}} \leq 1$  we deduce that  $\mu_{\mathbf{x}} \geq \kappa - \lambda_{\mathbf{x}} \geq \kappa - \sqrt{\lambda_{\mathbf{x}}}$  which implies that

$$\langle \mathbf{x}, Q_3(t) \mathbf{x} \rangle \leq 1 - \frac{1}{6}\mu_{\mathbf{x}} t^3 + \frac{2}{3}\sqrt{\lambda_{\mathbf{x}}} \|\sqrt{B_H} B^2\| t^3 \leq 1 - \frac{\kappa}{12} t^3, \quad \text{for } t \geq 0. \quad (2.30)$$

Case b: If  $\alpha \leq \sqrt{\lambda_{\mathbf{x}}} \leq \|\sqrt{B_H}\|$  then we restrict the time interval to obtain a similar estimate. Using  $\alpha$  from (2.29), define

$$t_1 := \frac{\alpha}{\sqrt{\|B^* B_H B\|}}. \quad (2.31)$$

For  $t \in [0, t_1]$ , we use  $0 \leq \mu_{\mathbf{x}} \leq \|B^* B_H B\|$  to derive

$$\sqrt{\lambda_{\mathbf{x}}} - \frac{\sqrt{\mu_{\mathbf{x}}}}{2} t \geq \alpha - \frac{\sqrt{\mu_{\mathbf{x}}}}{2} t_1 \geq \frac{\alpha}{2} > 0.$$

Consequently, the expression in (2.28) can be estimated as

$$\begin{aligned} \langle \mathbf{x}, Q_3(t)\mathbf{x} \rangle &\leq 1 - \frac{\alpha^2}{2} t + \frac{2}{3} \|\sqrt{B_H}\| \|\sqrt{B_H} B^2\| t^3 \\ &\leq 1 - t^3 \left( \frac{\alpha^2}{2} \frac{1}{t_1^2} - \frac{2}{3} \|\sqrt{B_H}\| \|\sqrt{B_H} B^2\| \right). \end{aligned}$$

Then, there exists  $t_* \in (0, t_1]$  such that

$$\langle \mathbf{x}, Q_3(t)\mathbf{x} \rangle \leq 1 - \frac{\kappa}{12} t^3, \quad \text{for } t \in [0, t_*]. \quad (2.32)$$

Together the estimates (2.30) and (2.32) as well as (2.25), (2.27) prove the upper bound (2.21b) with  $a = 3$  for  $m_{HC} = 1$ , e.g. with  $c_2 := \kappa/24$ . For  $m_{HC} \geq 2$ , see Lemma A.1 in Appendix A.

To sum up, due to (2.11a) and Lemma A.1, there exist constants  $c_1, c_2 > 0$  such that  $a = 2m_{HC} + 1$  and

$$1 - c_1 t^a + \mathcal{O}(t^{a+1}) \leq \|P(t)\|_2 \leq 1 - c_2 t^a + \mathcal{O}(t^{a+1}) \quad \text{as } t \rightarrow 0^+, \quad (2.33)$$

e.g. choosing  $c_1$  as in (2.19) and  $c_2$  as in Lemma A.1, respectively. Moreover, due to (2.10),  $\|P(t)\|_2$  is analytic on  $[0, t_0]$ . This implies that the propagator norm satisfies (2.8) for some  $c \in \mathbb{R}$ , which satisfies  $0 < c_2 \leq c \leq c_1$ .

**For the reverse direction** suppose that (2.8) is satisfied for some constant  $c > 0$  and an exponent  $a > 0$ . Then, as we shall show, for some least finite value of  $m \in \mathbb{N}_0$ ,  $T_m = \sum_{j=0}^m B_A^j B_H (B_A^*)^j > 0$  has to hold, or equivalently,

$$\tilde{T}_m = \sum_{j=0}^m (B^*)^j B_H B^j > 0.$$

Otherwise, for arbitrarily large values of  $\tilde{m} \in \mathbb{N}$ , we could find unit vectors  $\mathbf{x}_0$  such that  $\tilde{T}_{\tilde{m}-1} \mathbf{x}_0 = 0$ , and then by the first part of this proof, we would have for such a vector  $\|P(t)\mathbf{x}_0\|_2^2 \geq 1 - c_{\tilde{m}} t^{2\tilde{m}+1} + \mathcal{O}(t^{2\tilde{m}+2})$  with some  $c_{\tilde{m}} \geq 0$ , because of (2.17). But for sufficiently large  $\tilde{m}$ , this is incompatible with (2.8). Thus we conclude that, whenever (2.8) is valid for any  $a > 0$ ,  $B$  has a finite hypocoercivity index  $m_{HC} \in \mathbb{N}_0$ , and then necessarily,  $a \in 2\mathbb{N}_0 + 1$ .

Part (b): So far we have proved that the propagator norm satisfies (2.33), e.g. choosing  $c_1$  as in (2.19) and  $c_2$  as the lower bound of  $c$  in (A.2), respectively. Using the improved upper bound  $c_1$  in (A.34) and lower bound  $c_2$  in (A.2) for the multiplicative constant  $c$  we realize that  $c_2 = c = c_1$  such that (2.9) holds. In the final identity of (2.9) we used again (2.20) to reveal that  $c$  is proportional to  $B_A^{2m_{HC}}$ . This finishes the proof.  $\square$

For  $\epsilon$ -dependent ODE systems of the form (1.5), Theorem 2.7(b) implies the following result:

**Corollary 2.8.** Consider the  $\epsilon$ -dependent ODE (1.5) with system matrix  $B = \epsilon A + C$  where  $\epsilon \in \mathbb{R}$ . If  $B = \epsilon A + C$  is hypocoercive for  $\epsilon \neq 0$ , then the coefficient  $c = c_\epsilon$  in the Taylor expansion of the propagator norm (2.8) satisfies

$$c = c_\epsilon = \epsilon^{2m_{HC}} \frac{1}{(2m_{HC} + 1)! \binom{2m_{HC}}{m_{HC}}} \min_{\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}_0\|_2=1} \langle \mathbf{x}_0, (A^*)^{m_{HC}} C A^{m_{HC}} \mathbf{x}_0 \rangle. \quad (2.34)$$

**Remark 2.9.** As already briefly mentioned in Example 2 of [4], the hypocoercivity index  $m_{HC}$  is upper semicontinuous w.r.t. the matrix  $B$ : An arbitrarily small perturbation of  $B$  can lower but not increase the index. This is consistent with the result in (2.8): All Taylor coefficients are there of course continuous in  $B$ . But a small perturbation of  $B$  may lower, but not increase the number  $a$  of the first non-vanishing monomial in this Taylor series (beyond the constant 1).

**Remark 2.10.** Note that the leading exponent in (2.8) can only be odd. This is related to the local behavior of trajectories that decay the worst (in the vicinity of a stationary point  $t_0$  of  $\|\mathbf{x}(t)\|$ ), see the yellow curve in Fig. 1, left. Since the system (1.1) is assumed to be conservative-dissipative, such a trajectory, of course, cannot behave locally like  $1 - c(t - t_0)^a$  with  $a$  even.

**Remark 2.11.** Special cases of the above theorem were pointed out to us by Laurent Miclo: In §1 of [17] the short time decay behavior of the Goldstein-Taylor model (a linear transport equation with relaxation term) was determined as  $1 - \frac{t^3}{3} + o(t^3)$ . Actually, this model is a PDE. But since it is considered on a torus in  $x$ , each of its spatial Fourier modes (except of the 0-mode) satisfies a conservative-dissipative ODE system with hypocoercivity index 1 (see [1] for details of this modal decomposition). Hence, mode by mode, the result from [17] is an example for Theorem 2.7. For closely related BGK-models with hypocoercivity index 2 and 3 we refer to [2].

In [11] the short time decay behavior of a kinetic Fokker-Planck equation on the torus in  $x$  was computed as  $1 - \frac{t^3}{12} + o(t^3)$ . Again, in Fourier space and by using a Hermite function basis in velocity, this model can be written as an (infinite dimensional) conservative-dissipative system with hypocoercivity index 1 (see §2.1 of [11]). In that paper it was also mentioned that the decay exponent in (2.8) can be seen as some “order of hypocoercivity” of the generator.

For degenerate Fokker-Planck equations, the hypocoercivity index can also be related to the regularization rate for short times: In [18, Theorem A.12] the regularization of initial data from a weighted  $L^2$  space into a weighted  $H^1$  space is derived, and in [18, Theorem A.15], [8, Theorem 4.8] it is generalized to entropy functionals and their corresponding Fisher informations. In all these cases the regularization rate is  $t^{-a}$  with  $a = 2m_{HC} + 1$  (somewhat related to Theorem 2.7 above).

By definition, the propagator norm of an ODE (1.1) is given as the envelope of the norm of a family of solutions, see e.g. [6] and Fig. 1. But, maybe surprisingly, even its precise short time behavior is *not* given by the norm of any specific solution. This is illustrated in the following example.

**Example 2.12.** We consider ODE (1.1) with matrix

$$B = \begin{pmatrix} 1 & -3/10 \\ 3/10 & 0 \end{pmatrix}. \quad (2.35)$$

The eigenvalues of  $B$  are  $\lambda_1 = 1/10$  and  $\lambda_2 = 9/10$ , and the eigenvalues of  $B_H$  are 0 and 1. Thus, matrix  $B$  is hypocoercive with hypocoercivity index  $m_{HC} = 1$ . Following (the first part of) the proof of Theorem 2.7, solutions starting in  $\mathbf{x}_0$  satisfying (2.18) are used to establish the desired lower bound of the propagator norm. The kernel of  $B_H$  is one-dimensional and it is spanned by  $\mathbf{x}_0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ . The solution of (1.1) with initial condition  $\mathbf{x}(0) = \mathbf{x}_0$  is given by

$$\mathbf{x}(t) = \frac{1}{8} \begin{pmatrix} 3 & -3 \\ 9 & -1 \end{pmatrix} \begin{pmatrix} e^{-t/10} \\ e^{-9t/10} \end{pmatrix}$$

and its squared norm satisfies

$$\|\mathbf{x}(t)\|_2^2 = \frac{45}{32} e^{-t/5} - \frac{9}{16} e^{-t} + \frac{5}{32} e^{-9t/5} \sim 1 - 0.06 t^3 + \mathcal{O}(t^4) \quad \text{for } t \rightarrow 0^+.$$

By contrast, due to [6, Proposition 4.2], the squared propagator norm satisfies

$$\begin{aligned} \|e^{-Bt}\|_2^2 &\leq e^{-t} \frac{1}{16} \left( \sqrt{(25 \cosh(8t/10) - 9)^2 - 16^2} + 25 \cosh(8t/10) - 9 \right) \\ &\sim 1 - 0.015 t^3 + \mathcal{O}(t^4) \quad \text{for } t \rightarrow 0^+. \end{aligned} \tag{2.36}$$

Thus the propagator norm decays slower than the solution starting at the vector  $\mathbf{x}_0$  which satisfies (2.18) with  $m_{HC} = 1$ , see also Fig. 1.

In fact, the sharp constant  $c = 0.015$  in (2.36) cannot be obtained by any single trajectory, but rather by a family of trajectories starting at the one-parameter family of normalized initial conditions,  $\mathbf{x}_\tau$ ,  $\tau \geq 0$  emerging from  $\mathbf{x}_0$ : More concretely, using  $\mathbf{x}_\tau := \left( -\frac{3\tau}{20}, 1 \right)^\top / \sqrt{1 + \frac{9\tau^2}{400}}$  (as constructed in Lemma A.4 below) yields

$$\|e^{-Bt} \mathbf{x}_\tau\|_2^2 \sim 1 - 0.015 t^3 + \mathcal{O}(t^4) \quad \text{for } t \rightarrow 0^+.$$

Note that we used here the initial conditions  $\mathbf{x}_\tau$  (i.e.  $\tau = t$ ), and that we have  $\frac{d}{d\tau} \mathbf{x}_\tau(\tau = 0) = \left( -\frac{3}{20}, 0 \right)^\top = \frac{1}{2} B \mathbf{x}_0$ .

### 2.3. Numerical illustration of the short time decay and the waiting time $t_0$

Next we shall illustrate the decay behavior on two examples of dimension  $n = 4$ . In particular we shall consider the  $\epsilon$ -dependence of the three phases, the asymptotic phase close to  $t = 0$  (as characterized in Theorem 2.7), the intermediate phase (characterized by the waiting time  $t_0$  and the exponential decay for large time (see (1.3)).

**Example 2.13.** Consider the matrix family  $B_\epsilon := \epsilon A + C$ ,  $\epsilon \neq 0$  with

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}, \quad C = \text{diag}(0, 0, 1, 1), \tag{2.37}$$

which satisfies  $m_{HC}(B_\epsilon) = 1$ .

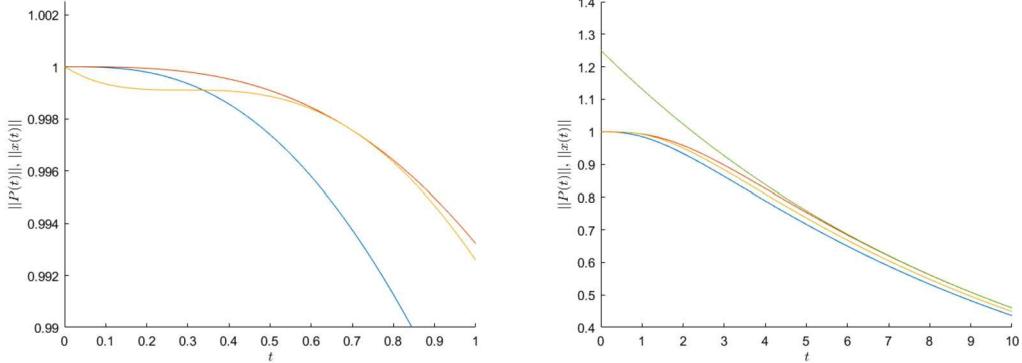


Fig. 1. Evolution of (1.1) with matrix  $B$  from (2.35): Comparison between the propagator norm (red line), its upper exponential envelope  $1.25 \exp(-t/10)$  (green line), and the norm of the solution with initial condition  $x(0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  (blue line) and that with initial condition  $x(0) = \begin{pmatrix} -0.1 \\ \sqrt{0.99} \end{pmatrix}$  (yellow line), all plotted on two time scales. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

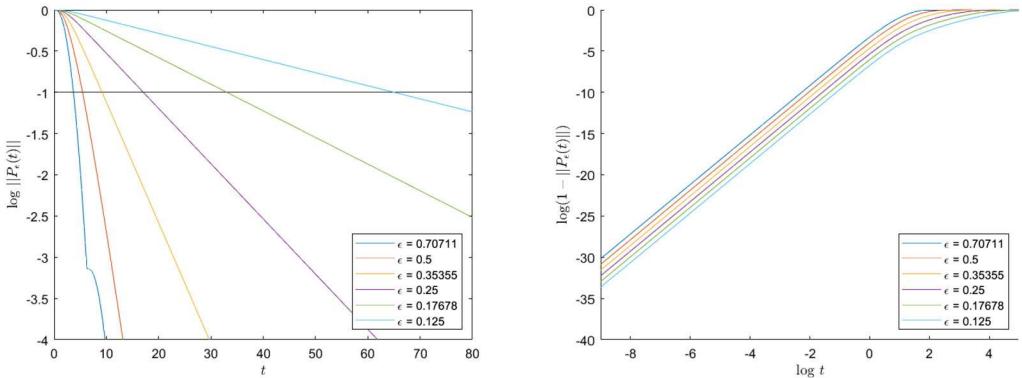


Fig. 2. The decay of  $\|P_\epsilon(t)\|_2$  is given for six values of  $\epsilon$ . Left: For  $t$  away from 0, this semigroup decays almost exponentially. With the logarithmic scale used here, the horizontal (black) line corresponds to  $1/e$ . The waiting times (defined as intersection with the line  $1/e$ ) behave like  $\mathcal{O}(\epsilon^{-2})$ . We remark that the kink in the leftmost (blue) curve is *not* a numerical artifact. Right: This double logarithmic plot shows  $1 - \|P_\epsilon(t)\|_2 \sim c_\epsilon t^3$  for small time, more precisely for  $t \in [e^{-9}, e^5]$ . The curves have slope 3, and  $c_\epsilon = \tilde{c} \epsilon^2$ . The plot also shows the quite sharp transition from the initial algebraic behavior  $1 - \tilde{c} \epsilon^2 t^3$  to the exponential behavior  $c_\epsilon^* e^{-\tilde{\mu} \epsilon^2 t}$ .

Fig. 2 shows the spectral norm of the semigroup  $P_\epsilon(t) := e^{-B_\epsilon t}$  as a function of time and for several values of  $\epsilon$  (differing from each other by the factor  $\sqrt{2}$ ). The numerically observed waiting times (to decrease the solution norm by the factor  $1/e$ ) are very close to  $t_0 \sim 1/\epsilon^2$ . They increase by a factor of about 2 when passing from  $\epsilon$  to  $\sqrt{2}\epsilon$ , at least in the asymptotic regime  $\epsilon \rightarrow 0$ . A close-up of the same figure around  $t = 0$  shows that the asymptotic behavior of the semigroup norm is like  $\|P_\epsilon(t)\|_2 \sim 1 - c_\epsilon t^3$  with  $c_\epsilon = \tilde{c} \epsilon^2$ , and by recalling that  $m_{HC}(B_\epsilon) = 1$  implies  $a = 3$  in (2.8). Following (2.34), the multiplicative factor is  $\tilde{c} = \frac{1}{12}$  (compare with e.g. [6]).

**Example 2.14.** This example is analogous to Example 2.13, but for the matrix family  $B_\epsilon := \epsilon A + C$ ,  $\epsilon \neq 0$  with

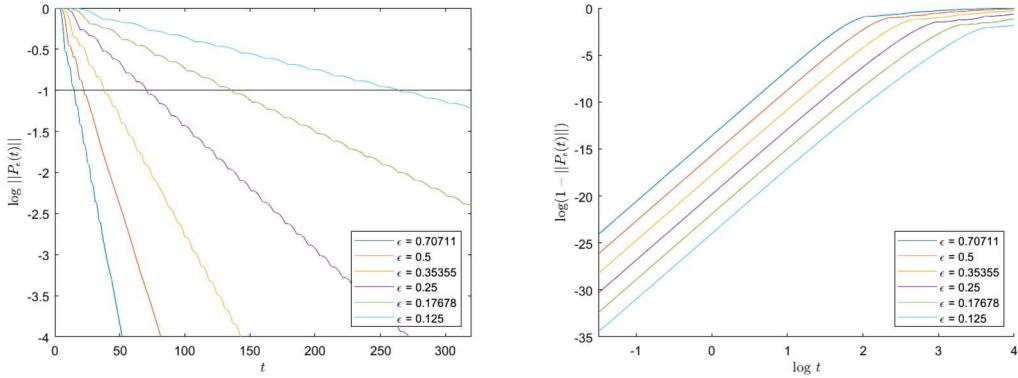


Fig. 3. The decay of  $\|P_\epsilon(t)\|_2$  is given for six values of  $\epsilon$ . Left: For  $t$  away from 0, this semigroup decays almost exponentially. With the logarithmic scale used here, the horizontal (black) line corresponds to  $1/e$ . The waiting times (defined as intersection with the line  $1/e$ ) behave like  $\mathcal{O}(\epsilon^{-2})$ . Right: This double logarithmic plot shows  $1 - \|P_\epsilon(t)\|_2 \sim c_\epsilon t^7$  for small time, more precisely for  $t \in [e^{-1.5}, e^4]$ . The curves have slope 7, and  $c_\epsilon = \tilde{c}\epsilon^6$ . The plot also shows the quite sharp transition from the initial algebraic behavior  $1 - \tilde{c}\epsilon^6 t^7$  to the exponential behavior  $c_\epsilon^* e^{-\tilde{\mu}\epsilon^2 t}$ .

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{pmatrix}, \quad C = \text{diag}(0, 0, 0, 1), \quad (2.38)$$

which satisfies  $m_{HC}(B_\epsilon) = 3$ .

Fig. 3 shows the spectral norm of the semigroup  $P_\epsilon(t) := e^{-B_\epsilon t}$ . The numerically observed waiting times (to decrease the solution norm by the factor  $1/e$ ) are very close to  $t_0 \sim 4/\epsilon^2$ . A close-up around  $t = 0$  shows that the asymptotic behavior of the semigroup norm is like  $\|P_\epsilon(t)\|_2 \sim 1 - c_\epsilon t^7$  with  $c_\epsilon = \tilde{c}\epsilon^6$ , and by recalling that  $m_{HC}(B_\epsilon) = 3$  implies  $a = 7$  in (2.8). Following (2.34), the multiplicative factor is  $\tilde{c} = \frac{1}{100800}$ .

## Data availability

No data was used for the research described in the article.

## Acknowledgment

The first author (FA) was supported by the Austrian Science Fund (FWF) via the FWF-funded SFB # F65. The second author (AA) was supported by the Austrian Science Fund (FWF), partially via the FWF-doctoral school “Dissipation and dispersion in non-linear partial differential equations” (# W1245) and the FWF-funded SFB # F65. The third author (EC) was partially supported by U.S. NSF grant DMS 2055282.

The second author (AA) acknowledges fruitful discussions with Laurent Miclo on the possible connection between the hypocoercivity index and the short time decay rate, which eventually led to Theorem 2.7. The authors are grateful to the anonymous referee, whose suggestions helped to simplify the proof of Theorem 2.7.

## Appendix A. Auxiliary results to prove short time decay of propagator norm

To finish the proof of Theorem 2.7, we shall prove the following upper bound for the propagator norm:

**Lemma A.1.** *Let the ODE system (1.1) be conservative-dissipative, and let the system matrix  $B$  be (hypo)coercive with hypocoercivity index  $m_{HC} \in \mathbb{N}$ . Then, there exist constants  $c_2, M, t_2 > 0$  such that*

$$\|e^{-Bt}\|_2 \leq 1 - c_2 t^a + M t^{a+1} \quad \text{for } t \in [0, t_2], \quad (\text{A.1})$$

where  $a = 2m_{HC} + 1$ . Moreover, the multiplicative factor  $c$  in (2.8) satisfies

$$c \geq c_2 \quad \text{with } c_2 := \frac{1}{(2m_{HC} + 1)! \binom{2m_{HC}}{m_{HC}}} \min_{\mathbf{x} \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}\|=1} \langle \mathbf{x}, (B^*)^{m_{HC}} B_H B^{m_{HC}} \mathbf{x} \rangle. \quad (\text{A.2})$$

**Proof of Lemma A.1.** First we note that the hypocoercivity of  $B$  implies  $B_H \neq 0$  due to (1.8). Following the proof of Theorem 2.7, we consider  $\|e^{-Bt}\|_2^2 = \lambda_{\max}(Q(t))$  for small  $t > 0$ , where

$$Q(t) = e^{-B^*t} e^{-Bt} = \sum_{j=0}^{\infty} \frac{t^j}{j!} U_j \quad (\text{A.3})$$

with

$$U_j = (-1)^j \sum_{k=0}^j \binom{j}{k} (B^*)^k B^{j-k}, \quad \text{satisfying } \|U_j\|_2 \leq (2\|B\|_2)^j, \quad j \in \mathbb{N}_0. \quad (\text{A.4})$$

To compute  $\lambda_{\max}(Q(t)) = \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^* Q(t) \mathbf{x}$ , we consider the  $t$ -dependent function  $g(\mathbf{x}; t) := \mathbf{x}^* Q(t) \mathbf{x} - 1$  with  $\mathbf{x}$  in the sphere  $\mathcal{S} := \{\mathbf{x} \in \mathbb{C}^n \mid \|\mathbf{x}\|_2 = 1\}$ . For  $a = 2m_{HC} + 1$ , we denote the Taylor series for  $Q(t)$  and  $g(t)$  truncated after the  $t^a/a!$  term with  $Q_a(t)$  and  $g_a(\mathbf{x}; t)$ , respectively. We recall that  $U_0 = I$  and  $U_1 = -2B_H$ .

Let the matrices  $U_j$ ,  $j \in \mathbb{N}_0$  denote the coefficients of  $t^j/j!$  in the Taylor expansion (2.24) and (A.3), such that

$$U_j := (-1)^j \sum_{k=0}^j \binom{j}{k} (B^*)^k B^{j-k}, \quad j \in \mathbb{N}_0, \quad (\text{A.5})$$

and note that each  $U_j$  is self-adjoint. By Pascal's identity  $\binom{j}{k} = \binom{j-1}{k-1} + \binom{j-1}{k}$  with the usual convention that  $\binom{j}{-1} = \binom{j}{j+1} = 0$ ,

$$\begin{aligned}
& \sum_{k=0}^j \binom{j}{k} (B^*)^k B^{j-k} \\
&= \sum_{k=0}^j \left( \binom{j-1}{k-1} (B^*)^k B^{j-k} + \binom{j-1}{k} (B^*)^k B^{j-k} \right) \\
&= \sum_{k=0}^j \left( \binom{j-1}{k-1} (B^*)^{k-1} B^* B^{j-k} + \binom{j-1}{k} (B^*)^k B B^{j-1-k} \right) \\
&= \sum_{k=0}^{j-1} \binom{j-1}{k} (B^*)^k (B^* + B) B^{j-1-k} = 2 \sum_{k=0}^{j-1} \binom{j-1}{k} (B^*)^k B_H B^{j-1-k}.
\end{aligned}$$

Consequently,

$$U_j = (-1)^j 2 \sum_{k=0}^{j-1} \binom{j-1}{k} (B^*)^k B_H B^{j-1-k}, \quad j \in \mathbb{N}. \quad (\text{A.6})$$

First we outline the strategy of the proof, say for the case  $m_{HC} = 1$ , i.e.  $a = 3$ : If  $\mathbf{x} \in \ker(U_1) = \ker(B_H)$  with  $\|\mathbf{x}\|_2 = 1$ , then (2.18) with  $\mathbf{x}_0 = \mathbf{x}$  holds. Since  $m_{HC} = 1$ ,  $B^* B_H B$  is positive definite on  $\ker(B_H)$ . Consequently,  $\mathbf{x}^* U_2 \mathbf{x} = 0$ ,  $\mathbf{x}^* U_3 \mathbf{x} < 0$ , such that  $g_a(\mathbf{x}; t) = \frac{t^3}{3!} \mathbf{x}^* U_3 \mathbf{x} < 0$  for  $t > 0$ . By contrast, if  $\mathbf{x} \notin \ker U_1$ , we have  $\mathbf{x}^* U_1 \mathbf{x} = -2\mathbf{x}^* B_H \mathbf{x} < 0$ . Hence, for  $\mathbf{x} \notin \ker U_1$ ,

$$g_a(\mathbf{x}; t) = -\hat{c}t + \mathcal{O}(t^2) \leq -\tilde{c}t^3 \quad \text{for } t \rightarrow 0^+$$

follows for some  $\hat{c}, \tilde{c} > 0$  that depend on  $\mathbf{x}$ . Since  $g_a(\mathbf{x}; t)$  depends continuously on  $\mathbf{x}$ , it is possible to combine these two estimates with a constant  $c$  that is independent of  $\mathbf{x} \in \mathcal{S}$ . Since  $(\ker U_1)^c \cap \mathcal{S}$  is not compact, we do not obtain a uniform estimate “automatically”. So, the key aspect is here to obtain a uniform decay estimate for  $\mathbf{x}$  “close to  $\ker U_1$ ”, in the sense that  $-\epsilon \leq \mathbf{x}^* U_1 \mathbf{x} \leq 0$ .

*Step 1.* *Matrices with hypocoercivity index  $m_{HC} = 1$ .* We suppose that matrix  $B$  has hypocoercivity index  $m_{HC} = 1$ , i.e. there exists  $\kappa > 0$  such that

$$\tilde{\tilde{T}}_{m_{HC}} = \sum_{j=0}^{m_{HC}} (B^*)^j B_H B^j = B_H + B^* B_H B \geq \kappa I > 0. \quad (\text{A.7})$$

Since  $m_{HC} = 1$ ,  $B^* B_H B$  is positive definite on  $\ker(B_H)$ . Our goal is to estimate  $g(\mathbf{x}; t)$  on  $\mathcal{S}$ . For  $\mathbf{x} \in \mathcal{S}$ , we define

$$\lambda_{\mathbf{x}} := \langle \mathbf{x}, B_H \mathbf{x} \rangle \geq 0, \quad \mu_{\mathbf{x}} := \langle \mathbf{x}, B^* B_H B \mathbf{x} \rangle \geq 0, \quad \text{such that } \lambda_{\mathbf{x}} + \mu_{\mathbf{x}} \geq \kappa > 0. \quad (\text{A.8})$$

*Step 1a.* Consider  $\boxed{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta}$  where  $\delta \in (0, \kappa)$  will be chosen later and  $\kappa > 0$  such that (A.7) holds. The key idea (to estimate  $g(\mathbf{x}; t)$  for  $t \in [0, 1]$ ) is to collect the terms  $t^j$  of order  $j$  less than  $a = 2m_{HC} + 1 = 3$  in a quadratic form which is non-positive. Therefore, we

use (A.6) and Lemma A.2 with  $U = (-B)^*$ ,  $V = U_1$ ,  $W = -B$  and  $m = m_{HC} - 1 = 0$ , to rewrite  $g(\mathbf{x}; t)$  as

$$\begin{aligned} g(\mathbf{x}; t) &= \mathbf{x}^* \left( \sum_{j=1}^{\infty} \frac{t^j}{j!} U_j \right) \mathbf{x} \\ &= \sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} ((-B)^k \mathbf{x})^* U_1 (-B)^{j-k-1} \mathbf{x} \\ &= t \left( \sum_{k=0}^{\infty} \frac{1}{(k+1)!} t^k ((-B)^k \mathbf{x})^* \right) U_1 \left( \sum_{\ell=0}^{\infty} \frac{1}{(\ell+1)!} t^{\ell} (-B)^{\ell} \mathbf{x} \right) \\ &\quad + \frac{t^3}{3!} \binom{2}{1} \frac{1}{4} ((-B) \mathbf{x})^* U_1 (-B) \mathbf{x} + \sum_{j=4}^{\infty} \frac{t^j}{j!} \sum_{k=1}^{j-2} \binom{j-1}{k} \Delta_{j,k}^{(1)} ((-B)^k \mathbf{x})^* U_1 (-B)^{j-k-1} \mathbf{x}. \end{aligned}$$

The first term is non-positive since  $U_1 = -2B_H \leq 0$ . The second term is retained. The third term can be estimated from above by  $M_4 t^4$  with  $M_4 := \sum_{j=4}^{\infty} \frac{1}{j!} (2\|B\|_2)^j$  by using (A.6), (A.4), and  $\Delta_{j,k}^{(1)} < 1$ . Altogether, we derive the estimate

$$g(\mathbf{x}; t) \leq \frac{1}{3! 2} ((B \mathbf{x})^* U_1 B \mathbf{x}) t^3 + M_4 t^4 = -\frac{\mu_{\mathbf{x}}}{3!} t^3 + M_4 t^4. \quad (\text{A.9})$$

To establish a uniform negative upper bound for  $(B \mathbf{x})^* U_1 B \mathbf{x}$  for  $\mathbf{x} \in \mathcal{S}$  with  $\lambda_{\mathbf{x}} \leq \delta$ , we use (A.8) to deduce  $\mu_{\mathbf{x}} \geq \kappa - \lambda_{\mathbf{x}} \geq \kappa - \delta > 0$  since  $\delta \in (0, \kappa)$ . For  $\delta \in (0, \kappa)$ , define

$$\mu_{\delta} := \min_{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta} \mu_{\mathbf{x}} = \min_{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta} \langle \mathbf{x}, B^* B_H B \mathbf{x} \rangle \quad (\text{A.10})$$

such that  $\mu_{\delta} \geq \kappa - \delta > 0$  and

$$\mu_0 := \lim_{\delta \rightarrow 0} \mu_{\delta} = \min_{\mathbf{x} \in \ker(\tilde{T}_0), \|\mathbf{x}\|=1} \langle \mathbf{x}, B^* B_H B \mathbf{x} \rangle. \quad (\text{A.11})$$

Then, we derive from (A.9) that

$$g(\mathbf{x}; t) \leq -\frac{\mu_{\delta}}{3!} t^3 + M_4 t^4 \quad \text{for all } \mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta \text{ and } t \in [0, 1]. \quad (\text{A.12})$$

Step 1b. Consider  $\boxed{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} > \delta}$  where  $\delta \in (0, \kappa)$  will be chosen later and  $\kappa > 0$  such that (A.7) holds. For  $t \in [0, 1]$ , we deduce

$$g(\mathbf{x}; t) := \mathbf{x}^* \left( \sum_{j=1}^{\infty} \frac{t^j}{j!} U_j \right) \mathbf{x} = t \mathbf{x}^* U_1 \mathbf{x} + \sum_{j=2}^{\infty} \frac{t^j}{j!} \mathbf{x}^* U_j \mathbf{x} \leq -2\lambda_{\mathbf{x}} t + M_2 t^2, \quad (\text{A.13})$$

since  $|\sum_{j=2}^{\infty} \frac{t^j}{j!} \mathbf{x}^* U_j \mathbf{x}| \leq t^2 M_2$  with  $M_2 := \sum_{j=2}^{\infty} \frac{1}{j!} (2\|B\|_2)^j$ . Then,  $-2\lambda_{\mathbf{x}} t + M_2 t^2 \leq -\lambda_{\mathbf{x}} t$  for all  $0 \leq t \leq \lambda_{\mathbf{x}}/M_2$ . For any given  $c > 0$ , the estimate  $-\lambda_{\mathbf{x}} t \leq -ct^3$  holds if  $0 \leq t \leq \sqrt{\lambda_{\mathbf{x}}/c}$ . Define  $t_{\delta} := \min\{\delta/M_2, \sqrt{\delta/c}, 1\}$ . Then, we derive

$$g(\mathbf{x}; t) \leq -ct^3 \quad \text{for all } \mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} > \delta \text{ and } t \in [0, t_{\delta}]. \quad (\text{A.14})$$

To sum up, choosing any  $\delta \in (0, \kappa)$ , the estimate (A.12) is derived. Then, for  $c := \mu_{\delta}/3!$ , there exists a (sufficiently small)  $t_{\delta} > 0$  (as defined in *Step 1b*) such that the estimate (A.14) holds. Consequently, we obtain

$$g(\mathbf{x}; t) \leq -\frac{\mu_{\delta}}{3!} t^3 + M_4 t^4 \quad \text{for all } \mathbf{x} \in \mathcal{S} \text{ and } t \in [0, t_{\delta}]. \quad (\text{A.15})$$

This shows (A.1) with  $c_2 := c/2 = \mu_{\delta}/(3! 2)$  and  $a = 2m_{HC} + 1 = 3$ .

*Step 1c.* To prove the second statement in Lemma A.1, we improve the estimate of  $c$  as follows: By definition, the time  $t_{\delta}$  satisfies  $\lim_{\delta \rightarrow 0} t_{\delta} = 0$ . To derive (the sharp) lower estimate (A.2) on the multiplicative factor  $c$ , we consider the Taylor expansion (2.8) of the propagator norm, use estimate (A.15), and take the limit  $\delta \rightarrow 0$ :

$$-2c = \lim_{\delta \rightarrow 0} \frac{\|e^{-B t_{\delta}}\|_2^2 - 1}{t_{\delta}^3} = \lim_{\delta \rightarrow 0} \frac{g(\mathbf{x}; t_{\delta})}{t_{\delta}^3} \leq \lim_{\delta \rightarrow 0} \left( -\frac{\mu_{\delta}}{3!} + M_4 t_{\delta} \right) = -\frac{\mu_0}{3!}. \quad (\text{A.16})$$

Hence, we identified a lower estimate for the multiplicative factor  $c$  in (2.8) as

$$c \geq \frac{\mu_0}{3! 2} = \frac{1}{3! 2} \min_{\substack{\mathbf{x} \in \ker(\tilde{T}_0), \\ \|\mathbf{x}\|=1}} \langle \mathbf{x}, B^* B_H B \mathbf{x} \rangle. \quad (\text{A.17})$$

This finishes the proof of the second statement in Lemma A.1 in the case  $m_{HC} = 1$ .

*Step 2. Matrices with hypocoercivity index  $m_{HC} \geq 2$ .* For matrices  $B$  with hypocoercivity index  $m_{HC} \geq 2$ , i.e. there exists  $\kappa > 0$  such that

$$\tilde{T}_{m_{HC}} = \sum_{j=0}^{m_{HC}} (B^*)^j B_H B^j \geq \kappa I > 0, \quad (\text{A.18})$$

we generalize this procedure as follows: We define, for  $\mathbf{x} \in \mathcal{S}$ ,

$$\lambda_{\mathbf{x}} := \langle \mathbf{x}, \tilde{T}_{m_{HC}-1} \mathbf{x} \rangle \geq 0, \quad \mu_{\mathbf{x}} := \langle \mathbf{x}, (B^*)^{m_{HC}} B_H B^{m_{HC}} \mathbf{x} \rangle \geq 0, \quad \text{such that } \lambda_{\mathbf{x}} + \mu_{\mathbf{x}} \geq \kappa > 0. \quad (\text{A.19})$$

*Step 2a.* Consider  $\boxed{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta}$  where  $\delta \in (0, \kappa)$  will be chosen later and  $\kappa > 0$  such that (A.18) holds. For  $t \in [0, 1]$ , we derive as in *Step 2b* (see (A.27) below):

$$\begin{aligned} g(\mathbf{x}; t) &\leq \frac{(B^{m_{HC}} \mathbf{x})^* U_1 B^{m_{HC}} \mathbf{x}}{(2m_{HC} + 1)! \binom{2m_{HC}}{m_{HC}}} t^{2m_{HC}+1} + M_{2m_{HC}+2} t^{2m_{HC}+2} \\ &= -\frac{2}{(2m_{HC} + 1)! \binom{2m_{HC}}{m_{HC}}} \mu_{\mathbf{x}} t^{2m_{HC}+1} + M_{2m_{HC}+2} t^{2m_{HC}+2}, \end{aligned} \quad (\text{A.20})$$

with  $M_{2m_{HC}+2} := \sum_{j=2m_{HC}+2}^{\infty} \frac{1}{j!} (2\|B\|_2)^j > 0$ . To establish a uniform negative upper bound for  $-\mu_{\mathbf{x}}$  for  $\mathbf{x} \in \mathcal{S}$  with  $\lambda_{\mathbf{x}} \leq \delta$ , we use (A.19) to deduce  $\mu_{\mathbf{x}} \geq \kappa - \lambda_{\mathbf{x}} \geq \kappa - \delta > 0$  since  $\delta \in (0, \kappa)$ . For  $\delta \in (0, \kappa)$ , define

$$\mu_{\delta} := \min_{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta} \mu_{\mathbf{x}} = \min_{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta} \langle \mathbf{x}, (B^*)^{m_{HC}} B_H B^{m_{HC}} \mathbf{x} \rangle \quad (\text{A.21})$$

such that  $\mu_{\delta} \geq \kappa - \delta > 0$  and

$$\mu_0 := \lim_{\delta \rightarrow 0} \mu_{\delta} = \min_{\mathbf{x} \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}\|=1} \langle \mathbf{x}, (B^*)^{m_{HC}} B_H B^{m_{HC}} \mathbf{x} \rangle. \quad (\text{A.22})$$

Then, we derive from (A.20) that

$$g(\mathbf{x}; t) \leq -c t^{2m_{HC}+1} + M_{2m_{HC}+2} t^{2m_{HC}+2} \quad \text{for all } \mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} \leq \delta \text{ and } t \in [0, 1], \quad (\text{A.23a})$$

where

$$c := \frac{2\mu_{\delta}}{(2m_{HC}+1)! \binom{2m_{HC}}{m_{HC}}}. \quad (\text{A.23b})$$

Step 2b. Next, we shall show the following statement: Consider  $\boxed{\mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} > \delta}$  where  $\kappa > 0$  such that (A.18) holds. For given  $c > 0$ , there exists  $t_{\delta} > 0$  such that

$$g(\mathbf{x}; t) \leq -c t^{2m_{HC}+1} \quad \text{for all } \mathbf{x} \in \mathcal{S} \text{ with } \lambda_{\mathbf{x}} > \delta \text{ and } t \in [0, t_{\delta}]. \quad (\text{A.24})$$

First, we decompose the sphere  $\mathcal{S}$  into the (non-disjoint) closed subsets

$$\begin{aligned} \mathcal{C}_0 &:= \{\mathbf{x} \in \mathcal{S} \mid \mathbf{x}^* U_1 \mathbf{x} \leq -\epsilon\}, \\ \mathcal{C}_1 &:= \{\mathbf{x} \in \mathcal{S} \mid -\epsilon \leq \mathbf{x}^* U_1 \mathbf{x} \wedge (B\mathbf{x})^* U_1 B\mathbf{x} \leq -\epsilon\}, \\ \mathcal{C}_2 &:= \{\mathbf{x} \in \mathcal{S} \mid -\epsilon \leq \mathbf{x}^* U_1 \mathbf{x} \wedge -\epsilon \leq (B\mathbf{x})^* U_1 B\mathbf{x} \wedge (B^2 \mathbf{x})^* U_1 B^2 \mathbf{x} \leq -\epsilon\}, \\ &\vdots \\ \mathcal{C}_m &:= \{\mathbf{x} \in \mathcal{S} \mid \forall k \in \{0, \dots, m-1\} : -\epsilon \leq (B^k \mathbf{x})^* U_1 B^k \mathbf{x} \wedge (B^m \mathbf{x})^* U_1 B^m \mathbf{x} \leq -\epsilon\}, \\ &\vdots \\ \mathcal{C}_{m_{HC}-1} &:= \{\mathbf{x} \in \mathcal{S} \mid \forall k \in \{0, \dots, m_{HC}-2\} : -\epsilon \leq (B^k \mathbf{x})^* U_1 B^k \mathbf{x} \\ &\quad \wedge (B^{m_{HC}-1} \mathbf{x})^* U_1 B^{m_{HC}-1} \mathbf{x} \leq -\epsilon\}, \end{aligned} \quad (\text{A.25})$$

as well as  $\mathcal{C}_{m_{HC}} := \{\mathbf{x} \in \mathcal{S} \mid \forall k \in \{0, \dots, m_{HC}-1\} : -\epsilon \leq (B^k \mathbf{x})^* U_1 B^k \mathbf{x}\}$ , for some positive parameter  $\epsilon$  to be determined next.

We show that there exists  $\epsilon > 0$  such that  $\mathcal{C}_{m_{HC}} \subseteq \{\mathbf{x} \in \mathcal{S} \mid \lambda_{\mathbf{x}} \leq \delta\}$ : Consider  $\mathbf{x} \in \mathcal{C}_{m_{HC}}$ . Then,  $\mathbf{x}$  satisfies

$$\epsilon/2 \geq \langle \mathbf{x}, (B^*)^k B_H B^k \mathbf{x} \rangle \quad \text{for } k = 0, \dots, m_{HC} - 1,$$

which (upon summing up) implies that

$$m_{HC}\epsilon/2 \geq \langle \mathbf{x}, \sum_{k=0}^{m_{HC}-1} (B^*)^k B_H B^k \mathbf{x} \rangle = \langle \mathbf{x}, \tilde{T}_{m_{HC}-1} \mathbf{x} \rangle = \lambda_{\mathbf{x}}.$$

Choosing

$$\epsilon := \frac{2\delta}{m_{HC}} \quad \text{implies } \mathcal{C}_{m_{HC}} \subseteq \{\mathbf{x} \in \mathcal{S} \mid \lambda_{\mathbf{x}} \leq \delta\}. \quad (\text{A.26})$$

Hence, the already established estimate (A.23) holds in particular for  $\mathbf{x} \in \mathcal{C}_{m_{HC}}$  and  $t \in [0, 1]$ . Using  $\bigcup_{j=0}^{m_{HC}-1} \mathcal{C}_j \supseteq \{\mathbf{x} \in \mathcal{S} \mid \lambda_{\mathbf{x}} > \delta\}$  (as the complementary inclusion of (A.26)), we are left to prove the estimate (A.24) for all  $\mathbf{x} \in \bigcup_{j=0}^{m_{HC}-1} \mathcal{C}_j$ :

For all  $\ell \in \{0, \dots, m_{HC} - 1\}$ ,  $\boxed{\mathbf{x} \in \mathcal{C}_\ell}$  and  $t \in [0, 1]$ , the key idea (to estimate  $g(\mathbf{x}; t)$ ) is to collect the terms  $t^j$  of order  $j$  less than  $2\ell + 1$  in a quadratic form which is non-positive. Therefore, we use again (A.6) and Lemma A.2 with  $U = -B^*$ ,  $V = U_1$ ,  $W = -B$  and  $m = \ell - 1$ , to rewrite  $g(\mathbf{x}; t)$  as

$$\begin{aligned} g(\mathbf{x}; t) &= \mathbf{x}^* \left( \sum_{j=1}^{\infty} \frac{t^j}{j!} U_j \right) \mathbf{x} \\ &= \sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} ((-B)^k \mathbf{x})^* U_1 (-B)^{j-k-1} \mathbf{x} \\ &= \sum_{j=0}^{\ell-1} \frac{t^{2j+1}}{(2j+1)!} \binom{2j}{j} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} t^k ((-B)^{k+j} \mathbf{x})^* \right) \\ &\quad \times U_1 \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} t^k (-B)^{k+j} \mathbf{x} \right) \\ &\quad + \frac{t^{2\ell+1}}{(2\ell+1)!} \binom{2\ell}{\ell} \Delta_{2\ell+1, \ell}^{(\ell)} ((-B)^\ell \mathbf{x})^* U_1 (-B)^\ell \mathbf{x} \\ &\quad + \sum_{j=2\ell+2}^{\infty} \frac{t^j}{j!} \sum_{k=\ell}^{j-\ell-1} \binom{j-1}{k} \Delta_{j, k}^{(\ell)} ((-B)^k \mathbf{x})^* U_1 (-B)^{j-k-1} \mathbf{x}. \end{aligned} \quad (\text{A.27a})$$

The first term is non-positive since  $U_1 = -2B_H \leq 0$ . The second term is estimated using the assumption  $\mathbf{x} \in \mathcal{C}_\ell$  and the identity  $\Delta_{2\ell+1, \ell}^{(\ell)} = \binom{2\ell}{\ell}^{-2}$ . The third term can be estimated from above by  $M_{2\ell+2} t^{2\ell+2}$  with  $M_{2\ell+2} := \sum_{j=2\ell+2}^{\infty} \frac{1}{j!} (2\|B\|_2)^j > 0$  by using (A.6), (A.4), and  $\Delta_{j, k}^{(\ell)} \leq 1$  for  $0 \leq \ell \leq k \leq j - \ell - 1$ . Altogether, we obtain the estimate

$$g(\mathbf{x}; t) \leq -\frac{\epsilon}{(2\ell+1)! \binom{2\ell}{\ell}} t^{2\ell+1} + M_{2\ell+2} t^{2\ell+2}. \quad (\text{A.27b})$$

For given  $c > 0$  (e.g. as in (A.23b)), there exists  $\tilde{t}_\ell > 0$  (depending on  $\epsilon$ , with  $\lim_{\epsilon \rightarrow 0} \tilde{t}_\ell = 0$ ) such that

$$g(\mathbf{x}; t) \leq -c t^{2m_{HC}+1} \quad \text{for all } \mathbf{x} \in \mathcal{C}_\ell \text{ and } t \in [0, \tilde{t}_\ell]. \quad (\text{A.28})$$

Choosing  $\epsilon = 2\delta/m_{HC}$  as in (A.26) (such that  $\bigcup_{j=0}^{m_{HC}-1} \mathcal{C}_j \supseteq \{\mathbf{x} \in \mathcal{S} \mid \lambda_{\mathbf{x}} > \delta\}$ ) and

$$t_\delta := \min\{\tilde{t}_\ell, \ell = 0, \dots, m_{HC} - 1\}, \quad (\text{A.29})$$

implies the estimate (A.24).

To sum up, for any fixed  $\delta \in (0, \kappa)$ , the estimate (A.23) with multiplicative constant  $c$  in (A.23b) is proven in *Step 2a*. Then, for  $\epsilon = 2\delta/m_{HC}$  and  $c$  in (A.23b), there exists a (sufficiently small)  $t_\delta > 0$  (as defined in (A.29)) such that the estimate (A.24) holds. Consequently, we obtain

$$g(\mathbf{x}; t) \leq -\frac{2\mu_\delta}{(2m_{HC}+1)!(\frac{2m_{HC}}{m_{HC}})} t^{2m_{HC}+1} + M_{2m_{HC}+2} t^{2m_{HC}+2} \quad \text{for all } \mathbf{x} \in \mathcal{S} \text{ with } t \in [0, t_\delta]. \quad (\text{A.30})$$

This shows (A.1) with  $c_2 := c/2$ ,  $c$  as defined in (A.23b) and  $a = 2m_{HC} + 1$ . This finishes the proof of the first statement in Lemma A.1 for  $m_{HC} \in \mathbb{N}$ .

Step 2c. To prove the second statement in Lemma A.1, we improve the estimate of  $c$  as follows: By definition, the time  $t_\delta$  depends on  $\delta$  (since  $\epsilon = 2\delta/m_{HC}$ ) such that  $\lim_{\delta \rightarrow 0} t_\delta = 0$ . To derive the (sharp) lower estimate (A.2) on the multiplicative factor  $c$ , we consider the Taylor expansion (2.8) of the propagator norm, use estimate (A.30), and take the limit  $\delta \rightarrow 0$  such that

$$\begin{aligned} -2c &= \lim_{\delta \rightarrow 0} \frac{\|e^{-B t_\delta}\|_2^2 - 1}{t_\delta^{2m_{HC}+1}} \leq \lim_{\delta \rightarrow 0} \left( -\frac{2\mu_\delta}{(2m_{HC}+1)!(\frac{2m_{HC}}{m_{HC}})} + M_{2m_{HC}+2} t_\delta \right) \\ &= -\frac{2\mu_0}{(2m_{HC}+1)!(\frac{2m_{HC}}{m_{HC}})}. \end{aligned} \quad (\text{A.31})$$

This proves the lower estimate (A.2) for the multiplicative factor  $c$  in (2.8).  $\square$

The proof of Lemma A.1 uses the following identity:

**Lemma A.2.** *Let  $U, V, W \in \mathbb{C}^{n \times n}$ . For all  $m \in \mathbb{N}_0$ , the following identity holds*

$$\begin{aligned} &\sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} U^k V W^{j-k-1} \\ &= \sum_{j=0}^m \frac{t^{2j+1}}{(2j+1)!} \binom{1}{\binom{2j}{j}} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} t^k U^{k+j} \right) V \\ &\quad \times \left( \sum_{\ell=0}^{\infty} \frac{(2j+1)!}{(\ell+2j+1)!} \binom{\ell+j}{j} t^\ell W^{\ell+j} \right) \end{aligned}$$

$$+ \sum_{j=2m+3}^{\infty} \frac{t^j}{j!} \sum_{k=m+1}^{j-m-2} \binom{j-1}{k} \Delta_{j,k}^{(m+1)} U^k V W^{j-k-1}, \quad (\text{A.32})$$

where  $\Delta_{j,k}^{(m)} := \frac{\binom{k}{m} \binom{j-k-1}{m}}{\binom{k+m}{m} \binom{j-k-1+m}{m}}$  for all  $m \leq k$  and  $m \leq j - k - 1$ .

**Proof.** We will prove the identity by induction. For  $m = 0$ , we have to prove the identity

$$\begin{aligned} & \sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} U^k V W^{j-k-1} \\ &= t \left( \sum_{k=0}^{\infty} \frac{1}{(k+1)!} t^k U^k \right) V \left( \sum_{\ell=0}^{\infty} \frac{1}{(\ell+1)!} t^{\ell} W^{\ell} \right) + \sum_{j=3}^{\infty} \frac{t^j}{j!} \sum_{k=1}^{j-2} \binom{j-1}{k} \frac{k(j-k-1)}{(k+1)(j-k)} U^k V W^{j-k-1} \end{aligned}$$

since  $\Delta_{j,k}^{(1)} = \frac{\binom{k}{1} \binom{j-k-1}{1}}{\binom{k+1}{1} \binom{j-k}{1}} = \frac{k(j-k-1)}{(k+1)(j-k)}$ . The first term on the right hand side can be written by the Cauchy product formula as

$$\begin{aligned} & t \left( \sum_{k=0}^{\infty} \frac{1}{(k+1)!} t^k U^k \right) V \left( \sum_{\ell=0}^{\infty} \frac{1}{(\ell+1)!} t^{\ell} W^{\ell} \right) \\ &= t \sum_{j=0}^{\infty} t^j \sum_{k=0}^j \frac{1}{(k+1)!} U^k V \frac{1}{(j-k+1)!} W^{j-k} = \sum_{j=0}^{\infty} t^{j+1} \sum_{k=0}^j \frac{1}{(k+1)!} U^k V \frac{1}{(j-k+1)!} W^{j-k} \\ &= \sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \frac{j!}{(k+1)!(j-k)!} U^k V W^{j-k-1} = \sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} \frac{j}{(k+1)(j-k)} U^k V W^{j-k-1} \\ &= \sum_{j=1}^2 \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} U^k V W^{j-k-1} + \sum_{j=3}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} \frac{j}{(k+1)(j-k)} U^k V W^{j-k-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} & t \left( \sum_{k=0}^{\infty} \frac{1}{(k+1)!} t^k U^k \right) V \left( \sum_{\ell=0}^{\infty} \frac{1}{(\ell+1)!} t^{\ell} W^{\ell} \right) + \sum_{j=3}^{\infty} \frac{t^j}{j!} \sum_{k=1}^{j-2} \binom{j-1}{k} \frac{k(j-k-1)}{(k+1)(j-k)} U^k V W^{j-k-1} \\ &= \sum_{j=1}^2 \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} U^k V W^{j-k-1} + \sum_{j=3}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} \frac{j}{(k+1)(j-k)} U^k V W^{j-k-1} \\ &+ \sum_{j=3}^{\infty} \frac{t^j}{j!} \sum_{k=1}^{j-2} \binom{j-1}{k} \frac{k(j-k-1)}{(k+1)(j-k)} U^k V W^{j-k-1} \end{aligned}$$

$$= \sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} U^k V W^{j-k-1}.$$

We assume that the formula holds for  $m \in \mathbb{N}_0$  and prove it for  $m + 1$ . First, we use again the Cauchy product formula to derive

$$\begin{aligned} & \frac{t^{2m+3}}{(2m+3)!} \frac{1}{\binom{2m+2}{m+1}} \left( \sum_{k=0}^{\infty} \frac{(2m+3)!}{(k+2m+3)!} \binom{k+m+1}{m+1} t^k U^{k+m+1} \right) V \left( \sum_{\ell=0}^{\infty} \frac{(2m+3)!}{(\ell+2m+3)!} \binom{\ell+m+1}{m+1} t^{\ell} W^{\ell+m+1} \right) \\ &= \frac{t^{2m+3}}{(2m+3)!} \frac{1}{\binom{2m+2}{m+1}} \sum_{j=0}^{\infty} t^j \sum_{k=0}^j \frac{(2m+3)!}{(k+2m+3)!} \binom{k+m+1}{m+1} U^{k+m+1} V \frac{(2m+3)!}{(j-k+2m+3)!} \binom{j-k+m+1}{m+1} W^{j-k+m+1} \\ &= \sum_{j=0}^{\infty} \frac{t^{j+2m+3}}{(j+2m+3)!} \sum_{k=0}^j \frac{(j+2m+3)!}{(2m+3)!(\binom{2m+2}{m+1})} \frac{(2m+3)!}{(k+2m+3)!} \binom{k+m+1}{m+1} U^{k+m+1} V \frac{(2m+3)!}{(j-k+2m+3)!} \binom{j-k+m+1}{m+1} W^{j-k+m+1} \\ &= \sum_{j=2m+3}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-2m-3} \frac{j!}{(2m+3)!(\binom{2m+2}{m+1})} \frac{(2m+3)!}{(k+2m+3)!} \binom{k+m+1}{m+1} U^{k+m+1} V \frac{(2m+3)!}{(j-k)!} \binom{j-k-m-2}{m+1} W^{j-k-m-2} \\ &= \sum_{j=2m+3}^{\infty} \frac{t^j}{j!} \sum_{k=m+1}^{j-m-2} \frac{j!}{(2m+3)!(\binom{2m+2}{m+1})} \frac{(2m+3)!}{(k+m+2)!} \binom{k}{m+1} \frac{(2m+3)!}{(j-k+m+1)!} \binom{j-k-1}{m+1} U^k V W^{j-k-1}. \end{aligned}$$

Therefore,

$$\begin{aligned} & \sum_{j=0}^{m+1} \frac{t^{2j+1}}{(2j+1)!} \frac{1}{\binom{2j}{j}} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} t^k U^{k+j} \right) V \\ & \quad \times \left( \sum_{\ell=0}^{\infty} \frac{(2j+1)!}{(\ell+2j+1)!} \binom{\ell+j}{j} t^{\ell} W^{\ell+j} \right) \\ &+ \sum_{j=2m+5}^{\infty} \frac{t^j}{j!} \sum_{k=m+2}^{j-m-3} \binom{j-1}{k} \Delta_{j,k}^{(m+2)} U^k V W^{j-k-1} \\ &= \sum_{j=0}^m \frac{t^{2j+1}}{(2j+1)!} \frac{1}{\binom{2j}{j}} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} t^k U^{k+j} \right) V \\ & \quad \times \left( \sum_{\ell=0}^{\infty} \frac{(2j+1)!}{(\ell+2j+1)!} \binom{\ell+j}{j} t^{\ell} W^{\ell+j} \right) \\ &+ \sum_{j=2m+3}^{\infty} \frac{t^j}{j!} \sum_{k=m+1}^{j-m-2} \frac{j!}{(2m+3)!(\binom{2m+2}{m+1})} \frac{(2m+3)!}{(k+m+2)!} \binom{k}{m+1} \frac{(2m+3)!}{(j-k+m+1)!} \binom{j-k-1}{m+1} U^k V W^{j-k-1} \\ &+ \sum_{j=2m+5}^{\infty} \frac{t^j}{j!} \sum_{k=m+2}^{j-m-3} \binom{j-1}{k} \Delta_{j,k}^{(m+2)} U^k V W^{j-k-1}. \end{aligned} \tag{A.33}$$

Using  $\Delta_{j,k}^{(m+2)} = \Delta_{j,k}^{(m+1)} \frac{(k-m-1)(j-k-m-2)}{(k+m+2)(j-k+m+1)}$  we deduce that (A.33) equals:

$$\begin{aligned}
&= \sum_{j=0}^m \frac{t^{2j+1}}{(2j+1)!} \frac{1}{\binom{2j}{j}} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} t^k U^{k+j} \right) V \left( \sum_{\ell=0}^{\infty} \frac{(2j+1)!}{(\ell+2j+1)!} \binom{\ell+j}{j} t^{\ell} W^{\ell+j} \right) \\
&\quad + \sum_{j=2m+3}^{\infty} \frac{t^j}{j!} \sum_{k=m+1}^{j-m-2} \binom{j-1}{k} \Delta_{j,k}^{(m+1)} \frac{(2m+3)j}{(k+m+2)(j-k+m+1)} U^k V W^{j-k-1} \\
&\quad + \sum_{j=2m+5}^{\infty} \frac{t^j}{j!} \sum_{k=m+2}^{j-m-3} \binom{j-1}{k} \Delta_{j,k}^{(m+1)} \frac{(k-m-1)(j-k-m-2)}{(k+m+2)(j-k+m+1)} U^k V W^{j-k-1} \\
&= \sum_{j=0}^m \frac{t^{2j+1}}{(2j+1)!} \frac{1}{\binom{2j}{j}} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} t^k U^{k+j} \right) V \left( \sum_{\ell=0}^{\infty} \frac{(2j+1)!}{(\ell+2j+1)!} \binom{\ell+j}{j} t^{\ell} W^{\ell+j} \right) \\
&\quad + \sum_{j=2m+3}^{\infty} \frac{t^j}{j!} \sum_{k=m+1}^{j-m-2} \binom{j-1}{k} \Delta_{j,k}^{(m+1)} U^k V W^{j-k-1} \\
&= \sum_{j=1}^{\infty} \frac{t^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} U^k V W^{j-k-1},
\end{aligned}$$

where we used the induction hypothesis, i.e. (A.32), in the final equality. This finishes the proof.  $\square$

To determine the optimal multiplicative factor  $c$  in (2.8) (see Theorem 2.7(b)), we shall next derive an improved upper estimate for  $c$ , compared to (2.19):

**Lemma A.3.** *Let the ODE system (1.1) be conservative-dissipative, and let the system matrix  $B$  be (hypo)coercive with hypocoercivity index  $m_{HC} \in \mathbb{N}$ . Then, the multiplicative factor  $c$  in (2.8) satisfies*

$$c \leq c_1 \quad \text{with } c_1 := \frac{1}{(2m_{HC}+1)! \binom{2m_{HC}}{m_{HC}}} \min_{\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1}), \|\mathbf{x}_0\|_2=1} \langle B^{m_{HC}} \mathbf{x}_0, B_H B^{m_{HC}} \mathbf{x}_0 \rangle. \quad (\text{A.34})$$

**Proof.** As illustrated by Example 2.12, see also Fig. 1, the propagator norm is in general not determined by the norm of one specific solution. Instead we consider a parameterized family of solutions pertaining to initial values  $\mathbf{x}_\tau$ ,  $\tau \in [0, 1]$ :

Due to Lemma A.4 below, for  $\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1})$  (with  $\|\mathbf{x}_0\|_2 = 1$ ), there exist real constants  $b_\ell$ ,  $\ell = 1, \dots, m_{HC}$  such that

$$\mathbf{x}_\tau := \mathbf{x}_0 + \sum_{\ell=1}^{m_{HC}} b_\ell \tau^\ell B^\ell \mathbf{x}_0$$

satisfies

$$g(\mathbf{x}_\tau; \tau) = -2c_1(\mathbf{x}_0) \tau^{2m_{HC}+1} + \mathcal{O}(\tau^{2m_{HC}+2}), \quad \text{for } \tau \in [0, 1], \quad (\text{A.35})$$

with  $c_1(\mathbf{x}_0)$  defined in (A.40) and  $\lim_{\tau \rightarrow 0} \mathbf{x}_\tau = \mathbf{x}_0$ . To normalize the family of vectors  $\mathbf{x}_\tau$ ,  $\tau \in [0, 1]$  we define

$$\tilde{\mathbf{x}}_\tau := \frac{\mathbf{x}_\tau}{\|\mathbf{x}_\tau\|_2}, \quad (\text{A.36})$$

still satisfying  $\lim_{\tau \rightarrow 0} \tilde{\mathbf{x}}_\tau = \mathbf{x}_0$ . For  $\tau \in [0, 1]$ , we estimate the propagator norm as

$$\begin{aligned} \|e^{-B\tau}\|_2^2 &\geq \|e^{-Bt}\tilde{\mathbf{x}}_\tau\|_2^2 \Big|_{t=\tau} = \frac{\|e^{-B\tau}\mathbf{x}_\tau\|_2^2}{\|\mathbf{x}_\tau\|_2^2} \\ &= \frac{\|\mathbf{x}_\tau\|_2^2 + g(\mathbf{x}_\tau; \tau)}{\|\mathbf{x}_\tau\|_2^2} = 1 - 2\frac{c_1(\mathbf{x}_0)}{\|\mathbf{x}_\tau\|_2^2} \tau^{2m_{HC}+1} + \frac{\mathcal{O}(\tau^{2m_{HC}+2})}{\|\mathbf{x}_\tau\|_2^2}, \end{aligned} \quad (\text{A.37})$$

where we used definition (A.20) and (A.35). To derive a bound for the multiplicative factor  $c$  in (2.8), we consider the Taylor expansion (2.8) of the propagator norm, use estimate (A.37) for  $\tau > 0$ , and take the limit  $\tau \rightarrow 0$  such that

$$\begin{aligned} -2c &= \lim_{\tau \rightarrow 0} \frac{\|e^{-B\tau}\|_2^2 - 1}{\tau^{2m_{HC}+1}} \geq \lim_{\tau \rightarrow 0} \frac{g(\mathbf{x}_\tau; \tau)}{\|\mathbf{x}_\tau\|_2^2 \tau^{2m_{HC}+1}} \\ &= \lim_{\tau \rightarrow 0} \left( -2\frac{c_1(\mathbf{x}_0)}{\|\mathbf{x}_\tau\|_2^2} + \frac{\mathcal{O}(\tau)}{\|\mathbf{x}_\tau\|_2^2} \right) = -2c_1(\mathbf{x}_0). \end{aligned}$$

Hence,  $c \leq c_1(\mathbf{x}_0)$  for all normalized vectors  $\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1})$ . Taking the minimum of  $c_1(\mathbf{x}_0)$  over all normalized vectors  $\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1})$  yields the upper bound for the multiplicative factor  $c$  as given in (A.34). This finishes the proof.  $\square$

The proof of Lemma A.3 uses the following construction of a vector function:

**Lemma A.4.** *Let the ODE system (1.1) be conservative-dissipative, and let the system matrix  $B$  be hypocoercive with hypocoercivity index  $m_{HC} \in \mathbb{N}$ . Then, for each  $\mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1})$ , there exists a polynomial vector function  $\mathbf{x}_\tau \in \mathbb{C}^n$ ,  $\tau \in [0, 1]$  of the form*

$$\mathbf{x}_\tau = \mathbf{x}_0 + \sum_{\ell=1}^{m_{HC}} b_\ell \tau^\ell B^\ell \mathbf{x}_0, \quad \text{for a suitable choice of } b_\ell \in \mathbb{R}, \quad \ell = 1, \dots, m_{HC}, \quad (\text{A.38})$$

such that

$$g(\mathbf{x}_\tau; \tau) := \mathbf{x}_\tau^* \left( \sum_{j=1}^{\infty} \frac{\tau^j}{j!} U_j \right) \mathbf{x}_\tau = -2c_1(\mathbf{x}_0) \tau^a + \mathcal{O}(\tau^{a+1}) \quad \text{for } \tau \in [0, 1], \quad (\text{A.39})$$

where  $a = 2m_{HC} + 1$  and

$$c_1(\mathbf{x}_0) := \frac{\|\sqrt{B_H} B^{m_{HC}} \mathbf{x}_0\|_2^2}{(2m_{HC} + 1)! \binom{2m_{HC}}{m_{HC}}}. \quad (\text{A.40})$$

**Proof.** For  $\mathbf{x} \in \mathbb{C}^n$  and  $\tau \in [0, 1]$ , consider  $g(\mathbf{x}; \tau) := \mathbf{x}^* \left( \sum_{j=1}^{\infty} \frac{\tau^j}{j!} U_j \right) \mathbf{x}$  using  $U_j$  in the form (A.6). Following Lemma A.2 with  $U = -B^*$ ,  $V = U_1$ ,  $W = -B$ , and  $m = m_{HC} - 1$ , we rewrite  $g(\mathbf{x}; \tau)$  as

$$\begin{aligned}
g(\mathbf{x}; \tau) &= \sum_{j=1}^{\infty} \frac{\tau^j}{j!} \sum_{k=0}^{j-1} \binom{j-1}{k} ((-B)^k \mathbf{x})^* U_1 (-B)^{j-k-1} \mathbf{x} \\
&= \sum_{j=0}^{m_{HC}-1} \frac{\tau^{2j+1}}{(2j+1)!} \frac{1}{\binom{2j}{j}} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} \tau^k ((-B)^{k+j} \mathbf{x})^* \right) U_1 \\
&\quad \times \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} \tau^k (-B)^{k+j} \mathbf{x} \right) \\
&\quad + \frac{\tau^{2m_{HC}+1}}{(2m_{HC}+1)!} \binom{2m_{HC}}{m_{HC}} \Delta_{2m_{HC}+1, m_{HC}}^{(m_{HC})} ((-B)^{m_{HC}} \mathbf{x})^* U_1 (-B)^{m_{HC}} \mathbf{x} \\
&\quad + \sum_{j=2m_{HC}+2}^{\infty} \frac{\tau^j}{j!} \sum_{k=m_{HC}}^{j-m_{HC}-1} \binom{j-1}{k} \Delta_{j, k}^{(m_{HC})} ((-B)^k \mathbf{x})^* U_1 (-B)^{j-k-1} \mathbf{x} \\
&= -2 \sum_{j=0}^{m_{HC}-1} \frac{\tau}{(2j+1)!} \frac{1}{\binom{2j}{j}} \left\| \sqrt{B_H} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} \tau^{k+j} (-B)^{k+j} \mathbf{x} \right) \right\|_2^2 \\
&\quad - 2 \frac{\tau^{2m_{HC}+1}}{(2m_{HC}+1)!} \binom{2m_{HC}}{m_{HC}} \left\| \sqrt{B_H} B^{m_{HC}} \mathbf{x} \right\|_2^2 + \mathcal{O}(\tau^{2m_{HC}+2}),
\end{aligned} \tag{A.41}$$

using that  $U_1 = -2B_H$  is a negative semi-definite Hermitian matrix, and the identity

$$\Delta_{2m_{HC}+1, m_{HC}}^{(m_{HC})} = \binom{2m_{HC}}{m_{HC}}^{-2},$$

to rewrite the first and second term, respectively. The third term in (A.41) can be bounded by  $M_{2m_{HC}+2} \tau^{2m_{HC}+2}$  with  $M_{2m_{HC}+2} := \sum_{j=2m_{HC}+2}^{\infty} \frac{1}{j!} (2\|B\|_2)^j > 0$ , using that  $\Delta_{j, k}^{(m_{HC})} \leq 1$  for  $0 \leq m_{HC} \leq k \leq j - m_{HC} - 1$  and  $\tau \in [0, 1]$ .

Step 1. To estimate the second term in the last identity of (A.41) for  $\mathbf{x} = \mathbf{x}_\tau$ , we use a polynomial ansatz for  $\mathbf{x}_\tau$ :

For  $\tau \in [0, 1]$ , we consider the ansatz

$$\begin{aligned}
\mathbf{x}_\tau &:= \sum_{\ell=0}^{m_{HC}} \tau^\ell \mathbf{x}_\ell \quad \text{with the given } \mathbf{x}_0 \in \ker(\tilde{T}_{m_{HC}-1}) \\
&\quad \text{and some } \mathbf{x}_\ell \in \mathbb{C}^n, \ell = 1, \dots, m_{HC} \text{ to be chosen.}
\end{aligned} \tag{A.42}$$

Then, we observe that

$$\|\sqrt{B_H} B^{m_{HC}} \mathbf{x}_\tau\|_2^2 = \|\sqrt{B_H} B^{m_{HC}} \sum_{\ell=0}^{m_{HC}} \tau^\ell \mathbf{x}_\ell\|_2^2 = \|\sqrt{B_H} B^{m_{HC}} \mathbf{x}_0\|_2^2 + \mathcal{O}(\tau), \quad (\text{A.43})$$

such that the second term in the last identity of (A.41) satisfies

$$\begin{aligned} & -2 \frac{\tau^{2m_{HC}+1}}{(2m_{HC}+1)! \binom{2m_{HC}}{m_{HC}}} \|\sqrt{B_H} B^{m_{HC}} \mathbf{x}_\tau\|_2^2 \\ & = -2 \frac{\tau^{2m_{HC}+1}}{(2m_{HC}+1)! \binom{2m_{HC}}{m_{HC}}} \|\sqrt{B_H} B^{m_{HC}} \mathbf{x}_0\|_2^2 + \mathcal{O}(\tau^{2m_{HC}+2}). \end{aligned} \quad (\text{A.44})$$

Step 2. To estimate the first term in the last identity of (A.41) for  $\mathbf{x} = \mathbf{x}_\tau$ , we refine the ansatz (A.42) for  $\mathbf{x}_\tau$  as follows:

Consider (A.42) with

$$\mathbf{x}_\ell := b_\ell B^\ell \mathbf{x}_0, \quad \text{with some } b_\ell \in \mathbb{R}, \quad \ell = 1, \dots, m_{HC} \text{ to be chosen.} \quad (\text{A.45})$$

We shall construct the coefficients  $b_\ell \in \mathbb{R}$ ,  $\ell = 1, \dots, m_{HC}$  and set  $b_0 := 1$  such that the first term in the last identity of (A.41) satisfies

$$\sum_{j=0}^{m_{HC}-1} \frac{\tau}{(2j+1)!} \frac{1}{\binom{2j}{j}} \left\| \sqrt{B_H} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} \tau^{k+j} (-B)^{k+j} \mathbf{x}_\tau \right) \right\|_2^2 = \mathcal{O}(\tau^{2m_{HC}+2}). \quad (\text{A.46})$$

Each term in the outer sum is non-negative. Therefore, for  $j = 0, \dots, m_{HC}-1$ , we consider each

$$S_j := \left\| \sqrt{B_H} \left( \sum_{k=0}^{\infty} \frac{(2j+1)!}{(k+2j+1)!} \binom{k+j}{j} \tau^{k+j} (-B)^{k+j} \sum_{\ell=0}^{m_{HC}} b_\ell \tau^\ell B^\ell \mathbf{x}_0 \right) \right\|_2^2 \quad (\text{A.47})$$

separately, and construct  $b_\ell$ ,  $\ell = 1, \dots, m_{HC}$  iteratively such that  $S_j = \mathcal{O}(\tau^{2m_{HC}+1})$ .

Starting with  $\ell = 1$ , we determine  $b_\ell = b_1$  by considering  $S_{m_{HC}-\ell} = S_{m_{HC}-1}$ : Using

$$\sqrt{B_H} \mathbf{x}_0 = \dots = \sqrt{B_H} B^{m_{HC}-1} \mathbf{x}_0 = 0, \quad (\text{A.48})$$

we can rewrite  $S_{m_{HC}-1}$  as

$$\begin{aligned} S_{m_{HC}-1} &= \left\| \sqrt{B_H} \left( \sum_{k=0}^{\infty} \frac{(2m_{HC}-1)!}{(k+2m_{HC}-1)!} \binom{k+m_{HC}-1}{m_{HC}-1} \tau^{k+m_{HC}-1} (-B)^{k+m_{HC}-1} \sum_{\ell=0}^{m_{HC}} b_\ell \tau^\ell B^\ell \mathbf{x}_0 \right) \right\|_2^2 \\ &= \left\| \sqrt{B_H} \left( \tau^{m_{HC}-1} (-B)^{m_{HC}-1} b_1 \tau B \mathbf{x}_0 \right. \right. \\ &\quad \left. \left. + \frac{(2m_{HC}-1)!}{(2m_{HC})!} \binom{m_{HC}}{m_{HC}-1} \tau^{m_{HC}} (-B)^{m_{HC}} \mathbf{x}_0 + \mathcal{O}(\tau^{m_{HC}+1}) \right) \right\|_2^2 \\ &= \tau^{2m_{HC}} \left\| \sqrt{B_H} \left( b_1 - \frac{(2m_{HC}-1)!}{(2m_{HC})!} \binom{m_{HC}}{m_{HC}-1} \right) (-B)^{m_{HC}} \mathbf{x}_0 \right\|_2^2 + \mathcal{O}(\tau^{2m_{HC}+1}). \end{aligned} \quad (\text{A.49})$$

Choosing

$$b_1 := \frac{(2m_{HC}-1)!}{(2m_{HC})!} \binom{m_{HC}}{m_{HC}-1} = \frac{1}{2} \quad (\text{A.50})$$

yields  $S_{m_{HC}-1} = \mathcal{O}(\tau^{2m_{HC}+1})$ .

Subsequently, for  $\ell = 2, \dots, m_{HC}$ , we determine  $b_\ell$  by considering  $S_{m_{HC}-\ell}$  and using (A.48):

$$\begin{aligned} S_{m_{HC}-\ell} &= \left\| \sqrt{B_H} \left( \sum_{k=0}^{\infty} \frac{(2m_{HC}-2\ell+1)!}{(k+2m_{HC}-2\ell+1)!} \binom{k+m_{HC}-\ell}{m_{HC}-\ell} \tau^{k+m_{HC}-\ell} (-B)^{k+m_{HC}-\ell} \sum_{p=0}^{m_{HC}} b_p \tau^p B^p \mathbf{x}_0 \right) \right\|_2^2 \\ &= \left\| \sqrt{B_H} \left( \sum_{k=0}^{\ell} \frac{(2m_{HC}-2\ell+1)!}{(k+2m_{HC}-2\ell+1)!} \binom{k+m_{HC}-\ell}{m_{HC}-\ell} \tau^{k+m_{HC}-\ell} (-B)^{k+m_{HC}-\ell} b_{\ell-k} \tau^{\ell-k} B^{\ell-k} \mathbf{x}_0 \right) \right\|_2^2 + \mathcal{O}(\tau^{2m_{HC}+1}) \\ &= \tau^{2m_{HC}} \left\| \sqrt{B_H} \left( \sum_{k=0}^{\ell} \frac{(2m_{HC}-2\ell+1)!}{(k+2m_{HC}-2\ell+1)!} \binom{k+m_{HC}-\ell}{m_{HC}-\ell} b_{\ell-k} (-1)^{\ell-k} \right) (-B)^{m_{HC}} \mathbf{x}_0 \right\|_2^2 + \mathcal{O}(\tau^{2m_{HC}+1}). \end{aligned}$$

Using  $b_p$ ,  $p = 0, \dots, \ell - 1$  from the previous steps, and choosing

$$b_\ell := -(-1)^\ell \sum_{k=1}^{\ell} \frac{(2m_{HC}-2\ell+1)!}{(k+2m_{HC}-2\ell+1)!} \binom{k+m_{HC}-\ell}{m_{HC}-\ell} b_{\ell-k} (-1)^{\ell-k} \quad (\text{A.51})$$

yields  $S_{m_{HC}-\ell} = \mathcal{O}(\tau^{2m_{HC}+1})$ . Choosing these  $b_\ell$ , we have verified (A.46). Thus, using the ansatz (A.38) for  $\mathbf{x}_\tau$  implies that  $g(\mathbf{x}_\tau; \tau)$  from (A.41) equals the r.h.s. of (A.44). This proves that the identity (A.39) holds for  $\tau \in [0, 1]$ .  $\square$

## References

- [1] F. Achleitner, A. Arnold, E.A. Carlen, On linear hypocoercive BGK models, in: From Particle Systems to Partial Differential Equations. III, in: Springer Proc. Math. Stat., vol. 162, Springer, Cham, 2016, pp. 1–37.
- [2] F. Achleitner, A. Arnold, E.A. Carlen, On multi-dimensional hypocoercive BGK models, Kinet. Relat. Models 11 (4) (2018) 953–1009.
- [3] F. Achleitner, A. Arnold, E.A. Carlen, The hypocoercivity index for the intermediate and large time behavior of ODEs, 2022, work in progress.
- [4] F. Achleitner, A. Arnold, V. Mehrmann, Hypocoercivity and controllability in linear semi-dissipative Hamiltonian ordinary differential equations and differential-algebraic equations, Z. Angew. Math. Mech. (2021) e202100171.
- [5] F. Achleitner, A. Arnold, V. Mehrmann, Hypocoercivity and hypocontractivity concepts for linear dynamical systems, Electron. J. Linear Algebra 39 (2023) 33–61.
- [6] F. Achleitner, A. Arnold, B. Signorello, On optimal decay estimates for ODEs and PDEs with modal decomposition, in: Stochastic Dynamics Out of Equilibrium, in: Springer Proc. Math. Stat., vol. 282, Springer, Cham, 2019, pp. 241–264.
- [7] F. Achleitner, A. Arnold, D. Stürzer, Large-time behavior in non-symmetric Fokker–Planck equations, Riv. Math. Univ. Parma (N. S.) 6 (1) (2015) 1–68.
- [8] A. Arnold, J. Erb, Sharp entropy decay for hypocoercive and non-symmetric Fokker–Planck equations with linear drift, preprint, arXiv:1409.5425, 2014.
- [9] A. Arnold, C. Schmeiser, B. Signorello, Propagator norm and sharp decay estimates for Fokker–Planck equations with linear drift, Comm. Math. Sc. 20 (4) (2022) 1047–1080.
- [10] J. Dolbeault, C. Mouhot, C. Schmeiser, Hypocoercivity for linear kinetic equations conserving mass, Trans. Am. Math. Soc. 367 (6) (2015) 3807–3828.
- [11] S. Gadat, L. Miclo, Spectral decompositions and  $L^2$ -operator norms of toy hypocoercive semi-groups, Kinet. Relat. Models 6 (2) (2013) 317–372.

- [12] D. Hinrichsen, A.J. Pritchard, Mathematical Systems Theory. I, Texts in Applied Mathematics, vol. 48, Springer-Verlag, Berlin, 2005.
- [13] L. Hörmander, Hypoelliptic second order differential equations, *Acta Math.* 119 (1967) 147–171.
- [14] T. Kato, Perturbation Theory for Linear Operators, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.
- [15] L. Kohaupt, Differential calculus for some  $p$ -norms of the fundamental matrix with applications, *J. Comput. Appl. Math.* 135 (1) (2001) 1–21.
- [16] C. Mehl, V. Mehrmann, P. Sharma, Stability radii for linear Hamiltonian systems with dissipation under structure-preserving perturbations, *SIAM J. Matrix Anal. Appl.* 37 (4) (2016) 1625–1654.
- [17] L. Miclo, P. Monmarché, Étude spectrale minutieuse de processus moins indécis que les autres, in: Séminaire de Probabilités XLV, in: Lecture Notes in Math., vol. 2078, Springer, Cham, 2013, pp. 459–481.
- [18] C. Villani, Hypocoercivity, *Mem. Am. Math. Soc.* 202 (950) (2009) iv+141.