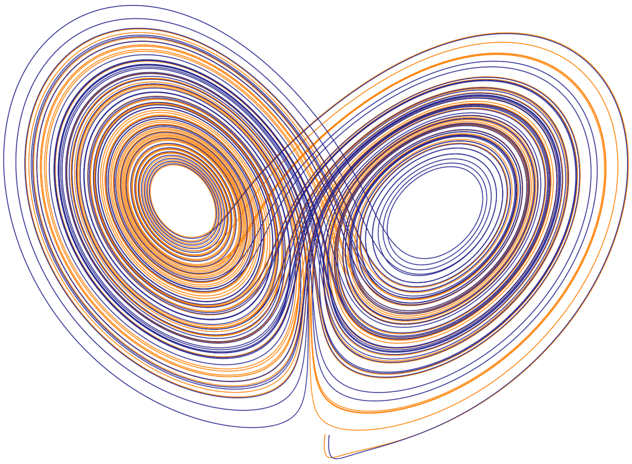
Computing covariant Lyapunov vectors

A convergence analysis of Ginelli's algorithm



A Dissertation submitted to the

Faculty of Mathematics, Informatics and Natural Sciences
Department of Mathematics of Universität Hamburg

in partial fulfillment of the requirements for the degree of

doctor rerum naturalium (Dr. rer. nat.)

by

Florian Noethen

Hamburg 2019

1. Reviewer	Prof. Dr. Reiner Lauterbach Department of Mathematics Universität Hamburg
2. Reviewer	Prof. Dr. Gary Froyland School of Mathematics and Statistics University of New South Wales
3. Reviewer	Dr. Cecilia González-Tokman School of Mathematics and Physics The University of Queensland
Supervisors	Prof. Dr. Reiner Lauterbach and Prof. Dr. Ingenuin Gasser
Date of disputation	March 4, 2020

Abstract

Covariant Lyapunov vectors (CLVs) are intrinsic modes that describe long-term linear perturbations of solutions of dynamical systems. Similar to eigenmode decompositions for steady states, they form a basis of the tangent space that links growth rates to directions invariant under the linearized flow along a given reference trajectory. In this sense CLVs play the role of eigenvectors for Lyapunov exponents.

Due to an increased interest in applications, several algorithms were developed to compute CLVs. The Ginelli algorithm is among the most commonly used. Although several properties of the algorithm have been analyzed, mathematical results are quite rare. This thesis combines first mathematically rigorous convergence results in an analysis of Ginelli's algorithm.

An important factor of our analysis is the multiplicative ergodic theorem, which provides existence of CLVs. We restrict our analysis to two different versions of the theorem, one for finite-dimensional and one for infinite-dimensional random dynamical systems. While the former assumes a fully invertible system, meaning that the base flow and the linear propagator are invertible, the latter only requires a semi-invertible setting in which the linear propagator may not be invertible. Using different approaches, we prove convergence of Ginelli's algorithm in these settings. The proof for finite dimensions links CLVs to singular vectors of the linear propagator and investigates so-called admissible input vectors. Through careful measure estimates of the set of admissible input vectors, we are able to prove convergence. Since estimates with respect to Lebesgue measure are not possible in infinite dimensions, we require different techniques to proof convergence. Among others, we derive an auxiliary result about the existence and the genericity of common complements for families of countably many subspaces.

The precise notion of convergence differs between discrete and continuous time. Namely, the discrete-time version of Ginelli's algorithm converges for almost every input, whereas the continuous-time version only converges in measure. Here, "almost everywhere" should be understood with respect to Lebesgue measure in finite dimensions and with respect to prevalence in infinite dimensions. In addition to the pure convergence statements, our theorems link the speed of convergence to Lyapunov exponents. It turns out that Ginelli's algorithm converges exponentially fast with a rate given by the spectral gap between Lyapunov exponents.

Zusammenfassung

Kovariante Lyapunov Vektoren (KLVs) sind intrinsische Modi, die das Langzeitverhalten von Störungen entlang von Lösungen dynamischer Systeme beschreiben. Ähnlich wie Eigenvektoren für stationären Lösungen formen KLVs eine Basis des Tangentialraums, welche den Wachstumsraten von Störungen Richtungen zuordnet, die invariant unter dem linearisierten Fluss entlang der Referenzlösung sind. In diesem Sinne kann man KLVs als eine Art von Eigenvektoren für Lyapunov Exponenten auffassen.

Aufgrund des hohen Interesses an KLVs bei Anwendern wurden mehrere Algorithmen zur Berechnung von KLVs entwickelt. Einer der meistgenutzten ist Ginellis Algorithmus. Obwohl einige Eigenschaften des Algorithmus bereits untersucht worden sind, sind mathematische Ergebnisse selten. Diese Dissertation verbindet erste mathematisch rigorose Ergebnisse zu einer Analyse von Ginellis Algorithmus.

Ein wichtiger Bestandteil der Analyse ist der Multiplikative Ergodensatz, der die Existenz von KLVs liefert. Wir beschränken unsere Analyse auf zwei Versionen des Satzes, eine für endlichdimensionale und eine für unendlichdimensionale dynamische Systeme mit Zufallsvariable. Während die erste Version annimmt, dass der grundlegende Fluss und der lineare Propagator invertierbar sind, kommt die zweite Version ohne die Annahme über Invertierbarkeit des linearen Propagators aus. Wir beweisen die Konvergenz von Ginellis Algorithmus in diesen Fällen mittels verschiedener Ansätze. Der Beweis im Endlichdimensionalen verbindet KLVs mit Singulärvektoren des linearen Propagators und untersucht sogenannte zulässige Inputvektoren. Durch vorsichtige Abschätzungen des Lebesgue-Maßes der Menge von zulässigen Vektoren sind wir in der Lage, die Konvergenz zu beweisen. Da solche Abschätzungen im Unendlichdimensionalen nicht möglich sind, brauchen wir für diesen Fall andere Mittel, um die Konvergenz zu zeigen. Unter anderem leiten wir ein Hilfsresultat über die Existenz und die Generizität von Unterräumen her, die komplementär zu einer Familie von abzählbar vielen gegebenen Unterräumen sind.

Die genau Form der Konvergenz unterscheidet sich im Falle diskreter und stetiger Zeit. Nämlich konvergiert die Version von Ginellis Algorithmus mit diskreter Zeit für fast jeden Input, während die Version mit stetiger Zeit nur im Maße konvergiert. Der Ausdruck “fast jeden” im Endlichdimensionalen ist im Sinne des Lebesgue-Maßes zu verstehen und im Unendlichdimensionalen im Sinne der Prävalenz. Zusätzlich zu den reinen Konvergenzaussagen verbinden unsere Resultate die Konvergenz-

geschwindigkeit mit den Lyapunov Exponenten. Es stellt sich heraus, dass Ginellis Algorithmus exponentiell schnell konvergiert mir einer Rate, die durch den spektralen Abstand zwischen Lyapunov Exponenten gegeben ist.

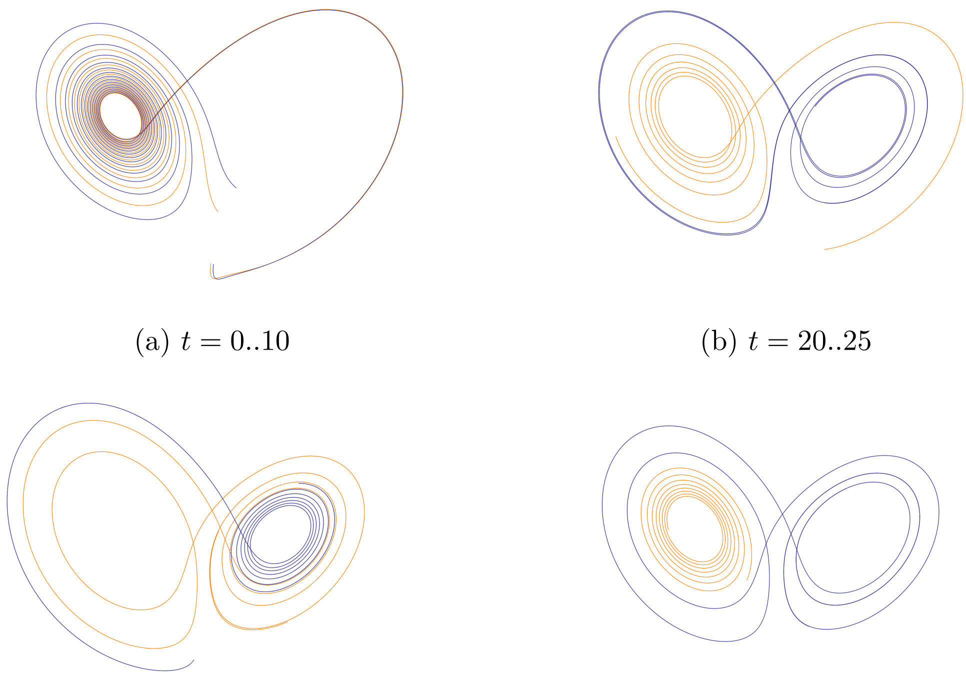
Contents

1 Introduction

"Does the flap of a butterfly's wings in Brazil set off a tornado in Texas?"
– Edward Norton Lorenz, 1972

During a conference in 1972 Lorenz argued that the atmosphere may be *unstable*, that is, the tiniest change, like the flap of a butterfly's wings, can drastically change the weather at a later point in time [41]. He came to this conclusion after two runs of the same weather simulation yielded completely different outcomes. The computed weathers coincided initially, but after two month there was seemingly no correlation anymore. While at first Lorenz suspected a problem with his computer, he soon discovered that the problem originated from the initial conditions. He had started the second run with round-off values instead of the precise initial conditions from the first run. This small perturbation caused an error that steadily doubled in size every few days of the simulation and ultimately resulted in two completely different weather forecasts [42] (see Fig. 1.1).

Such a behavior is not unique to Lorenz's model but can be found in various scenarios. A famous example is the double-rod pendulum [59, 60]. It consists of two pendulums such that the second pendulum is attached to the weight of the first. While the movement of a single pendulum is highly predictable, the double pendulum exhibits *chaotic* behavior. Even though the double pendulum moves according to a deterministic law, it is nearly impossible to predict its trajectory in experiments, as there is always a slight change in the initial position of the pendulum during the setup. Similar to Lorenz's model, the error caused by the initial perturbation grows exponentially until there is no visible correlation between the trajectories initiated with and without the perturbation. Even if we had a means to predict the future trajectory based on a given initial condition perfectly, a marginal input error would result in an entirely different trajectory than observed. Especially for long-term simulations it is imperative to understand how unavoidable input errors, e.g., due to the limiting accuracy of measurements, evolve in time and how to reduce them.



(a) $t = 0..10$ (b) $t = 20..25$

(c) $t = 45..50$ (d) $t = 60..65$

Figure 1.1: Two solutions of the Lorenz 63 model ($\dot{x}_1 = \sigma(x_2 - x_1)$, $\dot{x}_2 = x_1(\rho - x_3) - x_2$, $\dot{x}_3 = x_1x_2 - \beta x_3$; $\sigma = 10$, $\beta = 8/3$, $\rho = 28$), a simplified model for cellular convection, with slightly perturbed initial conditions computed in Maple 2016. The orange trajectory starts at $x(0) = (15, -15, 8)$ and the blue trajectory at $x(0) = (15.5, -15, 8)$. Both solutions were integrated until $t = 65$. While they are close initially, as it can be seen in (a), they eventually separate and visit different parts of the Lorenz attractor. Their full computed trajectories ($t = 0..65$) are displayed on the title page.

From a mathematical perspective the evolution of objects, such as the double pendulum, is described by *differential equations*. They relate the change in time of an object to the current state via an equation of the form

$$\dot{x} = f(t, x).$$

A solution $x(t)$ describing the state of an object at time t implicitly depends on states at other times via the above equation. In particular, under certain conditions on f, the evolution of an object is completely determined by its state at one point in time, just as the trajectory of a pendulum is determined the moment it is set into motion. We call such a state (t_0, x_0) *initial condition*. The corresponding solution should satisfy $x(t_0) = x_0$. To study in which way small perturbations of x_0 affect the solution we need the concept of *stability*.

1.1 Stability theory

Let us fix a reference trajectory $x(t)$ with respect to which we want to investigate stability and let $x(t)+v(t)$ be a perturbed solution with small error $v(t)$. We perform a first order approximation by linearizing f along x:

$$\dot{x} + \dot{v} = f(t, x + v) \approx f(t, x) + (D_x f)_{(t,x)}\, v.$$

The perturbation $v(t)$ approximately satisfies the *tangent linear equation*

$$\dot{v} = (D_x f)_{(t,x)}\, v.$$

By studying this equation we aim to understand more about the local behavior around our reference trajectory. For a fixed reference trajectory, the equation is a *matrix differential equation* of the form $\dot{v} = A(t)v$.

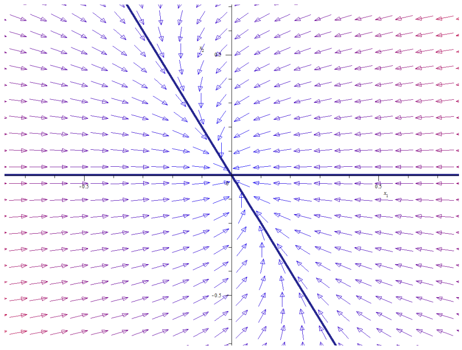
If our reference trajectory is a *steady state*, meaning that it does not change in time, then solutions of the tangent linear equation are of the form $v(t) = \exp(tA)v_0$. By analyzing eigenvalues and (generalized) eigenspaces of A, we can characterize the tangent linear dynamics (see Fig. 1.2). The eigenspaces form invariant subspaces for the linearized flow, and perturbations inside them grow or decay exponentially in size according to the real parts of the associated eigenvalues. In other words, we have

$$\lim_{t\to\infty} \frac{1}{t} \log \|v(t)\| = \lambda$$

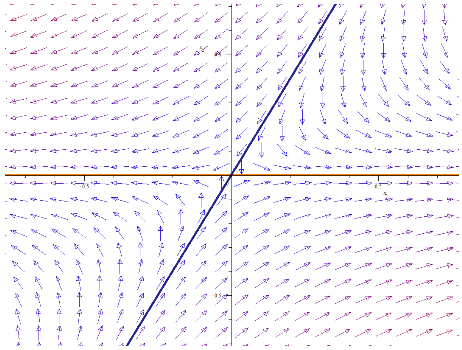
for a solution $v(t)$, where $v(0) = v_0$ is a (generalized) eigenvector of A corresponding to some eigenvalue with real part λ. The linearized system $\dot{v} = Av$ is called *(asymptotically) stable* if all eigenvalues have negative real parts. By the above observation, all solutions of a stable system decay exponentially fast and converge to the origin for $t \to \infty$. Contrariwise, if at least one eigenvalue has a positive real part, we call the system *unstable*. In this case almost every solution diverges from the origin for $t \to \infty$. Under rather general assumptions, stability properties of the tangent linear equation may be transferred to the original system according to the Hartman-Grobman theorem. It conjugates the dynamics via a local homeomorphism mapping the origin of the linear system to the steady state.

If our reference trajectory is *periodic* with period T, meaning that $T > 0$ is minimal with $x(t + T) = x(t)$ for all t, and if f does not depend on t explicitly, then $A(t)$ is T-periodic. According to Floquet theory, general solutions are of the form $v(t) = P(t)\exp(tB)v_0$, where $P(t)$ is invertible and T-periodic and B is a constant matrix. Since $\|P(t)\|$ and $\|P(t)^{-1}\|$ are bounded by constants, we may reduce the stability analysis to the (possibly complex) autonomous system $\dot{u} = Bu$. Notice that, unlike for steady states, the tangent vectors along the periodic reference trajectory are non-trivial and, hence, define a T-periodic solution of the tangent linear equation. In particular, B has at least one purely complex eigenvalue. Aside from this exception, we may deduce stability properties for the reference trajectory as in the case for steady states by looking at real parts of the eigenvalues of B. The corresponding eigenspaces lift to invariant bundles and define flow-invariant manifolds along the reference trajectory.

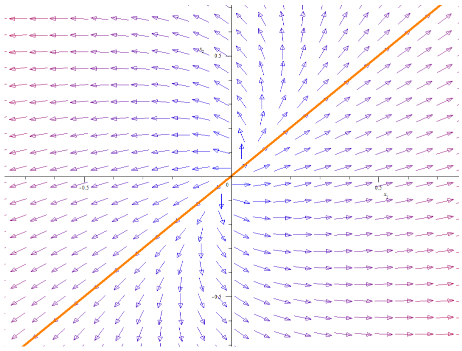
In general, the different rates of exponential growth or decay of linear perturbations are characterized by *Lyapunov exponents (LEs)*. If our reference solution is a steady state, they coincide with the real parts of the eigenvalues of A, while for periodic solutions they coincide with the real parts of the eigenvalues of B. Through LEs we hope to derive stability properties of the reference trajectory in the original system. However, note that it is a misconception that a negative largest LE indicates stability or that a positive largest LE indicates instability for general solutions



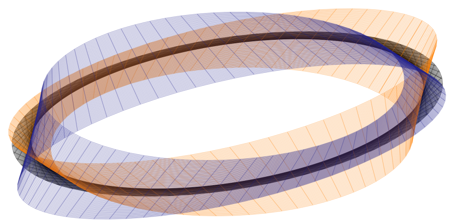
(a) $A = \begin{pmatrix} -3 & -1 \\ 0 & -1 \end{pmatrix}$



(b) $A = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}$



(c) $A = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix}$



(d) $A(t)$ periodic

Figure 1.2: A few examples depicting the linear stability of steady states and of periodic orbits. Images (a)-(c) show the flow of linear systems $\dot{v} = Av$. Blue lines indicate eigenspaces of negative eigenvalues, whereas orange lines stand for eigenspaces of positive eigenvalues. Only system (a) is stable. Systems (b) and (c) are unstable. Moreover, 2 is a double eigenvalue for (c). Its generalized eigenspace is the whole tangent space. (d) is a conceptual image of the tangent linear dynamics of a periodic solution. We see three bundles along the periodic orbit that are linearly independent at each point. The black bundle consists of tangent vectors of the trajectory.

of non-autonomous systems (see the Perron effect [38]). Nevertheless, for classes of random dynamical systems, which appear when the defining differential equation depends on an additional stochastic parameter, there is a local stability theory as for steady states and periodic orbits with probability 1. Indeed, for each LE, we find invariant bundles of the tangent linear flow that can be related to invariant manifolds via a stochastic version of the Hartman-Grobman theorem. A thorough treatment can be found in Arnold's book "Random Dynamical Systems" [1].

1.2 Covariant Lyapunov vectors

Along almost every trajectory of a random dynamical system, the bundles corresponding to different LEs constitute a splitting of the tangent space. We call this splitting *Oseledets splitting* and the associated spaces *Oseledets spaces* (see Fig. 1.3). Furthermore, we call a choice of normalized and covariant basis vectors subject to the splitting *covariant Lyapunov vectors (CLVs)*. Covariance means that CLVs are mapped to other CLVs by the linear propagator along trajectories up to normalizing factors. These factors grow or decay exponentially according to the corresponding LEs. Thus, CLVs generalize the concept of eigenmode decompositions of steady states to possibly chaotic trajectories and form an intrinsic basis of the tangent space that characterizes sensitivity to initial conditions.

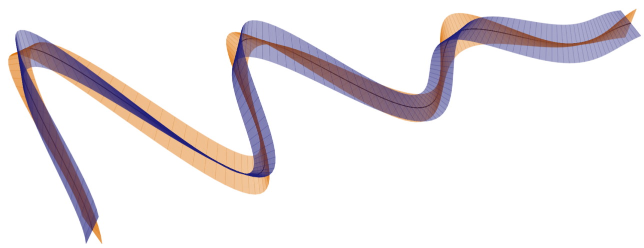


Figure 1.3: A conceptual picture of the tangent linear dynamics of a general trajectory. As in the case for periodic solutions, we imagine that the tangent space along the trajectory is split into invariant bundles consisting of different Oseledets spaces.

Due to their natural properties, CLVs have received strong resonance in applications during the last years. They have been described as the “physically relevant” modes in dissipative systems [61] and have been used to detect coherent structures, i.e., slow mixing sets, via the Perron-Frobenius operator [22, 23, 28], the dual of the Koopman operator. Recent research on coherent structures includes the analysis of large-scale features of the ocean and atmosphere relevant for climate [28, chapter 6]. Apart from techniques involving transfer operators, CLVs have been used directly to analyze instabilities in coupled models. Two examples are the assessment of long-term predictability in ocean-atmosphere models [15, 57, 66] and the decoupling of instabilities into modes associated to different timescales to analyze mixing in a two-scale Lorenz 96 model [12]. Other applications are the analysis of Navier-Stokes turbulence [32] and hard-disk systems [8, 9, 46, 64]. Furthermore, the angle between stable and unstable CLVs has been interpreted as a degree of hyperbolicity [13, 56, 68, 69] and has been used as an indicator for critical transitions in long-term behavior of solutions [4, 58].

On the theoretical side, the existence of CLVs and Oseledets spaces is ensured by the *multiplicative ergodic theorem (MET)*. While the original MET from 1968 is due to Oseledets [52], various versions emerged until today. They differ in their settings and proof techniques. Several versions follow Raghunathan’s approach [54], which uses a singular value decomposition (SVD) of the linear propagator and relates finite-time optimal growth rates given by singular values to LEs and singular vectors to CLVs. In [1] Arnold applies this approach to derive multiple METs. In particular, he proves an invertible and a non-invertible version of the theorem. The invertible version requires an invertible base flow and an invertible linear propagator. On the other hand, the non-invertible version requires neither of those. Though, this comes at a disadvantage. Instead of an Oseledets splitting, the non-invertible version only yields an *Oseledets filtration*. Consequently, there is no natural notion of CLVs.

Most METs can be categorized into invertible and non-invertible versions. However, recently there emerged a new kind of *semi-invertible* MET to adjust to the settings of transfer operators. It requires an invertible base flow while the linear propagator is allowed to be non-invertible. Semi-invertible METs still provide Oseledets splittings and, hence, also CLVs. The first semi-invertible MET was published by Froyland, Lloyd and Quas for finite-dimensional systems [22]. Semi-invertible versions for infinite-dimensional systems followed by Froyland, Lloyd, Quas and González-Tokman [23, 29, 30]. Since their METs are formulated on Banach spaces, they require new techniques that do not involve the use of SVDs. Instead, some form of compactness is assumed as in the first MET on Hilbert spaces by Ruelle [55] and in the first MET on Banach spaces by Mañé [44].

Aside from the mentioned METs, there are a lot more versions that developed since Oseledets' MET in 1968. We refer to two quite extensive historical overviews found in the introduction of [1, chapter 3] and in [28, section 4.2].

1.3 Ginelli's algorithm

In 1980 Benettin and others published an effective algorithm to compute LEs [5, 6]. The algorithm tracks the evolution of a set of randomly chosen initial perturbations by subsequently applying the linear propagator and an orthonormalization procedure. Orthonormalizing the propagated perturbations prevents that all vectors collapse onto the fastest growing direction and lets us compute growth rates of volumes of different dimensions. In particular, the i^{th} LE (counted with multiplicities) can be obtained as the difference between expansion rates of generic i-dimensional and $(i-1)$-dimensional volumes.

The propagated vectors in Benettin's algorithm converge to an orthonormal version of CLVs. Those vectors are sometimes referred to as *forward* or as *backward Lyapunov vectors* depending on which part of the trajectory is used to compute them. Since they encode either only the past or only the future of a trajectory, they lack critical information about the geometrical structure of the local flow. Contrary to CLVs, they are not covariant and depend on the chosen norm. Thus, CLVs form a much more natural basis of the local dynamics and are better suited for an extension of the linear stability theory of steady states and periodic orbits.

Effective algorithms to compute CLVs only appeared during the last years, e.g., in form of *Ginelli's algorithm* [26, 27], the algorithm by Wolfe and Samelson [67], the algorithm by Kuptsov and Parlitz [36], and others in [24]. Most algorithms either use a dynamical approach that combines propagation with some kind of orthonormalization procedure or rely on a SVD of the linear propagator, its inverse, or its adjoint if the linear propagator is not invertible. Additionally, the algorithm by Wolfe and Samelson computes CLVs as vectors lying in intersections of subspaces encoding past and future dynamics of the reference trajectory [67]. Moreover, CLV-algorithms differ in their choice of orthonormalization procedure. For example, the algorithm by Kuptsov and Parlitz uses an LU-factorization, whereas Ginelli's algorithm relies on a QR-decomposition.

Here, we focus on Ginelli's algorithm, which follows the dynamical approach. First, it propagates a set of randomly chosen vectors from the far past of a trajectory to the present while performing intermittent orthonormalizations. This part of the algorithm is not really different from the method by Benettin and others. We get approximations of LEs and of backward Lyapunov vectors. Next, the same procedure is continued to approximate backward Lyapunov vectors along the future part of the trajectory. Since the span of the first i backward Lyapunov vectors coincides with the span of the first i CLVs, we have an approximation of the span of the first i CLVs. Finally, the algorithm uses backward propagation from the far future to the present to extract the i^{th} CLV. Vectors for backward propagation are initialized inside spans of forward propagated vectors. Thus, backward propagation is restricted to approximately the span of the first i CLVs, which makes it possible to apply Ginelli's algorithm even in semi-invertible settings. Since the fastest growing direction in backward-time is the slowest growing direction in forward-time, that is, the i^{th} CLV, we have a means to compute CLVs. At least on a theoretical level, this is the basic idea behind Ginelli's algorithm.

Numerically, there are more subtleties to Ginelli's algorithm. For example, backward propagation is performed inside a coefficient space associated to forward-propagated vectors, and linear propagators on coefficient spaces are given by inverses of R-matrices from QR-decompositions obtained during the forward propagation. Other numerical features along with the original algorithm by Ginelli can be found in [26]. Another group of researchers around Froyland suggested an improvement of Ginelli's algorithm by choosing better suited initial vectors and compared different CLV-algorithms [24]. They showed that Ginelli's algorithm performs quite well for long time series and requires less memory than other compared algorithms during the three test cases in [24].

Despite their frequent use in applications, CLV-algorithms have received little attention from a viewpoint of rigorous mathematical analysis. First steps in this direction were done in [27]. Ginelli and others attempted a convergence proof of their algorithm by invoking a result of Ershov and Potapov [20], which handles convergence to backward Lyapunov vectors during the forward propagation of the algorithm. Ershov and Potapov performed their convergence analysis with respect to singular vectors of the linear propagator between the present and the far future and shifted their result to the propagator between the far past and the present. As runtime varies, the derived requirements on vectors initiated in the far past also vary. This plays an important role for the precise notion of convergence and has not been regarded in [20] and [27]. Moreover, the convergence proof in [27] assumes perfect convergence to backward Lyapunov vectors during the forward propagation when proving convergence of the backward propagation. Due to this, errors arising during the forward propagation do not carry over to the backward propagation. Those neglected errors come from projections of forward propagated vectors onto CLVs of higher order and form the fastest growing directions in backward-time. Hence, they have a potentially high influence on the speed of convergence. Indeed, [27] predicts an exponential speed of convergence to the i^{th} CLV proportional to the difference between the $(i-1)^{\text{th}}$ and i^{th} LEs. However, including errors from the forward propagation, it turns out that the exponential speed of convergence is actually proportional the spectral gap between the i^{th} LE and neighboring exponents [48]. In particular, it also depends on the $(i+1)^{\text{th}}$ LE.

1.4 This thesis

This thesis provides a rigorous mathematical analysis of Ginelli's algorithm in terms of convergence. We correct and expand upon the existing approach from [20, 27] to derive convergence theorems for Ginelli's algorithm in a wider class of systems. This includes theorems for finite-dimensional and a theorem for infinite-dimensional dynamical systems. While the former rely on a SVD of the propagator, as it appears in the MET from [1], the latter uses a different technique based on the semi-invertible MET from [29]. We try to be as general as possible so that the derived tools can be applied to other MET-settings as well.

Unlike in the previous attempt to prove convergence, our proofs are formulated for even degenerate Lyapunov spectra, which allow for higher multiplicities of LEs. Degeneracies naturally occur in, e.g., equivariant systems and result in internal dynamics of Oseledets spaces. Due to the internal dynamics, single outputs of the algorithm might not converge. Instead, one needs to regard convergence with respect to subspaces spanned by multiple output vectors.

Here, we give precise notions of convergence and relate the speed of convergence to the Lyapunov spectrum. As already predicted and observed in applications [20, 24, 27, 66], the algorithm convergences exponentially fast with a rate given by the spectral gap between LEs. We attribute different time-parameters to the past and future parts of the trajectory that are used in the algorithm. Thus, we allow for different amounts of past and future data during the convergence analysis. This might be helpful for future studies, when an optimal relation between past and future data in terms of convergence is sought. Such a relation clearly depends on the particular system and is not part of our general analysis here.

Three articles [48–50] were merged to create this thesis. [48] derives a projector-based convergence proof of Ginelli's algorithm for finite-dimensional dynamical systems, while the other two form a convergence proof on Hilbert spaces. Since both proofs require the definition of Ginelli's algorithm and heavily rely on versions of the MET, we devote the next two chapters to combine and compare relevant content from the aforementioned articles.

Chapter 2 introduces the class of random dynamical systems for which we state different versions of the MET. On the one hand, we present METs from [1] for finite-dimensional systems, on the other, we state the semi-invertible MET from [29], which treats systems on separable Banach spaces. In addition to the properties of the MET from [29], we need uniform bounds for growth rates of perturbations inside Oseledets spaces and inside spaces of the Oseledets filtration. Although such bounds are used in [29] and hints for their derivation are given, the actual proof is not carried out. We make up for the missing details by executing the suggested ideas in Appendix A.

Chapter 3 presents Ginelli's algorithm with special focus on its analytical kernel. A short discussion on the implementation and an application to the Lorenz 63 model is provided. Associated sample code can be found in Appendix B.

Chapters 4 and 5 contain the convergence analysis. In Chapter 4 we repeat the projector-based convergence proof from [48] for invertible finite-dimensional dynamical systems. It requires several tools. First, we introduce the *Lyapunov index notation* and define the space of subspaces, the Grassmannian of $\mathbb{R}^d$. Notions of distances and angles between subspaces in terms of projection operators are defined. Next, we emphasize on what we call *admissible tuples.* Those tuples represent special families of subspaces assuring that input vectors for Ginelli's algorithm stay close to singular vectors of the linear propagator. Singular vectors define directions of optimal growth or decay for finite time. Thus, they are linked with CLVs through a limit of finite-time scenarios. This link is described by the proof of the MET in [1] and plays a major role in our convergence analysis. We check alignment of propagated vectors with singular vectors in Ginelli's algorithm and derive estimates for forward and for backward propagation in terms of LEs. Ultimately, different parts of the algorithm are combined to form our convergence proof. It turns out that the precise notion of convergence differs between discrete and continuous time. Namely, the discrete-time version of Ginelli's algorithm converges for almost every input, whereas the continuous-time version only converges in measure - a slight difference that does not play a role in applications, but is important for a deeper understanding of the algorithm.

Chapter 5 presents the convergence proof from [49] for semi-invertible infinite-dimensional dynamical systems. The fundamental differences to the finite-dimensional versions are that estimates in terms of Lebesgue measure are no longer possible as there is no natural notion of Lebesgue measure for general Banach or Hilbert spaces, we do not have a SVD connected to the MET, and the linear propagator is not assumed to be invertible. In particular, we cannot simply obtain backward-time estimates via forward-time estimates of the time-reversed system. Moreover, the Lyapunov spectrum may consist of only a few, at most countably many exceptional LEs until a possibly non-discrete part of the spectrum is reached. This restricts our analysis to CLVs of the exceptional LEs.

As Chapter 4, we start Chapter 5 by introducing Grassmannians, this time for Banach spaces. We concentrate on closed complemented subspaces whose dimension or codimension is finite. Instead of input vectors for Ginelli's algorithm that stay close to singular vectors, we seek input vectors that stay far from spaces of the Oseledets filtration. To this end, we introduce the notion of *well-separating common complements*. Those are common complements for families of subspaces of finite codimension such that the *degree of transversality*, which describes the separation between complementary subspaces, decays at most subexponentially. In our proof the family of subspaces is given by spaces of the Oseledets filtration for different initial times. Following [50], we show that well-separating common complements are prevalent in Hilbert spaces. *Prevalence* is a generalized concept of "Lebesgue almost everywhere" for infinite-dimensional vectors spaces. Even though we may not be able to perform direct measure estimates, we use the concept of prevalence to prove convergence of Ginelli's algorithm for almost every choice of input vectors. The analysis is split into estimates for forward- and for backward-time. Finally, we assemble the derived estimates to form our convergence proof.

Chapter 6 summarizes the most important findings and discusses their implications.

2

Multiplicative Ergodic Theorem

The *multiplicative ergodic theorem (MET)* describes the asymptotic nature of linear perturbations along trajectories of dynamical systems. Its main contribution is the existence of a filtration or a splitting of the tangent space that relates directions to asymptotic growth rates via *Lyapunov exponents (LEs)*. Noninvertible versions of the MET prove existence of *Oseledets filtrations*, whereas invertible and more recent semi-invertible versions yield *Oseledets splittings* and *covariant Lyapunov vectors (CLVs)*.

As already mentioned in Section 1.2, there has been much development concerning METs since Oseledets' original version from 1968 [52]. The main objective of this chapter is to introduce different scenarios under which the MET guarantees an Oseledets splitting and CLVs. We mainly follow the book of Arnold on random dynamical systems [1], which gives a comprehensive treatment of METs and associated constructions on $\mathbb{R}^d$, and an article by González-Tokman and Quas deriving a semi-invertible MET on separable Banach spaces [29]. Both require the notion of random dynamical systems.

2.1 Random dynamical systems

Most METs are formulated in the context of random dynamical systems. Those systems arise from ordinary differential equations (ODEs) that depend on an additional stochastic parameter. Here, we take a purely dynamical approach. A more in-depth analysis of the relation between random differential equations and random dynamical systems can be found in Arnold's book [1].

We consider different cases of time for random dynamical systems: discrete one- and two-sided time $\mathbb{T} \in \{\mathbb{N}_0, \mathbb{Z}\}$ and continuous one- and two-sided time $\mathbb{T} \in \{\mathbb{R}_{\geq 0}, \mathbb{R}\}$. In [1] Arnold derives versions of the MET for all of these cases. In particular, he derives an Oseledets filtration for one-sided time and an Oseledets splitting for two-sided time. Since we want to compute Oseledets spaces, the analysis in Chapter 4 will be restricted to two-sided time. The MET on Banach spaces from [29] and our analysis in Chapter 5 require two-sided discrete time.

Definition 2.1.1 ([1])
A measurable dynamical system *consists of a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *with* $\Omega \neq \{0\}$ *and a family of transformations* $(\sigma_t)_{t\in\mathbb{T}}$ *on* Ω *such that*

1. $(t, \omega) \mapsto \sigma_t\omega$ *is measurable, where* $\mathbb{T}\times\Omega$ *is endowed with the product* σ*-algebra* $\mathcal{B}(\mathbb{T}) \times \mathcal{F}$*,*

2. $\sigma_0 = \mathrm{id}_\Omega$*, and*

3. $\sigma_{s+t} = \sigma_s \circ \sigma_t$ *for all* $s, t \in \mathbb{T}$ *((semi-)flow property).*

If additionally σ *preserves probabilities, i.e.,* $\mathbb{P}(\sigma_t^{-1}A) = \mathbb{P}(A)$ *for* $A \in \mathcal{F}$ *and* $t \in \mathbb{T}$*, then* $(\Omega, \mathcal{F}, \mathbb{P}, (\sigma_t)_{t\in\mathbb{T}})$ *is called* metric dynamical system*. Furthermore, such a system is called* ergodic *if all measurable, flow-invariant subsets of* Ω *have either probability* 0 *or* 1*.*

If we assume two-sided time, the flow property implies that transformations are invertible with $\sigma_t^{-1} = \sigma_{-t}$. Moreover, in the discrete case the flow is generated by its time-one-map $\sigma := \sigma_1$. It holds $\sigma_n = \sigma^n$ for every $n \in \mathbb{Z}$. Consequently, we may interchange the family $(\sigma_n)_{n\in\mathbb{Z}}$ with its *generator* σ while requiring measurability of σ and σ^{-1}.

A very useful result when constructing metric dynamical systems on compact metric spaces is the famous theorem by Kryloff and Bogoliouboff [35], which ensures the existence of a probability measure invariant under a given continuous group action of a family $(\sigma_t)_{t\in\mathbb{T}}$ on Ω. In particular, smooth autonomous ODEs of the general form $\dot{\omega} = f(\omega)$ on compact manifolds induce metric dynamical systems.

To analyze local dynamics near solutions of ODEs one usually studies linearizations. The evolution of linear perturbations is described by the tangent linear equation. Solving the tangent equation along a fixed reference trajectory, we get a family of linear operators $\mathcal{L}_\omega^{(t)} : \mathbb{R}^d \to \mathbb{R}^d$ (also called *linear propagators*) evolving linear perturbations. These operators satisfy the *cocycle property*, which we define more abstractly using the notion of *random dynamical systems*:

Definition 2.1.2 ([1])
Let $(\Omega, \mathcal{F}, \mathbb{P}, (\sigma_t)_{t\in\mathbb{T}})$ *be a metric dynamical system. A* (linear) cocycle (over σ) *is a measurable map*

$$\begin{aligned} \mathcal{L} : \mathbb{T} \times \Omega \times \mathbb{R}^d &\to \mathbb{R}^d \\ (t, \omega, x) &\mapsto \mathcal{L}_\omega^{(t)} x, \end{aligned}$$

such that $\mathcal{L}_\omega^{(t)} : \mathbb{R}^d \to \mathbb{R}^d$ *is linear,* $\mathcal{L}_\omega^{(\cdot)} : \mathbb{T} \times \mathbb{R}^d \to \mathbb{R}^d$ *is continuous and* $\mathcal{L}$ *satisfies the cocycle property:*

1. $\mathcal{L}_\omega^{(0)} = I$*,*

2. $\mathcal{L}_\omega^{(s+t)} = \mathcal{L}_{\sigma_t\omega}^{(s)} \circ \mathcal{L}_\omega^{(t)}$ *for all* $s, t \in \mathbb{T}$*.*

We call $(\Omega, \mathcal{F}, \mathbb{P}, (\sigma_t)_{t\in\mathbb{T}}, \mathcal{L})$ *a* (linear) random dynamical system *with time* $\mathbb{T}$*. Similarly, we may define a cocycle for one-sided time by restricting the cocycle property.*

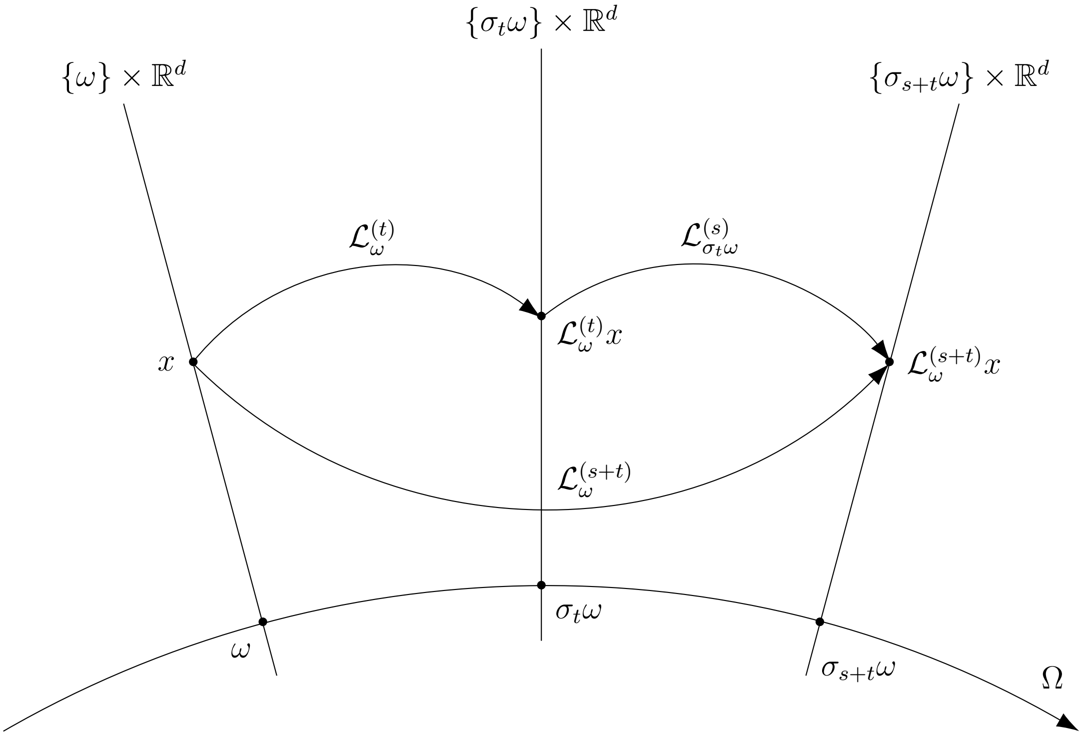


Figure 2.1: A random dynamical system can be seen as an action on the bundle $\Omega \times \mathbb{R}^d$.

The special case of one-sided random dynamical systems defined over a two-sided base σ is sometimes referred to as *semi-invertible* (first occurrence of the term in [23]). The term semi-invertible is due to σ being invertible, while the action of the cocycle via $\mathcal{L}_\omega^{(t)}$ may not be invertible. On the other hand, if we have a two-sided random dynamical system, the cocycle property implies that $\mathcal{L}_\omega^{(t)}$ is invertible with $(\mathcal{L}_\omega^{(t)})^{-1} = \mathcal{L}_{\sigma_t\omega}^{(-t)}$.

In a discrete-time setting the cocycle is *generated* by $\mathcal{L}_\omega := \mathcal{L}_\omega^{(1)}$ via $\mathcal{L}_\omega^{(n)} = \mathcal{L}_{\sigma_{n-1}\omega} \circ \cdots \circ \mathcal{L}_\omega$. Vice versa, each measurable map $\mathcal{L} : \Omega \to \mathbb{R}^{d\times d}$ generates a one-sided cocycle. If $\mathcal{L}_\omega$ is invertible for every ω such that $\mathcal{L}_\omega^{-1}$ depends measurably on ω, then we get a two-sided cocycle by the above observation. Thus, we may interchange cocycles with their generators.

There are several constructions to obtain new cocycles. The most important for us is the *time-reversed cocycle*, which we get in fully invertible settings. Via $\mathcal{L}_\omega^{-,(t)} := \mathcal{L}_\omega^{(-t)}$, we get a new cocycle over $\sigma_t^- := \sigma_{-t}$ describing the evolution in reversed time. Later on, we will use the associated time-reversed random dynamical system to obtain estimates for backward-time via estimates for forward-time. Other important constructions include cocycles for exterior powers $\wedge^k \mathcal{L}$, restrictions to subbundles, and cocycles on quotient spaces. Moreover, there are cocycles for manifolds and for groups [1, chapters 4-6]. In the second chapter of his book, Arnold even derives random dynamical systems from pathwise random differential equations of the form $\dot{x} = f(\sigma_t\omega, x)$ and, vice versa, he constructs random differential equations via $f(\omega, x) := d/dt(\mathcal{L}_\omega^{(t)}x)|_{t=0}$ given differentiability of the cocycle.

Random dynamical systems on Banach spaces can be constructed similarly to random dynamical systems on $\mathbb{R}^d$. We require the same cocycle properties as in

Definition 2.1.2, but we need additional assumptions for the semi-invertible MET from [29]. Since the analysis on Banach spaces is only for discrete time, we exchange base and cocycle with their generators.

Definition 2.1.3 ([39])
Let $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ be a metric dynamical system. A tuple $(\Omega, \mathcal{F}, \mathbb{P}, \sigma, X, \mathcal{L})$ with a separable Banach space X and a strongly measurable generator $\mathcal{L} : \Omega \to L(X)$ is called separable strongly measurable random dynamical system*. We say a map $\mathcal{L} : \Omega \to L(X)$ is* strongly measurable[1] *if it is measurable with respect to the strong σ-algebra $\mathcal{S}$ on $L(X)$. $\mathcal{S}$ is generated by sets of the form $W_{x,U} := \{T \mid Tx \in U\}$ for $U \subset X$ open.*

In [29, appendix A] it is shown that the strong σ-algebra is the Borel σ-algebra of the strong operator topology of X. In particular, it is a subset of the Borel σ-algebra $\mathcal{B}_X$ induced by the norm topology. Thus, being strongly measurable is a weaker requirement than being measurable. On the other hand, X being separable is a strong restriction.

2.2 METs on $\mathbb{R}^d$

The METs from [1] rely on a SVD of the cocycle. While Arnold connects singular values to LEs, he derives an Oseledets filtration as a limit of spaces spanned by singular vectors. His first MET is a deterministic version in which he assumes restrictions on growth rates of perturbations along the reference trajectory. Then, Arnold proceeds to show that the assumptions of the deterministic MET are satisfies for almost every trajectory of random dynamical systems with a certain integrability condition leading him to a MET for one-sided time. By applying the one-sided MET to the reversed cocycle, he obtains a second Oseledets filtration, which, when intersected with the first, yields an Oseledets slitting resulting in a MET for two-sided time. In this section we present these METs. For completeness' sake, let us also mention the semi-invertible MET on $\mathbb{R}^d$ [22], which yields an Oseledets filtration even for non-invertible linear propagators.

Later on, it will be convenient to group indices of tuples according to degeneracies of the Lyapunov spectrum. Let us assume degeneracies $d_i \geq 1$ with $d_1 + \cdots + d_p = d$. The case $p = d$ is called *nondegenerate*. We denote a d-tuple of vectors of $\mathbb{R}^d$ by

$$(x) := \Big(x_{1_1}, x_{1_2}, \dots, x_{1_{d_1}}, x_{2_1}, \dots, x_{2_{d_2}}, \dots, x_{p_{d_p}}\Big).$$

The usual index is related to our new notation via $x_{i_j} = x_{d_1 + \cdots + d_{i-1} + j}$. To further shorten notation, we write (Ax) for the d-tuple of vectors we get by applying a linear map A to each vector of (x).

Before stating the deterministic MET, we recall a few preliminary facts about SVDs.

[1] Equivalently, we could require that $\mathcal{L}_{(.)}x : \Omega \to X$ is $(\mathcal{F}, \mathcal{B}_X)$-measurable for every $x \in X$ [29, appendix A].

2.2.1 SVD

Definition 2.2.1
Let $A \in \mathbb{R}^{d\times d}$. The singular value decomposition (SVD) *of A is given by*

$$A = U\Sigma V^T,$$

where $\Sigma = \operatorname{diag}(\delta_{1_1}, \dots, \delta_{p_{d_p}})$ is the diagonal matrix of singular values $\delta_{i_j} \geq 0$ *and $U, V \in O(d, \mathbb{R})$ are orthogonal matrices. The columns $(u) := (Ue)$ of U are called* left singular vectors *and the columns $(v) := (Ve)$ of V are called* right singular vectors*. Here, (e) denotes the standard basis of $\mathbb{R}^d$.*

A connection between left and right singular vectors is established via

$$Av_{i_j} = \delta_{i_j} u_{i_j}.$$

In general the SVD is not unique. Given a linear map A we settle for a descending ordering of singular values:

$$\delta_{1_1} \geq \dots \geq \delta_{p_{d_p}} \geq 0. \tag{2.1}$$

When applied to cocycles, each group of singular values approximates a different LE. Hence, if the approximations are good enough, the inequalities between $\delta_{i_{d_i}}$ and $\delta_{(i+1)_1}$ are strict. In this case the spaces spanned by singular vectors of one group, i.e., $\operatorname{span}(u_{i_1}, \dots, u_{i_{d_i}})$ and $\operatorname{span}(v_{i_1}, \dots, v_{i_{d_i}})$, are uniquely determined independent of our choice of SVD with Eq. (2.1).

A SVD $\hat{U}\hat{\Sigma}\hat{V}^T$ for the inverse of some invertible linear map A is obtained by inverting $A = U\Sigma V^T$ and, heeding Eq. (2.1), reversing the order of singular values and of singular vectors. In other words, a SVD for the inverse is given by $(\hat{\delta}) = (1/\delta)^r$, $\hat{u} = (v)^r$, and $\hat{v} = (u)^r$ with $(\cdot)^r$ being the tuple in reversed order.

For convenience sake, we denote the smallest and the largest singular values in each group by

$$\delta_i^{\min} := \min_{j=1,\dots,d_i} \delta_{i_j} \quad \text{and} \quad \delta_i^{\max} := \max_{j=1,\dots,d_i} \delta_{i_j}.$$

2.2.2 Deterministic MET

Fix $\omega \in \Omega$. The deterministic MET from [1] only requires a sequence of matrices generating a cocycle. Here, we directly substitute the sequence of matrices by the generator on the positive part of the discrete orbit of ω. Hence, we may regard the assumptions on the sequence of matrices as assumptions on ω. In a nutshell, the deterministic MET assumes that changes during a single timestep do not matter on an exponential scale and that expansion rates of different volumes are well-defined and do not exceed exponential scales.

Proposition 2.2.2 (Deterministic MET for $\mathbb{T} = \mathbb{N}_0$ [1])
Assume that the cocycle along the positive part of the discrete orbit of ω satisfies

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_{\sigma_n\omega}\| \leq 0$$

and that the limit

$$\lim_{n\to\infty} \frac{1}{n} \log \| \wedge^k \mathcal{L}_\omega^{(n)}\| \in \mathbb{R} \cup \{-\infty\}$$

exists for all orders of the wedge product of $\mathcal{L}_\omega^{(n)}$*.*

Then, there exists a Lyapunov spectrum with a corresponding filtration of subspaces capturing different asymptotic growth rates:

1. *The* Lyapunov spectrum *consists of* Lyapunov exponents (LEs)

$$\infty > \lambda_1 > \cdots > \lambda_p \geq -\infty,$$

which are the distinct limits of singular values, together with degeneracies $d_1 + \cdots + d_p = d$*:*

$$\forall i_j : \lim_{n\to\infty} \frac{1}{n} \log \delta_{i_j}\left(\mathcal{L}_\omega^{(n)}\right) = \lambda_i.$$

2. *There is a filtration*

$$\mathbb{R}^d = V_1 \supset \cdots \supset V_p \supset V_{p+1} := \{0\}$$

called Oseledets filtration *given by subspaces*

$$V_i := \left\{ x \in \mathbb{R}^d \;\middle|\; \lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_\omega^{(n)} x\| \leq \lambda_i \right\}.$$

Limits in the definition of V_i *exist for all* $x \in \mathbb{R}^d$ *and take values in* $\{\lambda_1, \dots, \lambda_p\}$*. Moreover, it holds*

$$\dim V_i - \dim V_{i+1} = d_i.$$

The proof shows that $\Psi := \lim_{n\to\infty}((\mathcal{L}_\omega^{(n)})^T \mathcal{L}_\omega^{(n)})^{1/2n}$ exists. Its eigenvalues are of the form e^{λ_i} and its eigenspaces U_i form the Oseledets filtration via $V_i = U_p \oplus \cdots \oplus U_i$. In particular, the spaces spanned by right singular vectors of $\mathcal{L}_\omega^{(n)}$ converge to the eigenspaces of Ψ. The convergence is exponentially fast with a rate given by the distance between associated LEs. A more technical view on the link between MET and SVD will be given in Section 4.4.2, when proving convergence of Ginelli's algorithm on $\mathbb{R}^d$.

The eigenvectors of Ψ are sometimes referred to as *forward Lyapunov vectors*, since they encode information about the forward part of the trajectory. As Ψ is symmetric, forward Lyapunov vectors corresponding to different LEs are orthogonal. Our analysis in Chapter 4 shows that these vectors can be obtained by pushing back and orthonormalizing a set of randomly chosen basis vectors. Similarly, one gets the so-called *backward Lyapunov vectors* via to the time-reversed cocycle. In particular, backward Lyapunov vectors are associated to push-forwards of vectors along the past of the trajectory. This plays an important role in Ginelli's algorithm.

Let us remark that, while Ψ, U_i, and forward and backward Lyapunov vectors depend on the chosen scalar product and norm, the LEs and the Oseledets filtration are independent of the norm. Indeed, every norm on $\mathbb{R}^d$ is equivalent. Hence, changing the norm contributes an at most constant factor that vanishes in asymptotics on exponential scales. Since LEs and the Oseledets filtration have asymptotic characterizations, they remain the same when changing the norm. Thus, without loss of generality, we will settle for the euclidean norm in Chapter 4.

Adding two more assumptions to Proposition 2.2.2, namely invertibility and an extension of the first assumption, we get a version of the deterministic MET for one-sided continuous time:

Corollary 2.2.3 (Deterministic MET for $\mathbb{T} = \mathbb{R}_{\geq 0}$ [1])
Assume that the action of the one-sided continuous cocycle is invertible, satisfies

$$\limsup_{n\to\infty} \frac{1}{n} \sup_{s\in[0,1]} \log \|(\mathcal{L}^{(s)}_{\sigma_n\omega})^{\pm 1}\| \leq 0,$$

and that limits of the wedge product of $\mathcal{L}^{(t)}_\omega$ exist as in Proposition 2.2.2.

Then, the Lyapunov spectrum and the Oseledets filtration exist and their asymptotic characterizations hold with continuous instead of discrete time.

Proof. This statement is part of the proof of the MET for one-sided time in [1]. The additional assumptions and the cocycle property ensure that nothing happens in-between discrete timesteps. □

Without many additional assumptions we can extend the results to the whole trajectory:

Corollary 2.2.4
In the setting of Proposition 2.2.2 for cocycles with invertible action, the Lyapunov spectrum and the Oseledets filtration exist along the whole trajectory. Furthermore, $p(\omega)$, $\lambda_i(\omega)$ and $d_i(\omega)$ are invariant under σ, and the Oseledets filtration changes in an equivariant way:

$$\mathcal{L}^{(t)}_\omega V_i(\omega) = V_i(\sigma_t\omega).$$

Similar statements hold for continuous time in the setting of Corollary 2.2.3 if we additionally require that

$$\forall\, v \in \mathbb{R}_{\geq 0} : \sup_{s\in[0,1]} \|(\mathcal{L}^{(s)}_{\sigma_v\omega})^{\pm 1}\| < \infty. \tag{2.2}$$

Proof. The first assumption of Proposition 2.2.2 is trivially satisfied if we replace ω by $\sigma_u\omega$. To prove existence of limits of wedge products, we use the following properties:

1. $\| \wedge^k A\| = \delta_1 \ldots \delta_k$,
2. $\wedge^k I = I$, and
3. $\wedge^k(AB) = (\wedge^k A)(\wedge^k B)$

for $A, B \in \mathbb{R}^{d\times d}$. Now, the existence of

$$\lim_{n\to\infty} \frac{1}{n} \log \| \wedge^k \mathcal{L}^{(n)}_{\sigma_u\omega}\| < \infty$$

follows due to the cocycle property:

$$\begin{aligned}
&\Big(\frac{n+u}{n}\Big)\Big(\frac{1}{n+u} \log \| \wedge^k \mathcal{L}^{(n+u)}_\omega\|\Big) - \frac{1}{n} \log \| \wedge^k \mathcal{L}^{(u)}_\omega\| \\
&\leq \frac{1}{n} \log \| \wedge^k \mathcal{L}^{(n)}_{\sigma_u\omega}\| \\
&\leq \Big(\frac{n+u}{n}\Big)\Big(\frac{1}{n+u} \log \| \wedge^k \mathcal{L}^{(n+u)}_\omega\|\Big) + \frac{1}{n} \log \|(\wedge^k \mathcal{L}^{(u)}_\omega)^{-1}\|.
\end{aligned}$$

Thus, applying the proposition, we get existence of the Lyapunov spectrum and of the Oseledets filtration at $\sigma_u\omega$. In particular, the above shows that the limits of singular values for ω and for $\sigma_u\omega$ coincide on exponential scales. Hence, LEs and their multiplicities are the same for all states along the orbit of ω. Finally, equivariance of the Oseledets filtration follows from its asymptotic characterization.

The extra condition of Corollary 2.2.3 can be derived from

$$\mathcal{L}^{(s)}_{\sigma_{n+u}\omega} = \mathcal{L}^{(u)}_{\sigma_{s+n}\omega} \circ \mathcal{L}^{(s)}_{\sigma_n\omega} \circ \left(\mathcal{L}^{(u)}_{\sigma_n\omega}\right)^{-1}$$

by using Eq. (2.2) and similar inequalities as for the wedge product above. The rest follows as in the discrete case. □

Assuming two-sided time, similar statements can be derived for the time-reversed cocycle $\mathcal{L}^-$ over σ^-. We denote its Lyapunov spectrum by $(\lambda_i^-, d_i^-)_{i=1,\ldots,p^-}$ and the corresponding filtration spaces by $V_i^-(\omega)$.

In order to define an equivariant splitting of the tangent space that captures asymptotic growth rates in both forward- and backward-time, we require additional assumptions on the Lyapunov spectra and on the associated filtrations of $\mathcal{L}$ and $\mathcal{L}^-$:

1. $p = p^-$, $d_i^- = d_{p+1-i}$, $\lambda_i^- = -\lambda_{p+1-i}$, and
2. $V_{i+1}(\omega) \cap V^-_{p+1-i}(\omega) = \{0\}$.

A direct consequence is the finiteness of LEs. We set $\lambda_0 := \infty$ and $\lambda_{p+1} := -\infty$ for convenience sake.

Proposition 2.2.5 (Deterministic MET for two-sided time [1])
Assuming the above relations between the Lyapunov spectra and the Oseledets filtrations of $\mathcal{L}$ and $\mathcal{L}^-$, there exists a splitting $\mathbb{R}^d = Y_1(\omega) \oplus \cdots \oplus Y_p(\omega)$, called Oseledets splitting, *of the tangent space into so-called* Oseledets spaces

$$Y_i(\omega) := V_i(\omega) \cap V^-_{p+1-i}(\omega). \tag{2.3}$$

Oseledets spaces can be characterized via

$$y \in Y_i(\omega) \setminus \{0\} \qquad \Longleftrightarrow \qquad \lim_{t\to\pm\infty} \frac{1}{|t|} \log \|\mathcal{L}^{(t)}_\omega y\| = \pm\lambda_i, \tag{2.4}$$

where convergence is uniform with respect to y in the unit sphere of $Y_i(\omega)$. Furthermore, they are equivariant:

$$\mathcal{L}^{(t)}_\omega Y_i(\omega) = Y_i(\sigma_t\omega),$$

and satisfy $\dim Y_i(\omega) = d_i$.

Proof. The proof is purely algebraic and can be found along the lines of the proof of the MET for two-sided time in [1]. □

The Oseledets filtrations of $\mathcal{L}$ and $\mathcal{L}^-$ can be reconstructed from the Oseledets splitting via

$$V_i(\omega) = \bigoplus_{j=i}^{p} Y_j(\omega) \quad \text{and} \quad V^-_{p+1-i}(\omega) = \bigoplus_{j=1}^{i} Y_j(\omega).$$

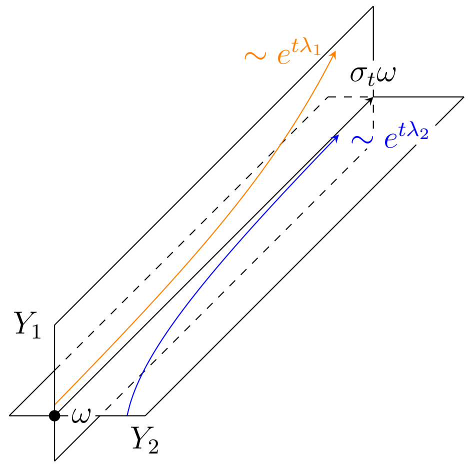


Figure 2.2: Linear asymptotics of a diagonal cocycle in dimension 2.

For random dynamical systems satisfying a certain integrability condition, the cocycle along almost every trajectory of the system admits an Oseledets splitting. We state the corresponding METs in the next subsection. Moreover, in an ergodic setting the Lyapunov spectrum is constant $\mathbb{P}$-almost everywhere. Therefore, in applications it is often assumed that the underlying system is ergodic at least near an interesting structure.[2] Via CLVs one hopes to better understand the local flow near such a structure.

Definition 2.2.6
Normalized basis vectors, which are covariant and chosen subject to the Oseledets splitting, are called covariant Lyapunov vectors (CLVs)*. We call a family of unit vectors* $u : \Omega \to \mathbb{R}^d$ covariant *if* $\mathcal{L}_\omega^{(t)} u(\omega)$ *and* $u(\sigma_t\omega)$ *coincide up to normalization for almost every* ω *and for all* $t \in \mathbb{T}$*.*

By Eq. (2.4) CLVs form a basis that describes different asymptotic rates of growth or decay. This is a direct consequence of them being chosen subject to the Oseledets splitting. Moreover, since Oseledets filtrations and thus also Oseledets splittings are norm-independent, changing the norm does not change the directions of CLVs. Thus, unlike the orthogonal forward and backward Lyapunov vectors, CLVs are norm-independent and not necessarily orthogonal.

2.2.3 METs for random dynamical systems

In Chapter 4 we will show convergence of Ginelli's algorithm to compute CLVs, or more generally Oseledets spaces, in the deterministic setting. Our proof only requires a sequence of matrices with the restrictions of Proposition 2.2.5 which ensures existence and uniqueness of Oseledets spaces. However, for the sake of completeness we now state the main METs from [1].

Assuming an integrability condition on the generator, the MET for random dynamical systems provides a Lyapunov spectrum and an Oseledets filtration for almost every trajectory:

[2] See the concept of SRB-measures for attractors [10].

Theorem 2.2.7 (MET for one-sided time [1])
Let $(\Omega, \mathcal{F}, \mathbb{P}, (\sigma_t)_{t\in\mathbb{T}_{\geq 0}}, \mathcal{L})$ be a random dynamical system. The following statements hold:

(A) *Non-invertible case $\mathbb{T} = \mathbb{N}_0$: If the generator $\mathcal{L} : \Omega \to \mathbb{R}^{d\times d}$ satisfies*

$$\log^+ \|\mathcal{L}\| \in L^1(\Omega, \mathcal{F}, \mathbb{P}),$$

where $\log^+(\cdot) := \max(\log(\cdot), 0)$, then a Lyapunov spectrum and an Oseledets filtration as in Proposition 2.2.2 exist on a σ-invariant set of full $\mathbb{P}$-measure. They depend measurably on ω. Moreover, the Lyapunov spectrum is invariant under σ and the Oseledets filtration is equivariant, i.e., it satisfies $\mathcal{L}_\omega V_i(\omega) \subset V_i(\sigma\omega)$.

If the underlying metric dynamical system is ergodic, then the Lyapunov spectrum is constant $\mathbb{P}$-almost everywhere.

(B) *Invertible case $\mathbb{T} = \mathbb{N}_0$: If the generator $\mathcal{L} : \Omega \to Gl(d, \mathbb{R})$ satisfies*

$$\log^+ \|\mathcal{L}^{\pm 1}\| \in L^1(\Omega, \mathcal{F}, \mathbb{P}),$$

then, in addition to (A), the smallest LE is real and equivariance of the filtration holds with equality: $\mathcal{L}_\omega V_i(\omega) = V_i(\sigma\omega)$.

(C) *Invertible case $\mathbb{T} = \mathbb{R}_{\geq 0}$: Let $\mathcal{L}^{(t)}_\omega \in Gl(d, \mathbb{R})$. Assume that*

$$\sup_{s\in[0,1]} \log^+ \|(\mathcal{L}^{(s)}_\omega)^{\pm 1}\| \in L^1(\Omega, \mathcal{F}, \mathbb{P}).$$

Then, all statements of (B) hold with $\mathbb{N}_0$ replaced by $\mathbb{R}_{\geq 0}$.

The following two-sided MET gives us existence of an Oseledets splitting for almost every trajectory. Remember that actions of cocycles over two-sided time are invertible automatically.

Theorem 2.2.8 (MET for two-sided time [1])
Let $(\Omega, \mathcal{F}, \mathbb{P}, (\sigma_t)_{t\in\mathbb{T}}, \mathcal{L})$ be a random dynamical system. The following statements hold:

(A) *$\mathbb{T} = \mathbb{Z}$: If the generator satisfies*

$$\log^+ \|\mathcal{L}^{\pm 1}\| \in L^1(\Omega, \mathcal{F}, \mathbb{P}),$$

then there is a σ-invariant set of full $\mathbb{P}$-measure on which (B) of Theorem 2.2.7 holds and on which an Oseledets splitting as in Proposition 2.2.5 exists. Moreover, the Oseledets splitting depends measurably on ω.

(B) *$\mathbb{T} = \mathbb{R}$: Assume that*

$$\sup_{s\in[0,1]} \log^+ \|(\mathcal{L}^{(s)}_\omega)^{\pm 1}\| \in L^1(\Omega, \mathcal{F}, \mathbb{P}).$$

Then, all statements of (A) hold with $\mathbb{Z}$ replaced by $\mathbb{R}$.

In [1] Arnold also derives versions of the MET for various constructions of cocycles. Those include time-reversed and adjoint cocycles, exterior powers, tensor products, linear subbundles, quotient spaces, and manifolds.

2.3 Semi-invertible MET on separable Banach spaces

The first MET on Hilbert spaces was published by Ruelle in 1982 [55]. One year later, Mañé's paper generalized the MET to Banach spaces [44]. Extending Thieullen's work [62], the first infinite-dimensional semi-invertible MET for quasi-compact operators on Banach spaces followed by Froyland, Lloyd and Quas in 2013 [23]. In this section, we present the semi-invertible MET by González-Tokman and Quas from 2014 [29]. The proof of their MET inspired some ideas for our convergence proof on Hilbert spaces. Especially their technique of pushing forward so-called *good complements* of the Oseledets filtration from the past to the present in order to obtain the Oseledets splitting shows promising potential in the analysis of Ginelli's algorithm.

The MET from [29] is formulated for strongly measurable random dynamical systems on separable Banach spaces with discrete two-sided time as in Definition 2.1.3. Compared to the finite-dimensional case, systems on Banach spaces exhibit Lyapunov spectra with possibly non-discrete parts. In fact, an Oseledets filtration and an Oseledets splitting exist only for the first, at most countably many *exceptional LEs* that are isolated from the rest of the spectrum. To discern the exceptional LEs, we need the notion of *quasi-compactness*.

Let $(X, \|.\|)$ be a Banach space. Write $B \subset X$ for the unit ball and $S \subset X$ for the unit sphere in X. Given a bounded linear operator $A \in L(X)$ on X, we define the *index of compactness* of A as

$$\|A\|_{ic(X)} := \inf\{r > 0 \mid A(B) \text{ can be covered by finitely many balls of radius } r\}.$$

The index gives us a measure of how close A is to being a compact operator. In fact, the index of compact operators, such as operators on $\mathbb{R}^d$ or operators with finite range, is always zero. The following result extends the index of compactness to cocycles:

Proposition 2.3.1 ([29])
Let $\mathcal{R} = (\Omega, \mathcal{F}, \mathbb{P}, \sigma, X, \mathcal{L})$ be a separable strongly measurable random dynamical system such that $\log^+ \|\mathcal{L}\| \in L^1(\Omega, \mathcal{F}, \mathbb{P})$.

For $\mathbb{P}$-almost every $\omega \in \Omega$, the maximal Lyapunov exponent

$$\lambda(\omega) := \lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_\omega^{(n)}\|$$

and the index of compactness of the cocycle [62]

$$\kappa(\omega) := \lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_\omega^{(n)}\|_{ic(X)}$$

exist. Furthermore, λ and κ are measurable and σ-invariant.

If σ is ergodic, then λ and κ are constant $\mathbb{P}$-almost everywhere. Denote these constants by λ^* and κ^*. It holds $\kappa* \leq \lambda* < \infty$.

The next theorem states that there are only exceptional LEs between κ^* and λ^* and that Oseledets spaces exist for all of these exponents. Since cocycles consisting of compact operators fulfill $\kappa^* = -\infty$, their Lyapunov spectrum has only exceptional LEs and possibly $-\infty$. Under additional assumptions, such cocycles have been analyzed in Hilbert spaces with respect to stability of LEs and of Oseledets spaces [21]. In the context of stability, let us also mention a recent positive result by Crimmins, who proved stability for a class of semi-invertible random dynamical systems on separable Banach spaces and takes into account numerical representations of the linear operators [14].

We call a separable strongly measurable random dynamical system with ergodic base *quasi compact* if $\kappa^* < \lambda^*$. For such a system, Doan derives the existence of an Oseledets filtration [17] as a corollary of the two-sided MET by Lian and Lu [39]. With the additional assumption that the base is invertible, [29] proves a semi-invertible MET with a splitting that is similar to the Oseledets splitting obtained in fully invertible METs:

Theorem 2.3.2 (Semi-invertible MET [29])
Let $\mathcal{R} = (\Omega, \mathcal{F}, \mathbb{P}, \sigma, X, \mathcal{L})$ be a separable strongly measurable random dynamical system over an ergodic invertible base such that $\log^+ \|\mathcal{L}(\omega)\| \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Furthermore, assume that $\mathcal{R}$ is quasi-compact.

There exist $1 \leq p \leq \infty$ exceptional Lyapunov exponents $\lambda^* = \lambda_1 > \cdots > \lambda_p > \kappa^*$ (or if $p = \infty$: $\lambda_1 > \lambda_2 > \cdots > \kappa^*$ and $\lim_{i\to\infty} \lambda_i = \kappa^*$), multiplicities $d_1, \ldots, d_p \in \mathbb{N}$, and a unique measurable splitting of X into closed subspaces

$$X = \bigoplus_{i=1}^{p} Y_i(\omega) \oplus V(\omega)$$

defined on a σ-invariant subset $\Omega' \subset \Omega$ of full $\mathbb{P}$-measure such that the following hold for $\omega \in \Omega'$:

1. *the splitting is equivariant, i.e., $\mathcal{L}(\omega)V(\omega) \subset V(\sigma\omega)$ and $\mathcal{L}(\omega)Y_i(\omega) = Y_i(\sigma\omega)$,*
2. *$\dim Y_i(\omega) = d_i$,*
3. *$\lim_{n\to\infty}(1/n)\log\|\mathcal{L}_\omega^{(n)} y\| = \lambda_i$ for $y \in Y_i(\omega) \setminus \{0\}$,*
4. *$\limsup_{n\to\infty}(1/n)\log\|\mathcal{L}_\omega^{(n)} v\| \leq \kappa^*$ for $v \in V(\omega)$,*
5. *the norms of the projections associated to the splitting are tempered with respect to σ, where a function $f : \Omega \to \mathbb{R}$ is called tempered if $\lim_{n\to\pm\infty}(1/n)\log|f(\sigma_n\omega)| = 0$ for $\mathbb{P}$-almost every ω.*

We call the above splitting *Oseledets splitting* and the spaces $Y_i(\omega)$ *Oseledets spaces*. The *Oseledets filtration* $X = V_1(\omega) \supset \cdots \supset V_p(\omega) \supset V_{p+1}(\omega)$ from Doan's theorem can be reconstructed via $V_{p+1}(\omega) = V(\omega)$ and

$$V_i(\omega) = \bigoplus_{j=i}^{p} Y_j(\omega) \oplus V(\omega) \tag{2.5}$$

for $1 \leq i \leq p$.

Note that, contrary to the fully invertible finite-dimensional cases, the above Oseledets splitting has no asymptotic characterization for reversed time. This is due to the semi-invertible setting in which the cocycle is not necessarily injective. Nonetheless, its properties assure that the new Oseledets splitting coincides with the old one when applied to random dynamical systems on $\mathbb{R}^d$.

In Section 3.1 we generalize the concept of Ginelli's algorithm to compute Oseledets spaces in Hilbert spaces for a fixed $\omega \in \Omega'$. So far, the restriction to Hilbert spaces is of a technical nature and may be lifted in the future. Our convergence analysis in Chapter 5 requires cocycle data along the trajectory of ω and basic asymptotic properties that appear, e.g., in the METs from [17] and [29]. That is, we need uniform upper bounds for asymptotics of the Oseledets filtration and uniform lower bounds for asymptotics of the Oseledets splitting. A detailed derivation can be found in Appendix A. While bounds for the Oseledets filtration are recovered from Doan's work [17]:

$$\lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_\omega^{(n)}|_{V_i(\omega)}\| = \lambda_i \tag{2.6}$$

for $1 \le i \le p$ and

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_\omega^{(n)}|_{V(\omega)}\| \le \kappa^*, \tag{2.7}$$

bounds for the Oseledets splitting are due to [29]. By choosing a suitable basis, González-Tokman and Quas reduce the cocycle along $Y_i(\omega)$ to a cocycle of matrices (similar to [23, lemma 19]) for which uniform estimates are known. By applying the same arguments to the sum of Oseledets spaces $Y_1(\omega) \oplus \cdots \oplus Y_i(\omega)$, we get uniform lower bounds of growth rates inside sums of Oseledets spaces:

$$\liminf_{n\to\infty} \inf_{y\in Y_1(\omega)\oplus\cdots\oplus Y_i(\omega)\cap S} \frac{1}{n} \log \|\mathcal{L}_\omega^{(n)} y\| = \lambda_i. \tag{2.8}$$

In addition to the bounds for $\mathcal{L}_\omega^{(n)}$, we need similar bounds for $\mathcal{L}_{\sigma_{-n}\omega}^{(n)}$. These can be obtained via [22, lemma 8.2]. We get

$$\lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_{\sigma_{-n}\omega}^{(n)}|_{V_i(\sigma_{-n}\omega)}\| = \lambda_i \tag{2.9}$$

for $1 \le i \le p$ and

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_{\sigma_{-n}\omega}^{(n)}|_{V(\sigma_{-n}\omega)}\| \le \kappa^* \tag{2.10}$$

for $\mathbb{P}$-almost every ω. Uniform lower bounds for the Oseledets splitting are again obtained from reduced systems via matrix cocycles. We have

$$\liminf_{n\to\infty} \inf_{y\in Y_1(\sigma_{-n}\omega)\oplus\cdots\oplus Y_i(\sigma_{-n}\omega)\cap S} \frac{1}{n} \log \|\mathcal{L}_{\sigma_{-n}\omega}^{(n)} y\| = \lambda_i. \tag{2.11}$$

The uniform estimates for $\mathcal{L}_\omega^{(n)}$ and for $\mathcal{L}_{\sigma_{-n}\omega}^{(n)}$ are then used in [29] to prove temperedness of projections as stated in Theorem 2.3.2.

Observe that $\ker \mathcal{L}_\omega^{(n)} \subset V(\omega)$ and $\ker \mathcal{L}_{\sigma_{-n}\omega}^{(n)} \subset V(\sigma_{-n}\omega)$ for every $n \in \mathbb{N}$. Indeed, $\ker \mathcal{L}_\omega^{(n)} \subset V(\omega)$ follows from the different growth rates of vectors inside the Oseledets spaces. Since $\ker \mathcal{L}_\omega^{(n)} \subset V(\omega)$ holds on a σ-invariant subset of Ω, we get $\ker \mathcal{L}_{\sigma_{-n}\omega}^{(n)} \subset V(\sigma_{-n}\omega)$.

Besides the uniform estimates, our convergence proof only needs the properties stated in Theorem 2.3.2. We remark that these properties are present in most semi-invertible and invertible versions of the MET. Hence, by adjusting the notation, Chapter 5 can be generalized to various MET-scenarios.

3

Ginelli's Algorithm

In this chapter we present a method to compute CLVs, or more generally Oseledets spaces, that arise in the MET. Due to newly developed algorithms, CLVs were made available and have gained increased interest in applications during the last years, see Sections 1.2 and 1.3. Among the most famous algorithms is the *Ginelli algorithm* [26], which uses a dynamical approach to approximate CLVs. Other algorithms, like the one by Wolfe and Samelson [67], combine the use of dynamical techniques with SVDs of the cocycle, its adjoint, or its inverse. Here, we focus on a purely dynamical approach since convergence of singular values and vectors is already well-described by the proof of the MET (at least in finite dimensions). By combining our techniques from Chapter 4 with the proof of Proposition 2.2.2 one may extend convergence results to other algorithms besides the one by Ginelli.

We start by describing the concept of Ginelli's algorithm on an analytical level. A precise mathematical formulation is derived for Hilbert spaces. As an example, we implement the algorithm in MATLAB R2019a and apply it to compute CLVs along a trajectory of the Lorenz attractor.

3.1 Concept

Ginelli's algorithm requires cocycle data along a given trajectory for which Oseledets spaces exist. The cocycle may be part of a random dynamical system or simply a sequence of matrices as in Proposition 2.2.2. Independent of the setting the fundamental idea behind Ginelli's algorithm is that almost every vector has a non-vanishing projection (subject to the Oseledets splitting) onto the first Oseledets space. Since vectors inside the first Oseledets space have the highest exponential growth rate, almost every vector will align with the first Oseledets space asymptotically in forward-time. Similarly, we expect the linear span of $k = d_1 + \cdots + d_i$ randomly chosen vectors to align with the fastest expanding k-dimensional subspace, the sum of the first i Oseledets spaces, in forward-time. Reversing time, the fastest growing direction inside $Y_1 \oplus \cdots \oplus Y_i$ is the slowest growing direction in forward-time, that is, Y_i. Thus, we have a means to compute Oseledets spaces.

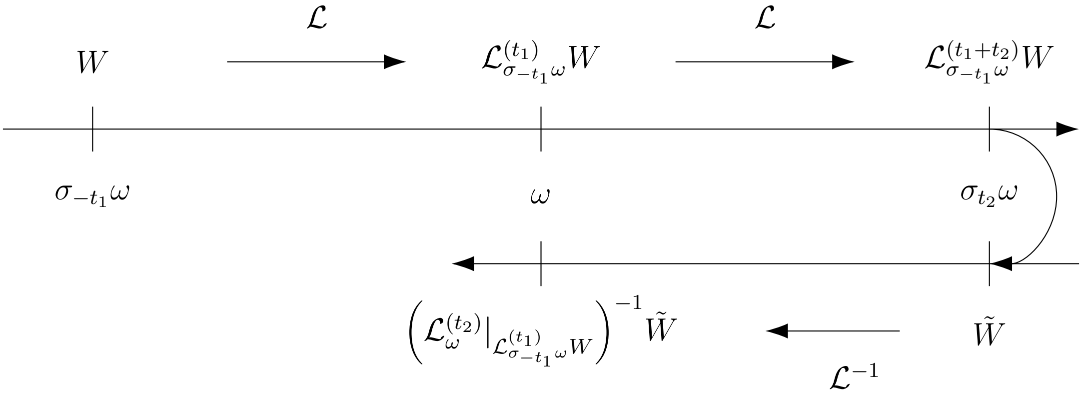


Figure 3.1: Ginelli's algorithm at the level of Grassmannians.

At a more abstract level, Ginelli's algorithm starts with a randomly chosen subspace $W \subset X$ of dimension $d_1 + \cdots + d_i$, which is propagated from the far past to the present via $\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}$ to get an approximation of $Y_1(\omega) \oplus \cdots \oplus Y_i(\omega)$ for large t_1. Then, $\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}W$ is propagated further via $\mathcal{L}^{(t_2)}_{\omega}$ to approximate $Y_1 \oplus \cdots \oplus Y_i$ in the far future. Next, the algorithm randomly chooses a subspace $\tilde{W} \subset \mathcal{L}^{(t_1+t_2)}_{\sigma_{-t_1}\omega}W$ of dimension d_i. This subspace is propagated backwards to approximate $Y_i(\omega)$ for large t_1, t_2. (see Fig. 3.1)

In practice we express W in terms of a basis $(x_1, \ldots, x_k)$. By propagating these vectors, we can track the evolution of W. Similarly, we express $\tilde{W}$ in terms of a basis. The corresponding vectors can be described as coefficients of the propagated vectors of W. Hence, the backward propagation can be done completely inside a finite-dimensional coefficient space.

Let $X = H$ be a Hilbert space. To avoid that all vectors $x_1, \ldots, x_k$ collapse onto the first Oseledets space, which renders them numerically indistinguishable, Ginelli and others suggest to orthonormalize them between smaller propagation steps. While this procedure does not change the outcome of Ginelli's algorithm analytically, as the involved spaces remain the same, it helps with numerical stability. In particular, they use a QR-decomposition to store orthonormalized vectors in a matrix Q and the cocycle on coefficient space in a matrix R for each propagation step. The upper triangular R-matrices can easily be inverted to perform the backward propagation in coefficient space. Using the identification between vectors and coefficients, we substitute initial vectors for the backward propagation by an upper triangular matrix representing their coefficients. For more details on the implementation see Section 3.2, Appendix B, or [26, 27].

Definition 3.1.1
Taking the above into account, we define (the analytical kernel[1] *of)* Ginelli's algorithm *on Hilbert spaces as*

$$G^{t_1,t_2}_{\omega,k} : H^k \times \mathbb{R}^{k\times k}_{ru} \to H^k,$$

[1]We leave out numerical details of the implementation since they do not affect the output of Ginelli's algorithm analytically.

where $\omega \in \Omega$ defines the trajectory, $k \in \mathbb{N}$ is the number of CLVs we wish to compute, $t_1 \in \mathbb{T}_{\geq 0}$ is the amount of past data, $t_2 \in \mathbb{T}_{\geq 0}$ is the amount of future data, and $\mathbb{R}^{k\times k}_{ru}$ denotes the set of upper triangular $k \times k$-matrices. $G^{t_1,t_2}_{\omega,k}$ operates on $((x_1,\dots,x_k),(r_{ij})^k_{i,j=1})$ via the following steps:

1. *forward propagation from $\sigma_{-t_1}\omega$ to ω:*

$$\left(x^1_1,\dots,x^1_k\right) := \left(\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}x_1,\dots,\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}x_k\right).$$

2. *forward propagation from ω to $\sigma_{t_2}\omega$:*

$$\left(x^2_1,\dots,x^2_k\right) := \left(\mathcal{L}^{(t_2)}_{\omega}x^1_1,\dots,\mathcal{L}^{(t_2)}_{\omega}x^1_k\right).$$

3. *orthonormalization[2]:*

$$\left(x^3_1,\dots,x^3_k\right) := \mathrm{orth}\left(x^2_1,\dots,x^2_k\right).$$

4. *initialization of vectors for backward propagation:*

$$\left(y^1_1,y^1_2,\dots,y^1_k\right) := \left(r_{11}x^3_1, r_{12}x^3_1 + r_{22}x^3_2,\dots,\sum_{j=1}^k r_{jk}x^3_j\right).$$

5. *backward propagation from $\sigma_{t_2}\omega$ to ω:*

$$\left(y^2_1,\dots,y^2_k\right) := \left(\left(\mathcal{L}^{(t_2)}_{\omega}|_{W^1}\right)^{-1}y^1_1,\dots,\left(\mathcal{L}^{(t_2)}_{\omega}|_{W^1}\right)^{-1}y^1_k\right),$$

   where $W^1 := \mathrm{span}(x^1_1,\dots,x^1_k)$.

6. *normalization:*

$$\left(y^3_1,\dots,y^3_k\right) := \left(\frac{y^2_1}{\|y^2_1\|},\dots,\frac{y^2_k}{\|y^2_k\|}\right).$$

We set $G^{t_1,t_2}_{\omega,k}((x_1,\dots,x_k),(r_{ij})^k_{i,j=1}) := (y^3_1,\dots,y^3_k)$ as our approximation of the first k CLVs at ω.

Ginelli's algorithm requires two types of inputs: a tuple of vectors $(x_1,\dots,x_k)$ as initial vectors for forward propagation and a coefficient matrix $(r_{ij})^k_{i,j=1}$ to initialize vectors for backward propagation. We denote new vectors obtained during the different propagation steps by x and y with corresponding indices. In the above definition, x-vectors are used during forward propagation and y-vectors are used during backward propagation.

Let us remark that whenever $G^{t_1,t_2}_{\omega,k+1}((x_1,\dots,x_{k+1}),(r_{ij})^{k+1}_{i,j=1})$ is well-defined, its first k components coincide with $G^{t_1,t_2}_{\omega,k}((x_1,\dots,x_k),(r_{ij})^k_{i,j=1})$. Thus, it suffices to assume $k = d_1 + \cdots + d_i$ for our convergence analysis in Chapters 4 and 5. In fact, we will set $k = d$ (or equivalently $i = p$) when investigating the finite-dimensional case $H = \mathbb{R}^d$.

[2] Any orthonormalization procedure respecting the order of the tuple is feasible. For example, this includes the QR-decomposition and the Gram-Schmidt procedure.

The asymptotic expansion rate $\lambda_1 + \cdots + \lambda_i$ of $Y_1 \oplus \cdots \oplus Y_i$ can be computed as a byproduct of the forward phase of the algorithm by looking at renormalization factors, e.g., in case of QR-decompositions these are products of diagonal elements of R-matrices. Thus, we can derive the involved LEs.[3]

In Chapters 4 and 5 we provide convergence proofs of the algorithm as $\min(t_1, t_2) \to \infty$. The speed of convergence turns out to be exponential in relation to the minimal distance between LEs. Furthermore, the kind of convergence differs between discrete and continuous time. The discrete version with $t_1, t_2 \in \mathbb{N}_0$ converges for almost every initial tuple, whereas the continuous version with $t_1, t_2 \in \mathbb{R}_{\geq 0}$ only converges in measure. Details on convergence will be discussed in the respective chapters.

3.2 Implementation

Before beginning the convergence analysis in the next chapter, we implement Ginelli's algorithm and apply it to the Lorenz 63 model [40] given by

$$\begin{aligned}
\dot{x}_1 &= \sigma(x_2 - x_1)\\
\dot{x}_2 &= x_1(\rho - x_3) - x_2\\
\dot{x}_3 &= x_1 x_2 - \beta x_3,
\end{aligned}$$

where $\sigma = 10$, $\beta = 8/3$, and $\rho = 28$ are the classical parameter values. The model has three fixed points: the origin, which is a saddle, and two unstable spirals $(\pm\sqrt{\beta(\rho-1)}, \pm\sqrt{\beta(\rho-1)}, \rho - 1)$. Moreover, the equations admit a strange attractor called *Lorenz attractor*. With strange attractor we mean an invariant set, which is neither a steady state nor a periodic orbit, towards which almost every solution of the system is evolving and which has sensitive dependence on initial conditions (a positive first LE). The Lorenz attractor is robust, i.e., it persists for small perturbations of the parameters, and admits a unique *SRB-measure*[4] [45, 65]. The measure is invariant under the flow and its support coincides with the attractor. In particular, LEs and Oseledets spaces exist for almost every point of the Lorenz attractor. Due to ergodicity[5] the LEs are constant almost everywhere. Here, "almost every" is understood with respect to the SRB-measure.

Even though the MET gives us existence of CLVs on a set of full measure with respect to the SRB-measure, it is potentially a set of Lebesgue measure zero. When applying Ginelli's algorithm, we hope that the chosen trajectory will be close enough to the attractor, so that we may approximate LEs and CLVs of the attractor. Note that this is only a heuristic argument since Theorem 2.2.8 only ensures measurable dependence of LEs and CLVs on the system state. In fact, due to the asymptotic nature of Oseledets spaces, stability properties are nontrivial [21, 31, 51].

[3]This concept was already used in 1980 by Benettin [5, 6] to compute the Lyapunov spectrum.

[4]The defining properties of an SRB-measure μ are invariance under the flow and equality of time and space averages for Lebesgue almost every point in the basin of attraction: $\lim_{n\to\infty}(1/n)\sum_{i=0}^{n-1}\varphi(\sigma_n x) = \int \varphi(x)d\mu$, where σ denotes the flow and φ is a continuous observable defined on the basin of attraction.

[5]The SRB-measure μ of the Lorenz attractor even has the *mixing property*, which is stronger than ergodicity: $\lim_{n\to\infty}\mu(\sigma_n(A)\cap B) = \mu(A)\mu(B)$, where A, B are μ-measurable subsets of the basin of attraction [43].

In our implementation the Lorenz system is combined with its linearization and integrated using a fourth order Runge-Kutta method with fixed timestep starting at an arbitrarily chosen initial state (Fig. 3.2a). We obtain a nonlinear background trajectory and the cocycle along the trajectory. After half of the integration time we fix the corresponding point on the background trajectory and separate the cocycle into two parts, one for past and one for future data (Fig. 3.2b). Then, Ginelli's algorithm with intermittent QR-decompositions is applied to the data. We obtain an approximation of LEs and of CLVs at the specific point (Fig. 3.2c). Moreover, we compare the computed CLVs with CLVs obtained using less integration time. The rate of convergence of each CLV is related to the distance between computed LEs (Figs. 3.2d to 3.2f). Associated figures can be found at the end of this subsection and the full implementation is given in Appendix B.

We observe that with increasingly higher integration time the computed CLVs seem to converge to the reference CLVs. Moreover, the convergence is exponentially fast with a rate given by the spectral gap between associated LEs (at least with respect to the reference LEs). In particular, the convergence rate of the third CLV does not depend on the first LE. Since $|\lambda_2 - \lambda_3|$ is larger than $|\lambda_1 - \lambda_2|$, the third CLV reaches system accuracy much faster than the other two CLVs. The small fluctuations on subexponential scales that can be seen in Figs. 3.2d to 3.2f are due to the nonlinearity of the underlying system, numerical errors, and different randomly chosen initial conditions for each run of the algorithm.

Without much computational effort we may compute CLVs at other states along the trajectory. Indeed, by covariance it is enough to push CLVs forwards and backwards with the linear propagator. If there is long enough transient time for past and future data, the approximations on a fixed time interval around the chosen state will still be good. In particular, by topological transitivity[6] of the Lorenz attractor [65], one may compute CLVs for finitely many points covering a large portion of the attractor to get properties of the whole attractor. However, this is only a heuristic argument as CLVs depend measurably on the system state. Without further stability analysis we cannot rigorously derive CLVs at other points of the attractor by this argument.

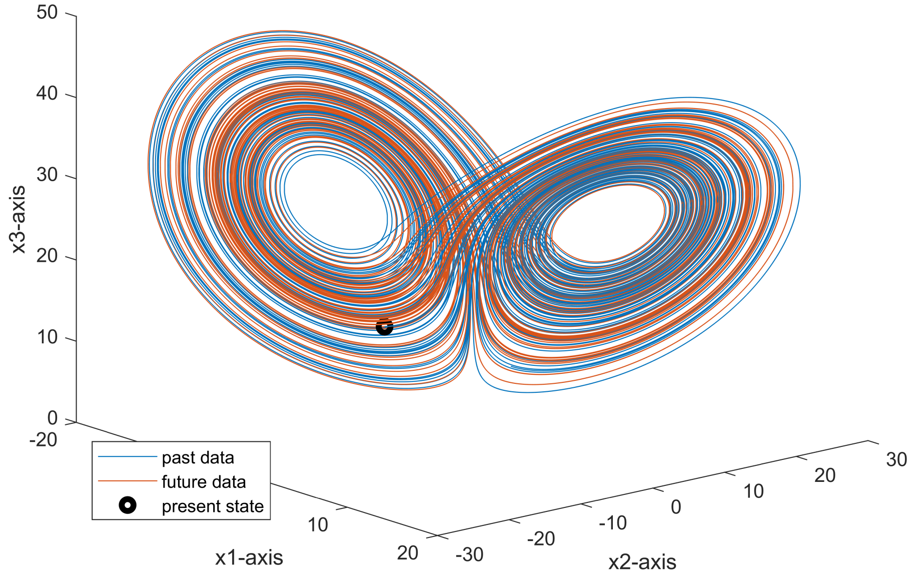
We remark that our implementation does not compute CLVs at a given state of the system, but at the half way mark of a trajectory initiated at a given state. Due to the chaotic nature of the Lorenz attractor, we cannot predetermine this half way mark. Moreover, backward integration should be considered carefully as the attractor would become a repellor, which causes numerical errors to blow up very quickly. Without knowing an analytical solution we have to rely on heuristic arguments as above to get CLVs at a desired point on the attractor.

[6] A σ-invariant subset is called *topologically transitive* if for every two non-empty open subsets U, V there is $n \in \mathbb{N}$ with $\sigma_n(U) \cap V \neq \emptyset$.

```
--- Input data ---

System:
Lorenz 63 (rho = 28, sigma = 10, beta = 8/3)

Number of timesteps for past data:
100000

Number of timesteps for future data:
100000

Stepsize:
0.001

Number of LEs to compute:
3

Initial state of past data orbit:
[ 1.73000 3.23000 8.01000 ]
```

(a) input

Figure 3.2: Ginelli's algorithm applied to the Lorenz attractor. We integrated the Lorenz 63 model with classical parameter values in MATLAB R2019a and applied Ginelli's algorithm to the cocycle (see Appendix B for the code). The highlighted system state in Fig. 3.2b is the half way mark of the computed trajectory. Using the full computed trajectory and cocycle, we approximated LEs and CLVs. The rounded output is displayed in Fig. 3.2c. For Figs. 3.2d to 3.2f, we computed CLVs with less of the trajectory and cocycle data centered around the highlighted system state in Fig. 3.2b and compared the result to the CLVs from Fig. 3.2c computed with the full cocycle data. Red dots in Figs. 3.2d to 3.2f represent the logarithmic errors between CLVs in Fig. 3.2c and CLVs of the same system state computed with integration time t for both the past and the future parts of the trajectory. Missing dots indicate a numerical error identical to zero, which corresponds to a logarithmic error of $-\infty$.

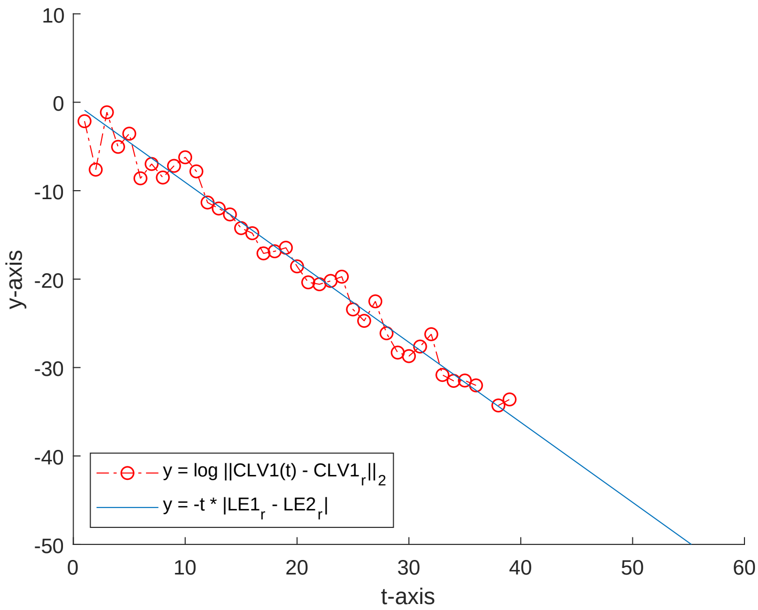


(b) trajectory

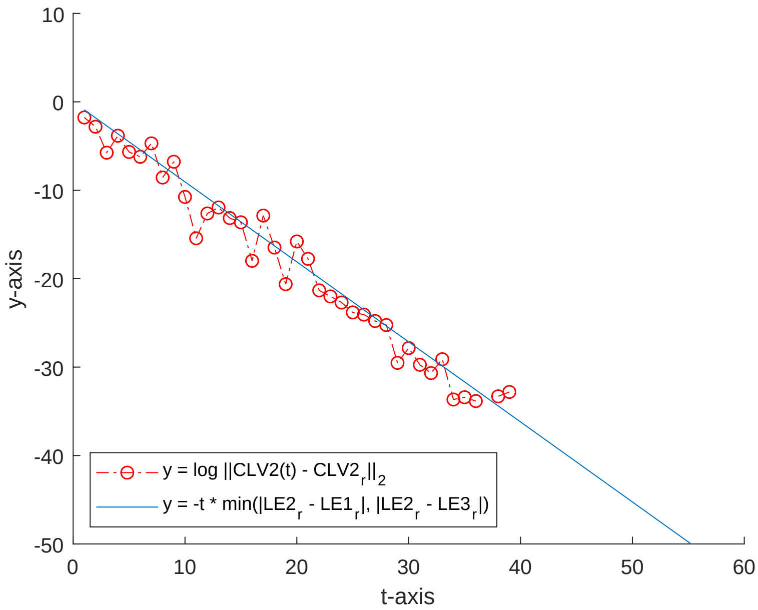
```
--- Output data ---

Lyapunov exponents:
0.90545
0.00048586
-14.573

Covariant Lyapunov vectors:
[ -0.42991 -0.66370 0.61211 ]
[ 0.48058 0.87246 0.08861 ]
[ -0.70515 0.66447 -0.24746 ]

System state:
[ -4.06992 -7.07230 12.86971 ]
```

(c) output

Figure 3.2: Ginelli's algorithm applied to the Lorenz attractor (cont.).

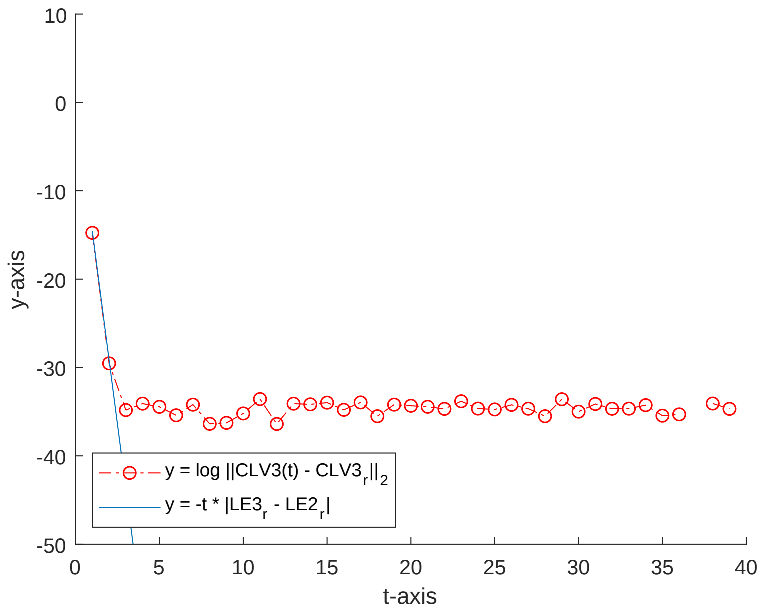


(d) convergence rate of first CLV



(e) convergence rate of second CLV

Figure 3.2: Ginelli's algorithm applied to the Lorenz attractor (cont.).



(f) convergence rate of third CLV

Figure 3.2: Ginelli's algorithm applied to the Lorenz attractor (cont.).

4

Convergence of Ginelli's Algorithm on $\mathbb{R}^d$

This chapter follows [48]. We investigate convergence of Ginelli's algorithm in a deterministic setting on $\mathbb{R}^d$ as given by Section 2.2.2. Although there already has been an attempt to prove convergence [20, 27], it harbors conceptual difficulties and is only valid for nondegenerate Lyapunov spectra, see Section 1.3. Here, we provide a mathematically rigorous convergence proof that even applies to degenerate Lyapunov spectra. Contrary to the previous approach, we use more compact *projector-based* techniques to capture evolution of subspaces rather than single vectors. It turns out that this is the right framework when including degeneracies. Moreover, we point out the subtle differences between convergence for continuous and for discrete time and relate the speed of convergence to LEs.

We start by introducing the *Lyapunov index notation*. On the one hand, the notation helps us to get familiar with exponential scales, on the other, it shortens the analysis of our convergence proof. The proof itself focuses on the evolution of subspaces. Hence, we derive notions of distances and angles between subspaces in terms of projection operators in Section 4.2. Section 4.3 introduces the concept of *admissibility*. A tuple of vectors is called admissible with respect to another tuple if their corresponding filtrations are close enough. We use this concept to describe how tuples of vectors evolve compared to singular vectors, which are connected to CLVs via the proof of the MET. In fact, they provide directions of optimal growth for finite-time, whereas CLVs describe the asymptotic limit. The link between finite-time scenarios and asymptotics plays a major role in the proof of the MET from [1] and is substantial to our convergence analysis in Section 4.4. After stating the convergence theorems we devote the rest of Section 4.4 to their proofs. Next to the link between MET and SVD, we derive estimates for forward and for backward propagation and combine them to form the convergence proofs. The analysis heavily depends on our concept of admissibility, which turns out to be the right choice to describe initial vectors for Ginelli's algorithm. In particular, to get the correct notion of convergence we require Lebesgue measure estimates of the set of admissible tuples. It turns out that the precise notion of convergence for discrete and for continuous time differs. Namely, the discrete version converges for almost every input, whereas the continuous version only converges in measure.

4.1 Lyapunov index

When analyzing an algorithm, one of the main aspects to consider is the speed of convergence. It is defined as the rate of change of the distance between a current and a sough-after state as a parameter, such as time, is increased. In our case time can be either discrete ($\mathbb{T} = \mathbb{Z}$) or continuous ($\mathbb{T} = \mathbb{R}$). Moreover, the nature of the problem or features of the algorithm might already prescribe certain timescales. In fact, LEs and CLVs describe properties on exponential scales, which can be captured by the Lyapunov index notation.

Definition 4.1.1 ([1])
The Lyapunov index $\lambda(f) \in \mathbb{R} \cup \{\pm\infty\}$ *of a function* $f : \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ *is defined as the limit*

$$\lambda(f) := \limsup_{t\to\infty} \frac{1}{t} \log f(t).$$

Roughly speaking, the function f behaves similar to $e^{t\lambda(f)}$ on exponential scales. For example, a negative Lyapunov index implies exponential decay. However, one should note that variations on smaller scales are not included in this notation[1], but very well may be of importance for limited-time scenarios such as numerical computations.

Next, we list some useful properties of the Lyapunov index, which can be found in Arnold's book and are easily verified:

Proposition 4.1.2 ([1])
Let $f, g : \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$*. The following are true:*

1. $\lambda(0) = -\infty$,
2. $\lambda(c) = 0$ *for* $c > 0$ *constant,*
3. $\lambda(\alpha f) = \lambda(f)$ *for* $\alpha > 0$,
4. $\lambda(f^\alpha) = \alpha\lambda(f)$ *for* $\alpha > 0$,
5. $f \leq g \implies \lambda(f) \leq \lambda(g)$,
6. $\lambda(f + g) \leq \max(\lambda(f), \lambda(g))$, *and*
7. $\lambda(f \cdot g) \leq \lambda(f) + \lambda(g)$ *(if the right-hand side makes sense).*

As the Ginelli algorithm consists of two subsequent phases, a forward and a backward phase, the Lyapunov index is not enough to discuss the algorithm. Each phase has its own runtime that influences the resulting approximation. For a good approximation, both runtimes need to be increased. Certainly, there are circumstances and rules that prescribe a favoring relation between those runtimes. However, we will not discuss them here. Instead, we settle for a formulation that allows two different runtimes. For this purpose, we extend the notion of the Lyapunov index to a formulation depending on two parameters:

[1]For example, e^{-t} and $\sin(t)t^2e^{-t}$ have the same Lyapunov index.

Definition 4.1.3
The extended Lyapunov index $\overline{\lambda}(f) \in \mathbb{R} \cup \{\pm\infty\}$ *of a function* $f : \mathbb{T}_{\geq 0} \times \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ *is defined as the limit*

$$\overline{\lambda}(f) := \limsup_{T \to \infty} \sup_{t_1, t_2 \geq T} \frac{1}{\min(t_1, t_2)} \log f(t_1, t_2).$$

In contrast to the standard Lyapunov index, the new quantity describes behavior on exponential scales as $\min(t_1, t_2)$ is increased. Especially, when fixing a certain relation between both parameters, an upper bound of the speed of convergence is given by the extended Lyapunov index.[2] In fact, the extended version exhibits properties similar to the usual Lyapunov index:

Proposition 4.1.4
Rules 1-7 of Proposition 4.1.2 hold true with λ replaced by $\overline{\lambda}$. Furthermore, if we extend a function $f : \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ to $\overline{f} : \mathbb{T}_{\geq 0} \times \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ by setting $\overline{f}(t_1, t_2) := f(t_1)$, then

8. $\lambda(f) < 0 \implies \overline{\lambda}(\overline{f}) = \lambda(f)$.

Proof. Rules 1,2,4,5 and 7 follow directly from the definition. To show rule 3, we have $f \leq \alpha f$ for $\alpha \geq 1$, and hence

$$\overline{\lambda}(f) \leq \overline{\lambda}(\alpha f) \leq \overline{\lambda}(\alpha) + \overline{\lambda}(f) = \overline{\lambda}(f).$$

The case $0 < \alpha < 1$ follows by looking at $\beta := 1/\alpha$ and $g := \alpha f$. Moreover, it is easily verified that

$$\overline{\lambda}(f + g) \leq \overline{\lambda}(2 \max(f, g)) = \overline{\lambda}(\max(f, g)) = \max(\overline{\lambda}(f), \overline{\lambda}(g)).$$

Now, let $\overline{f}$ be the extension of some function $f : \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ as above. The relation $\lambda(f) \leq \overline{\lambda}(\overline{f})$ is always satisfied. To show equality, we remark that $\lambda(f) < 0$ implies the existence of some $T > 0$ with $\log f(t) < 0$ for all $t \geq T$. In particular, it holds

$$\sup_{t_1, t_2 \geq t} \frac{1}{\min(t_1, t_2)} \log f(t_1) \leq \sup_{t_1 \geq t} \frac{1}{t_1} \log f(t_1)$$

with a right-hand side converging to $\lambda(f)$ for $t \to \infty$. □

We demonstrate two exceptional cases where the function is either growing subexponentially or decaying superexponentially:

Example 4.1.5
Let $f(t_1, t_2) := \lceil \min(t_1, t_2)^2 \rceil$ and $g(t_1, t_2) := \alpha^{2f(t_1, t_2) - 1}$ for $0 < \alpha < 1$. We compute

$$0 = \overline{\lambda}(1) \leq \overline{\lambda}(f) \leq \overline{\lambda}(\min(t_1, t_2)^2 + 1) \leq \max(2\overline{\lambda}(\min(t_1, t_2)), 0) = 0$$

and

$$\overline{\lambda}(g) = \overline{\lambda}\Big(\frac{1}{\alpha}\big(\alpha^f\big)^2\Big) = 2\overline{\lambda}\big(\alpha^f\big) \leq 2\overline{\lambda}\big(\alpha^{\min(t_1, t_2)^2}\big) = -\infty.$$

[2] For example, given the relation $t_1 = 2t_2$ we have $\lambda(f(2t, t)) \leq \overline{\lambda}(f)$.

4.2 Grassmannians

The *Grassmannian* $\mathcal{G}(\mathbb{R}^d)$ is the space of all subspaces of $\mathbb{R}^d$. A more general definition for Banach spaces will be given in Section 5.1. Here, we have the advantage that all subspaces are of finite dimension and finite codimension. Moreover, the euclidean metric allows us to associate subspaces with their orthogonal projections. We denote the orthogonal projection onto a subspace $M \subset \mathbb{R}^d$ by P_M. Through this identification we can define distances and angles between subspaces, or even speak of converging sequences of subspaces. In this context we present some essential results from [34, chapter 1.6], [25], and [16].

Since our analysis focuses on the euclidean norm $\|.\|_2$, let us drop the subscript and simply write $\|.\|$ during this chapter. In fact, since all norms on $\mathbb{R}^d$ are equivalent, quantities that are defined on exponential scales remain the same. This includes LEs and CLVs as well as our estimates of the speed of convergence in Section 4.4.1.

Definition 4.2.1
The distance *between two subspaces $M, N \subset \mathbb{R}^d$ is defined as*

$$d(M,N) := \|P_M - P_N\|.$$

We state a collection of handy properties mostly from [25]:

Proposition 4.2.2 ([25])
The distance d is a metric on the set of subspaces. Moreover, the following holds for all subspaces $M, N \subset \mathbb{R}^d$:

1. $0 \leq d(M,N) \leq 1$,
2. $d(M,N) = d(M^\perp, N^\perp)$, *and*
3. $d(M,N) < 1 \implies \dim(M) = \dim(N)$.

In case that $\dim(M) = \dim(N)$ *we have:*

4. $d(M,N) = \|P_M P_{N^\perp}\|$, *and*
5. $d(M,N) = 1 \iff M \cap N^\perp \neq \{0\}$.

If $V \in O(d,\mathbb{R})$ is an orthogonal transformation, then

6. $d(V(M), V(N)) = d(M,N)$.

Every invertible linear map induces a Lipschitz-continuous transformation of the set of subspaces:

Corollary 4.2.3
For each $A \in Gl(d,\mathbb{R})$ and all subspaces $M, N \subset \mathbb{R}^d$, we have

$$d(A(M), A(N)) \leq \|A\|\,\|A^{-1}\|\,d(M,N).$$

Proof. Fix an invertible linear map A. For subspaces of different dimension, the inequality is trivially satisfied. So, let M and N be of the same dimension. We compute

$$\begin{aligned}
d(A(M), A(N)) &= \|P_{A(M)} P_{(A(N))^\perp}\| \\
&= \|P_{A(M)} P_{(A^T)^{-1} N^\perp}\| \\
&= \max_{x \in M \setminus \{0\},\, y \in N^\perp \setminus \{0\}} \frac{|\langle Ax, (A^T)^{-1} y \rangle|}{\|Ax\| \, \|(A^T)^{-1} y\|} \\
&= \max_{x \in M \setminus \{0\},\, y \in N^\perp \setminus \{0\}} \frac{|\langle x, y \rangle|}{\|x\| \, \|y\|} \frac{\|A^{-1}(Ax)\|}{\|Ax\|} \frac{\|A^T((A^T)^{-1} y)\|}{\|(A^T)^{-1} y\|} \\
&\leq \max_{x \in M \setminus \{0\},\, y \in N^\perp \setminus \{0\}} \frac{|\langle x, y \rangle|}{\|x\| \, \|y\|} \|A^{-1}\| \, \|A^T\| \\
&= \|A\| \, \|A^{-1}\| \, \|P_M P_{N^\perp}\| \\
&= \|A\| \, \|A^{-1}\| \, d(M, N).
\end{aligned}$$

□

The next concept needed is the (minimal) angle between two subspaces. A lot on this topic can be found in [16].

Definition 4.2.4 ([16])
The cosine of the angle *between M and N is given by*

$$c(M, N) := \max \left\{ \frac{|\langle x, y \rangle|}{\|x\| \, \|y\|} \; : \; x \in M \cap (M \cap N)^\perp, y \in N \cap (M \cap N)^\perp, \; x, y \neq 0 \right\}$$

and the cosine of the minimal angle *between M and N is defined as*

$$c_0(M, N) := \max \left\{ \frac{|\langle x, y \rangle|}{\|x\| \, \|y\|} \; : \; x \in M, y \in N, \; x, y \neq 0 \right\},$$

where we set $\max(\emptyset) := 0$.

Both definitions agree if $M \cap N = \{0\}$. However, they are different in general. We state a few important properties in order to work with these quantities:

Proposition 4.2.5 ([16])
The following statements are true for all subspaces $M, N \subset \mathbb{R}^d$:

1. $0 \leq c(M, N) \leq c_0(M, N) \leq 1$,
2. $c(M, N) < 1$,
3. $c_0(M, N) < 1 \iff M \cap N = \{0\}$,
4. $c(M, N) = c(N, M)$ *and* $c_0(M, N) = c_0(N, M)$,
5. $c(M, N) = c(M^\perp, N^\perp)$,
6. $c_0(M, N) = \|P_M P_N\|$, *and*
7. $c(M, N) = \|P_M P_N - P_{M \cap N}\|$.

One can easily check that $P_M P_N$ is the orthogonal projection onto $M \cap N$ if, and only if, P_M and P_N commute. Nevertheless, if they do not commute, it is still possible to describe $P_{M\cap N}$ via P_M and P_N through the *method of alternating projections*, which is due to von Neumann [47]:

Theorem 4.2.6 ([47])
For each two subspaces M and N, the method of alternating projections converges:

$$\lim_{k\to\infty} \|(P_M P_N)^k - P_{M\cap N}\| = 0.$$

A discussion on the speed of convergence can be found in [16]. The following estimate will be enough for our purposes:

Proposition 4.2.7 ([16])
For each two subspaces M and N, it holds

$$\forall k : \|(P_M P_N)^k - P_{M\cap N}\| \leq c(M,N)^{2k-1}.$$

Utilizing the method of alternating projections, we can relate the distance of two intersections to the distance of intersecting subspaces:

Proposition 4.2.8
Let $M, N \subset \mathbb{R}^d$ be two subspaces, and set $\delta := c_0(M^\perp, N^\perp)$.
For all subspaces $M', N' \subset \mathbb{R}^d$ with

$$d(M',M) + d(N',N) \leq \frac{1-\delta}{2},$$

we have

$$d(M' \cap N', M \cap N) \leq \delta^{2k-1} + \left(\frac{1+\delta}{2}\right)^{2k-1} + k\left(d(M',M) + d(N',N)\right)$$

with arbitrary $k \in \mathbb{N}$.

Proof. Assume M, N, δ and M', N' as above. Using the method of alternating projections, we estimate for arbitrary $k \in \mathbb{N}$:

$$\begin{aligned}
\|P_{M'\cap N'} - P_{M\cap N}\| &\leq \|P_{M'\cap N'} - (P_{M'}P_{N'})^k\| + \|(P_{M'}P_{N'})^k - (P_M P_N)^k\| \\
&\quad + \|(P_M P_N)^k - P_{M\cap N}\| \\
&\leq c(M',N')^{2k-1} + \|(P_{M'}P_{N'})^k - (P_M P_N)^k\| \\
&\quad + c(M,N)^{2k-1}.
\end{aligned}$$

Since the minimal angle depends continuously on its subspaces, we have

$$\begin{aligned}
c(M',N') &= c\Big((M')^\perp, (N')^\perp\Big) \\
&\leq c_0\Big((M')^\perp, (N')^\perp\Big) \\
&= \|P_{(M')^\perp} P_{(N')^\perp}\| \\
&\leq \|P_{(M')^\perp} P_{(N')^\perp} - P_{M^\perp} P_{(N')^\perp}\| + \|P_{M^\perp} P_{(N')^\perp} - P_{M^\perp} P_{N^\perp}\| \\
&\quad + \|P_{M^\perp} P_{N^\perp}\| \\
&\leq \|P_{(M')^\perp} - P_{M^\perp}\| + \|P_{(N')^\perp} - P_{N^\perp}\| + \|P_{M^\perp} P_{N^\perp}\| \\
&= \|P_{M'} - P_M\| + \|P_{N'} - P_N\| + \delta \\
&\leq \frac{1+\delta}{2}.
\end{aligned}$$

For the middle summand in the estimate of $\|P_{M'\cap N'} - P_{M\cap N}\|$, we deduce

$$\begin{aligned}
&\|(P_{M'}P_{N'})^k - (P_M P_N)^k\| \\
&\le \sum_{l=0}^{k-1} \|(P_M P_N)^l (P_{M'}P_{N'})^{k-l} - (P_M P_N)^l P_M P_{N'} (P_{M'}P_{N'})^{k-(l+1)}\| \\
&\quad + \|(P_M P_N)^l P_M P_{N'} (P_{M'}P_{N'})^{k-(l+1)} - (P_M P_N)^{l+1} (P_{M'}P_{N'})^{k-(l+1)}\| \\
&\le \sum_{l=0}^{k-1} \|P_{M'} - P_M\| + \|P_{N'} - P_N\| \\
&= k\,(\|P_{M'} - P_M\| + \|P_{N'} - P_N\|).
\end{aligned}$$

For the last summand, we remark

$$c(M,N) = c(M^\perp, N^\perp) \le c_0(M^\perp, N^\perp) = \delta.$$

Combining the above yields the desired estimate. □

Now, assume we are given two converging sequences of subspaces $(M_t)_{t\in\mathbb{T}_{\geq 0}}$ and $(N_t)_{t\in\mathbb{T}_{\geq 0}}$ with transversal[3] limits M and N. As an immediate consequence of Proposition 4.2.8 with the right choice of $k = k(t)$, we see that the sequence of intersections $(M_t \cap N_t)_{t\in\mathbb{T}_{\geq 0}}$ converges to the intersection of the limits $M \cap N$. Moreover, we show that the speed of convergence on exponential scales can be preserved in a uniform manner:

Corollary 4.2.9
Let $M, N \subset \mathbb{R}^d$ be two transversal subspaces. Moreover, assume $(\mathcal{M}_t)_{t\in\mathbb{T}_{\geq 0}}$ and $(\mathcal{N}_t)_{t\in\mathbb{T}_{\geq 0}}$ are two sequences of collections of subspaces that converge to M, resp. N, exponentially fast:

$$\lambda_M := \lambda\Big(\sup_{M'\in\mathcal{M}_t} d(M', M)\Big) < 0 \quad \textit{and} \quad \lambda_N := \lambda\Big(\sup_{N'\in\mathcal{N}_t} d(N', N)\Big) < 0.$$

Then,

$$\overline{\lambda}\Big(\sup_{M'\in\mathcal{M}_{t_1}} \sup_{N'\in\mathcal{N}_{t_2}} d(M'\cap N', M\cap N)\Big) \le \max(\lambda_M, \lambda_N).$$

Proof. Let $\delta := c_0(M^\perp, N^\perp) < 1$. Since we have $\lambda_M, \lambda_N < 0$ (exponential decay of distances), there is $T > 0$ with

$$\sup_{M'\in\mathcal{M}_{t_1}} \sup_{N'\in\mathcal{N}_{t_2}} d(M', M) + d(N', N) \le \frac{1-\delta}{2}$$

for all $t_1, t_2 \geq T$. Invoking Proposition 4.2.8, we get

$$\begin{aligned}
&\sup_{M'\in\mathcal{M}_{t_1}} \sup_{N'\in\mathcal{N}_{t_2}} d(M'\cap N', M\cap N) \\
&\le \delta^{2k-1} + \left(\frac{1+\delta}{2}\right)^{2k-1} + k\left(\sup_{M'\in\mathcal{M}_{t_1}} d(M', M) + \sup_{N'\in\mathcal{N}_{t_2}} d(N', N)\right)
\end{aligned}$$

[3]Two subspaces M and N are called *transversal* if $M + N = \mathbb{R}^d$. Since $(M+N)^\perp = M^\perp \cap N^\perp$, transversality is equivalent to $c_0(M^\perp, N^\perp) < 1$.

for arbitrary $k \in \mathbb{N}$. With $k = k(t_1, t_2) := \lceil \min(t_1, t_2)^2 \rceil$ and by means of Proposition 4.1.4 and of Example 4.1.5 we compute

$$\begin{aligned}
&\overline{\lambda}\Big(\sup_{M' \in \mathcal{M}_{t_1}} \sup_{N' \in \mathcal{N}_{t_2}} d(M' \cap N', M \cap N) \Big) \\
&\quad \leq \max\left(\overline{\lambda}\Big(\delta^{2k(t_1,t_2)-1}\Big), \overline{\lambda}\left(\left(\frac{1+\delta}{2}\right)^{2k(t_1,t_2)-1} \right), \right. \\
&\qquad \left. \overline{\lambda}(k(t_1,t_2)) + \max\Big(\overline{\lambda}\Big(\sup_{M' \in \mathcal{M}_{t_1}} d(M', M) \Big), \overline{\lambda}\Big(\sup_{N' \in \mathcal{N}_{t_2}} d(N', N) \Big) \Big) \right) \\
&\quad = \max(\lambda_M, \lambda_N).
\end{aligned}$$

□

Corollary 4.2.9 helps us to combine the forward with the backward phase of Ginelli's algorithm. Indeed, linear spans of forward propagated vectors (past to present) approximate the Oseledets filtration of the reversed cocycle, whereas spans of backward propagated vectors (future to present) give us approximations of the Oseledets filtration of the cocycle. Since intersections of filtration spaces for the cocycle and for the reversed cocycle yield Oseledets spaces (see Eq. (2.3)), we can combine estimates for forward and for backward propagation to get estimates for our approximations of Oseledets spaces. Later on, the sets $\mathcal{M}_{t_1}$ and $\mathcal{N}_{t_2}$ will contain subspaces spanned by admissible tuples of input vectors that ensure convergence of Ginelli's algorithm.

4.3 Admissible tuples

Ultimately, the MET provides an asymptotic link between singular vectors (resp. singular values) and Oseledets spaces (resp. LEs). Hence, in order to investigate how a tuple of vectors evolves under subsequent applications of linear maps and of orthonormalizations, we relate the tuple to singular vectors. The relation is represented by a single parameter α. It describes how strong the corresponding filtrations are correlated. A value of 0 means no correlation and a value of 1 implies equality. Thus, we call tuples that have a positive value of α *admissible*. A special task will be to understand how many tuples fulfill a certain level of admissibility. For this purpose, we denote by μ the Lebesgue measure of the respective dimension.

To shorten our notation we set

$$U_i^{(b)} := \operatorname{span}\Big(b_{i_1}, \ldots, b_{i_{d_i}}\Big)$$

for a given d-tuple (b) and for degeneracies given by LEs. The filtration corresponding to (b) is given by $\{0\} \subset \overline{U}_1^{(b)} \subset \cdots \subset \overline{U}_p^{(b)} = \mathbb{R}^d$, where

$$\overline{U}_i^{(b)} := \bigoplus_{j=1}^{i} U_j^{(b)} = \operatorname{span}\Big(b_{1_1}, \ldots, b_{i_{d_i}}\Big).$$

We denote the associated orthogonal projections by

$$P_i^{(b)} := P_{U_i^{(b)}} \quad \text{and} \quad \overline{P}_i^{(b)} := P_{\overline{U}_i^{(b)}}.$$

Note that $P_i^{(b)}$ is defined differently in [48]. However, the notation is only used for orthogonal tuples for which both definitions coincide. Moreover, if (b) is orthogonal, it holds

$$\overline{P}_i^{(b)} = \sum_{j=1}^{i} P_j^{(b)}.$$

Definition 4.3.1
*Let $0 < \alpha \leq 1$ and a basis (c) of $\mathbb{R}^d$ be given. A d-tuple (b) is called α-*admissible *with respect to (c) if it is linearly independent and*

$$\forall\, i < p: \; d\Big(\overline{U}_i^{(b)}, \overline{U}_i^{(c)}\Big)^2 \leq 1 - \alpha^2.$$

We denote the set of all α-admissible tuples by $\mathcal{A}d^{(c)}(\alpha)$ and the set of all tuples that are admissible for some $\alpha > 0$ by $\mathcal{A}d^{(c)}$.

As admissibility is described by distances of filtration spaces, we are allowed to interchange the involved tuples with orthonormalized versions. So, let us assume that (c) is an orthonormal basis from now on. Moreover, the invariance of distances under orthogonal transformations implies that α-admissibility of (b) w.r.t. (c) is equivalent to α-admissibility of (Ab) w.r.t. (Ac) for all $A \in O(d, \mathbb{R})$. Hence, $A^d\Big(\mathcal{A}d^{(c)}(\alpha)\Big)$ and $\mathcal{A}d^{(Ac)}(\alpha)$ coincide.

Next, let us proceed with an alternative characterization of admissibility:

Lemma 4.3.2
A basis (b) is α-admissible w.r.t. (c) if, and only if, for all $i < p$ and $x \in \overline{U}_i^{(b)}$ with $\|x\| = 1$, we have

$$\sum_{j=1}^{i} \sum_{k} |\langle x, c_{j_k}\rangle|^2 \geq \alpha^2.$$

Proof. We reformulate the distance between filtration spaces as follows:

$$\begin{aligned}
\left\| \Big(I - \overline{P}_i^{(c)}\Big)\overline{P}_i^{(b)} \right\|^2 &= \max_{x \in S} \left\| \Big(I - \overline{P}_i^{(c)}\Big)\overline{P}_i^{(b)} x \right\|^2 \\
&= \max_{x \in \overline{U}_i^{(b)} \cap S} \left\| \Big(I - \overline{P}_i^{(c)}\Big) x \right\|^2 \\
&= 1 - \min_{x \in \overline{U}_i^{(b)} \cap S} \left\| \overline{P}_i^{(c)} x \right\|^2 \\
&= 1 - \min_{x \in \overline{U}_i^{(b)} \cap S} \sum_{j=1}^{i} \sum_{k} |\langle x, c_{j_k}\rangle|^2,
\end{aligned}$$

where $S \subset \mathbb{R}^d$ denotes the unit sphere. $\square$

Now, we are able to relate the evolution of a tuple under a linear map to singular vectors. As it turns out, the relation is sensitive to the admissibility parameter. In fact, being able to control the following estimate was a major reason to introduce the concept of admissibility.

Proposition 4.3.3
Let $A = U\Sigma V^T$ be the SVD of an invertible matrix and $0 < \alpha \leq 1$. For all $(b) \in \mathcal{A}d^{(v)}(\alpha)$, it holds

$$\forall i : d\left(\overline{U}_i^{(Ab)}, \overline{U}_i^{(u)}\right) \leq \frac{1}{\alpha} \frac{\delta_{i+1}^{\max}}{\delta_i^{\min}},$$

where $(u), (v)$ are the singular vectors corresponding to singular values (δ).

Proof. First, express $x \in \mathbb{R}^d$ using right singular vectors:

$$x = \sum_{j_k} \langle x, v_{j_k} \rangle v_{j_k}.$$

Applying the linear map $A = U\Sigma V^T$, we get

$$Ax = \sum_{j_k} \langle x, v_{j_k} \rangle \delta_{j_k} u_{j_k} \qquad \Longrightarrow \qquad \|Ax\|^2 = \sum_{j_k} |\langle x, v_{j_k} \rangle|^2 \delta_{j_k}^2.$$

For $x \in \overline{U}_i^{(b)}$ with $\|x\| = 1$, this means

$$\|Ax\|^2 \geq \sum_{j=1}^{i} \sum_{k} |\langle x, v_{j_k} \rangle|^2 \delta_{j_k}^2 \geq \left(\delta_i^{\min}\right)^2 \sum_{j=1}^{i} \sum_{k} |\langle x, v_{j_k} \rangle|^2 \geq \alpha^2 \left(\delta_i^{\min}\right)^2$$

by admissibility of (b). Moreover, the following holds for $x \in \mathbb{R}^d$ with $\|x\| = 1$:

$$\left\| \left(I - \overline{P}_i^{(u)}\right) Ax \right\|^2 = \sum_{j>i} \sum_{k} |\langle x, v_{j_k} \rangle|^2 \delta_{j_k}^2 \leq \left(\delta_{i+1}^{\max}\right)^2.$$

Now, we compute

$$\begin{aligned} d\left(\overline{U}_i^{(Ab)}, \overline{U}_i^{(u)}\right) &= \left\| \left(I - \overline{P}_i^{(u)}\right) \overline{P}_i^{(Ab)} \right\| \\ &= \max_{y \in \overline{U}_i^{(Ab)} \setminus \{0\}} \frac{\left\| \left(I - \overline{P}_i^{(u)}\right) y \right\|}{\|y\|} \\ &= \max_{x \in \overline{U}_i^{(b)} \setminus \{0\}} \frac{\left\| \left(I - \overline{P}_i^{(u)}\right) Ax \right\|}{\|Ax\|} \\ &= \max_{x \in \overline{U}_i^{(b)} \cap S} \frac{\left\| \left(I - \overline{P}_i^{(u)}\right) Ax \right\|}{\|Ax\|} \\ &\leq \frac{1}{\alpha} \frac{\delta_{i+1}^{\max}}{\delta_i^{\min}}. \end{aligned}$$

□

The above proposition only describes behavior of admissible tuples. However, it turns out that almost every tuple is admissible. Indeed, for admissibility to be generic, the complement of the open set

$$\mathcal{A}d^{(c)} = \left\{ (b) \text{ basis} \;\middle|\; \forall i : d\left(\overline{U}_i^{(b)}, \overline{U}_i^{(c)}\right) < 1 \right\} \subset \left(\mathbb{R}^d\right)^d$$

must be a set of measure zero. Using Proposition 4.2.2, we can rewrite the condition as follows:

$$d\Big(\overline{U}_i^{(b)}, \overline{U}_i^{(c)}\Big) < 1 \quad \iff \quad \overline{U}_i^{(b)} \oplus \Big(\overline{U}_i^{(c)}\Big)^{\perp} = \mathbb{R}^d.$$

Since (c) is an orthonormal basis, we yet have another equivalent formulation on the level of basis vectors:

$$d\Big(\overline{U}_i^{(b)}, \overline{U}_i^{(c)}\Big) < 1 \quad \iff \quad \det\Big(b_{1_1}, \ldots, b_{i_{d_i}}, c_{(i+1)_1}, \ldots, c_{p_{d_p}}\Big) \neq 0.$$

This form easily reveals the following:

Proposition 4.3.4
The set of nonadmissible tuples $\left(\mathbb{R}^d\right)^d \setminus \mathcal{A}d^{(c)}$ *has Lebesgue measure zero.*

Proof. In the above expression write vectors of (b) as coefficients in terms of (c). Now, the claim is a direct consequence of the fact that $\det^{-1}(0) \subset \mathbb{R}^{k\times k}$ is a subset of measure zero for each $k \geq 1$. □

Restricted to a domain of finite measure, the last proposition tells us that the measure of non-α-admissible tuples converges to zero as α goes to zero:

Corollary 4.3.5
For each subset $\mathcal{F} \subset \left(\mathbb{R}^d\right)^d$ *of finite Lebesgue measure, it holds*

$$\lim_{\alpha \searrow 0} \mu\Big(\mathcal{F} \setminus \mathcal{A}d^{(c)}(\alpha)\Big) = 0.$$

Proof. This is a direct consequence of the previous result and of the continuity of the Lebesgue measure:

$$\lim_{\alpha \searrow 0} \mu\Big(\mathcal{F} \setminus \mathcal{A}d^{(c)}(\alpha)\Big) = \mu\left(\bigcap_{0<\alpha\leq 1} \mathcal{F} \setminus \mathcal{A}d^{(c)}(\alpha)\right) = \mu\Big(\mathcal{F} \setminus \mathcal{A}d^{(c)}\Big) = 0.$$

□

We now have a better understanding of how most tuples of vectors evolve compared to singular vectors. In particular, Proposition 4.3.3 can be used to describe forward and backward propagation. However, backward propagation requires additional treatment since it depends on the forward phase. Vectors for the backward phase are initiated inside spaces spanned by forward propagated vectors. Thus, there is a restriction on the domain of initial vectors for backward propagation.

In the following we extend the concept of admissibility to adjust to the above situation. Let (b) be a linearly independent d-tuple such that

$$b_{i_1}, \ldots, b_{i_{d_i}} \in \operatorname{span}\Big(c_{i_1}, \ldots, c_{p_{d_p}}\Big) = U_i^{(c)} \oplus \cdots \oplus U_p^{(c)}$$

for each i. Instead of admissibility, it will be enough that $b_{i_1}, \ldots, b_{i_{d_i}}$ can be extended to an admissible tuple of the form

$$\Big(*, \ldots, *, b_{i_1}, \ldots, b_{i_{d_i}}, *, \ldots, *\Big) \in \mathcal{A}d^{(c)}(\alpha).$$

The set of all (b) satisfying this extension property will be denoted by $\mathcal{A}d_{\text{ext}}^{(c)}(\alpha)$. We write $\mathcal{A}d_{\text{ext}}^{(c)}$ for the union of these sets over $0 < \alpha \leq 1$.

As before, one readily checks that $A^d(\mathcal{A}d_{\text{ext}}^{(c)}(\alpha)) = \mathcal{A}d_{\text{ext}}^{(Ac)}(\alpha)$ for $A \in O(d, \mathbb{R})$. Moreover, we again conclude that almost every tuple satisfies *extended admissibility*:

Proposition 4.3.6
The set

$$\left(\left(U_1^{(c)} \oplus \cdots \oplus U_p^{(c)}\right)^{d_1} \times \left(U_2^{(c)} \oplus \cdots \oplus U_p^{(c)}\right)^{d_2} \times \cdots \times \left(U_p^{(c)}\right)^{d_p}\right) \setminus \mathcal{A}d_{\mathrm{ext}}^{(c)}$$

has Lebesgue measure zero.

Proof. For each i, we show that the set of tuples

$$\left(b_{i_1}, \ldots, b_{i_{d_i}}\right) \in \left(U_i^{(c)} \oplus \cdots \oplus U_p^{(c)}\right)^{d_i}$$

not satisfying the extension property has Lebesgue measure zero.

The idea is to apply Proposition 4.3.4 to a reduced setting for fixed i. To this end, look at $\mathbb{R}^{d'}$ with degeneracies $d' = d'_1 + \cdots + d'_{p'}$ given by $d'_j := d_{i-1+j}$ for all $j = 1, \ldots, p' := p + 1 - i$, and let (e') be its standard basis. We get

$$\mu\left(\left(\mathbb{R}^{d'}\right)^{d'} \setminus \mathcal{A}d^{(e')}\right) = 0.$$

In particular, this implies

$$\mu\left(\left(\mathbb{R}^{d'}\right)^{d'_1} \setminus \left\{\left(b'_{1_1}, \ldots, b'_{1_{d'_1}}\right) \text{ has admissible extension}\right\}\right) = 0.$$

Now, we transfer the result from $\mathbb{R}^{d'}$ to $U_i^{(c)} \oplus \cdots \oplus U_p^{(c)}$ by identifying (e') with $(c_{i_1}, \ldots, c_{p_{d_p}})$. As an identification between orthonormal bases, Lebesgue measure, distance between subspaces, and admissibility are preserved. Hence, for almost every given tuple $(b_{i_1}, \ldots, b_{i_{d_i}}) \in (U_i^{(c)} \oplus \cdots \oplus U_p^{(c)})^{d_i}$, we find $0 < \alpha \leq 1$ and $g_{(i+1)_1}, \ldots, g_{p_{d_p}} \in U_i^{(c)} \oplus \cdots \oplus U_p^{(c)}$ such that

$$d\left(\operatorname{span}\left(b_{i_1}, \ldots, b_{i_{d_i}}\right), U_i^{(c)}\right)^2 \leq 1 - \alpha^2$$

and

$$\forall\, j > i : \, d\left(\operatorname{span}\left(b_{i_1}, \ldots, b_{i_{d_i}}, g_{(i+1)_1}, \ldots, g_{j_{d_j}}\right), U_i^{(c)} \oplus \cdots \oplus U_j^{(c)}\right)^2 \leq 1 - \alpha^2.$$

We can extend such a tuple

$$\left(b_{i_1}, \ldots, b_{i_{d_i}}, g_{(i+1)_1}, \ldots, g_{p_{d_p}}\right)$$

to an α-admissible tuple (g) by setting $g_{j_k} := c_{j_k}$ for $j < i$. This concludes the proof. ☐

As a consequence, we get the following corollary:

Corollary 4.3.7
Given a subset $\mathcal{F} \subset (U_1^{(c)} \oplus \cdots \oplus U_p^{(c)})^{d_1} \times \cdots \times (U_p^{(c)})^{d_p}$ of finite Lebesgue measure, it holds

$$\lim_{\alpha \searrow 0} \mu\left(\mathcal{F} \setminus \mathcal{A}d_{\mathrm{ext}}^{(c)}(\alpha)\right) = 0.$$

In the discrete-time convergence proof of Ginelli's algorithm, a more precise measure-estimate of non-α-admissible tuples will be necessary. However, it will be sufficient to know the case, where $\mathcal{F}$ is a products of balls. The rest of Section 4.3 is devoted to a rather technical derivation of explicit estimates needed only for the proof of Theorem 4.4.4.

Proposition 4.3.8
Let $d > 1$. There is a constant $\eta = \eta(d, M) > 0$ such that

$$\mu\Big(B_d(0,M)^d \setminus \mathcal{A}d^{(c)}(\alpha)\Big) \leq \eta\alpha^{\frac{1}{d-1}},$$

where $B_d(0,M)$ denotes the ball of radius M in $(\mathbb{R}^d, \|.\|_2)$ centered at the origin.

Two lemmata on how to construct admissible tuples will guide us to the above proposition. Since admissible tuples for the nondegenerate case are admissible for all possible degenerate cases, it is enough to find an estimate for the nondegenerate case.

Lemma 4.3.9
Let (f) be an orthonormal basis of $\mathbb{R}^d$. Fix $1 < i < d$ and $0 < \alpha_1, \alpha_2 \leq 1$. If

$$\left\|P_{\mathrm{span}(f_1,\ldots,f_{i-1},c_{i+1},\ldots,c_d)} f_i\right\|^2 \leq 1-\alpha_1^2 \quad \textit{and} \quad \left\|\overline{P}^{(f)}_{i-1}\Big(I-\overline{P}^{(c)}_i\Big)\right\|^2 \leq 1-\alpha_2^2,$$

then

$$d\Big(\overline{U}^{(f)}_i, \overline{U}^{(c)}_i\Big)^2 \leq 1-(\alpha_1\alpha_2)^2.$$

Proof. First, we reduce the problem to the case $i = 2$ and $d = 3$: There are unit vectors $f'_1 \in \mathrm{span}(f_1, \ldots, f_{i-1})$ and $c'_3 \in \mathrm{span}(c_{i+1}, \ldots, c_d)$ such that

$$\left\|\overline{P}^{(f)}_i\Big(I-\overline{P}^{(c)}_i\Big)\right\|^2 = \left\|\overline{P}^{(f)}_i c'_3\right\|^2 = \left\|\overline{P}^{(f)}_{i-1} c'_3\right\|^2 + |\langle f_i, c'_3\rangle|^2 = |\langle f'_1, c'_3\rangle|^2 + |\langle f'_2, c'_3\rangle|^2$$

with $f'_2 := f_i$. Furthermore, the assumptions yield

$$\left\|P_{\mathrm{span}(f'_1,c'_3)} f'_2\right\|^2 \leq \left\|P_{\mathrm{span}(f_1,\ldots,f_{i-1},c_{i+1},\ldots,c_d)} f_i\right\|^2 \leq 1-\alpha_1^2$$

and

$$|\langle f'_1, c'_3\rangle|^2 \leq \left\|\overline{P}^{(f)}_{i-1} c'_3\right\|^2 \leq \left\|\overline{P}^{(f)}_{i-1}\Big(I-\overline{P}^{(c)}_i\Big)\right\|^2 \leq 1-\alpha_2^2.$$

In particular, f'_1, f'_2 and c'_3 are linearly independent. Thus, the problem reduces to finding the right estimate of

$$d\Big(\overline{U}^{(f')}_2, \overline{U}^{(c')}_2\Big)^2 = \left\|\overline{P}^{(f')}_2 c'_3\right\|^2 = |\langle f'_1, c'_3\rangle|^2 + |\langle f'_2, c'_3\rangle|^2$$

inside $\mathrm{span}(f'_1, f'_2, c'_3) \cong \mathbb{R}^3$, where (f') and (c') are some orthonormal bases of $\mathrm{span}(f'_1, f'_2, c'_3)$ extending (f'_1, f'_2) and c'_3.

The case $i = 2$ and $d = 3$ can be shown by a short calculation. It holds

$$\begin{aligned}\left\|P_{\mathrm{span}(f'_1,c'_3)} f'_2\right\|^2 &= |\langle f'_1, f'_2\rangle|^2 + \left|\left\langle \frac{c'_3 - \langle f'_1, c'_3\rangle f'_1}{\|c'_3 - \langle f'_1, c'_3\rangle f'_1\|}, f'_2\right\rangle\right|^2 \\ &= \frac{|\langle c'_3, f'_2\rangle|^2}{\|c'_3 - \langle f'_1, c'_3\rangle f'_1\|^2} \\ &= \frac{|\langle c'_3, f'_2\rangle|^2}{1-|\langle f'_1, c'_3\rangle|^2}.\end{aligned}$$

Thus, by our assumptions:

$$|\langle f_2', c_3'\rangle|^2 = \left\|P_{\operatorname{span}(f_1',c_3')} f_2'\right\|^2 (1 - |\langle f_1', c_3'\rangle|^2) \le (1-\alpha_1^2)(1 - |\langle f_1', c_3'\rangle|^2).$$

We estimate

$$\begin{aligned} |\langle f_1', c_3'\rangle|^2 + |\langle f_2', c_3'\rangle|^2 &\le |\langle f_1', c_3'\rangle|^2 + (1-\alpha_1^2)(1 - |\langle f_1', c_3'\rangle|^2) \\ &= 1 - \alpha_1^2 + \alpha_1^2 \, |\langle f_1', c_3'\rangle|^2 \\ &\le 1 - \alpha_1^2 + \alpha_1^2 (1 - \alpha_2^2) \\ &= 1 - (\alpha_1 \alpha_2)^2. \end{aligned}$$

□

The previous lemma can be used to give a sufficient condition for a tuple to be α-admissible:

Lemma 4.3.10
If a basis (b) *satisfies*

$$\forall\, i < d : \left\|P_{\operatorname{span}(f_1,\dots,f_{i-1},c_{i+1},\dots,c_d)} f_i\right\|^2 \le 1 - \left(\alpha^{\frac{1}{d-1}}\right)^2,$$

where $(f) := \operatorname{orth}(b)$*, then* (b) *is* α*-admissible.*

Proof. We prove the result by induction over i showing that

$$d\left(\overline{U}_i^{(b)}, \overline{U}_i^{(c)}\right)^2 = d\left(\overline{U}_i^{(f)}, \overline{U}_i^{(c)}\right)^2 \le 1 - \left(\alpha^{\frac{i}{d-1}}\right)^2 \le 1 - \alpha^2.$$

For $i = 1$, we have

$$d\left(\overline{U}_1^{(f)}, \overline{U}_1^{(c)}\right)^2 = \left\|\left(I - \overline{P}_1^{(c)}\right) f_1\right\|^2 = \left\|P_{\operatorname{span}(c_2,\dots,c_d)} f_1\right\|^2 \le 1 - \left(\alpha^{\frac{1}{d-1}}\right)^2.$$

Let $1 < i < d$ and assume the induction hypothesis is true for $i-1$, which implies that

$$\left\|\overline{P}_{i-1}^{(f)}\left(I - \overline{P}_i^{(c)}\right)\right\|^2 \le \left\|\overline{P}_{i-1}^{(f)}\left(I - \overline{P}_{i-1}^{(c)}\right)\right\|^2 = d\left(\overline{U}_{i-1}^{(f)}, \overline{U}_{i-1}^{(c)}\right)^2 \le 1 - \left(\alpha^{\frac{i-1}{d-1}}\right)^2.$$

Simply apply Lemma 4.3.9 to close the induction step. □

Proof of Proposition 4.3.8. Set $\tilde{\alpha} := \alpha^{\frac{1}{d-1}}$ and let

$$\mathcal{N} := \left\{ (b) \in B_d(0,M)^d \;\middle|\; \exists i : \det(b_1,\dots,b_i,c_{i+1},\dots,c_d) = 0 \right\}$$

be the set of all nonadmissible tuples inside $B_d(0,M)^d$. From Proposition 4.3.4 we know that $\mathcal{N}$ has measure zero. On its complement we define a continuous mapping into the d-fold product of spheres:

$$w : B_d(0,M)^d \setminus \mathcal{N} \to S^d$$

with components $w_i(b_1,\dots,b_d) := \operatorname{orth}_d(b_1,\dots,b_{i-1},c_{i+1},\dots,c_d,c_i)$, where orth_d is the last component of an orthonormalization procedure respecting the order of the tuple (e.g., use the Gram-Schmidt procedure). By construction $w_i = w_i(b_1,\dots,b_d)$

is the up to sign unique unit vector orthogonal to $\operatorname{span}(b_1, \dots, b_{i-1}, c_{i+1}, \dots, c_d)$, and only depends on the first $i-1$ vectors of (b). w will help us to measure sets of admissible vectors.

The tuple $(f) = \operatorname{orth}(b)$ is constructed by setting $f_i := b_i'/\|b_i'\|$ with $b_i' := (I - \overline{P}_{i-1}^{(b)})b_i$. Assuming $|\langle w_i, b_i\rangle| \geq M\tilde{\alpha}$, we get

$$\begin{aligned}
\left\|P_{\operatorname{span}(f_1,\dots,f_{i-1},c_{i+1},\dots,c_d)} f_i\right\|^2 &= \left\|P_{\operatorname{span}(b_1,\dots,b_{i-1},c_{i+1},\dots,c_d)} f_i\right\|^2 \\
&= 1 - |\langle w_i, f_i\rangle|^2 \\
&= 1 - \frac{|\langle w_i, b_i'\rangle|^2}{\|b_i'\|^2} \\
&= 1 - \frac{|\langle w_i, b_i\rangle|^2}{\|b_i'\|^2} \\
&\leq 1 - \frac{|\langle w_i, b_i\rangle|^2}{\|b_i\|^2} \\
&\leq 1 - \frac{|\langle w_i, b_i\rangle|^2}{M^2} \\
&\leq 1 - \tilde{\alpha}^2.
\end{aligned}$$

Hence, if $(b) \in B_d(0, M)^d \setminus \mathcal{N}$ satisfies

$$\forall\, i < d : \, |\langle w_i, b_i\rangle| \geq M\tilde{\alpha},$$

then (b) is α-admissible by Lemma 4.3.10. In particular, the subset of all non-α-admissible tuples is contained in the subset of all (b), which either do not fulfill the above condition or which are elements of the set of measure zero $\mathcal{N}$. Therefore, a measure-estimate on tuples not fulfilling the condition is enough for the claim:

$$\begin{aligned}
&\mu\Big(B_d(0, M)^d \setminus \mathcal{A}d^{(c)}(\alpha)\Big) \\
&\quad\leq \mu\Big(\Big\{\,(b) \in B_d(0, M)^d \setminus \mathcal{N} \;\Big|\; \exists\, i < d : \, |\langle w_i, b_i\rangle| < M\tilde{\alpha}\Big\}\Big) \\
&\quad\leq \sum_{i<d} \mu\Big(\Big\{(b) \in B_d(0, M)^d \setminus \mathcal{N} \;:\; |\langle w_i, b_i\rangle| < M\tilde{\alpha}\Big\}\Big) \\
&\quad= \sum_{i<d} \mu\Big(\Big\{\,(b) \in B_d(0, M)^d \;\Big|\; \det(b_1, \dots, b_{i-1}, c_i, \dots, c_d) \neq 0 \text{ and } |\langle w_i, b_i\rangle| < M\tilde{\alpha}\Big\}\Big) \\
&\quad= \sum_{i<d} (\mu(B_d(0, M)))^{d-i} \int_{\{\,(b_1,\dots,b_{i-1}) \in B_d(0,M)^{i-1} \;|\; \det(b_1,\dots,b_{i-1},c_i,\dots,c_d) \neq 0\}} \\
&\qquad \int_{\{b_i \in B_d(0,M) \;:\; |\langle w_i, b_i\rangle| < M\tilde{\alpha}\}} 1 \, db_i \, d(b_1, \dots, b_{i-1}) \\
&\quad\underset{(\star)}{=} \sum_{i<d} (\mu(B_d(0, M)))^{d-i} \int_{\{\,(b_1,\dots,b_{i-1}) \in B_d(0,M)^{i-1} \;|\; \det(b_1,\dots,b_{i-1},c_i,\dots,c_d) \neq 0\}} \\
&\qquad \int_{\{b_i \in B_d(0,M) \;:\; |\langle e_1, b_i\rangle| < M\tilde{\alpha}\}} 1 \, db_i \, d(b_1, \dots, b_{i-1}) \\
&\quad= \sum_{i<d} (\mu(B_d(0, M)))^{d-1} \mu\Big(B_d(0, M) \cap \Big((-M\tilde{\alpha}, M\tilde{\alpha}) \times \mathbb{R}^{d-1}\Big)\Big) \\
&\quad\leq (d-1)(\mu(B_d(0, M)))^{d-1} (2M)^d \tilde{\alpha}.
\end{aligned}$$

We used Fubini's theorem to measure components separately. In $(\star)$ we rotated w_i to the first vector of the standard basis. Afterwards, we enlarged $B_d(0, M) \cap ((-M\tilde{\alpha}, M\tilde{\alpha}) \times \mathbb{R}^{d-1})$ to $(-M\tilde{\alpha}, M\tilde{\alpha}) \times (-M, M)^{d-1}$ for a simpler estimate.

Now, setting $\eta := (d-1)(\mu(B_d(0,M)))^{d-1}(2M)^d$ yields the desired estimate. $\square$

A similar estimate will be necessary for non-α-admissible tuples with respect to extended admissibility.

Proposition 4.3.11
Let $d > 1$. There is a constant $\eta = \eta(d, M) > 0$ such that

$$\mu\Big(B(M) \setminus \mathcal{A}d^{(c)}_{\mathrm{ext}}(\alpha)\Big) \le \eta\alpha^{\frac{1}{d-1}},$$

where $B(M)$ is given by a product of balls of radius M inside the special domain:

$$B(M) := B_d(0,M)^{d_1} \times \cdots \times B_{d_p}(0,M)^{d_p} \subset \Big(U_1^{(c)} \oplus \cdots \oplus U_p^{(c)}\Big)^{d_1} \times \cdots \times \Big(U_p^{(c)}\Big)^{d_p}.$$

Proof. The proof is similar to the one of Proposition 4.3.6. Again, it is enough to find such a bound for the set of all tuples in

$$B_{d_i+\cdots+d_p}(0,M)^{d_i} \subset \Big(U_i^{(c)} \oplus \cdots \oplus U_p^{(c)}\Big)^{d_i}$$

that cannot be extended to an α-admissible tuple.

Using the same identification as before, we reduce the problem to finding such an estimate for the set

$$B_{d'}(0,M)^{d'_1} \setminus \Big\{\Big(b'_{1_1}, \ldots, b'_{1_{d'_1}}\Big) \text{ has an } \alpha\text{-admissible extension}\Big\}.$$

Proposition 4.3.8 yields η' only depending on d' and M with

$$\mu\Big(B_{d'}(0,M)^{d'} \setminus \mathcal{A}d^{(e')}(\alpha)\Big) \le \eta'\alpha^{\frac{1}{d'-1}}.$$

This implies

$$\begin{aligned}&\mu\Big(B_{d'}(0,M)^{d'_1} \setminus \Big\{\Big(b'_{1_1}, \ldots, b'_{1_{d'_1}}\Big) \text{ has an } \alpha\text{-admissible extension}\Big\}\Big)\\ &\le \left(\frac{1}{\mathrm{vol}(B_{d'}(0,M))}\right)^{d'-d'_1} \eta'\alpha^{\frac{1}{d'-1}}.\end{aligned}$$

Finally, an estimate η only depending on M and d is achieved by taking the maximum over estimates for all possible combinations of degeneracies. $\square$

4.4 Convergence results

Finally, we have gathered enough background knowledge to state and to prove our convergence results for Ginelli's algorithm on $\mathbb{R}^d$. However, before stating the convergence theorems we motivate our results via two simple examples.

Example 4.4.1 (diagonal cocycle)
Assume $\Omega = \{\omega\}$ with trivial flow $\sigma_t\omega = \omega$. For given $\lambda_1 > \cdots > \lambda_p$, let $\mathcal{L}_\omega :=$ $\mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_p})$. The generator defines a cocycle whose CLVs (at ω) coincide with the standard basis (e) of $\mathbb{R}^d$.

Now, fix a vector $b_1 \in \mathbb{R}^d$ with $|\langle b_1, e_1\rangle| > 0$. We have

$$\langle \mathcal{L}_\omega^{(t)} b_1, e_i\rangle = \langle b_1, e_i\rangle e^{t\lambda_i}.$$

Thus, we compute

$$\left|\left\langle \frac{\mathcal{L}_\omega^{(t)} b_1}{\|\mathcal{L}_\omega^{(t)} b_1\|}, e_i\right\rangle\right|^2 = \frac{|\langle b_1, e_i\rangle|^2 e^{2t\lambda_i}}{\sum_j |\langle b_1, e_j\rangle|^2 e^{2t\lambda_j}} = \frac{|\langle b_1, e_i\rangle|^2 e^{-2t|\lambda_1-\lambda_i|}}{\sum_j |\langle b_1, e_j\rangle|^2 e^{-2t|\lambda_1-\lambda_j|}}.$$

The last nominator takes values between $|\langle b_1, e_1\rangle|^2$ and $\|b_1\|^2$. In particular, it can be treated as a positive constant for the Lyapunov index notation:

$$\begin{aligned}
\lambda\Big(d\Big(\overline{U}_1^{(\mathcal{L}_\omega^{(t)} b)}, \overline{U}_1^{(e)}\Big)\Big) &= \frac{1}{2}\lambda\left(\left\|\Big(I - \overline{P}_1^{(e)}\Big)\overline{P}_1^{(\mathcal{L}_\omega^{(t)} b)}\right\|^2\right) \\
&= \frac{1}{2}\lambda\left(\sum_{i\neq 1}\left|\left\langle \frac{\mathcal{L}_\omega^{(t)} b_1}{\|\mathcal{L}_\omega^{(t)} b_1\|}, e_i\right\rangle\right|^2\right) \\
&\leq \max_{i\neq 1} \frac{1}{2}\lambda\left(\left|\left\langle \frac{\mathcal{L}_\omega^{(t)} b_1}{\|\mathcal{L}_\omega^{(t)} b_1\|}, e_i\right\rangle\right|^2\right) \\
&\leq \max_{i\neq 1} -|\lambda_1 - \lambda_i| \\
&= -|\lambda_1 - \lambda_2|.
\end{aligned}$$

In general, the upcoming convergence analysis will show that

$$\lambda\Big(d\Big(\overline{U}_i^{(\mathcal{L}_\omega^{(t)} b)}, \overline{U}_i^{(e)}\Big)\Big) \leq -|\lambda_i - \lambda_{i+1}|$$

for all tuples (b) that are admissible w.r.t. (e).

Ginelli's algorithm starts with a random choice of initial vectors to prevent non-admissible configurations. One such configuration would be the unlikely case where the first vector lies in the second Oseledets space. As Oseledets spaces are equivariant, the first vector would stay inside the second Oseledets space when propagated.[4] Consequently, it would not be a good approximation of a CLV from the first Oseledets space. The next example shows that this problems might occur for every tuple when initiated at a wrong time in the continuous version of Ginelli's algorithm.

Example 4.4.2 (rotating Oseledets spaces)
Let $\Omega := S^1 \cong \mathbb{R}/\mathbb{Z}$ be a periodic trajectory with homogeneous flow $\sigma_t\omega := \omega + t$. Furthermore, let $R : \mathbb{R} \to SO(2)$ be the parameterization of $SO(2)$ by 2×2 rotation matrices

$$R(\omega) := \begin{pmatrix} \cos(2\pi\omega) & -\sin(2\pi\omega) \\ \sin(2\pi\omega) & \cos(2\pi\omega) \end{pmatrix},$$

so that $R(0) = R(1) = I$ and $R(s+t) = R(s)R(t)$. Moreover, we set $D := \operatorname{diag}(e^{\lambda_1}, e^{\lambda_2})$ for some $\lambda_1 > \lambda_2$, and define the cocycle to be

$$\mathcal{L}_\omega^{(t)} := R(\sigma_t\omega) D^t R(-\omega).$$

[4] This statement is true analytically. Numerically, errors would kick the propagated vector out of the second Oseledets space.

One readily checks that $\mathcal{L}$ indeed is a cocycle over σ.

Next, we use the characterization of Oseledets spaces via asymptotic growth rates:

$$\lim_{t\to\pm\infty} \frac{1}{t} \log \left\| \mathcal{L}_\omega^{(t)} R(\omega) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| = \begin{cases} \lambda_1 & x_1 \neq 0 \text{ and } x_2 = 0 \\ \lambda_2 & x_1 = 0 \text{ and } x_2 \neq 0 \end{cases}$$

to see that

$$Y_1(\omega) = \operatorname{span}\left(R(\omega) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) \quad \text{and} \quad Y_2(\omega) = \operatorname{span}\left(R(\omega) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right).$$

In particular, both Oseledets spaces are rotating uniformly with ω. Hence, for every fixed vector $b_1 \in \mathbb{R}^2$ and $T > 0$, we find $t_1 \in \mathbb{R}_{>0}$ bigger than T with $b_1 \in Y_2(\sigma_{-t_1}\omega)$. This implies that, for all fixed choices of b_1, the continuous version of Ginelli's algorithm does not converge. Instead, it is shown later that the continuous version converges in measure, i.e., if b_1 is chosen randomly.

In the discrete case, however, the set $\bigcup_{t_1\in\mathbb{N}} Y_2(\sigma_{-t_1}\omega)$ has Lebesgue measure zero indicating that the above problem occurs only on a set of measure zero. In fact, we will show convergence for almost every initial tuple in the discrete-time case.

Setting $D = \operatorname{diag}(e^{\lambda_1}, e^{\lambda_1})$ in the previous example yields a trivial Oseledets space $Y_1(\omega) = \mathbb{R}^2$ with inner rotation. In general, Oseledets spaces can have complicated internal dynamics that prevent single propagated vectors from converging. Additionally, CLVs are not unique in the presence of degeneracies. Therefore, objects of interest should not be the propagated vectors themselves, but rather the spaces spanned by them subject to degeneracies.[5]

4.4.1 Theorems

We now present convergence theorems for Ginelli's algorithm in terms of the map $G_{\omega,k}^{t_1,t_2}$ from Definition 3.1.1. Our convergence results are formulated for the case $k = d$, i.e., when all CLVs are computed, but can easily be adjusted if less CLVs are needed. As seen in the above examples, we cannot expect convergence for almost every input in a continuous-time setting. However, a weaker form of convergence holds:

Theorem 4.4.3 (Convergence in measure of Ginelli's algorithm on $\mathbb{R}^d$)
Assume the setting of Proposition 2.2.5, so that LEs and Oseledets spaces are well-defined. Moreover, set $\lambda_0 := \infty$ and $\lambda_{p+1} := -\infty$.

For each compact set of inputs $\mathcal{K} \subset (\mathbb{R}^d)^d \times \mathbb{R}_{ru}^{d\times d}$ and $\epsilon > 0$, Ginelli's algorithm converges in measure exponentially fast:[6]

[5] In practice, degeneracies can be derived from growth rates of propagated vectors during the forward phase of Ginelli's algorithm. Moreover, they might be forced by symmetries (e.g., in equivariant systems), whereas for some classes of systems degenerate scenarios are the exception [2].

[6] We remind the reader of the index notation i_j introduced in the beginning of Section 2.2, which counts indices with respect to degeneracies of the Lyapunov spectrum.

$$\begin{aligned}
\lim_{T\to\infty} \sup_{t_1,t_2\geq T} \mu\Big(\mathcal{K} \setminus \Big\{((b),R) \;\Big|\; &(b) \textit{ linearly independent and } \forall i: \\
&\frac{1}{\min(t_1,t_2)} \log d\Big(\operatorname{span}\Big\{ \Big(G^{t_1,t_2}_{\omega,d}((b),R)\Big)_{i_j} \;\Big|\; j=1,\ldots,d_i\Big\}, Y_i(\omega)\Big) \\
&\leq -\min(|\lambda_i-\lambda_{i-1}|,|\lambda_i-\lambda_{i+1}|)+\epsilon\Big\}\Big) \\
= 0.&
\end{aligned}$$

A stronger form of convergence holds for discrete time:

Theorem 4.4.4 (Convergence a.e. of Ginelli's algorithm on $\mathbb{R}^d$ with $\mathbb{T}=\mathbb{Z}$)
Assume the setting of Proposition 2.2.5 and set $\lambda_0 := \infty$ and $\lambda_{p+1} := -\infty$.

For Lebesgue almost every input $((b),R) \in (\mathbb{R}^d)^d \times \mathbb{R}^{d\times d}_{ru}$, Ginelli's algorithm converges exponentially fast:

$$\begin{aligned}
\limsup_{N\to\infty} \sup_{n_1,n_2\geq N} &\frac{1}{\min(n_1,n_2)} \log d\Big(\operatorname{span}\Big\{ \Big(G^{n_1,n_2}_{\omega,d}((b),R)\Big)_{i_j} \;\Big|\; j=1,\ldots,d_i\Big\}, Y_i(\omega)\Big) \\
&\leq -\min(|\lambda_i-\lambda_{i-1}|,|\lambda_i-\lambda_{i+1}|).
\end{aligned}$$

The theorems tell us that the output vectors of Ginelli's algorithm span subspaces which approximate Oseledets spaces. Since CLVs are normalized basis vectors subject to the Oseledets splitting, the output vectors approximate CLVs. In particular, the convergence is exponentially fast with a rate given by the gap between associated LEs. Hence, the total speed of convergence for computing all CLVs can be bounded from above (up to subexponential prefactors) by $\exp(-\min(n_1,n_2)\min_i|\lambda_i-\lambda_{i-1}|)$.

In applications one usually wants to compute CLVs at more than just one point along the trajectory. In fact, we can push forward and backward computed vectors to approximate CLVs near ω along the trajectory, i.e., on a set of the form $\{\sigma_t\omega \mid t \in [a,b]\}$. Since Oseledets spaces are equivariant and since applying the propagator for this time interval only yields a constant factor that vanishes on exponential scales (see Corollary 4.2.3), similar statements about convergence hold uniformly on bounded time intervals. Thus, it is enough to run the algorithm once to compute nearby CLVs along the trajectory.

Our convergence results apply to almost every trajectory in random dynamical systems. Indeed, since the deterministic MET holds $\mathbb{P}$-almost everywhere in the setting of Theorem 2.2.8, our results apply to almost every trajectory. Though, it is much harder to predict how the choice of the background trajectory affects the speed of convergence. A more detailed sensitivity analysis would require stability results for LEs and for Oseledets spaces.

In the following three subsections we prove Theorem 4.4.3 and Theorem 4.4.4. Since most arguments hold for discrete and for continuous time, we do not distinguish between both cases until after we have shown convergence in measure. Most statements are formulated using the Lyapunov index notation, thus, providing us with a direct link to the speed of convergence on exponential scales.

4.4.2 Link between MET and SVD

Let $\mathcal{L}^{(t)}_\omega = U(t)\Sigma(t)(V(t))^T$ be a SVD of the cocycle for $t \geq 0$, where singular values are ordered as in Eq. (2.1). Using right singular vectors, Arnold shows that the filtration $V_1(t) \supset \cdots \supset V_p(t)$ given by

$$V_i(t) := \left(\overline{U}_{i-1}^{(v(t))}\right)^{\perp}$$

converges exponentially fast to the Oseledets filtration $V_1(\omega) \supset \cdots \supset V_p(\omega)$ [1, proof of prop. 3.4.2]. Distances between filtrations are measured in a special metric. Unraveling the notation, we end up with

$$\forall\, i \neq j :\ \lambda\Big(\Big\|P_i^{(v(t))} P_j\Big\|\Big) \leq -|\lambda_i - \lambda_j|,$$

where $P_p + \cdots + P_i$ is the orthogonal projection onto $V_i(\omega)$ for each i.

Lemma 4.4.5
It holds

$$\forall i :\ \lambda\Big(d\Big(\overline{U}_i^{(v(t))}, (V_{i+1}(\omega))^{\perp}\Big)\Big) \leq -|\lambda_i - \lambda_{i+1}|.$$

Proof. We compute

$$\begin{aligned}
\lambda\Big(d\Big(\overline{U}_i^{(v(t))}, (V_{i+1}(\omega))^{\perp}\Big)\Big) &= \lambda\Big(\Big\|\overline{P}_i^{(v(t))} P_{V_{i+1}(\omega)}\Big\|\Big) \\
&\leq \lambda\left(\sum_{\substack{k,j \\ k \leq i < j}} \Big\|P_k^{(v(t))} P_j\Big\|\right) \\
&\leq \max_{\substack{k,j \\ k \leq i < j}} -|\lambda_k - \lambda_j| \\
&= -|\lambda_i - \lambda_{i+1}|.
\end{aligned}$$

□

A similar result holds for the time-reversed cocycle $\mathcal{L}_\omega^{-,(t)} = \mathcal{L}_\omega^{(-t)}$ over $\sigma^-(t) = \sigma_{-t}$ with SVD $U^-(t)\Sigma^-(t)(V^-(t))^T$ for $t \geq 0$, where singular values are ordered as in Eq. (2.1). Note that, for the time-reversed cocycle, we need to consider degeneracies in reversed order: $d_1^- = d_p, \ldots, d_p^- = d_1$. To distinguish between both types of degeneracies we equip the notation introduced in the beginning of Section 4.3 with a minus sign following the subindex whenever we count with respect to reversed degeneracies.

Lemma 4.4.6
It holds

$$\forall i :\ \lambda\Big(d\Big(\overline{U}_{i,-}^{(v^-(t))}, \big(V_{i+1}^-(\omega)\big)^{\perp}\Big)\Big) \leq -|\lambda_i^- - \lambda_{i+1}^-|.$$

Ginelli's algorithm starts by propagating vectors from past to present via $\mathcal{L}_{\sigma_{-t}\omega}^{(t)} = (\mathcal{L}_\omega^{(-t)})^{-1}$, and ends with propagating vectors from future to present via $\mathcal{L}_{\sigma_t\omega}^{(-t)} = (\mathcal{L}_\omega^{(t)})^{-1}$. Thus, it is important to keep track of singular vectors for inverted actions of cocycles as well. We use the same notation as in Section 2.2.1 to denote singular values and singular vectors of inverted maps.

Lemma 4.4.7
It holds

$$\forall i :\ \lambda\Big(d\Big(\overline{U}_{i,-}^{(\hat{u}(t))}, V_{p+1-i}(\omega)\Big)\Big) \leq -|\lambda_{p-i} - \lambda_{p+1-i}|.$$

Proof. This is a consequence of Lemma 4.4.5, since

$$\begin{aligned} d\Big(\overline{U}_{i,-}^{(\hat{u}(t))}, V_{p+1-i}(\omega)\Big) &= d\Big(\overline{U}_{i,-}^{(v(t))^r}, V_{p+1-i}(\omega)\Big) \\ &= d\bigg(\Big(\overline{U}_{p-i}^{(v(t))}\Big)^{\perp}, V_{p+1-i}(\omega)\bigg) \\ &= d\Big(\overline{U}_{p-i}^{(v(t))}, (V_{p+1-i}(\omega))^{\perp}\Big), \end{aligned}$$

where $(.)^r$ denotes the tuple in reversed order. Here, we used the identity

$$\overline{U}_{i,-}^{(c)^r} = \Big(\overline{U}_{p-i}^{(c)}\Big)^{\perp},$$

which is true for all orthonormal bases (c). $\square$

Again, we derive a similar result for reversed time:

Lemma 4.4.8
It holds

$$\forall i:\ \lambda\bigg(d\Big(\overline{U}_i^{(\hat{u}^-(t))}, V_{p+1-i}^-(\omega)\Big)\bigg) \le -|\lambda_i - \lambda_{i+1}|.$$

4.4.3 Forward-time estimates

The first step of Ginelli's algorithm in Definition 3.1.1 propagates vectors from past to present. It turns out that admissible tuples yield good approximations of $V_{p+1-i}^-(\omega) = Y_1(\omega) \oplus \cdots \oplus Y_i(\omega)$. Moreover, changes of the admissibility parameter on subexponential scales do not influence the exponential speed of convergence of the algorithm:

Lemma 4.4.9
Let $0 < \alpha(t) < 1$ *be a sequence with* $\lambda(1/\alpha) = 0$*. We have*

$$\lambda\left(\sup_{(b)\in\mathcal{A}d^{(\hat{v}^-(t))}(\alpha(t))} d\left(\overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t}\omega}^{(t)} b\right)}, V_{p+1-i}^-(\omega)\right)\right) \le -|\lambda_i - \lambda_{i+1}|.$$

Proof. First use the triangle inequality, then apply Proposition 4.3.3 to the map $A = (\mathcal{L}_\omega^{(-t)})^{-1}$, and finally use Lemma 4.4.8 to obtain

$$\lambda\left(\sup_{(b)\in\mathcal{A}d^{(\hat{v}^-(t))}(\alpha(t))} d\left(\overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t}\omega}^{(t)} b\right)}, V_{p+1-i}^-(\omega)\right)\right)$$

$$
\begin{aligned}
&\leq \max\left(\lambda\left(\sup_{(b)\in\mathcal{A}d^{(\hat{v}^-(t))}(\alpha(t))} d\left(\overline{U}_i^{\left(\mathcal{L}^{(t)}_{\sigma_{-t}\omega}b\right)}, \overline{U}_i^{(\hat{u}^-(t))}\right)\right),\right.\\
&\quad \left.\lambda\Big(d\Big(\overline{U}_i^{(\hat{u}^-(t))}, V^-_{p+1-i}(\omega)\Big)\Big)\right)\\
&\leq \max\left(\lambda\left(\frac{1}{\alpha(t)}\frac{\left(\hat{\delta}^-_{i+1}(t)\right)^{\max}}{\left(\hat{\delta}^-_i(t)\right)^{\min}}\right), -|\lambda_i-\lambda_{i+1}|\right)\\
&\leq \max\left(\lambda\left(\frac{\left(\delta^-_{p+1-i}(t)\right)^{\max}}{\left(\delta^-_{p-i}(t)\right)^{\min}}\right), -|\lambda_i-\lambda_{i+1}|\right)\\
&= -|\lambda_i-\lambda_{i+1}|.
\end{aligned}
$$

□

To continue using our tools in the second step of Definition 3.1.1 we need to retain admissibility for tuples propagated in the first step.

Lemma 4.4.10
Let $0<\alpha(t)<1$ with $\lambda(1/\alpha)=0$. There are $0<\epsilon<1$ and $T>0$ such that admissible tuples in step 1 get mapped to admissible tuples for step 2, i.e., it holds

$$
\left(\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}\right)^d\Big(\mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))\Big) \subset \mathcal{A}d^{(v(t_2))}(\epsilon)
$$

for all $t_1, t_2 \geq T$.

Proof. Choose $0<\epsilon<1$ with

$$
d\Big(V^-_{p+1-i}(\omega), (V_{i+1}(\omega))^{\perp}\Big) \leq \sqrt{1-\epsilon^2}-2\epsilon.
$$

This is possible due to Proposition 4.2.2, since we assumed $V^-_{p+1-i}(\omega)\cap V_{i+1}(\omega)=\{0\}$. Now, Lemma 4.4.9 gives us the existence of $T_1>0$ such that

$$
d\left(\overline{U}_i^{\left(\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}b\right)}, V^-_{p+1-i}(\omega)\right) \leq \epsilon
$$

for all $t_1 \geq T_1$ and all $(b) \in \mathcal{A}d^{\left(\hat{v}^-(t_1)\right)}(\alpha(t_1))$. Moreover, Lemma 4.4.5 yields $T_2>0$ with

$$
d\Big((V_{i+1}(\omega))^{\perp}, \overline{U}_i^{(v(t_2))}\Big) \leq \epsilon
$$

for all $t_2 \geq T_2$. Set $T:=\max(T_1,T_2)$ and combine the previous three estimates to get

$$
\begin{aligned}
d\left(\overline{U}_i^{\left(\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}b\right)}, \overline{U}_i^{(v(t_2))}\right) &\leq d\left(\overline{U}_i^{\left(\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega}b\right)}, V^-_{p+1-i}(\omega)\right)\\
&\quad + d\Big(V^-_{p+1-i}(\omega), (V_{i+1}(\omega))^{\perp}\Big)\\
&\quad + d\Big((V_{i+1}(\omega))^{\perp}, \overline{U}_i^{(v(t_2))}\Big)\\
&\leq \sqrt{1-\epsilon^2}.
\end{aligned}
$$

This concludes the proof. □

The following lemma combines steps 1 and 2 of Ginelli's algorithm into a characterization of the forward phase:

Lemma 4.4.11
Let $0 < \alpha(t) < 1$ with $\lambda(1/\alpha) = 0$. There is $T > 0$ such that

$$\lambda\left(\sup_{t_1 \geq T} \sup_{(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))} d\left(\overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} b\right)}, \overline{U}_i^{(u(t_2))}\right)\right) \leq -|\lambda_i - \lambda_{i+1}|$$

holds, where the limit of the Lyapunov index is taken with respect to t_2.

Proof. Write

$$\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} = \mathcal{L}_{\omega}^{(t_2)} \circ \mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1)}.$$

By Lemma 4.4.10 we find $T > 0$ and $0 < \epsilon < 1$ such that for all $t_1, t_2 \geq T$ and $(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))$ the tuple $(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1)} b)$ is ϵ-admissible w.r.t. $v(t_2)$. Now, apply Proposition 4.3.3 with $A = \mathcal{L}_{\omega}^{(t_2)}$ to see that

$$d\left(\overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} b\right)}, \overline{U}_i^{(u(t_2))}\right) \leq \frac{1}{\epsilon} \frac{\delta_{i+1}^{\max}(t_2)}{\delta_i^{\min}(t_2)}.$$

Since the estimate is independent of $t_1 \geq T$ and singular values converge to LEs, the claim is proved. $\square$

4.4.4 Backward-time estimates

Initial tuples for the backward propagation are obtained from spaces spanned by vectors of the forward phase. Thus, it appears more practical to describe admissibility in terms of forward propagated vectors instead of $(\hat{v}(t_2))$.

Lemma 4.4.12
Let $0 < \alpha(t) < 1/\sqrt{2}$ with $\lambda(1/\alpha) = 0$ be given. There is $T > 0$ such that for all $t_1, t_2 \geq T$ and all $(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))$ we have

$$\mathcal{A}d_-^{(f)^r}\left(\sqrt{2}\alpha(t_2)\right) \subset \mathcal{A}d_-^{(\hat{v}(t_2))}(\alpha(t_2)),$$

where $(f) := \operatorname{orth}(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} b)$ and admissibility holds with respect to reversed degeneracies.

Proof. Let $(f) := \operatorname{orth}(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} b)$ for $(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))$ be given, and let $(g) \in \mathcal{A}d_-^{(f)^r}(\sqrt{2}\alpha(t_2))$ be an admissible tuple. We estimate

$$\begin{aligned} d\left(\overline{U}_{i,-}^{(g)}, \overline{U}_{i,-}^{(\hat{v}(t_2))}\right) &\leq d\left(\overline{U}_{i,-}^{(g)}, \overline{U}_{i,-}^{(f)^r}\right) + d\left(\overline{U}_{i,-}^{(f)^r}, \overline{U}_{i,-}^{(\hat{v}(t_2))}\right) \\ &\leq \sqrt{1 - 2\alpha(t_2)^2} + d\left(\left(\overline{U}_{p-i}^{(f)}\right)^\perp, \left(\overline{U}_{p-i}^{(u(t_2))}\right)^\perp\right) \\ &= \sqrt{1 - 2\alpha(t_2)^2} + d\left(\overline{U}_{p-i}^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} b\right)}, \overline{U}_{p-i}^{(u(t_2))}\right). \end{aligned}$$

The last summand is bounded by

$$d(t_2) := \sup_{t_1 \geq T} \sup_{(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))} d\left(\overline{U}_{p-i}^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} b\right)}, \overline{U}_{p-i}^{(u(t_2))}\right)$$

for $t_2 \geq T$ with T as in Lemma 4.4.11. In particular, it holds $\lambda(d(t_2)) < 0$. Now, for (g) to be $\alpha(t_2)$-admissible w.r.t. $(\hat{v}(t_2))$, it suffices to show that

$$\sqrt{1 - 2\alpha(t_2)^2} + d(t_2) \leq \sqrt{1 - \alpha(t_2)^2}$$

for t_2 large enough, which in turn is equivalent to

$$1 - 2\alpha(t_2)^2 + 2\sqrt{1 - 2\alpha(t_2)^2} d(t_2) + d(t_2)^2 \leq 1 - \alpha(t_2)^2$$

and to

$$\frac{d(t_2)\left(2\sqrt{1 - 2\alpha(t_2)^2} + d(t_2)\right)}{\alpha(t_2)^2} \leq 1.$$

The latter is true for t_2 large enough since we have

$$\lambda\left(\frac{d(t_2)\left(2\sqrt{1 - 2\alpha(t_2)^2} + d(t_2)\right)}{\alpha(t_2)^2}\right) \leq \lambda\left(\frac{d(t_2)(2 + d(t_2))}{\alpha(t_2)^2}\right) < 0.$$

□

Next, we combine our characterization of the forward phase with backward propagation. During the backward phase, it is enough to restrict ourselves to tuples that have admissible extensions. A few tools from the forward phase can be applied to the time-reversed cocycle.

Lemma 4.4.13
Let $0 < \alpha(t) < 1/\sqrt{2}$ with $\lambda(1/\alpha) = 0$ be given. It holds

$$\begin{aligned}\overline{\lambda}\Bigg(&\sup_{(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))} \sup_{(b') \in \left(\mathcal{A}d_{\mathrm{ext}\_}^{(f)^r}\left(\sqrt{2}\alpha(t_2)\right)\right)^r} d\left(U_i^{\left(\mathcal{L}_{\sigma_{t_2}\omega}^{(-t_2)} b'\right)}, Y_i(\omega)\right)\Bigg) \\ &\leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|),\end{aligned}$$

where $(f) := \mathrm{orth}(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)} b)$. Here, $(.)^r$ applied to a set means that the order of every tuple in the set is reversed.

Proof. Applying Lemma 4.4.9 to $\mathcal{L}$ and to $\mathcal{L}^-$, we get

$$\lambda\left(\sup_{(b) \in \mathcal{A}d^{(\hat{v}^-(t))}(\alpha(t))} d\left(\overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t}\omega}^{(t)} b\right)}, V_{p+1-i}^-(\omega)\right)\right) \leq -|\lambda_i - \lambda_{i+1}|$$

and

$$\lambda\left(\sup_{(g) \in \mathcal{A}d_-^{(\hat{v}(t))}(\alpha(t))} d\left(\overline{U}_{i,-}^{\left(\mathcal{L}_{\sigma_t\omega}^{(-t)} g\right)}, V_{p+1-i}(\omega)\right)\right) \leq -|\lambda_i^- - \lambda_{i+1}^-|.$$

By switching indices we can rewrite the latter as

$$\lambda\left(\sup_{(g)\in\mathcal{A}d_{-}^{(\hat{v}(t))}(\alpha(t))} d\left(\overline{U}_{(p+1-i),-}^{\left(\mathcal{L}_{\sigma_t\omega}^{(-t)}g\right)}, V_i(\omega)\right)\right) \leq -|\lambda_i-\lambda_{i-1}|.$$

In short, we have exponentially fast converging approximations of $V_{p+1-i}^{-}(\omega)$ and of $V_i(\omega)$, which are transversal subspaces with intersection $Y_i(\omega)$ (see Eq. (2.3)). Thus, we can apply Corollary 4.2.9 with

$$\mathcal{M}_t := \left\{ \overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t}\omega}^{(t)}b\right)} \;\middle|\; (b)\in\mathcal{A}d^{(\hat{v}^{-}(t))}(\alpha(t)) \right\}$$

and

$$\mathcal{N}_t := \left\{ \overline{U}_{(p+1-i),-}^{\left(\mathcal{L}_{\sigma_t\omega}^{(-t)}g\right)} \;\middle|\; (g)\in\mathcal{A}d_{-}^{(\hat{v}(t))}(\alpha(t)) \right\}$$

to get an estimate of the rate of convergence of intersections:

$$\begin{aligned}\overline{\lambda}&\left(\sup_{(b)\in\mathcal{A}d^{(\hat{v}^{-}(t_1))}(\alpha(t_1))}\;\sup_{(g)\in\mathcal{A}d_{-}^{(\hat{v}(t_2))}(\alpha(t_2))} d\left(\overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1)}b\right)}\cap\overline{U}_{(p+1-i),-}^{\left(\mathcal{L}_{\sigma_{t_2}\omega}^{(-t_2)}g\right)}, Y_i(\omega)\right)\right)\\ &\leq -\min(|\lambda_i-\lambda_{i-1}|, |\lambda_i-\lambda_{i+1}|).\end{aligned}$$

By Lemma 4.4.12 we can take the supremum over

$$(g)\in\mathcal{A}d_{-}^{(f)^r}\left(\sqrt{2}\alpha(t_2)\right)$$

instead, while maintaining the estimate. In particular, this is true for all admissible extensions (g) of

$$(b')^r\in\mathcal{A}d_{\mathrm{ext}_-}^{(f)^r}\left(\sqrt{2}\alpha(t_2)\right).$$

Now, to prove the lemma it suffices to show that each admissible extension (g) of

$$\left((b')^r_{(p+1-i)_1},\ldots,(b')^r_{(p+1-i)_{d^-_{p+1-i}}}\right) = \left(b'_{i_{d_i}},\ldots,b'_{i_1}\right)$$

satisfies

$$U_i^{\left(\mathcal{L}_{\sigma_{t_2}\omega}^{(-t_2)}b'\right)} = \overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1)}b\right)}\cap\overline{U}_{(p+1-i),-}^{\left(\mathcal{L}_{\sigma_{t_2}\omega}^{(-t_2)}g\right)}.$$

We clearly have

$$U_i^{(b')} = U_{(p+1-i),-}^{(b')^r} = U_{(p+1-i),-}^{(g)} \subset \overline{U}_{(p+1-i),-}^{(g)}$$

and hence

$$U_i^{\left(\mathcal{L}_{\sigma_{t_2}\omega}^{(-t_2)}b'\right)} \subset \overline{U}_{(p+1-i),-}^{\left(\mathcal{L}_{\sigma_{t_2}\omega}^{(-t_2)}g\right)}$$

for an admissible extension (g). Moreover, the definition of extended admissibility requires that

$$\begin{aligned}(b')^r_{(p+1-i)_1},\ldots,(b')^r_{(p+1-i)_{d^-_{p+1-i}}} &\in U_{(p+1-i),-}^{(f)^r}\oplus\cdots\oplus U_{p^-,-}^{(f)^r}\\ &= U_i^{(f)}\oplus\cdots\oplus U_1^{(f)}\\ &= \overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1+t_2)}b\right)}\\ &= \mathcal{L}_{\omega}^{(t_2)}\overline{U}_i^{\left(\mathcal{L}_{\sigma_{-t_1}\omega}^{(t_1)}b\right)},\end{aligned}$$

or equivalently, it holds

$$U_i^{\left(\mathcal{L}^{(-t_2)}_{\sigma_{t_2}\omega} b'\right)} \subset \overline{U}_i^{\left(\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega} b\right)}.$$

Thus, we have

$$U_i^{\left(\mathcal{L}^{(-t_2)}_{\sigma_{t_2}\omega} b'\right)} \subset \overline{U}_i^{\left(\mathcal{L}^{(t_1)}_{\sigma_{-t_1}\omega} b\right)} \cap \overline{U}_{(p+1-i),-}^{\left(\mathcal{L}^{(-t_2)}_{\sigma_{t_2}\omega} g\right)}.$$

Since admissible tuples are linearly independent, the left-hand side has dimension d_i. The right-hand side must have the same dimension for t_1, t_2 large enough, because the intersection converges to $Y_i(\omega)$. Hence, we have equality of subspaces, which concludes the proof. □

Let us remark that the proof of Lemma 4.4.13 also works if we regard extended admissibility with respect to a nondegenerate spectrum, i.e., if we set $p = p^- = d$ in the definition of extended admissibility. In fact, this detail is necessary for our convergence proof since we require upper triangular coefficient matrices in Definition 3.1.1. The original proof in [48] does not make this assumption. It allows for upper triangular coefficient matrices with block-diagonal structure.

4.4.5 Proof of theorems

Lemma 4.4.13 describes how admissible tuples fare in Ginelli's algorithm. The remaining work lies in connecting the lemma to measurement results from Section 4.3.

Proof of Theorem 4.4.3. Fix $\epsilon > 0$. We identify $\mathbb{R}^{d\times d}_{ru}$ with $\mathbb{R} \times \mathbb{R}^2 \times \cdots \times \mathbb{R}^d$ column by column. Write

$$B'(M) := B_1(0, M) \times \cdots \times B_d(0, M)$$

for the subset consisting of product of balls of radius M. By compactness of $\mathcal{K}$ we find some $M > 0$ with $\mathcal{K} \subset B_d(0, M)^d \times B'(M)$. It is enough to prove the claim for the product of balls instead of $\mathcal{K}$.

We set $\alpha(t) := 1/\max(t, 2\sqrt{2})$, so that $\lambda(1/\alpha) = 0$. Now, use α in Lemma 4.4.13 to obtain

$$\frac{1}{\min(t_1, t_2)} \log d\left(U_i^{\left(\mathcal{L}^{(-t_2)}_{\sigma_{t_2}\omega} b'\right)}, Y_i(\omega)\right) \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|) + \epsilon$$

for all $(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))$ and $(b') \in (\mathcal{A}d^{(f)^r}_{\text{ext}_-}(\sqrt{2}\alpha(t_2)))^r$ with t_1 and t_2 large enough. Here, we assume extended admissibility with respect to the nondegenerate case. Since Ginelli's algorithm uses coefficient matrices instead of initial vectors for the backward phase, we relate both through forward propagated vectors as it is done in step 4 of Definition 3.1.1. A matrix of coefficients R with columns (r) gets mapped to a tuple of vectors (b') by the orthogonal transformation A sending (e) to (f). Via this identification, we can shift results for backward initial tuples to the coefficient space, which is independent of runtimes and of inputs. Hence, in the above convergence estimate we may exchange (b') for (Ar) with

$$(r) \in \left(A^{-1}\right)^d \left(\mathcal{A}d^{(f)^r}_{\text{ext}_-}\left(\sqrt{2}\alpha(t_2)\right)\right)^r = \left(\mathcal{A}d^{(e)^r}_{\text{ext}_-}\left(\sqrt{2}\alpha(t_2)\right)\right)^r.$$

Now, we show that the set of tuples $((b),(r))$ in $B_d(0,M)^d \times B'(M)$ such that (b) or (r) is nonadmissible has measure zero for $\min(t_1,t_2) \to \infty$:

$$\begin{aligned}
&\mu\Big(\Big(B_d(0,M)^d \times B'(M)\Big) \setminus \Big(\mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1)) \times \Big(\mathcal{A}d^{(e)^r}_{\text{ext}_-}\Big(\sqrt{2}\alpha(t_2)\Big)\Big)^r\Big)\Big)\\
&\quad \le \mu\Big(B_d(0,M)^d \setminus \mathcal{A}d^{(\hat{v}^-(t_1))}(\alpha(t_1))\Big)\mu(B'(M))\\
&\qquad + \mu\Big(B_d(0,M)^d\Big)\mu\Big(B'(M) \setminus \Big(\mathcal{A}d^{(e)^r}_{\text{ext}_-}\Big(\sqrt{2}\alpha(t_2)\Big)\Big)^r\Big)\\
&\quad = \mu\Big(B_d(0,M)^d \setminus \mathcal{A}d^{(e)}(\alpha(t_1))\Big)\mu(B'(M))\\
&\qquad + \mu\Big(B_d(0,M)^d\Big)\mu\Big((B'(M))^r \setminus \mathcal{A}d^{(e)^r}_{\text{ext}_-}\Big(\sqrt{2}\alpha(t_2)\Big)\Big),
\end{aligned}$$

Here, we used invariance under orthogonal transformations of $B_d(0,M)$ to switch from $(\hat{v}^-(t_1))$ to (e). By Corollary 4.3.5 and Corollary 4.3.7 the final estimate converges to zero as $\min(t_1,t_2)$ is increased. Hence, we get the desired convergence result. $\square$

The discrete-time version can be proved in a similar fashion.

Proof of Theorem 4.4.4. Assume discrete time $\mathbb{T} = \mathbb{Z}$ and $d > 1$. We define $\alpha_\epsilon(n) := (\epsilon/(\sqrt{2}n^2))^{d-1}$ as our admissibility parameter satisfying $\lambda(1/\alpha_\epsilon) = 0$ for each $0 < \epsilon < 1$. Using α_ϵ, we invoke Lemma 4.4.13 to find that

$$\overline{\lambda}\left(d\left(U_i^{\left(\mathcal{L}^{(-n_2)}_{\sigma_{n_2}\omega} b'\right)}, Y_i(\omega)\right)\right) \le -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|)$$

for $(b') = (Ar)$ with A as in the last proof, whenever

$$((b),(r)) \in \bigcap_{n_1,n_2 \in \mathbb{N}} \mathcal{A}d^{(\hat{v}^-(n_1))}(\alpha_\epsilon(n_1)) \times \Big(\mathcal{A}d^{(e)^r}_{\text{ext}_-}\Big(\sqrt{2}\alpha_\epsilon(n_2)\Big)\Big)^r.$$

This is true independent of our choice of ϵ. Hence, it suffices to show that the complement of

$$\bigcup_{0<\epsilon<1} \bigcap_{n_1,n_2 \in \mathbb{N}} \mathcal{A}d^{(\hat{v}^-(n_1))}(\alpha_\epsilon(n_1)) \times \Big(\mathcal{A}d^{(e)^r}_{\text{ext}_-}\Big(\sqrt{2}\alpha_\epsilon(n_2)\Big)\Big)^r \tag{4.1}$$

has measure zero[7], which can be proved by exhausting the domain of $((b),(r))$ with products of balls:

[7] Note that the statement is not true in general for continuous time. In fact, in Example 4.4.2 no tuple (b) is admissible w.r.t. $(\hat{v}^-(t_1))$ for all $t_1 \in \mathbb{R}_{>0}$ simultaneously. Hence, in this case the set in Eq. (4.1) would be empty.

Let $B(M) := (B'(M))^r$. It holds

$$\begin{aligned}
&\mu\Bigg(\Big(B_d(0,M)^d \times B'(M)\Big)\setminus \\
&\qquad \bigcup_{0<\epsilon<1}\ \bigcap_{n_1,n_2\in\mathbb{N}} \mathcal{A}d^{(\hat{v}^-(n_1))}(\alpha_\epsilon(n_1)) \times \Big(\mathcal{A}d^{(e)^r}_{\text{ext}_-}\Big(\sqrt{2}\alpha_\epsilon(n_2)\Big)\Big)^r\Bigg) \\
&\leq \inf_{0<\epsilon<1}\Bigg(\sum_{n_1\in\mathbb{N}} \mu\Big(B_d(0,M)^d \setminus \mathcal{A}d^{(e)}(\alpha_\epsilon(n_1))\Big)\mu(B'(M)) \\
&\qquad + \sum_{n_2\in\mathbb{N}} \mu\Big(B_d(0,M)^d\Big)\mu\Big(B(M)\setminus \mathcal{A}d^{(e)^r}_{\text{ext}_-}\Big(\sqrt{2}\alpha_\epsilon(n_2)\Big)\Big)\Bigg) \\
&\leq \inf_{0<\epsilon<1}\Bigg(\sum_{n_1\in\mathbb{N}} \eta_1(\alpha_\epsilon(n_1))^{\frac{1}{d-1}}\mu(B'(M)) + \sum_{n_2\in\mathbb{N}} \mu\Big(B_d(0,M)^d\Big)\eta_2\Big(\sqrt{2}\alpha_\epsilon(n_2)\Big)^{\frac{1}{d-1}}\Bigg) \\
&= \inf_{0<\epsilon<1}\Bigg(\sum_{n\in\mathbb{N}} \frac{\eta_1\mu(B'(M)) + \eta_2\sqrt{2}^{\frac{1}{d-1}}\mu\Big(B_d(0,M)^d\Big)}{\sqrt{2}n^2}\Bigg)\epsilon \\
&= 0
\end{aligned}$$

for all $M > 0$. Here, it was crucial to use Proposition 4.3.8 and Proposition 4.3.11 to get more precise measure estimates on nonadmissible tuples. □

4.5 Summary and discussion

We analyzed convergence of Ginelli's algorithm to compute CLVs, or more generally Oseledets spaces, for cocycles with invertible actions in finite dimensions. The existence of CLVs was provided by a deterministic version of the MET from [1]. Moreover, the proof of the theorem handed us an interface able to link CLVs with a limit of finite-time scenarios in which Ginelli's algorithm is applied to initial vectors. It turned out that certain tuples of initial vectors perform better than others given the same runtime, whereas in some cases the algorithm would not even converge - a problem that did not receive enough attention in previous attempts to prove convergence.

As a measure to tackle this problem, we introduced the concept of admissibility. A tuple of initial vectors is called admissible if it is not too far from the right singular vectors of the propagator. The term “not too far” was made more precise by a parameter α. In our formulation, values of α close to 1 imply a good correlation, whereas values of α close to 0 stand for greater distances between initial vectors and right singular vectors.

In [20] it is shown that tuples with positive α will align with left singular vectors when propagated from present to future. While the admissibility condition depends on the chosen runtime, according to the MET right singular vectors defining admissibility at the present state converge. In the limit it is possible to show that almost every initial tuple will yield a good approximation of left singular vectors when propagated long enough from the present state. However, in Ginelli's algorithm vectors are initiated at past states. Thus, the admissibility condition varies with the runtime and the set of admissible vectors generally does not converge, which was not regarded in [20] and motivates our new analysis. In fact, we presented an example where no fixed initial tuple is admissible for all past states simultaneously. Consequently, the continuous-time version of Ginelli's algorithm cannot be expected to converge for fixed initial tuples in general. Instead, we have shown convergence in measure of the continuous-time version by connecting admissibility and finite-time estimates before deriving asymptotic estimates via the proof of the MET. Moreover, due to suitable measure estimates of sets of admissible vectors, we were able to prove convergence for almost every initial tuple in the discrete-time case. In the presence of degeneracies of the Lyapunov spectrum convergence holds with respect to spaces spanned by propagated vectors rather than the vectors themselves. Indeed, internal dynamics of Oseledets spaces might prevent single output vectors from converging.

The convergence results for both cases of time relate the speed of convergence to LEs. Using the Lyapunov index notation, we were able to prove that Ginelli's algorithm converges exponentially fast with a rate given by the minimum distance between LEs. However, the Lyapunov index notation neglects system-dependent prefactors of the speed of convergence on subexponential timescales, which may very well be of importance for limited-time scenarios. Yet, if enough data is available, subexponential factors, e.g., from choosing different initial tuples, can be ignored. Moreover, nonadmissible initial tuples will in general turn admissible due to numerical noise. Hence, the concept of admissibility and the different notions of convergence do not play a noticeable role in practice. They can be seen rather as tools or as products of a precise mathematical proof.

Finally, let us mention that the tools obtained during the proof can be used to investigate other algorithms, such as Wolfe-Samelson's algorithm [67], as well. Indeed, the connection between dynamical features and the SVD of the cocycle enables the analysis of many constructions in the context of CLV-algorithms.

5

Convergence of Ginelli's Algorithm on Hilbert Spaces

This chapter follows [49, 50]. We analyze convergence of Ginelli's algorithm in a semi-invertible setting on Hilbert spaces as in Section 2.3. While our final convergence theorem looks similar to that from the last chapter, i.e., convergence for almost every input, the involved techniques are quite different. Jumping from finite to infinite dimensions, we cannot perform estimates with respect to Lebesgue measure and we do not have a SVD of the cocycle in general. Instead of admissibility, we introduce *well-separating common complements* that play a similar role during our convergence proof. It turns out that vectors spanning well-separating common complements are the right choice of input vectors for exponentially fast convergence of Ginelli's algorithm. Using the concept of *prevalence*, we show that almost every tuple of input vectors spans well-separating common complements to get a convergence theorem similar to Theorem 4.4.4.

As in the previous chapter, we regard convergence with respect to subspaces rather than isolated vectors. To this end, we generalize the concept of Grassmannians to Banach spaces in Section 5.1. Our focus lies on classes of subspaces of finite dimension or of finite codimension. They contain spaces of the Oseledets splitting and of the Oseledets filtration. Again, we define notions of distances and angles between subspaces. In Section 5.2 we treat well-separating common complements. We show that these special kind of subspaces are prevalent for Hilbert spaces. Finally, in Section 5.3 we derive a convergence theorem by investigating the evolution of well-separating common complements in Ginelli's algorithm for forward- and for backward-time. Note that, due to the semi-invertible setting, backward propagation can only be done restricted to subspaces on which the action of the cocycle is invertible.

In our convergence analysis we try to be as general as possible, so that the derived results may potentially be transferred to other settings as well. Our analysis is not tailored to the proof of the particular semi-invertible MET. In fact, we only require asymptotic characterizations of LEs and of Oseledets spaces as stated in most versions of the MET. Moreover, since many of our arguments do not need Hilbert spaces, we formulate them at the level of Banach spaces.

5.1 Grassmannians

Definition 5.1.1
Let $(X, \|.\|)$ be a Banach space. The Grassmannian $\mathcal{G}(X)$ *is the set of closed complemented subspaces of X, i.e., closed subspaces $V \subset X$ such that there is a closed subspace $W \subset X$ with $X = V \oplus W$. It contains $\mathcal{G}_k(X)$, the set of k-dimensional subspaces, and $\mathcal{G}^k(X)$, the set of closed subspaces of codimension k.*

The Grassmannian $\mathcal{G}(X)$ can be equipped with a metric $d_{\mathcal{G}}(V, W)$ via the Hausdorff distance between $V \cap B$ and $W \cap B$, where B denotes the closed unit ball in X [29, appendix B]:

$$\begin{aligned} d_{\mathcal{G}}(V,W) &:= d_H(V \cap B, W \cap B) \\ &= \max\left(\sup_{v \in V \cap B} d(v, W \cap B), \sup_{w \in W \cap B} d(w, V \cap B) \right) \\ &= \max\left(\sup_{v \in V \cap B} \inf_{w \in W \cap B} \|v - w\|, \sup_{w \in W \cap B} \inf_{v \in V \cap B} \|w - v\| \right) \end{aligned}$$

for $V, W \in \mathcal{G}(X)$. Another metric $\hat{d}_{\mathcal{G}}$ is given by exchanging B with the unit sphere S in the above definition. In fact, Kato shows that $\mathcal{G}(X)$ equipped with $\hat{d}_{\mathcal{G}}$ is a complete metric space [34, chapter IV, §2.1]. Moreover, he relates $\hat{d}_{\mathcal{G}}$ to the *gap* between subspaces, which is defined as

$$\hat{\delta}(V,W) := \max\left(\sup_{v \in V \cap S} d(v, W), \sup_{w \in W \cap S} d(w, V) \right).$$

In general the gap is not a metric. However, if X is a Hilbert space, it coincides with the metric defined in Section 4.2: $\hat{\delta}(V,W) = \|P_V - P_W\|$. All three concepts of distances on $\mathcal{G}(X)$ are related via $\hat{\delta} \leq d_{\mathcal{G}} \leq \hat{d}_{\mathcal{G}} \leq 2\hat{\delta}$. This follows from

$$\sup_{v \in V \cap S} d(v, W) \leq \sup_{v \in V \cap B} d(v, W \cap B) \leq \sup_{v \in V \cap S} d(v, W \cap S) \leq 2 \sup_{v \in V \cap S} d(v, W), \quad (5.1)$$

see [34]. In particular, $d_{\mathcal{G}}$ and $\hat{d}_{\mathcal{G}}$ induce the same topology. Hence, $\mathcal{G}(X)$ is complete with respect to $d_{\mathcal{G}}$.

The symmetry of $d_{\mathcal{G}}$ is an immediate consequence of its definition. We cannot reduce the definition to only one term since $\sup_{v \in V \cap B} d(v, W \cap B)$ and $\sup_{w \in W \cap B} d(w, V \cap B)$ are different in general. However, if one term is small, then so is the other [29, lemma B.7]:

Lemma 5.1.2 ([29])
If $V, W \in \mathcal{G}_k(X)$ are subspaces of dimension k, then

$$\sup_{v \in V \cap B} d(v, W \cap B) =: r < 3^{-k}/4 \quad \Longrightarrow \quad d_{\mathcal{G}}(V,W) < 4 \cdot 3^k r.$$

If $V, W \in \mathcal{G}^k(X)$ are closed subspaces of codimension k, then

$$\sup_{v \in V \cap B} d(v, W \cap B) =: r < 3^{-k}/8 \quad \Longrightarrow \quad d_{\mathcal{G}}(V,W) < 8 \cdot 3^k r.$$

Thus, when investigating convergence inside $\mathcal{G}_k(X)$ or $\mathcal{G}^k(X)$, it is enough to estimate only one of the two terms in the definition of $d_{\mathcal{G}}$.

Ultimately, we want to approximate Oseledets spaces, which are finite-dimensional complements to spaces of the Oseledets filtration. Hence, we will be working with tuples of the set

$$\mathrm{Comp}_k(X) := \left\{ (C, V) \in \mathcal{G}_k(X) \times \mathcal{G}^k(X) \;\middle|\; X = C \oplus V \right\}$$

for $k \in \mathbb{N}$. Given such a tuple, each $x \in X$ can be written uniquely as $x = c + v$ according to the associated splitting. In particular, we get two projections $\Pi_{C||V} : X \to C$ and $\Pi_{V||C} : X \to V$, which are bounded linear operators by the closed graph theorem. It can be shown that they are stable with respect to perturbations of the tuple (C, V) [29, lemma B.18]:

Lemma 5.1.3 ([29])
The mapping $\mathrm{Comp}_k(X) \to L(X)$ *given by* $(C, V) \mapsto \Pi_{C||V}$ *is continuous, where* $\mathrm{Comp}_k(X)$ *has the product topology induced by* $\mathcal{G}(X)$ *and where the space* $L(X)$ *of bounded linear operators on* X *is equipped with the norm topology.*

As in the finite-dimensional case, we need to keep track of angles between subspaces.

Definition 5.1.4
Let $C, V \subset X$ *be two subspaces. The sine of the* minimal angle *from* C *to* V *is defined as* $\inf_{c \in C \cap S} d(c, V)$.

The new notion of minimal angle generalizes the one from Definition 4.2.4. Indeed, let C and V be two subspaces of $(\mathbb{R}^d, \|.\|_2)$. By Proposition 4.2.5 we have

$$c_0(C, V)^2 = \|P_V P_C\|^2 = \sup_{c \in C \cap S} \|P_V c\|^2 = 1 - \inf_{c \in C \cap S} \|P_{V^\perp} c\|^2 = 1 - \inf_{c \in C \cap S} d(c, V)^2.$$

If $\theta_1 \in [0, \pi/2]$ denotes the old minimal angle and $\theta_2 \in [0, \pi/2]$ the new one, then

$$\cos(\theta_1)^2 = 1 - \sin(\theta_2)^2 = \cos(\theta_2)^2.$$

Thus, θ_1 and θ_2 coincide. However, note that the new minimal does not retain all properties of the old one if X deviates from the standard euclidean space. For example, the new minimal angle from C to V is generally not the same as the minimal angle from V to C.

Usually, we will have $(C, V) \in \mathrm{Comp}_k(X)$, i.e., C will be a complement to V. In this case, we call the sine of the minimal angle *degree of transversality*. It is equal to $1/\|\Pi_{C||V}\|$ [7] and describes the quality of the splitting $X = C \oplus V$. The degree of complementing subspaces is always positive, since $\inf_{c \in C \cap S} d(c, V) = 0$ would imply that $C \cap V \neq \{0\}$. On the other hand, if X is a Hilbert space, a degree of 1 implies $C = V^\perp$. Thus, we prefer complements with a high degree of transversality (close to 1) as they are better separated.

5.2 Well-separating common complements

In the convergence proof for finite dimensions we used the concept of admissibility to measure distances between initial vectors and singular vectors of the cocycle. More precisely, admissibility compared the associated filtration spaces. The concept was applied to find tuples whose filtration spaces stay close to those of singular vectors. By the proof of the MET for finite dimensions, the filtration spaces of singular vectors at ω converge to sums of the first Oseledets spaces which complement the Oseledets filtration. In this sense, we sought tuples of initial vectors with a filtration that is well-separated from the Oseledets filtration. Section 4.4 applied this idea to different phases of Ginelli's algorithm by letting the admissibility condition depend on the initial time. We required an admissibility parameter decaying to zero at a subexponential speed, so that it did not influence other factors of convergence. Then, we proved for discrete time that almost every tuple satisfies this criterion of separation from the Oseledets filtration at different initial times.

Here, we take a similar approach. We look for subspaces such that the degree of transversality to spaces of the Oseledets filtration at different initial times decays at most subexponentially. More abstractly, given a sequence of subspaces $(V_n)_{n\in\mathbb{N}} \subset \mathcal{G}^k(X)$, we ask for common complements, i.e., subspaces $C \subset X$ with $(C, V_n) \in \mathrm{Comp}_k(X)$ for all n, such that the degree of transversality of (C, V_n) decays at most subexponentially with n:

Definition 5.2.1
Let $(V_n)_{n\in\mathbb{N}} \subset \mathcal{G}^k(X)$ be given. A common complement $C \in \mathcal{G}_k(X)$ for $(V_n)_{n\in\mathbb{N}}$ is called well-separating *with respect to $(V_n)_{n\in\mathbb{N}}$ if*

$$\lim_{n\to\infty} \frac{1}{n} \log \inf_{c\in C\cap S} d(c, V_n) = 0.$$

Well-separating common complements can be used without interfering on exponential scales that are important for our convergence proof. Natural questions are the existence and the genericity of well-separating common complements. While the existence of common complements for two subspaces has already been studied in various scenarios [18, 19, 37], results on the genericity of common complements are rare. There is one result for finite dimensions saying that any countable family of subspaces of the same codimension has uncountable many common complements [63]. However, as far as the author knows, there is no similar statement for well-separating common complements yet.

After a short discussion on common complements for finitely many hyperplanes, we turn towards well-separating common complements. In particular, we investigate existence and genericity of well-separating common complements. It turns out that existence is guaranteed for Hilbert spaces. Moreover, we show that existence of one well-separating common complement already implies that they are generic.

5.2.1 Common complements for finitely many hyperplanes

Using simple geometric tools, we find common complements for finitely many hyperplanes when $X = \mathbb{R}^d$ or when X is an arbitrary Banach space. Those complements have dimension 1 and can be identified with an up to sign unique unit vector. The distance between this vector and the hyperplanes determines the quality of the complement. Besides the existence of common complements, we investigate their quality and motivate why subexponential decay of the degree of transversality is a natural assumption for common complements for families of countably many subspaces.

We start with a geometric tool for the existence of common complements in $\mathbb{R}^d$:

Lemma 5.2.2
Let $Q \subset (\mathbb{R}^d, \|.\|_2)$ be a compact, convex d-polytope with faces $(F_i)_{i=1}^m$ and normals $(f_i)_{i=1}^m$. Moreover, let $V \subset \mathbb{R}^d$ be a hyperplane with normal v.

The volume of the orthogonal projection of Q onto V satisfies

$$\mathrm{vol}_{d-1}(P_V Q) = \frac{1}{2}\sum_{i=1}^m \mathrm{vol}_{d-1}(F_i)|\langle f_i, v\rangle|.$$

Proof. This is a known result (e.g., see [11]). The basic ideas are that $P_V Q = P_V \partial Q$ and that the interior of $P_V \partial Q$ is covered twice by the projection of the hull ∂Q. Now, one only needs to check that $\mathrm{vol}_{d-1}(P_V F_i) = \mathrm{vol}_{d-1}(F_i)|\langle f_i, v\rangle|$ for each face. □

Corollary 5.2.3
Let $(V_n)_{n=1}^N$ be hyperplanes of $(\mathbb{R}^d, \|.\|_2)$. There exists a unit vector $x \in \mathbb{R}^d$ with $d(x, V_n) \geq 1/(2Nd)$ for all n.

Proof. Let $Q = [-1,1]^d$ and let v_n be the normal of V_n. We have

$$\mathrm{vol}_{d-1}(P_{V_n} Q) = 2^{d-1}\sum_{i=1}^d |\langle e_i, v_n\rangle| = 2^{d-1}\|v_n\|_1 \leq 2^{d-1}\sqrt{d}\,\|v_n\|_2 = 2^{d-1}\sqrt{d}.$$

Now, let $\delta := 1/(2N\sqrt{d})$. Using the Lebesgue measure μ on $\mathbb{R}^d$, we estimate

$$\begin{aligned}\mu(\{y \in Q \mid \exists n : |\langle y, v_n\rangle| \leq \delta\}) &\leq \sum_{n=1}^N \mu(\{y \in Q \ : \ |\langle y, v_n\rangle| \leq \delta\}) \\ &\leq 2\delta \sum_{n=1}^N \mathrm{vol}_{d-1}(P_{V_n} Q) \\ &\leq 2\delta N 2^{d-1}\sqrt{d} \\ &= 2^{d-1}.\end{aligned}$$

Since $\mathrm{vol}(Q) = 2^d$, there must be an element $y \in Q$ with $|\langle y, v_n\rangle| \geq \delta$ for all n. Writing $x := y/\|y\|_2$ yields $d(x, V_n) = |\langle x, v_n\rangle| \geq \delta/\|y\|_2 \geq 1/(2Nd)$. □

A lower bound better than $1/(2Nd)$ for arbitrary hyperplanes is possible by looking at intersections of the unit ball and hyperplanes instead of polytopes and hyperplanes. In the case $V_n = \{(x_1, \ldots, x_d) \mid x_n = 0\}$ with $N = d$ the best possible lower bound is $1/\sqrt{N}$, which corresponds to unit vectors lying in diagonal lines $\{(\pm t, \ldots, \pm t) \mid t \in \mathbb{R}\}$.

The next theorem is a well-known result in the context of the Banach-Mazur compactum. As a consequence of John's theorem [33] about ellipsoids, the (multiplicative) distance of any Banach space of dimension d to the standard euclidean space $(\mathbb{R}^d, \|.\|_2)$ is at most $\sqrt{d}$:

Theorem 5.2.4
Let X be a Banach space of dimension d. There exists an isomorphism $T : (X, \|.\|) \to (\mathbb{R}^d, \|.\|_2)$ such that $\|T\|\|T^{-1}\| \leq \sqrt{d}$.

By scaling such an isomorphism with a positive constant, we can assure that either $\|T\| = 1$ or $\|T^{-1}\| = 1$ holds additionally.

Corollary 5.2.5
Let $(V_n)_{n=1}^N$ be hyperplanes of a Banach space X. There exists a unit vector $x \in X$ such that $d(x, V_n) \geq 1/(4N^{\frac{5}{2}})$ for all n.

Proof. Set $V := V_1 \cap \cdots \cap V_N$ and $Y := X/V$. As a quotient space, Y is a Banach space of dimension $d \leq N$. The quotient map $\pi : X \to Y$ sends $(V_n)_{n=1}^N$ to hyperplanes of Y. Now, by Theorem 5.2.4 there is an isomorphism T mapping $(Y, \|.\|_Y)$ to $(\mathbb{R}^d, \|.\|_2)$ such that $\|T\| \leq \sqrt{d}$ and $\|T^{-1}\| \leq 1$. By Corollary 5.2.3 we find $z \in \mathbb{R}^d$ with $\|z\|_2 = 1$ and $d(z, T\pi V_n) \geq 1/(2Nd)$ for all n. Let $y := T^{-1}z$. It holds $\|y\|_Y \leq 1$ and

$$\frac{1}{2Nd} \leq \inf_{v_n \in V_n} \|z - T\pi v_n\|_2 \leq \inf_{v_n \in V_n} \|T\| \|y - \pi v_n\|_Y \leq \sqrt{d}\, d(y, \pi V_n).$$

Take $x' \in X$ with $\pi x' = y$. Since $\inf_{v \in V} \|x' - v\| = \|y\|_Y \leq 1$, we find $v' \in V$ with $\|x' - v'\| \leq 2$. Set $x := (x' - v')/\|x' - v'\|$. One readily checks that $d(y, \pi V_n) = d(x', V_n) = \|x' - v'\|\, d(x, V_n)$. The claim follows. □

Given N hyperplanes of a Banach space, Corollary 5.2.5 implies that there exists a common complement such that the degree of transversality between each pair is bounded from below by $1/(4N^{\frac{5}{2}})$. On the other hand, there are cases where $1/\sqrt{N}$ is the best that can be archived. Hence, as the number of hyperplanes is increased to infinity, we cannot hope for a common complement with a degree of transversality bounded away from zero in general. Instead, we ask for complements such that the degree of transversality decays at most subexponentially.

5.2.2 Existence

In this subsection we prove the existence of well-separating common complements in Hilbert spaces. So far, an existence result for Banch spaces has not been archived. The only remaining hurdle for a similar result as Theorem 5.2.6 would be to generalize Lemma 5.2.8 to Banach spaces.[1]

Theorem 5.2.6
Let H be a Hilbert space and let $(V_n)_{n\in\mathbb{N}} \subset \mathcal{G}^k(H)$. There exists a well-separating common complement $C \in \mathcal{G}_k(H)$ for $(V_n)_{n\in\mathbb{N}}$.

If $\dim H < \infty$, the claim of Theorem 5.2.6 for $k = 1$ follows from Proposition 5.2.10. For the case $\dim H = \infty$, we need the following two lemmata:

Lemma 5.2.7
Let $(v_n)_{n=1}^d \subset (\mathbb{R}^d, \|.\|_2)$ be unit vectors such that $v_n \in \mathbb{R}^n \times \{0\}$. There are an absolute constant $c > 0$ and $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$ with $|x_i| \leq 1/i^2$ and $|\langle x, v_n\rangle| \geq c/n^5$.

[1] If X is separable, then the problem reduces to solving Lemma 5.2.8 for $X = l^1$. Indeed, every separable Banach space is isomorphic to a quotient of l^1. Now, let $\pi : l^1 \to l^1/A$ be a quotient map. Then, π induces a map $\mathcal{G}^1(l^1/A) \to \mathcal{G}^1(l^1)$ by $V \mapsto \pi^{-1}V$. It holds $d(x, \pi^{-1}V) = d(\pi x, V)$ for $x \in l^1$. Hence, well-separating common complements in l^1 project onto well-separating common complements in l^1/A.

Proof. Let $Q = \prod_{i=1}^{d}[-i^{-2}, i^{-2}]$ and let V_n be the hyperplane orthogonal to v_n. By Lemma 5.2.2 we have

$$\begin{aligned}
\mathrm{vol}_{d-1}(P_{V_n}Q) &= \sum_{i=1}^{d}\left(\prod_{j=1, j\neq i}^{d} 2j^{-2}\right)|\langle e_i, v_n\rangle| \\
&= \frac{1}{2}\mathrm{vol}_d(Q)\sum_{i=1}^{d} i^2 |\langle e_i, v_n\rangle| \\
&= \frac{1}{2}\mathrm{vol}_d(Q)\sum_{i=1}^{n} i^2 |\langle e_i, v_n\rangle| \\
&\leq \frac{1}{2}\mathrm{vol}_d(Q) n^3.
\end{aligned}$$

Now, let $\delta_n := 3/(\pi^2 n^5)$. We estimate

$$\begin{aligned}
\mu(\{y \in Q \mid \exists n : |\langle y, v_n\rangle| \leq \delta_n\}) &\leq \sum_{n=1}^{d} \mu(\{y \in Q \ : \ |\langle y, v_n\rangle| \leq \delta_n\}) \\
&\leq \sum_{n=1}^{d} 2\delta_n \mathrm{vol}_{d-1}(P_{V_n}Q) \\
&\leq \mathrm{vol}_d(Q)\sum_{n=1}^{d} \delta_n n^3 \\
&\leq \mathrm{vol}_d(Q)\frac{3}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2} \\
&= \frac{1}{2}\mathrm{vol}_d(Q).
\end{aligned}$$

Thus, there must be an element $y \in Q$ with $|\langle y, v_n\rangle| \geq \delta_n$ for all n. Since $\|y\|_2^2 \leq \sum_{n=1}^{\infty} 1/n^4 = \pi^4/90$, writing $x := y/\|y\|_2$ yields $|\langle x, v_n\rangle| \geq \delta_n/\|y\|_2 \geq c/n^5$ with $c := 3\sqrt{90}/\pi^4$. □

Lemma 5.2.8
Let H be a Hilbert space of infinite dimension and let $(\varphi_n)_{n\in\mathbb{N}} \subset H'$ be a sequence of bounded linear functionals of norm 1. There exist a sequence $(\delta_n)_{n\in\mathbb{N}} \subset \mathbb{R}_{>0}$ and a unit vector $x \in H$ such that

$$\lim_{n\to\infty}\frac{1}{n}\log\delta_n = 0$$

and

$$|\varphi_n(x)| \geq \delta_n$$

for all n.

Proof. By Riesz's representation theorem, we can write $\varphi_n = \langle v_n, \cdot\rangle$ for unit vectors $v_n \in H$. Now, take an orthonormal set $(c_n)_{n\in\mathbb{N}} \subset H$ with $v_n \in \mathrm{span}(c_1, \ldots, c_n)$. We get maps $\pi_n : H \to \mathbb{R}^n$ defined through

$$\pi_n(x) := \begin{pmatrix}\langle x, c_1\rangle \\ \vdots \\ \langle x, c_n\rangle\end{pmatrix}.$$

By construction $(\pi_n(v_i))_{i=1}^n \subset \mathbb{R}^n$ are unit vectors such that $\pi_n(v_i) \in \mathbb{R}^i \times \{0\}$. In particular, Lemma 5.2.7 gives us the existence of an element $\alpha \in \prod_{k=1}^n [-k^{-2}, k^{-2}]$ with $|\langle \alpha, \pi_n(v_i)\rangle| \geq c/i^5 =: \tilde{\delta}_i$. Let A_n be the set of all such α:

$$A_n := \left\{ \alpha \in \prod_{k=1}^n [-k^{-2}, k^{-2}] \;\middle|\; \forall\, i \leq n : \left| \sum_{k=1}^n \alpha_k \langle v_i, c_k\rangle \right| \geq \tilde{\delta}_i \right\}.$$

We know that A_n is a nonempty, closed subset of $\mathbb{R}^n$. For $\alpha \in A_n$, we can define $y := \sum_{k=1}^n \alpha_k c_k$. Since $\|y\|^2 \leq \sum_{k=1}^\infty 1/k^4 = \pi^4/90$, it holds

$$|\varphi_i(x)| = \|y\|^{-1} |\langle v_i, y\rangle| = \|y\|^{-1} \left| \sum_{k=1}^n \alpha_k \langle v_i, c_k\rangle \right| \geq \sqrt{90}\pi^{-2} \tilde{\delta}_i =: \delta_i$$

for $i \leq n$, where $x := y/\|y\|$. Thus, every $\alpha \in A_n$ induces an element $x \in H$ fulfilling the claim for $\varphi_1, \ldots, \varphi_n$. The remainder of this proof treats the transition $n \to \infty$.

By Tychonoff's theorem the space $\mathcal{B} := \prod_{k=1}^\infty [-k^{-2}, k^{-2}]$ equipped with the product topology is compact. Since the product topology is the coarsest topology such that the canonical projections $\mathrm{pr}_k : \mathcal{B} \to [-k^{-2}, k^{-2}]$ are continuous, we find that $B_n := (\mathrm{pr}_1 \times \cdots \times \mathrm{pr}_n)^{-1} A_n$ are nonempty, closed subsets of $\mathcal{B}$. The sets B_n can be written as

$$B_n = \left\{ \alpha \in \mathcal{B} \;\middle|\; \forall\, i \leq n : \left| \sum_{k=1}^\infty \alpha_k \langle v_i, c_k\rangle \right| \geq \tilde{\delta}_i \right\}.$$

From this form is becomes obvious that $B_1 \supset B_2 \supset \ldots$ is a decreasing sequence of nonempty, closed subsets of $\mathcal{B}$. In particular, $(B_n)_{n\in\mathbb{N}}$ has the finite intersection property, i.e., finite intersection are nonempty. As $\mathcal{B}$ is compact, the intersection of all B_n must be nonempty. Thus, we find some α in

$$\bigcap_{n=1}^\infty B_n = \left\{ \alpha \in \mathcal{B} \;\middle|\; \forall i \in \mathbb{N} : \left| \sum_{k=1}^\infty \alpha_k \langle v_i, c_k\rangle \right| \geq \tilde{\delta}_i \right\}.$$

Similar to above, we set $y := \sum_{k=1}^\infty \alpha_k c_k$. Again, it holds $\|y\|^2 \leq \pi^4/90$. Defining $x := y/\|y\|$, we get

$$|\varphi_n(x)| = \|y\|^{-1} |\langle v_n, y\rangle| = \|y\|^{-1} \left| \sum_{k=1}^\infty \alpha_k \langle v_n, c_k\rangle \right| \geq \sqrt{90}\pi^{-2} \tilde{\delta}_n = \delta_n.$$

for $n \in \mathbb{N}$. □

The proof shows that δ_n can be chosen as c/n^5 for some constant $c > 0$. Improvements of the exponent of n are possible. For instance, one may use $1/n^{1+\epsilon}$ instead of $1/n^2$ to define the polytope in Lemma 5.2.7. However, since our goal is only to find an at most polynomially decaying lower bound, we aimed for better readability at the cost of a worse estimate.

So far it is not known to the author if Lemma 5.2.8 is true for Banach spaces instead of Hilbert spaces. However, the remainder of this section holds for arbitrary Banach spaces $(X, \|.\|)$.

proof of Theorem 5.2.6 for $k = 1$. By Hahn-Banach there are bounded linear functionals $(\varphi_n)_{n\in\mathbb{N}} \subset X'$ of norm 1 such that $\ker \varphi_n = V_n$. Assume that we find $(\delta_n)_{n\in\mathbb{N}}$ and $x \in X$ as described by Lemma 5.2.8. Since $|\varphi_n(x)| = d(x, V_n)$, the subspace spanned by x is a δ-well-separating common complement for $(V_n)_{n\in\mathbb{N}}$. □

To prove Theorem 5.2.6 for arbitrary k we need the following lemma:

Lemma 5.2.9
Let $(X, \|.\|)$ be a Banach space. Furthermore, assume $x_1, x_2 \in B_X(0,1)$ are vectors with $\|x_1\| \geq \mu_1$ and $d(x_2, \operatorname{span}(x_1)) \geq \mu_2$ for some numbers $0 < \mu_1, \mu_2 \leq 1$.

It holds

$$\inf_{t\in\mathbb{R}} \|tx_1 + (1-t)x_2\| \geq \frac{1}{2\sqrt{5}}\mu_1\mu_2.$$

Proof. The argument can be restricted to $\operatorname{span}(x_1, x_2)$. Thus, assume that $\dim X = 2$. First, we look at $X = \mathbb{R}^2$ equipped with $\|.\|_2$. After a rotation we may assume $x_1 = (\alpha_1, 0)$ with $\alpha_1 \geq \mu_1$. Now, the assumption on x_2 implies that its second coordinate has at least size μ_2. Let L be the line passing through x_1 and x_2 (see Fig. 5.1). We want to estimate the distance between L and the origin. Clearly, the distance becomes smallest if L intersects the unit circle at $(-\sqrt{1-\mu_2^2}, \pm\mu_2)$. Hence, the task reduces to finding δ in Fig. 5.2. After applying Pythagoras' theorem to find the diagonal d of the big triangle and comparing ratios between catheti opposite to α and the hypotenuses, we get

$$\delta = \frac{\mu_1\mu_2}{d} = \frac{\mu_1\mu_2}{\sqrt{\mu_2^2 + \left(\sqrt{1-\mu_2^2} + \mu_1\right)^2}} \geq \frac{1}{\sqrt{5}}\mu_1\mu_2.$$

Thus, the claim holds for the euclidean case.

Now, let X be any 2-dimensional Banach space. By Theorem 5.2.4 there exists an isomorphism T from $(X, \|.\|)$ to $(\mathbb{R}^2, \|.\|_2)$ with $\|T\| \leq 1$ and $\|T^{-1}\| \leq \sqrt{2}$. Let $x_1, x_2 \in X$ be as in the claim. It holds $Tx_1, Tx_2 \in B_2(0,1)$, $\|Tx_1\|_2 \geq \mu_1/\sqrt{2}$ and $d(Tx_2, \operatorname{span}(Tx_1)) \geq \mu_2/\sqrt{2}$. From the euclidean case we get

$$\inf_{t\in\mathbb{R}} \|tx_1 + (1-t)x_2\| \geq \inf_{t\in\mathbb{R}} \|tTx_1 + (1-t)Tx_2\|_2 \geq \frac{1}{2\sqrt{5}}\mu_1\mu_2.$$

$\square$

proof of Theorem 5.2.6 for arbitrary k. The proof is done by induction over k. Assume that the claim holds for $k \geq 1$. Let $(V_n)_{n\in\mathbb{N}} \subset \mathcal{G}^{k+1}(X)$ be as in the claim and define $\pi_n : X \to X/V_n$ to be the associated quotient maps. We embed $(V_n)_{n\in\mathbb{N}}$ into two different sequences of closed complemented subspaces of X, one having codimension k and the other having codimension 1. Summing their well-separating common complements will yield a well-separating common complement for our initial sequence.

First, take any $(V_n^1)_{n\in\mathbb{N}} \subset \mathcal{G}^k(X)$ with $V_n^1 \supset V_n$. According to the codimension k case we find a δ^1-well-separating common complement $C_1 \in \mathcal{G}_k(X)$ for $(V_n^1)_{n\in\mathbb{N}}$. It holds $\|\pi_n x_1\| \geq d(x_1, V_n^1) \geq \delta_n^1$ for all $x_1 \in C_1$ of norm 1.

Next, let $V_n^2 := V_n \oplus C_1$. Then, $(V_n^2)_{n\in\mathbb{N}} \subset \mathcal{G}^1(X)$ is a sequence of closed, complemented subspaces of codimension 1. Hence, we find a δ^2-well-separating common complement $C_2 \in \mathcal{G}_1(X)$. Let x_2 be one of the two unit vectors of C_2. We have $d(\pm\pi_n x_2, \pi_n C^1) \geq \delta_n^2$.

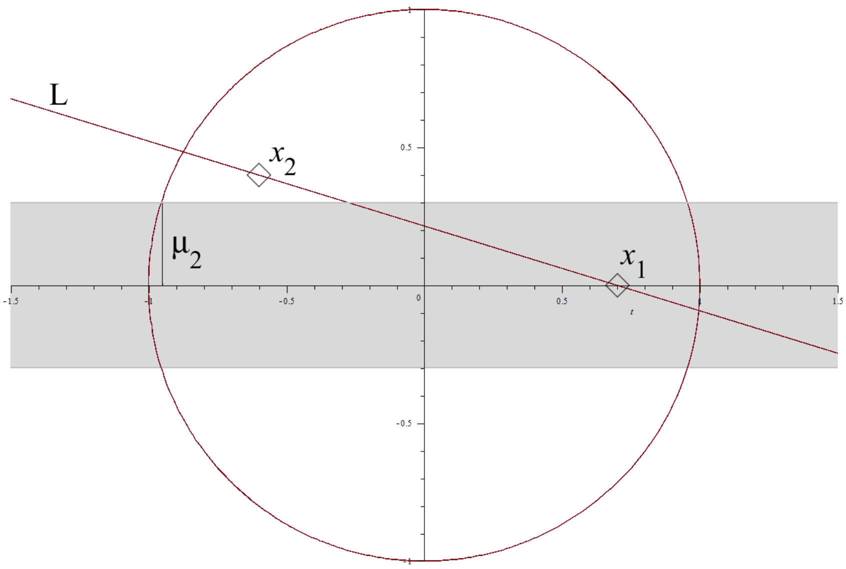


Figure 5.1: The simplified planar case from the proof of Lemma 5.2.9

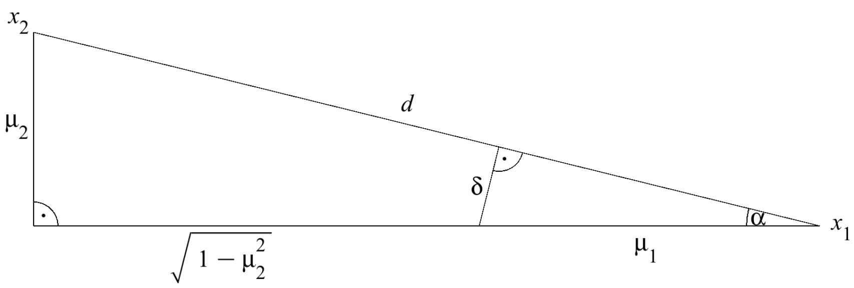


Figure 5.2: The triangle reduction from the proof of Lemma 5.2.9

Let $C := C_1 \oplus C_2$. To check if C is well-separating, we need to find a lower bound of $\|\pi_n x\|$ with $x \in C$ of norm 1. We scale x so that it intersects with an element of the boundary of the double cone

$$\Delta := \left\{ c \in C \;\middle|\; c = tx_1 + (1-t)(\pm x_2),\ t \in [0,1],\ x_1 \in B_{C_1}(0,1) \right\},$$

which is contained in $B_C(0,1)$ (see Fig. 5.3). The boundary $\partial\Delta$ is made up of line segments connecting unit vectors $x_1 \in C_1$ with one of the two apexes $\pm x_2 \in C_2$. By Lemma 5.2.9 the image of each line segment under π_n is far enough from the origin, i.e., we have

$$\inf_{t\in[0,1]} \|t\pi_n x_1 + (1-t)(\pm\pi_n x_2)\| \geq \frac{1}{2\sqrt{5}} \delta_n^1 \delta_n^2 =: \delta_n.$$

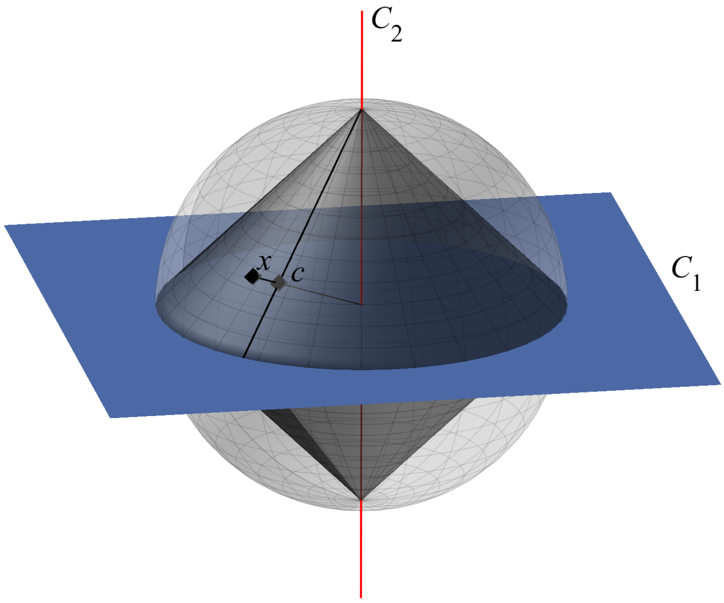


Figure 5.3: The double cone from the proof of Theorem 5.2.6

Since every $x \in C$ of norm 1 can be written as $x = \lambda c$ for some $\lambda \geq 1$ and $c \in \partial\Delta$, it holds $d(x, V_n) = \|\pi_n x\| = \lambda\|\pi_n c\| \geq \lambda\delta_n \geq \delta_n$. Thus, C is a δ-well-separating common complement for $(V_n)_{n\in\mathbb{N}}$. □

5.2.3 Genericity

In finite dimensions it is a simple task to show that almost every vector spans a well-separating common complement for a given family of countably many hyperplanes:

Proposition 5.2.10
Let X be a Banach space of finite dimension and let $(V_n)_{n\in\mathbb{N}} \subset \mathcal{G}^1(X)$ be hyperplanes. Almost every $x \in X$ spans a well-separating common complement for $(V_n)_{n\in\mathbb{N}}$.

Proof. Since well-separating common complements are retained when changing to an equivalent norm, we may assume $(X, \|.\|) = (\mathbb{R}^d, \|.\|_2)$. Furthermore, we can restrict ourselves to $x \in B_d(0,1)$. Define $\delta_n^\epsilon := \epsilon/n^2$ for $\epsilon > 0$. We estimate

$$\begin{aligned}\mu(\{x \in B_d(0,1) \mid \exists n : d(x, V_n) \leq \delta_n^\epsilon\}) &\leq \sum_{n=1}^{\infty} \mu(\{x \in B_d(0,1) \mid d(x, V_n) \leq \delta_n^\epsilon\}) \\ &\leq \sum_{n=1}^{\infty} 2\delta_n^\epsilon \mathrm{vol}_{d-1}(B_{d-1}(0,1)) \\ &= \epsilon\frac{\pi^2}{3}\mathrm{vol}_{d-1}(B_{d-1}(0,1)) \\ &\xrightarrow[\epsilon\to 0]{} 0.\end{aligned}$$

Hence, for almost every $x \in B_d(0,1)$, there is an $\epsilon > 0$ such that $\mathrm{span}(x)$ is a δ^ϵ-well-separating common complement for $(V_n)_{n\in\mathbb{N}}$ (see Fig. 5.4 for a conceptual representation of the proof). □

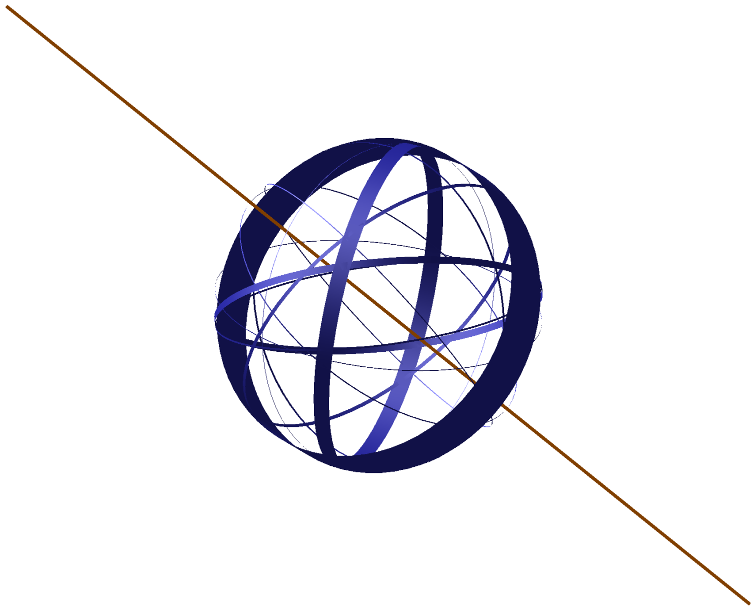


Figure 5.4: A common complement for ten randomly chosen hyperplanes of $\mathbb{R}^3$. The hyperplanes are represented by unit circles resulting from intersecting the hyperplanes with the unit sphere. Each circle on the unit sphere is equipped with a neighborhoods of size ϵ/n^2, where $n = 1, \ldots, 10$ is the number of the corresponding hyperplane. In the proof of Proposition 5.2.10 we seek common complements, here plotted as an orange line, which do not intersect the blue set for some $\epsilon > 0$. Since the blue set becomes arbitrary small for $\epsilon \to 0$, almost every 1-dimensional subspace fulfills this assumption.

Since there is no equivalent of the Lebesgue measure for arbitrary Banach or Hilbert spaces, the proof of Proposition 5.2.10 does not generalize to infinite dimensions. Even the notion of "almost every" in the claim is not clear a priori. Instead of "Lebesgue almost every" we will use the concept of prevalence:

Definition 5.2.11 ([53])
A Borel subset $E \subset X$ of a Banach space is called prevalent *if there exists a Borel measure μ on X such that*

1. *$0 < \mu(C) < \infty$ for some compact set $C \subset X$, and*

2. *$E + x$ has full μ-measure for all $x \in X$.*

A general subset $F \subset X$ is called prevalent *if it contains a prevalent Borel set. We say that* almost every *element $x \in X$ lies in F.*

Prevalence satisfies the following genericity axioms:

Proposition 5.2.12 ([53])
The following are true:

1. *F prevalent $\implies$ F dense in X,*

2. *$L \supset G$, G prevalent $\implies$ L prevalent,*

3. *countable intersections of prevalent sets are prevalent,*
4. *translations of prevalent sets are prevalent, and*
5. $G \subset \mathbb{R}^d$ *is prevalent if, and only if,* G *has full Lebesgue measure, i.e., its complement has Lebesgue measure zero.*

The last point implies that the notions of "almost every" in the sense of Lebesgue and in the sense of prevalence coincide in finite-dimensional Banach spaces.

To identify prevalent sets in infinite-dimensional spaces it is convenient to use probe spaces. A *probe* is a finite-dimensional subspace $P \subset X$ of a Banach space. By identification with the standard euclidean space we can equip P with a Borel measure λ_P. This measure induces a Borel measure μ_P on X by $\mu_P(A) := \lambda_P(A \cap P)$ for Borel sets $A \subset X$. Using μ_P in Definition 5.2.11 yields the following definition:

Definition 5.2.13 ([53])
A finite-dimensional subspace $P \subset X$ *is called a* probe *for* $F \subset X$ *if there exists a Borel set* $E \subset F$ *such that* $E + x$ *has full* μ_P*-measure for every* $x \in X$.

Proposition 5.2.14 ([53])
The existence of a probe for $F \subset X$ *implies that* F *is prevalent.*

With the additional terminology we are ready to state a result about the genericity of well-separating common complements in Hilbert spaces:

Theorem 5.2.15
Let H *be a Hilbert space and let* $(V_n)_{n\in\mathbb{N}} \subset \mathcal{G}^k(H)$. *The set of all* $(x_1, \dots, x_k) \in H^k$, *such that* $\operatorname{span}(x_1, \dots, x_k)$ *is a well-separating common complement for* $(V_n)_{n\in\mathbb{N}}$, *is prevalent.*

We show that the existence of one well-separating common complement already implies that they are prevalent. In particular, this proves Theorem 5.2.15. However, before beginning with the proof we need a few elementary and technical lemmata.

Lemma 5.2.16
Let X *be a Banach space and* $U \subset X$ *an open subset. If* $f : U \times (\mathbb{R}^k \setminus \{0\}) \to \mathbb{R}$ *is continuous, then the mapping* $g : U \to \mathbb{R}$ *defined by*

$$g(x) := \min_{\|\alpha\|_2=1} f(x, \alpha)$$

is continuous as well.

Proof. Let $\epsilon > 0$ be given. For each $(x, \alpha) \in U \times (\mathbb{R}^k \setminus \{0\})$, we find $\delta_{(x,\alpha)} > 0$ such that

$$\|(x, \alpha) - (y, \beta)\| < \delta_{(x,\alpha)} \quad \implies \quad |f(x, \alpha) - f(y, \beta)| < \epsilon$$

for $(y, \beta) \in U \times (\mathbb{R}^k \setminus \{0\})$. Fix $x \in U$. Since the set $\{x\} \times \{\|\alpha\|_2 = 1\}$ is compact, it is covered by finitely many balls of radius $\delta_{(x,\alpha)}$ with α from $\{\|\alpha\|_2 = 1\}$. Thus, we find $\delta_x > 0$ such that

$$\|(x, \alpha) - (y, \beta)\| < \delta_x \quad \implies \quad |f(x, \alpha) - f(y, \beta)| < \epsilon$$

for $(y, \beta) \in U \times (\mathbb{R}^k \setminus \{0\})$ with $\|\alpha\|_2 = 1$. Now, if $\|x - y\| < \delta_x$, then

$$g(x) \leq \min_{\|\alpha\|_2=1} \left(f(y, \alpha) + |f(x, \alpha) - f(y, \alpha)|\right) \leq g(y) + \epsilon.$$

□

Lemma 5.2.17
The set of all tuples spanning well-separating common complements for $(V_n)_{n\in\mathbb{N}}$ *is a Borel subset of* X^k.

Proof. First, define the map $s : X^k \to \mathbb{R}$ by

$$s(c) := \min_{\|\alpha\|_2=1} \left\| \sum_{i=1}^{k} \alpha_i c_i \right\|.$$

With the help of Lemma 5.2.16 it is easily seen that s is continuous. In particular, the set $U := s^{-1}(0,\infty)$ of all linearly independent tuples is open in X^k. Next, let $\pi_n : X \to X/V_n$ be the quotient map associated to V_n. We apply Lemma 5.2.16 again to see that the maps $g_n : U \to \mathbb{R}$ given by

$$g_n(c) := \min_{\|\alpha\|_2=1} \frac{\| \sum_{i=1}^{k} \alpha_i \pi_n c_i \|}{\| \sum_{i=1}^{k} \alpha_i c_i \|}$$

are continuous. Slightly rewriting g_n reveals that

$$g_n(c) = \inf_{x\in \mathrm{span}(c_1,\ldots,c_k)\cap S} d(x, V_n)$$

has the form as in Definition 5.2.1. In particular, $\mathrm{span}(c_1, \ldots, c_k)$ is a well-separating common complement if, and only if, $c \in U$, $g_n(c) > 0$, and

$$\lim_{n\to\infty} \frac{1}{n} \log g_n(c) = 0.$$

Let $f_n : U_n \to \mathbb{R}$ be given by $f_n(c) := (1/n) \log g_n(c)$, where $U_n := g_n^{-1}(0,\infty) \subset X^k$ is open. Then, f_n is continuous and bounded from above by zero (since $g_n(c)$ is bounded from above by 1). Finally, the set of tuples spanning well-separating common complements can be expressed as

$$\bigcap_{l>0} \bigcup_{N\in\mathbb{N}} \bigcap_{n\geq N} \left\{ c \in U_n \,\middle|\, f_n(c) > -\frac{1}{l} \right\},$$

which is a Borel set. □

Lemma 5.2.18
Let $(A_n)_{n\in\mathbb{N}} \subset \mathbb{R}^{k\times k}$ *be a sequence of matrices. For almost every* $A \in \mathbb{R}^{k\times k}$*, there exists* $\epsilon > 0$ *such that*

$$\forall n \in \mathbb{N} : \; |\det(A + A_n)| \geq \frac{\epsilon}{n^2}.$$

Proof. Let $M > 0$ and $\tilde{A}_n := (1/M)A_n$. Assume the claim holds for almost every $\tilde{A} \in B(0,1)^k$ with respect to the sequence $(\tilde{A}_n)_{n\in\mathbb{N}}$, where $B(0,1)^k \subset \mathbb{R}^{k\times k}$ is a product of unit balls in $(\mathbb{R}^k, \|.\|_2)$ each representing a column. Setting $A := M\tilde{A}$ for any such $\tilde{A}$ yields

$$|\det(A + A_n)| = M^k |\det(\tilde{A} + \tilde{A}_n)| \geq M^k \frac{\tilde{\epsilon}}{n^2}$$

for some $\tilde{\epsilon} > 0$. In particular, almost every $A \in B(0,M)^k$ fulfills the required estimate with respect to $(A_n)_{n\in\mathbb{N}}$. Exhausting $\mathbb{R}^{k\times k}$ with $B(0,M)^k$ for $M \to \infty$

implies that the claim holds for almost every $A \in \mathbb{R}^{k\times k}$. Thus, it remains to prove that the claim holds for almost every $A \in B(0,1)^k$.

For $A = (a_1, \dots, a_k)$, it holds

$$\begin{aligned}
|\det A| &= \Big|\det\Big(a_1, P_{\operatorname{span}(a_1)^\perp} a_2, \dots, P_{\operatorname{span}(a_1,\dots,a_{k-1})^\perp} a_k\Big)\Big| \\
&= \operatorname{vol}_k\Big(Q\Big(a_1, P_{\operatorname{span}(a_1)^\perp} a_2, \dots, P_{\operatorname{span}(a_1,\dots,a_{k-1})^\perp} a_k\Big)\Big) \\
&= \|a_1\|_2 \, \|P_{\operatorname{span}(a_1)^\perp} a_2\|_2 \dots \|P_{\operatorname{span}(a_1,\dots,a_{k-1})^\perp} a_k\|_2,
\end{aligned}$$

where $Q(v_1, \dots, v_k) \subset \mathbb{R}^k$ denotes the parallelepiped spanned by vectors $v_1, \dots, v_k \in \mathbb{R}^k$. Using this representation, we will derive an estimate of the form

$$\mu\Big(\Big\{A \in B(0,1)^k \ : \ |\det(A+\tilde{A})| \le \eta\Big\}\Big) \le c\eta \tag{5.2}$$

for all $\eta > 0$ independent of $\tilde{A}$, where $c > 0$ is a constant only depending on k.

To this end fix $\tilde{A} = (\tilde{a}_1, \dots, \tilde{a}_k)$ and define

$$\begin{aligned}
t_i(a_1, \dots, a_i) := \|a_1 + \tilde{a}_1\|_2 \, &\|P_{\operatorname{span}(a_1+\tilde{a}_1)^\perp}(a_2 + \tilde{a}_2)\|_2 \\
&\dots \|P_{\operatorname{span}(a_1+\tilde{a_1},\dots,a_{i-1}+\tilde{a}_{i-1})^\perp}(a_i + \tilde{a}_i)\|_2
\end{aligned}$$

for $i = 1, \dots, k$. Set $t_0 := 1$. To arrive at an estimate as in Eq. (5.2) we split the integral

$$\int_{B(0,1)^k} \chi_{\{A \,:\, |\det(A+\tilde{A})| \le \eta\}}(A)\, dA$$

using Fubini's theorem column by column. The inner integral becomes

$$I := \int_{B(0,1)} \chi_{\{a_k \,:\, \|P_{\operatorname{span}(a_1+\tilde{a}_1,\dots,a_{k-1}+\tilde{a}_{k-1})^\perp}(a_k+\tilde{a}_k)\|_2 \le \eta t_{k-1}^{-1}\}}(a_k)\, da_k,$$

where t_{k-1}^{-1} depends on $a_1, \dots, a_{k-1}$ and might be ∞. If it is ∞, then the inner integral is $\operatorname{vol}_k(B(0,1))$. In the other case $a_1 + \tilde{a}_1, \dots, a_{k-1} + \tilde{a}_{k-1}$ must be linearly independent. Hence, their linear span is of dimension $k-1$ and we find an orthogonal transformation T that maps $e_1, \dots, e_{k-1}$ into their span and maps e_k into the orthogonal complement. After applying the transformation to I, we have

$$I = \int_{B(0,1)} \chi_{\{b_k \,:\, \|P_{\operatorname{span}(a_1+\tilde{a}_1,\dots,a_{k-1}+\tilde{a}_{k-1})^\perp}(Tb_k+\tilde{a}_k)\|_2 \le \eta t_{k-1}^{-1}\}}(b_k)\, db_k.$$

Writing $b_k = (\beta_{1k}, \dots, \beta_{kk})^T$ and $\tilde{b}_k = (\tilde{\beta}_{1k}, \dots, \tilde{\beta}_{kk})^T$ for $\tilde{b}_k := T^{-1}\tilde{a}_k$, we get

$$\begin{aligned}
I &= \int_{B(0,1)} \chi_{\{b_k \,:\, |\beta_{kk}+\tilde{\beta}_{kk}| \le \eta t_{k-1}^{-1}\}}(b_k)\, db_k \\
&\le 2^{k-1} \int_{-1}^{1} \chi_{\{\beta_{kk} \,:\, |\beta_{kk}+\tilde{\beta}_{kk}| \le \eta t_{k-1}^{-1}\}}(\beta_{kk})\, d\beta_{kk} \\
&\le 2^{k-1} \int_{-1}^{1} \chi_{\{\beta_{kk} \,:\, |\beta_{kk}| \le \eta t_{k-1}^{-1}\}}(\beta_{kk})\, d\beta_{kk} \\
&= 2^k \min(1, \eta t_{k-1}^{-1}) \\
&\le 2^k \eta t_{k-1}^{-1}.
\end{aligned}$$

For the first inequality, we embedded $B(0,1)$ into $[-1,1]^k$. Now, we have an estimate of I depending on $a_1, \dots, a_{k-1}$ that also holds when $t_{k-1}^{-1} = \infty$. In the following we show that

$$\int_{B(0,1)^{k-1}} t_{k-1}^{-1}\, d(a_1, \dots, a_{k-1}) \le c' \tag{5.3}$$

for some constant c' by proving that

$$\int_{B(0,1)} t_{k-i}^{-1}\, da_{k-i} \leq c_i' t_{k-(i+1)}^{-1} \tag{5.4}$$

for some constants c_i' for $i = 1, \ldots, k-1$. Ultimately, it follows that we can set $c' := c_1' \ldots c_{k-1}'$ and $c := 2^k c'$ to reach the desired estimate in Eq. (5.2).

So, let us prove the above inductive formula Eq. (5.4). We write

$$\begin{aligned}&\int_{B(0,1)} t_{k-i}^{-1}\, da_{k-i} \\ &\quad = t_{k-(i+1)}^{-1} \int_{B(0,1)} \|P_{\operatorname{span}(a_1+\tilde{a}_1,\ldots,a_{k-(i+1)}+\tilde{a}_{k-(i+1)})^\perp}(a_{k-i}+\tilde{a}_{k-i})\|_2^{-1}\, da_{k-i}.\end{aligned}$$

As before, we distinguish between the cases $t_{k-(i+1)}^{-1} = \infty$ and $t_{k-(i+1)}^{-1} < \infty$. In the first case, the inductive formula Eq. (5.4) is obviously satisfied. In the second case, we again apply a transformation T which rotates the first $k-(i+1)$ vectors of the standard basis to $\operatorname{span}(a_1+\tilde{a}_1, \ldots, a_{k-(i+1)}+\tilde{a}_{k-(i+1)})$ and the remaining basis vectors to its orthogonal complement. Similar to before, writing $b_{k-i} = (\beta_{1(k-i)}, \ldots, \beta_{k(k-i)})^T$ and $\tilde{b}_{k-i} = (\tilde{\beta}_{1(k-i)}, \ldots, \tilde{\beta}_{k(k-i)})^T$ for $\tilde{b}_{k-i} := T^{-1}\tilde{a}_{k-i}$, we get

$$\begin{aligned}&\int_{B(0,1)} \|P_{\operatorname{span}(a_1+\tilde{a}_1,\ldots,a_{k-(i+1)}+\tilde{a}_{k-(i+1)})^\perp}(a_{k-i}+\tilde{a}_{k-i})\|_2^{-1}\, da_{k-i} \\ &\quad = \int_{B(0,1)} \|P_{\operatorname{span}(a_1+\tilde{a}_1,\ldots,a_{k-(i+1)}+\tilde{a}_{k-(i+1)})^\perp}(Tb_{k-i}+\tilde{a}_{k-i})\|_2^{-1}\, db_{k-i} \\ &\quad = \int_{B(0,1)} \|(\beta_{(k-i)(k-i)}+\tilde{\beta}_{(k-i)(k-i)}, \ldots, \beta_{k(k-i)}+\tilde{\beta}_{k(k-i)})^T\|_2^{-1}\, db_{k-i}.\end{aligned}$$

Let $\beta_{k-i} := (\beta_{(k-i)(k-i)}, \ldots, \beta_{k(k-i)})^T$ and $\tilde{\beta}_{k-i} := (\tilde{\beta}_{(k-i)(k-i)}, \ldots, \tilde{\beta}_{k(k-i)})^T$. Embedding $B(0,1) \subset \mathbb{R}^k$ into $[-1,1]^{k-(i+1)} \times B(0,1) \subset \mathbb{R}^{k-(i+1)} \times \mathbb{R}^{i+1}$ shows that the above integral can be estimated by

$$2^{k-(i+1)} \int_{B(0,1)} \|\beta_{k-i}+\tilde{\beta}_{k-i}\|_2^{-1}\, d\beta_{k-i} \leq 2^{k-(i+1)} \int_{B(0,1)} \|\beta_{k-i}\|_2^{-1}\, d\beta_{k-i} =: c_i' < \infty.$$

Tracing back the steps, this concludes the proof of Eq. (5.4), which in turn gives us Eq. (5.3) and Eq. (5.2). Having Eq. (5.2), we set $\eta := \epsilon/n^2$ and $\tilde{A} := A_n$. It holds

$$\begin{aligned}&\mu\Big(\Big\{A \in B(0,1)^k \mid \exists n : |\det(A+A_n)| \leq \epsilon n^{-2}\Big\}\Big) \\ &\quad \leq \sum_{n=1}^{\infty} \mu\Big(\Big\{A \in B(0,1)^k \ : \ |\det(A+A_n)| \leq \epsilon n^{-2}\Big\}\Big) \\ &\quad \leq \sum_{n=1}^{\infty} c\epsilon n^{-2} \\ &\quad = \epsilon \frac{c\pi^2}{6} \\ &\quad \xrightarrow[\epsilon\to 0]{} 0.\end{aligned}$$

Hence, for almost every $A \in B(0,1)^k$, there is $\epsilon > 0$ such that we have $|\det(A + A_n)| \geq \epsilon/n^2$ for all $n \in \mathbb{N}$. □

Lemma 5.2.19
Let $(A_n)_{n\in\mathbb{N}} \subset \mathbb{R}^{k\times k}$ *be a sequence of matrices such that* $\|A_n\|_2 \leq 1/\delta_n$ *with* $0 < \delta_n \leq 1$. *For almost every* $A \in \mathbb{R}^{k\times k}$, *there is* $\epsilon > 0$ *with*

$$\forall n \in \mathbb{N} : \ \|(A + A_n)^{-1}\|_2^{-1} \geq \epsilon n^{-2}\delta_n^{k-1}.$$

Proof. Let A be as in Lemma 5.2.18. Using the adjugate, we write

$$(A + A_n)^{-1} = \det(A + A_n)^{-1}(A + A_n)^{\mathrm{ad}}.$$

Hence, we have

$$\|(A + A_n)^{-1}\|_2^{-1} = |\det(A + A_n)| \, \|(A + A_n)^{\mathrm{ad}}\|_2^{-1}.$$

According to Lemma 5.2.18 the determinant part can be estimated from below by $\tilde{\epsilon}/n^2$. For the adjugate part, we remark that the spectral norm and the max norm on $\mathbb{R}^{k\times k}$ are equivalent. Thus, there are constants $c_1, c_2 > 0$ with $c_1\|.\|_{\max} \leq \|.\|_2 \leq c_2\|.\|_{\max}$. Moreover, the entries of the adjugate consist of determinants of $(k-1)\times(k-1)$-matrices with entries from $A + A_n$. As a simple corollary of Hadamard's inequality, we can estimate these determinants using the max norm to obtain

$$\begin{aligned}
\|(A + A_n)^{\mathrm{ad}}\|_2 &\leq c_2\|(A + A_n)^{\mathrm{ad}}\|_{\max} \\
&\leq c_2\|A + A_n\|_{\max}^{k-1}(k-1)^{\frac{k-1}{2}} \\
&\leq c_2(k-1)^{\frac{k-1}{2}} c_1^{-(k-1)}\|A + A_n\|_2^{k-1} \\
&\leq c_2(k-1)^{\frac{k-1}{2}} c_1^{-(k-1)}(\|A\|_2 + \|A_n\|_2)^{k-1} \\
&\leq c_2(k-1)^{\frac{k-1}{2}} c_1^{-(k-1)}(\|A\|_2 + \delta_n^{-1})^{k-1} \\
&\leq c_2(k-1)^{\frac{k-1}{2}} c_1^{-(k-1)}(\|A\|_2 + 1)^{k-1}\delta_n^{-(k-1)} \\
&=: c\delta_n^{-(k-1)}.
\end{aligned}$$

Now, we set $\epsilon := \tilde{\epsilon}/c$ to obtain the result. ☐

Proposition 5.2.20
Let X *be a Banach space. Assume there exists a well-separating common complement for* $(V_n)_{n\in\mathbb{N}} \subset \mathcal{G}^k(X)$. *Then, the set of all* $(x_1, \dots, x_k) \in X^k$, *such that* $\mathrm{span}(x_1, \dots, x_k)$ *is a well-separating common complement for* $(V_n)_{n\in\mathbb{N}}$, *is prevalent.*

Proof. Let C be a δ-well-separating common complement for $(V_n)_{n\in\mathbb{N}}$. To prove prevalence, we show that the set

$$\Big\{(c_1, \dots, c_k) \in C^k \ \Big| \ \mathrm{span}(c_1 + x_1, \dots, c_k + x_k) \text{ is a well-separating common complement for } (V_n)_{n\in\mathbb{N}}\Big\}$$

has full Lebesgue measure in the probe space C^k for every translation by $(x_1, \dots, x_k)$ in X^k. To get a notion of Lebesgue measure on C^k we identify a basis $(b_1, \dots, b_k)$ of C with the standard basis $(e_1, \dots, e_k)$ of $\mathbb{R}^k$. Let us denote this isomorphism by $I : C \to \mathbb{R}^k$. We naturally get an isomorphism $I^k : C^k \to \mathbb{R}^{k\times k}$ mapping elements of C^k to matrices column by column. Thus, we need to check for the measure of all coefficient matrices yielding well-separating common complements. At this point, let us note that the norm on X^k is given by $\|(x_1, \dots, x_k)\|_{X^k} := \|x_1\| + \dots + \|x_k\|$.

Fix a translation $(x_1, \dots, x_k)$. For each $n \in \mathbb{N}$, we can write $x_i = c'_{i,n} + v'_{i,n}$ according to the splitting $X = C \oplus V_n$. The translation contributed by $(c'_{1,n}, \dots, c'_{k,n})$ boils down to a translation on $\mathbb{R}^{k \times k}$ by $A_n := I^k(c'_{1,n}, \dots, c'_{k,n})$. We are interested in the extend of this translation with increasing n. To find an upper bound of the norm of A_n, we first assume that $\|c'_{i,n}\| > 0$. It holds

$$\frac{\|x_i\|}{\|c'_{i,n}\|} = \left\| \frac{c'_{i,n}}{\|c'_{i,n}\|} + \frac{v'_{i,n}}{\|c'_{i,n}\|} \right\| \geq d\left(\frac{c'_{i,n}}{\|c'_{i,n}\|}, V_n \right) \geq \delta_n.$$

Thus, we have $\|c'_{i,n}\| \leq (1/\delta_n) \max_i \|x_i\|$ even if $\|c'_{i,n}\| = 0$. Switching to the coefficient space, we get $\|A_n\|_2 \leq \|I^k\| k\, (1/\delta_n) \max_i \|x_i\|$, which can be estimated further by $1/\tilde{\delta}_n := \max\left(1, \|I^k\| k\, (1/\delta_n) \max_i \|x_i\|\right)$.

Now, let A be as in Lemma 5.2.19 with respect to the sequence $(A_n)_{n \in \mathbb{N}}$. We will show that A induces a well-separating common complement. Let $(c_1, \dots, c_k) := (I^k)^{-1} A$ and let $c \in \operatorname{span}(c_1 + x_1, \dots, c_k + x_k)$ with $\|c\| = 1$. We express c in terms of coefficients

$$c = \sum_{i=1}^{k} \gamma_i (c_i + x_i) = \sum_{i=1}^{k} \gamma_i \sum_{j=1}^{k} (\alpha_{ji} + \alpha_{ji,n}) b_j + \sum_{i=1}^{k} \gamma_i v'_{i,n}, \tag{5.5}$$

where $A = (\alpha_{ij})_{ij}$ and $A_n = (\alpha_{ij,n})_{ij}$. Since $A + A_n$ is invertible by Lemma 5.2.19 and $(b_1, \dots, b_k)$ is a basis, the vectors $\sum_{j=1}^{k} (\alpha_{ji} + \alpha_{ji,n}) b_j$ for $i = 1, \dots, k$ form a basis of C. In particular, the double sum in Eq. (5.5) does not vanish. Using the fact that C is δ-well-separating, we compute

$$\begin{aligned}
d(c, V_n) &= d\left(\sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) b_j, V_n \right) \\
&= \left\| \sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) b_j \right\| d\left(\frac{\sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) b_j}{\left\| \sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) b_j \right\|}, V_n \right) \\
&\geq \left\| \sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) b_j \right\| \delta_n.
\end{aligned}$$

We transfer further norm estimates onto the coefficient space. It holds

$$\left\| \sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) e_j \right\|_2 \leq \|I\| \left\| \sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) b_j \right\|.$$

Let $\gamma := (\gamma_1, \dots, \gamma_k)^T$. Using the identity $\gamma = (A + A_n)^{-1}(A + A_n)\gamma$, we get

$$\begin{aligned}
\left\| \sum_{i,j} \gamma_i (\alpha_{ji} + \alpha_{ji,n}) e_j \right\|_2 &= \|(A + A_n)\gamma\|_2 \\
&\geq \|(A + A_n)^{-1}\|_2^{-1} \|\gamma\|_2 \\
&\geq \epsilon n^{-2} \tilde{\delta}_n^{k-1} \|\gamma\|_2
\end{aligned}$$

with $\epsilon > 0$ from Lemma 5.2.19. As $(c_i + x_i)_{i=1}^{k}$ are linearly independent, the norm of γ such that $\|\sum_{i=1}^{k} \gamma_i (c_i + x_i)\| = 1$ for fixed c_i and x_i is bounded from below by a positive constant $\eta > 0$.

Let $\delta'_n := \epsilon n^{-2}\tilde{\delta}_n^{k-1}\eta\|I\|^{-1}\delta_n$. Putting everything together, we have shown that

$$\inf_{c\in\mathrm{span}(c_1+x_1,\ldots,c_k+x_k)\cap S} d(c,V_n) \geq \delta'_n$$

for all $n\in\mathbb{N}$, which tells us that $\mathrm{span}(c_1+x_1,\ldots,c_k+x_k)$ is a δ'-well-separating common complement for $(V_n)_{n\in\mathbb{N}}$. Hence, given an arbitrary translation by $(x_1,\ldots,x_k)\in X^k$, almost every $A\in\mathbb{R}^{k\times k}$ induces a well-separating common complement. □

Tracking δ' in the Hilbert space setting reveals that almost every tuple yields a common complement such that the degree of transversality decays at most polynomially with c/n^{5k^2+2} for some $c>0$ depending on the tuple. A better general rate of decay can be obtained by carefully refining the proofs.

5.3 Convergence results

We now derive a convergence theorem for Ginelli's algorithm in the setting of Theorem 2.3.2. The main statement is similar to the one for finite dimensions: exponentially fast convergence of the algorithm for almost every input. This time, however, convergence is analyzed using well-separating common complements instead of admissibility. Since we know about existence and prevalence of well-separating common complements in Hilbert spaces and since we used an orthonormalization procedure in the definition of Ginelli's algorithm[2], the upcoming convergence theorem is only for Hilbert spaces.

5.3.1 Theorem

Let us state the main result of this chapter:

Theorem 5.3.1 (Convergence a.e. of Ginelli's algorithm on Hilbert spaces with $\mathbb{T}=\mathbb{Z}$)
Let $\mathcal{R}=(\Omega,\mathcal{F},\mathbb{P},\sigma,H,\ \mathcal{L})$ satisfy the assumptions of Theorem 2.3.2 and let $k=d_1+\cdots+d_l$ for some finite $l\leq p$. Moreover, set $\lambda_0:=\infty$ and $\lambda_{p+1}:=\kappa^*$.

On a subset $\Omega'\subset\Omega$ of full $\mathbb{P}$-measure, Ginelli's algorithm converges for almost every input. That is, fixing $\omega\in\Omega'$, for almost every tuple $(x_1,\ldots,x_k)\in H^k$, for almost every $R\in\mathbb{R}^{k\times k}_{ru}$, and for all $i\leq l$, it holds[3]

$$\begin{aligned}\limsup_{N\to\infty}\sup_{n_1,n_2\geq N}\frac{1}{\min(n_1,n_2)}\log d_{\mathcal{G}}\Big(\mathrm{span}\Big\{\Big(G^{n_1,n_2}_{\omega,k}\Big)_{i_j}\ \Big|\ j=1,\ldots,d_i\Big\},Y_i(\omega)\Big)\\ \leq -\min(|\lambda_i-\lambda_{i-1}|,|\lambda_i-\lambda_{i+1}|)\end{aligned}\tag{5.6}$$

at $((x_1,\ldots,x_k),R)$.

There are three concepts of "almost every" in the statement of the theorem. Firstly, the algorithm fixes ω from a set of full $\mathbb{P}$-measure to determine the trajectory along which Ginelli's algorithm is applied. Secondly and thirdly, the algorithm

[2]On an abstract level, Ginelli's algorithm can be defined for Grassmannians (see Section 3.1) and does not require an orthonormalization procedure.

[3]We remind the reader of the index notation i_j introduced in the beginning of Section 2.2, which counts indices with respect to degeneracies of the Lyapunov spectrum.

requires a tuple $(x_1, \ldots, x_k) \in H^k$ and an upper triangular matrix $R \in \mathbb{R}^{k\times k}_{ru}$ as inputs. "Almost every" with respect to the tuple is understood in terms of prevalence, whereas "almost every" with respect to the matrix is meant in the usual Lebesgue sense. If H is finite-dimensional, the two previous notions coincide and we get the same statement as in Chapter 4 for discrete time.

Similar to the finite-dimensional case Theorem 5.3.1 tells us that, generically, output vectors of Ginelli's algorithm span subspaces that are exponentially close to the Oseledets spaces. Hence, the algorithm approximates CLVs. To get a good approximation, it is necessary to increase both n_1 and n_2. In other words, the algorithm needs sufficient data along the past and the future of the trajectory. Moreover, Theorem 5.3.1 reveals that the speed of convergence to the i^{th} Oseledets space $Y_i(\omega)$ is at least exponentially fast in proportion to the spectral gap between λ_i and neighboring LEs.

5.3.2 Forward-time estimates

The forward- and backward-time estimates are proved for general Banach spaces $(X, \|.\|)$. Our first result investigates how certain subspaces evolve in the presence of an equivariant splitting under a given map. The estimates consist of terms that are well-understood when the splitting is the Oseledets splitting. As before, we write $B \subset X$ for the unit ball and $S \subset X$ for the unit sphere in X.

Lemma 5.3.2
Let $(Y, V), (Y', V') \in \mathrm{Comp}_k(X)$ be two pairs of closed complemented subspaces. Assume we have a bounded linear map $\mathcal{L} \in L(X)$ respecting the splittings, i.e., $\mathcal{L}Y \subset Y'$ and $\mathcal{L}V \subset V'$, such that $\ker \mathcal{L} \subset V$.

If $W \in \mathcal{G}_k(X)$ is a complement to V such that the degree of transversality satisfies

$$\inf_{w\in W\cap S} d(w, V) \geq 2\|\Pi_{V||Y}\| \frac{\|\mathcal{L}|_V\|}{\inf_{y\in Y\cap S} \|\mathcal{L}y\|}, \tag{5.7}$$

then

$$\sup_{w'\in \mathcal{L}W\cap B} d(w', Y'\cap B) \leq 4 \frac{\|\Pi_{V||Y}\|}{\inf_{w\in W\cap S} d(w, V)} \frac{\|\mathcal{L}|_V\|}{\inf_{y\in Y\cap S} \|\mathcal{L}y\|}. \tag{5.8}$$

Proof. If $\mathcal{L}|_V = 0$, then $\ker \mathcal{L} = V$. Thus, $\mathcal{L}$ restricts to an isomorphism between any complement W to V and Y'. In this case the claim is trivially satisfies.

Now, assume $\mathcal{L}|_V \neq 0$. Let W be a complement as in the claim. For $w \in W \cap S$, it holds

$$\|\mathcal{L}\Pi_{V||Y} w\| \leq \|\mathcal{L}|_V\| \, \|\Pi_{V||Y}\|$$

and

$$\begin{aligned}
\|\mathcal{L}\Pi_{Y||V} w\| &\geq \inf_{y\in Y\cap S} \|\mathcal{L}y\| \, \|\Pi_{Y||V} w\| \\
&= \inf_{y\in Y\cap S} \|\mathcal{L}y\| \, \|w - \Pi_{V||Y} w\| \\
&\geq \inf_{y\in Y\cap S} \|\mathcal{L}y\| \, d(w, V) \\
&\geq 2\|\Pi_{V||Y}\| \, \|\mathcal{L}|_V\| > 0.
\end{aligned}$$

Combining both estimates, we get

$$\frac{\|\mathcal{L}\Pi_{V||Y}w\|}{\|\mathcal{L}\Pi_{Y||V}w\|} \leq \frac{1}{2}. \tag{5.9}$$

To derive Eq. (5.8) it is enough to estimate $d(\mathcal{L}w/\|\mathcal{L}w\|, Y' \cap B)$ for $w \in W \cap S$. Write $w = y + v$ according to the decomposition $X = Y \oplus V$. We have

$$\begin{aligned} d\left(\frac{\mathcal{L}w}{\|\mathcal{L}w\|}, Y' \cap B\right) &\leq \left\|\frac{\mathcal{L}w}{\|\mathcal{L}w\|} - \frac{\mathcal{L}y}{\|\mathcal{L}y\|}\right\| \\ &= \left\|\frac{\mathcal{L}v}{\|\mathcal{L}w\|} - \left(\frac{1}{\|\mathcal{L}y\|} - \frac{1}{\|\mathcal{L}w\|}\right)\mathcal{L}y\right\| \\ &\leq \frac{\|\mathcal{L}v\|}{\|\mathcal{L}w\|} + \left|1 - \frac{\|\mathcal{L}y\|}{\|\mathcal{L}w\|}\right|. \end{aligned}$$

Since $y \neq 0$ and by Eq. (5.9), we estimate the first term as follows:

$$\frac{\|\mathcal{L}v\|}{\|\mathcal{L}w\|} \leq \frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\| - \|\mathcal{L}v\|} = \frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\|}\left(1 - \frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\|}\right)^{-1} \leq 2\frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\|}.$$

For the other term, we distinguish between two cases. If $\|\mathcal{L}y\|/\|\mathcal{L}w\| \leq 1$, then

$$1 - \frac{\|\mathcal{L}y\|}{\|\mathcal{L}w\|} \leq 1 - \frac{\|\mathcal{L}y\|}{\|\mathcal{L}y\| + \|\mathcal{L}v\|} = \frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\| + \|\mathcal{L}v\|} \leq \frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\|}.$$

If $\|\mathcal{L}y\|/\|\mathcal{L}w\| \geq 1$, then

$$\frac{\|\mathcal{L}y\|}{\|\mathcal{L}w\|} - 1 = \frac{\|\mathcal{L}y\| - \|\mathcal{L}w\|}{\|\mathcal{L}w\|} \leq \frac{\|\mathcal{L}v\|}{\|\mathcal{L}w\|} \leq 2\frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\|}.$$

In total we get

$$d\left(\frac{\mathcal{L}w}{\|\mathcal{L}w\|}, Y' \cap B\right) \leq 4\frac{\|\mathcal{L}v\|}{\|\mathcal{L}y\|}.$$

Since $v = \Pi_{V||Y}w$ and $y = \Pi_{Y||V}w$, the claim follows from the estimates in the beginning. ☐

Corollary 5.3.3
In the setting of Lemma 5.3.2 it holds

$$\|\Pi_{V'||Y'}|_{\mathcal{L}W}\| \leq 2\,\frac{\|\Pi_{V||Y}\|}{\inf_{w \in W \cap S} d(w, V)}\,\frac{\|\mathcal{L}|_V\|}{\inf_{y \in Y \cap S}\|\mathcal{L}y\|}.$$

Proof. The corollary follows from

$$\|\Pi_{V'||Y'}|_{\mathcal{L}W}\| = \sup_{w \in W \cap S}\left\|\Pi_{V'||Y'}\frac{\mathcal{L}w}{\|\mathcal{L}w\|}\right\| = \sup_{w \in W \cap S}\frac{\|\mathcal{L}\Pi_{V||Y}w\|}{\|\mathcal{L}w\|}$$

and from the estimate of $\|\mathcal{L}v\|/\|\mathcal{L}w\|$ in the proof of Lemma 5.3.2. ☐

Next, we derive two lemmata that handle sequences of maps acting on equivariant splittings with different asymptotic growth rates. The first lemma is concerned with propagation from present to future, whereas the second lemma treats propagation from past to present.

Lemma 5.3.4
Let $(Y, V) \in \mathrm{Comp}_k(X)$ *and* $(Y(n), V(n)) \in \mathrm{Comp}_k(X)$ *for* $n \in \mathbb{N}$. *Assume we have bounded linear maps* $\mathcal{L}(n) \in L(X)$ *respecting the splittings, i.e.,* $\mathcal{L}(n)Y \subset Y(n)$ *and* $\mathcal{L}(n)V \subset V(n)$*, such that* $\ker \mathcal{L}(n) \subset V$. *Furthermore, assume there are numbers* $\infty > \lambda_Y > \lambda_V \geq -\infty$ *such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}(n)|_V\| \leq \lambda_V$$

and

$$\liminf_{n\to\infty} \inf_{y\in Y\cap S} \frac{1}{n} \log \|\mathcal{L}(n)y\| \geq \lambda_Y.$$

Then, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log d_{\mathcal{G}}(\mathcal{L}(n)W, Y(n)) \leq -|\lambda_Y - \lambda_V|$$

for any complement W *to* V.

Proof. According to the assumptions we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{\|\mathcal{L}(n)|_V\|}{\inf_{y\in Y\cap S} \|\mathcal{L}(n)y\|} \leq -|\lambda_Y - \lambda_V| < 0,$$

i.e., the quotient $\|\mathcal{L}(n)|_V\|/(\inf_{y\in Y\cap S} \|\mathcal{L}(n)y\|)$ decays exponentially fast with n. Thus, for any complement W to V, there is $N > 0$ such that Eq. (5.7) of Lemma 5.3.2 is satisfied for all $n \geq N$. Applying the lemma, we get

$$\limsup_{n\to\infty} \frac{1}{n} \log \sup_{w'\in\mathcal{L}(n)W\cap B} d(w', Y(n) \cap B) \leq -|\lambda_Y - \lambda_V|.$$

The claim follows from Lemma 5.1.2. □

Lemma 5.3.4 implies that complements to spaces of the Oseledets filtration will align with Oseledets spaces asymptotically (at an exponential speed). Moreover, the lemma tells us that any two complements to V will align asymptotically if they have a uniformly higher growth rate than V. Interestingly, we do not need the existence of an Oseledets splitting. In fact, the lemma may be applied to systems with a possibly non-invertible base (e.g, see [3, theorem 2] or [7]).

Lemma 5.3.5
Let $(Y, V) \in \mathrm{Comp}_k(X)$ *and* $(Y(-n), V(-n)) \in \mathrm{Comp}_k(X)$ *for* $n \in \mathbb{N}$. *Assume we have bounded linear maps* $\mathcal{L}(-n) \in L(X)$ *respecting the splittings, i.e.,* $\mathcal{L}(-n)Y(-n) \subset Y$ *and* $\mathcal{L}(-n)V(-n) \subset V$*, such that* $\ker \mathcal{L}(-n) \subset V(-n)$. *Furthermore, assume that*

$$\lim_{n\to\infty} \frac{1}{n} \log \|\Pi_{V(-n)\|Y(-n)}\| = 0$$

and that there are numbers $\infty > \lambda_Y > \lambda_V \geq -\infty$ *such that*

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}(-n)|_{V(-n)}\| \leq \lambda_V$$

and

$$\liminf_{n\to\infty} \inf_{y\in Y(-n)\cap S} \frac{1}{n} \log \|\mathcal{L}(-n)y\| \geq \lambda_Y.$$

Then, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log d_{\mathcal{G}}(\mathcal{L}(-n)W, Y) \leq -|\lambda_Y - \lambda_V|$$

for any well-separating common complement W for $(V(-n))_{n\in\mathbb{N}}$.

Proof. As in Lemma 5.3.4, we see that

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{\|\mathcal{L}(-n)|_{V(-n)}\|}{\inf_{y\in Y(-n)\cap S} \|\mathcal{L}(-n)y\|} \leq -|\lambda_Y - \lambda_V|.$$

By our assumption on the associated projections we get

$$\limsup_{n\to\infty} \frac{1}{n} \log\left(2\|\Pi_{V(-n)\|Y(-n)}\| \frac{\|\mathcal{L}(-n)|_{V(-n)}\|}{\inf_{y\in Y(-n)\cap S} \|\mathcal{L}(-n)y\|}\right) \leq -|\lambda_Y - \lambda_V| < 0.$$

In particular, by Definition 5.2.1 any well-separating common complement for $(V(-n))_{n\in\mathbb{N}}$ fulfills Eq. (5.7) for n large enough as the degree of transversality decays only subexponentially. The claim may be derived as in the proof of Lemma 5.3.4. □

Corollary 5.3.6
In the setting of Lemma 5.3.5, we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\Pi_{V\|Y}|_{\mathcal{L}(-n)W}\| \leq -|\lambda_Y - \lambda_V|$$

for any well-separating common complement W for $(V(-n))_{n\in\mathbb{N}}$.

Proof. Since Lemma 5.3.2 and Corollary 5.3.3 give the same estimate up to a factor of 2, the proof of Corollary 5.3.6 is the same as for Lemma 5.3.5. □

The following theorem gives us convergence of certain subspaces of Banach spaces to the sum of the first Oseledets spaces in forward-time:

Theorem 5.3.7
Let $\mathcal{R}$ be as in Theorem 2.3.2 and $\omega \in \Omega$ such that the Oseledets splitting exists. Write $\lambda_{p+1} := \kappa^*$ and fix some finite $i \leq p$.

If Eqs. (2.6) *to* (2.8) *hold*[4], *then*

$$\limsup_{n\to\infty} \frac{1}{n} \log d_{\mathcal{G}}\Big(\mathcal{L}_{\omega}^{(n)} W, Y_1(\sigma_n\omega) \oplus \cdots \oplus Y_i(\sigma_n\omega)\Big) \leq -|\lambda_i - \lambda_{i+1}|$$

for any complement W to $V_{i+1}(\omega)$.

If Eqs. (2.9) *to* (2.11) *hold, then*

$$\limsup_{n\to\infty} \frac{1}{n} \log d_{\mathcal{G}}\Big(\mathcal{L}_{\sigma_{-n}\omega}^{(n)} W, Y_1(\omega) \oplus \cdots \oplus Y_i(\omega)\Big) \leq -|\lambda_i - \lambda_{i+1}|$$

for any well-separating common complement W for $(V_{i+1}(\sigma_{-n}\omega))_{n\in\mathbb{N}}$.

[4] We remark that Eqs. (2.6) to (2.8) and Eqs. (2.9) to (2.11) hold for $\mathbb{P}$-almost every $\omega \in \Omega$.

Proof. The proof is a direct application of Lemma 5.3.4 and Lemma 5.3.5 to the splittings $(Y, V) = (Y_1(\omega) \oplus \cdots \oplus Y_i(\omega), V_{i+1}(\omega))$, $(Y(n), V(n)) := (Y_1(\sigma_n\omega) \oplus \cdots \oplus Y_i(\sigma_n\omega), V_{i+1}(\sigma_n\omega))$ for $n \in \mathbb{Z}$, and to the maps $\mathcal{L}(n) := \mathcal{L}_\omega^{(n)}$ and $\mathcal{L}(-n) := \mathcal{L}_{\sigma_{-n}\omega}^{(n)}$ for $n \in \mathbb{N}$. □

In view of Theorem 5.2.15, Theorem 5.3.7 for Hilbert spaces implies that we can compute the sum of the first Oseledets spaces $Y_1 \oplus \cdots \oplus Y_i$ at ω or asymptotically by pushing forwards a set of $d_1 + \cdots + d_i$ randomly chosen vectors. The convergence is exponentially fast with a rate given by the spectral gap between the consecutive LEs λ_i and λ_{i+1}.

5.3.3 Backward-time estimates

In this subsection we investigate backward propagation of certain subspaces. Since we did not assume a cocycle with invertible action, we cannot simply apply our results about forward propagation to a time-reversed system as it is done in the previous chapter. Instead, we derive new estimates for forward propagation to deduce properties for backward propagation.

Lemma 5.3.8
Let $(Y_1, V_1) \in \mathrm{Comp}_{k_1}(X)$ and $(Y_2, V_2) \in \mathrm{Comp}_{k_2}(V_1)$, so that $X = Y_1 \oplus V_1$ and $V_1 = Y_2 \oplus V_2$. Moreover, let W_i be a complement to V_i in X for $i = 1, 2$ such that $W_1 \subset W_2$. Assume we have a map $\mathcal{L} \in L(X)$ with $\ker \mathcal{L} \subset V_2$.

If $\tilde{W} \in \mathcal{G}_{k_2}(W_2)$ is a complement to W_1 in W_2 and if $\tilde{w} \in \tilde{W} \cap S$ satisfies

$$d(\tilde{w}, Y_2) \geq (2\|\Pi_{V_1||W_1}\| + \|\Pi_{V_1||Y_1}\| \, \|\Pi_{W_1||V_1}\|) \frac{\|\mathcal{L}|_{V_1}\|}{\inf_{y \in Y_1 \cap S} \|\mathcal{L}y\|} + \|\Pi_{V_2||Y_1 \oplus Y_2}|_{W_2}\|, \tag{5.10}$$

then

$$d\left(\frac{\mathcal{L}\tilde{w}}{\|\mathcal{L}\tilde{w}\|}, \mathcal{L}W_1\right) \leq \frac{2\|\mathcal{L}|_{V_1}\| \, \|\Pi_{V_1||W_1}\|}{(\inf_{y \in Y_1 \cap S} \|\mathcal{L}y\|)\Big(d(\tilde{w}, Y_2) - \|\Pi_{V_2||Y_1 \oplus Y_2}|_{W_2}\|\Big) - \|\mathcal{L}|_{V_1}\| \, \|\Pi_{V_1||Y_1}\| \, \|\Pi_{W_1||V_1}\|}. \tag{5.11}$$

Proof. Since $V_1 = Y_2 \oplus V_2$ is a splitting with $Y_2 \neq \{0\}$ and $\ker \mathcal{L} \subset V_2$, it holds $\mathcal{L}|_{V_1} \neq 0$.

Let $\tilde{w} \in \tilde{W} \cap S$ be as in the claim, so that Eq. (5.10) is satisfied. We estimate

$$\|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\| \leq \|\mathcal{L}|_{V_1}\| \, \|\Pi_{V_1||W_1}\|$$

and

$$\begin{aligned}
\|\mathcal{L}\Pi_{W_1||V_1}\tilde{w}\| &= \|\mathcal{L}(\Pi_{Y_1||V_1} + \Pi_{V_1||Y_1})\Pi_{W_1||V_1}\tilde{w}\| \\
&\geq \|\mathcal{L}\Pi_{Y_1||V_1}\Pi_{W_1||V_1}\tilde{w}\| - \|\mathcal{L}\Pi_{V_1||Y_1}\Pi_{W_1||V_1}\tilde{w}\| \\
&\geq \Big(\inf_{y \in Y_1 \cap S} \|\mathcal{L}y\|\Big)\|\Pi_{Y_1||V_1}\Pi_{W_1||V_1}\tilde{w}\| - \|\mathcal{L}|_{V_1}\| \, \|\Pi_{V_1||Y_1}\| \, \|\Pi_{W_1||V_1}\|.
\end{aligned}$$

The term with two consecutive projections applied to $\tilde{w}$ can be estimates further via

$$\begin{aligned}
\|\Pi_{Y_1||V_1}\Pi_{W_1||V_1}\tilde{w}\| &= \|\Pi_{W_1||V_1}\tilde{w} - \Pi_{V_1||Y_1}\Pi_{W_1||V_1}\tilde{w}\| \\
&= \|\tilde{w} - \Pi_{V_1||W_1}\tilde{w} - \Pi_{V_1||Y_1}\Pi_{W_1||V_1}\tilde{w}\| \\
&= \|\tilde{w} - \Pi_{V_1||Y_1}(\Pi_{V_1||W_1} + \Pi_{W_1||V_1})\tilde{w}\| \\
&= \|\tilde{w} - \Pi_{V_1||Y_1}\tilde{w}\| \\
&= \|\tilde{w} - \Pi_{Y_2||V_2}\Pi_{V_1||Y_1}\tilde{w} - \Pi_{V_2||Y_2}\Pi_{V_1||Y_1}\tilde{w}\| \\
&\geq \|\tilde{w} - \Pi_{Y_2||V_2}\Pi_{V_1||Y_1}\tilde{w}\| - \|\Pi_{V_2||Y_2}\Pi_{V_1||Y_1}\tilde{w}\| \\
&\geq d(\tilde{w}, Y_2) - \|\Pi_{V_2||Y_2}\Pi_{V_1||Y_1}\tilde{w}\| \\
&= d(\tilde{w}, Y_2) - \|\Pi_{V_2||Y_1\oplus Y_2}\tilde{w}\| \\
&\geq d(\tilde{w}, Y_2) - \|\Pi_{V_2||Y_1\oplus Y_2}|_{W_2}\|.
\end{aligned}$$

Note that $\Pi_{V_2||Y_2}$ and $\Pi_{Y_2||V_2}$ are projections defined on V_1. By Eq. (5.10) we have

$$\|\Pi_{Y_1||V_1}\Pi_{W_1||V_1}\tilde{w}\| \geq (2\|\Pi_{V_1||W_1}\| + \|\Pi_{V_1||Y_1}\|\,\|\Pi_{W_1||V_1}\|)\frac{\|\mathcal{L}|_{V_1}\|}{\inf_{y\in Y_1\cap S}\|\mathcal{L}y\|}.$$

Hence, we get

$$\|\mathcal{L}\Pi_{W_1||V_1}\tilde{w}\| \geq 2\|\Pi_{V_1||W_1}\|\,\|\mathcal{L}|_{V_1}\| > 0$$

and

$$\frac{\|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\|}{\|\mathcal{L}\Pi_{W_1||V_1}\tilde{w}\|} \leq \frac{1}{2}.$$

Finally, it holds

$$\begin{aligned}
d\left(\frac{\mathcal{L}\tilde{w}}{\|\mathcal{L}\tilde{w}\|}, \mathcal{L}W_1\right) &\leq \left\|\frac{\mathcal{L}\tilde{w}}{\|\mathcal{L}\tilde{w}\|} - \frac{\mathcal{L}\Pi_{W_1||V_1}\tilde{w}}{\|\mathcal{L}\tilde{w}\|}\right\| \\
&= \frac{\|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\|}{\|\mathcal{L}\tilde{w}\|} \\
&\leq \frac{\|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\|}{\|\mathcal{L}\Pi_{W_1||V_1}\tilde{w}\| - \|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\|} \\
&= \frac{\|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\|}{\|\mathcal{L}\Pi_{W_1||V_1}\tilde{w}\|}\left(1 - \frac{\|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\|}{\|\mathcal{L}\Pi_{W_1||V_1}\tilde{w}\|}\right)^{-1} \\
&\leq 2\frac{\|\mathcal{L}\Pi_{V_1||W_1}\tilde{w}\|}{\|\mathcal{L}\Pi_{W_1||V_1}\tilde{w}\|}.
\end{aligned}$$

Estimating the numerator and the denominator as in the beginning of the proof, we arrive at Eq. (5.11). □

Corollary 5.3.9
Let Y_i, V_i, W_i for $i = 1, 2$ and $\mathcal{L}$ be as in Lemma 5.3.8.
If $\tilde{W} \subset W_2$ is a complement to W_1 in W_2 satisfying

$$\inf_{\tilde{w}'\in\mathcal{L}\tilde{W}\cap S} d(\tilde{w}', \mathcal{L}W_1) \geq \delta$$

for some $0 < \delta \leq 1$, then

$$\begin{aligned}
&\sup_{\tilde{w}\in\tilde{W}\cap B} d(\tilde{w}, Y_2 \cap B) \\
&\quad\leq 2\left(\frac{2}{\delta}\|\Pi_{V_1||W_1}\| + \|\Pi_{V_1||Y_1}\|\,\|\Pi_{W_1||V_1}\|\right)\frac{\|\mathcal{L}|_{V_1}\|}{\inf_{y\in Y_1\cap S}\|\mathcal{L}y\|} + 2\|\Pi_{V_2||Y_1\oplus Y_2}|_{W_2}\|.
\end{aligned}$$

Proof. Assume $\tilde{w} \in \tilde{W} \cap S$ fulfills

$$d(\tilde{w}, Y_2) > \left(\frac{2}{\delta}\|\Pi_{V_1||W_1}\| + \|\Pi_{V_1||Y_1}\|\,\|\Pi_{W_1||V_1}\|\right)\frac{\|\mathcal{L}|_{V_1}\|}{\inf_{y\in Y_1\cap S}\|\mathcal{L}y\|} + \|\Pi_{V_2||Y_1\oplus Y_2}|_{W_2}\|,$$

then by Lemma 5.3.8

$$\begin{aligned}\delta &\leq d\left(\frac{\mathcal{L}\tilde{w}}{\|\mathcal{L}\tilde{w}\|}, \mathcal{L}W_1\right)\\ &\leq \frac{2\|\mathcal{L}|_{V_1}\|\,\|\Pi_{V_1||W_1}\|}{(\inf_{y\in Y_1\cap S}\|\mathcal{L}y\|)\Big(d(\tilde{w}, Y_2) - \|\Pi_{V_2||Y_1\oplus Y_2}|_{W_2}\|\Big) - \|\mathcal{L}|_{V_1}\|\,\|\Pi_{V_1||Y_1}\|\,\|\Pi_{W_1||V_1}\|}.\end{aligned}$$

However, the former would be strictly smaller than δ by our assumption on $d(\tilde{w}, Y_2)$. Thus, we must have

$$\begin{aligned}&\sup_{\tilde{w}\in\tilde{W}\cap S} d(\tilde{w}, Y_2)\\ &\quad\leq \left(\frac{2}{\delta}\|\Pi_{V_1||W_1}\| + \|\Pi_{V_1||Y_1}\|\,\|\Pi_{W_1||V_1}\|\right)\frac{\|\mathcal{L}|_{V_1}\|}{\inf_{y\in Y_1\cap S}\|\mathcal{L}y\|} + \|\Pi_{V_2||Y_1\oplus Y_2}|_{W_2}\|.\end{aligned}$$

The claim follows from Eq. (5.1). □

From Corollary 5.3.9 we can derive an upper bound of the distance between $\tilde{W}$ and Y_2 from a lower bound of the degree of transversality of $(\mathcal{L}\tilde{W}, \mathcal{L}W_1)$ in $\mathcal{L}W_2$. Hence, the corollary describes backward propagation.

Next, we use the spaces W_1, W_2 to connect estimates from Section 5.3.2 to backward propagation, ultimately giving us an understanding of Ginelli's algorithm at the level of maps:

Lemma 5.3.10
Let $(Y_1, V_1) \in \mathrm{Comp}_{k_1}(X)$, $(Y_2, V_2) \in \mathrm{Comp}_{k_2}(V_1)$, and $\infty > \lambda_{Y_1} > \lambda_{V_1} = \lambda_{Y_2} > \lambda_{V_2} \geq -\infty$.

For the past data, let $(Y_1(-n), V_1(-n)) \in \mathrm{Comp}_{k_1}(X)$ and $(Y_2(-n), V_2(-n)) \in \mathrm{Comp}_{k_2}(V_1(-n))$ for $n \in \mathbb{N}$. Assume we have bounded linear maps $\mathcal{L}(-n) \in L(X)$ respecting the splittings, i.e., $\mathcal{L}(-n)Y_i(-n) \subset Y_i$ for $i = 1, 2$ and $\mathcal{L}(-n)\ V_2(-n) \subset V_2$, such that $\ker\mathcal{L}(-n) \subset V_2(-n)$ for $n \in \mathbb{N}$. Moreover, assume that

1. $\lim_{n\to\infty}(1/n)\log\|\Pi_{V_1(-n)||Y_1(-n)}\| = 0$,
2. $\lim_{n\to\infty}(1/n)\log\|\Pi_{V_2(-n)||Y_1(-n)\oplus Y_2(-n)}\| = 0$,
3. $\limsup_{n\to\infty}(1/n)\log\|\mathcal{L}(-n)|_{V_i(-n)}\| \leq \lambda_{V_i}$ *for* $i = 1, 2$,
4. $\liminf_{n\to\infty}\inf_{y\in Y_1(-n)\cap S}(1/n)\log\|\mathcal{L}(-n)y\| \geq \lambda_{Y_1}$, *and*
5. $\liminf_{n\to\infty}\inf_{y\in Y_1(-n)\oplus Y_2(-n)\cap S}(1/n)\log\|\mathcal{L}(-n)y\| \geq \lambda_{Y_2}$.

For the future data, let $(Y_1(n), V_1(n)) \in \mathrm{Comp}_{k_1}(X)$ and $(Y_2(n), V_2(n)) \in \mathrm{Comp}_{k_2}(V_1(n))$ for $n \in \mathbb{N}$. Assume we have bounded linear maps $\mathcal{L}(n) \in L(X)$ respecting the splittings, i.e., $\mathcal{L}(n)Y_i \subset Y_i(n)$ for $i = 1, 2$ and $\mathcal{L}(n)V_2 \subset V_2(n)$, such that $\ker\mathcal{L}(n) \subset V_2$ for $n \in \mathbb{N}$. Moreover, assume that

6. $\limsup_{n\to\infty}(1/n)\log\|\mathcal{L}(n)|_{V_1}\| \leq \lambda_{V_1}$ *and*

7. $\liminf_{n\to\infty} \inf_{y\in Y_1\cap S} (1/n) \log \|\mathcal{L}(n)y\| \geq \lambda_{Y_1}$.

Let W_i be a well-separating common complement for $(V_i(-n))_{n\in\mathbb{N}}$ for $i = 1, 2$ such that $W_1 \subset W_2$. If $(\tilde{W}(n_1, n_2))_{n_1,n_2\in\mathbb{N}}$ is a family of subspaces such that $\mathcal{L}(n_2)\mathcal{L}(-n_1)W_1 \oplus \tilde{W}(n_1, n_2) = \mathcal{L}(n_2)\mathcal{L}(-n_1)W_2$, and if

$$\inf_{\tilde{w}\in\tilde{W}(n_1,n_2)\cap S} d(\tilde{w}, \mathcal{L}(n_2)\mathcal{L}(-n_1)W_1) \geq \delta \tag{5.12}$$

for some constant $0 < \delta \leq 1$, then

$$\begin{aligned}\limsup_{N\to\infty} \sup_{n_1,n_2\geq N} \frac{1}{\min(n_1, n_2)} \log d_{\mathcal{G}}\Big(\big(\mathcal{L}(n_2)|_{\mathcal{L}(-n_1)W_2}\big)^{-1}\tilde{W}(n_1, n_2), Y_2\Big)& \\ \leq -\min(|\lambda_{Y_2} - \lambda_{Y_1}|, |\lambda_{Y_2} - \lambda_{V_2}|).& \end{aligned} \tag{5.13}$$

Proof. Let W_1 and W_2 be as in the claim. We apply Lemma 5.3.5 to $(Y, V) = (Y_1, V_1)$ for $W = W_1$ and to $(Y, V) = (Y_1 \oplus Y_2, V_2)$ for $W = W_2$ with their respective spaces and mappings at $-n$. It follows that

$$\limsup_{n\to\infty} \frac{1}{n} \log d_{\mathcal{G}}(\mathcal{L}(-n)W_1, Y_1) \leq -|\lambda_{Y_1} - \lambda_{V_1}|$$

and

$$\limsup_{n\to\infty} \frac{1}{n} \log d_{\mathcal{G}}(\mathcal{L}(-n)W_2, Y_1 \oplus Y_2) \leq -|\lambda_{Y_2} - \lambda_{V_2}|.$$

Thus, we have good approximations of Y_1 and $Y_1 \oplus Y_2$ from the past data. Moreover, by Corollary 5.3.6 we have

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\Pi_{V_2||Y_1\oplus Y_2}|_{\mathcal{L}(-n)W_2}\| \leq -|\lambda_{Y_2} - \lambda_{V_2}|.$$

Since $\mathcal{L}(-n)W_1$ converges to Y_1, the projections $\Pi_{\mathcal{L}(-n)W_1||V_1}$ converge to $\Pi_{Y_1||V_1}$ by Lemma 5.1.3. In particular, $\|\Pi_{\mathcal{L}(-n)W_1||V_1}\|$ and $\|\Pi_{V_1||\mathcal{L}(-n)W_1}\|$ are bounded from above by a constant independent of n.

The growth rate assumptions on future data imply

$$\limsup_{n\to\infty} \frac{1}{n} \log \frac{\|\mathcal{L}(n)|_{V_1}\|}{\inf_{y\in Y_1\cap S} \|\mathcal{L}(n)y\|} \leq -|\lambda_{Y_1} - \lambda_{V_1}|.$$

Now, apply Corollary 5.3.9 to (Y_1, V_1), (Y_2, V_2), the two complements $\mathcal{L}(-n_1)W_1$ to V_1 and $\mathcal{L}(-n_1)W_2$ to V_2, $\mathcal{L} = \mathcal{L}(n_2)$, and $\tilde{W} = \big(\mathcal{L}(n_2)|_{\mathcal{L}(-n_1)W_2}\big)^{-1}\tilde{W}(n_1, n_2)$. We get

$$\begin{aligned}&\sup_{\tilde{w}\in\left(\mathcal{L}(n_2)|_{\mathcal{L}(-n_1)W_2}\right)^{-1}\tilde{W}(n_1,n_2)\cap B} d(\tilde{w}, Y_2 \cap B) \\ &\quad\leq 2\Big(\frac{2}{\delta}\|\Pi_{V_1||\mathcal{L}(-n_1)W_1}\| + \|\Pi_{V_1||Y_1}\|\,\|\Pi_{\mathcal{L}(-n_1)W_1||V_1}\|\Big)\frac{\|\mathcal{L}(n_2)|_{V_1}\|}{\inf_{y\in Y_1\cap S} \|\mathcal{L}(n_2)y\|} \\ &\qquad + 2\|\Pi_{V_2||Y_1\oplus Y_2}|_{\mathcal{L}(-n_1)W_2}\|.\end{aligned} \tag{5.14}$$

In view of Lemma 5.1.2, all that remains to prove Eq. (5.13) is to insert respective asymptotics into the terms of Eq. (5.14). Indeed, the terms inside the large bracket are bounded from above by a constant, and the other terms can be estimated as above. Now, Eq. (5.13) is an easy application of Proposition 4.1.4. □

Lemma 5.3.10 provides an appropriate tool to study convergence of Ginelli's algorithm in infinite dimensions. Since the algorithm initiates vectors for backward propagation inside spaces from the forward propagation, which vary with the chosen runtime, the domain for initial vectors is not constant. Hence, $\tilde{W}$ varies with n_1 and n_2. This poses a problem when talking about convergence with respect to initial conditions. One way to solve this problem is to express initial vectors of the backward propagation in terms of runtime-independent coefficients. If $X = H$ is a Hilbert space, then we may identify an orthonormal basis of $\mathcal{L}(n_1)\mathcal{L}(-n_2)W_2$ with the standard basis of $(\mathbb{R}^{k_1+k_2}, \|.\|_2)$ as it is done in our definition of Ginelli's algorithm on Hilbert spaces (see Definition 3.1.1). The identification defines an isometry leaving distances and angles invariant. In particular, we may represent $\tilde{W}$ in terms of runtime-independent coefficients and check Eq. (5.12) on the coefficient space.

5.3.4 Proof of theorem

We now combine our tools to prove Theorem 5.3.1:

proof of Theorem 5.3.1. Fix an element ω of the subset $\Omega' \subset \Omega$ of full $\mathbb{P}$-measure on which the Oseledets splitting is defined and on which Eqs. (2.6) to (2.11) hold. We show convergence of Ginelli's algorithm at ω for almost every input.

Let $\mathcal{F}_i \subset H^{d_1+\dots+d_i}$ be the subset of all tuples inducing well-separating common complements for $(V_{i+1}(\sigma_{-n}\omega))_{n\in\mathbb{N}}$ for $i = 1, \dots, l$. Then, the set

$$\mathcal{F} := \left(\mathcal{F}_1 \times H^{d_2+\dots+d_l}\right) \cap \left(\mathcal{F}_2 \times H^{d_3+\dots+d_l}\right) \cap \dots \cap \mathcal{F}_l \subset H^k$$

consists of tuples $(x_{1_1}, \dots, x_{l_{d_l}})$ such that $\operatorname{span}(x_{1_1}, \dots, x_{i_{d_i}})$ is a well-separating common complement for $(V_{i+1}(\sigma_{-n}\omega))_{n\in\mathbb{N}}$ for each $i = 1, \dots, l$. In particular, since products and intersections of prevalent sets are prevalent, Theorem 5.2.15 implies that $\mathcal{F}$ is prevalent. We use elements of $\mathcal{F}$ as initial vectors for the forward phase of Ginelli's algorithm.

Let $\mathcal{B} \subset \mathbb{R}^{k\times k}_{ru}$ be the subset of upper triangular matrices with non-zero diagonal elements, i.e., the subset of invertible upper triangular matrices. $\mathcal{B}$ has full Lebesgue measure and is used in our proof for initial vectors for the backward phase of Ginelli's algorithm.

Now, let $((x_1, \dots, x_k), R) \in \mathcal{F} \times \mathcal{B}$ be an input for Ginelli's algorithm. According to Theorem 5.3.7 the first set of vectors $(x_{1_1}, \dots, x_{1_{d_1}})$ gives an approximation of $Y_1(\omega)$ via the first step of Ginelli's algorithm. The remaining steps of Ginelli's algorithm do not change this approximation. In fact, the first set of output vectors $(G^{n_1,n_2}_{\omega,k})_{1_j}$ for $j = 1, \dots, d_1$ at $((x_1, \dots, x_k), R)$ spans the same space as $\left(\mathcal{L}^{(n_1)}_{\sigma_{-n_1}\omega} x_{1_1}, \dots, \mathcal{L}^{(n_1)}_{\sigma_{-n_1}\omega} x_{1_{d_1}}\right)$. Thus, we have

$$\begin{aligned}
&\limsup_{N\to\infty} \sup_{n_1,n_2\geq N} \frac{1}{\min(n_1,n_2)} \log d_{\mathcal{G}}\Big(\operatorname{span}\Big\{ \left(G^{n_1,n_2}_{\omega,k}\right)_{1_j} \;\Big|\; j = 1, \dots, d_1 \Big\}, Y_1(\omega)\Big) \\
&\quad \leq -|\lambda_1 - \lambda_2| \\
&\quad = -\min(|\lambda_1 - \lambda_0|, |\lambda_1 - \lambda_2|)
\end{aligned}$$

at $((x_1, \dots, x_k), R)$.

Convergence of the remaining spaces is due to Lemma 5.3.10. Fix some $1 < i \leq l$. We set $Y_1 = Y_1(\omega) \oplus \cdots \oplus Y_{i-1}(\omega)$, $V_1 = V_i(\omega)$, $Y_2 = Y_i(\omega)$, $V_2 = V_{i+1}(\omega)$, $\mathcal{L}(-n) = \mathcal{L}^{(n)}_{\sigma_{-n}\omega}$, $\mathcal{L}(n) = \mathcal{L}^{(n)}_{\omega}$, and spaces $Y_j(\pm n)$ and $V_j(\pm n)$ for $j = 1, 2$ accordingly. The growth rates in Lemma 5.3.10 are given by uniform bounds obtained from Theorem 2.3.2 and its proof. Furthermore, let $W_1 = \operatorname{span}(x_{1_1}, \ldots, x_{(i-1)_{d_{i-1}}})$ and $W_2 = \operatorname{span}(x_{1_1}, \ldots, x_{i_{d_i}})$ be the well-separating common complements, which approximate Y_1 and Y_2 in the first step of Ginelli's algorithm. The family of spaces $(\tilde{W}(n_1, n_2))_{n_1, n_2 \in \mathbb{N}}$ is given by $\operatorname{span}(y^1_{i_1}, \ldots, y^1_{i_{d_i}})$ via vectors of the fourth step of the algorithm. Indeed, the i_1^{th} to $i_{d_i}^{\text{th}}$ column

$$[r_{i_1}| \ldots |r_{i_{d_i}}] := \begin{bmatrix} * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \\ r_{i_1,i_1} & \cdots & r_{i_1,i_{d_i}} \\ & \ddots & \vdots \\ 0 & & r_{i_{d_i},i_{d_i}} \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

of R give us coefficients with which we may express $y^1_{i_1}, \ldots, y^1_{i_{d_i}}$ in terms of the orthonormalized vectors

$$\operatorname{orth}(\mathcal{L}(n_2)\mathcal{L}(-n_1)x_{1_1}, \ldots, \mathcal{L}(n_2)\mathcal{L}(-n_1)x_{i_{d_i}}) = \operatorname{orth}\Big(\mathcal{L}^{(n_1+n_2)}_{\sigma_{-n_1}\omega} x_{1_1}, \ldots, \mathcal{L}^{(n_1+n_2)}_{\sigma_{-n_1}\omega} x_{i_{d_i}}\Big),$$

that emerge in the third step of Ginelli's algorithm. By means of this orthogonal transformation between coefficients and initial vectors, Eq. (5.12) may be checked on coefficient space. Since $\mathcal{L}(n_2)\mathcal{L}(-n_1)W_1$ is mapped to $\mathbb{R}^{d_1+\cdots+d_{i-1}} \times \{0\} \subset \mathbb{R}^k$ and $\mathcal{L}(n_2)\mathcal{L}(-n_1)W_2$ to $\mathbb{R}^{d_1+\cdots+d_i} \times \{0\} \subset \mathbb{R}^k$, we need to check that

$$\inf_{r \in \operatorname{span}\left(r_{i_1}, \ldots, r_{i_{d_i}}\right) \cap S} \|\operatorname{pr}_i r\|_2 > 0,$$

where $\operatorname{pr}_i : \mathbb{R}^k \to \{0\} \times \mathbb{R}^{d_i} \times \{0\}$ is the projection onto the i_1^{th} to $i_{d_i}^{\text{th}}$ coordinates. This is easily verified, since R is an upper triangular matrix with non-zero elements on the diagonal. Thus, we may apply Lemma 5.3.10 to see that the linear span of the i_1^{th} to $i_{d_i}^{\text{th}}$ vectors from the fifth step of Ginelli's algorithm approximates $Y_i(\omega)$ at the desired speed. This concludes the proof.[5] □

5.4 Summary and discussion

With the emergence of semi-invertible METs, the concept of CLVs has been opened up to new settings. In particular, various infinite-dimensional versions of the MET have been proved. In this chapter we analyzed convergence of Ginelli's algorithm to

[5] The last step of Ginelli's algorithm only normalizes computed vectors. It does not change their linear span and, thus, plays no role in Eq. (5.6). However, the step is a necessary part of the algorithm, since CLVs are defined as normalized basis vectors of $Y_i(\omega)$.

compute CLVs in the setting of the semi-invertible MET from [29]. Our main result is a convergence proof of the algorithm in the context of Hilbert spaces. The proof not only generalizes the analysis of the last chapter to an infinite-dimensional setting, but also treats the case of non-invertible linear propagators. We formulated most arguments for maps on Banach spaces before connecting them to basic asymptotic properties of the Oseledets splitting. Since those properties appear in most versions of the MET, our convergence proof can be translated to other settings as well.

A major part of our convergence proof was the analysis of so-called *well-separating common complements*. Those are subspaces that stay well-separated from a given sequence of subspaces. In particular, we applied this concept to the case where the sequence of subspaces was given by the Oseledets filtration at different initial times. Then, we used vectors subject to the obtained well-separating common complements as input vectors for Ginelli's algorithm. Since almost every tuple of input vectors spans a well-separating common complement, describing convergence with respect to such complements is sufficient in Hilbert spaces.

The actual convergence proof was split into estimates for forward and for backward propagation. During forward propagation, almost every complement to spaces of the Oseledets filtration asymptotically aligns with the sum of the first Oseledets spaces. The fact that complements generically align in forward-time even holds if we only have an Oseledets filtration. For backward propagation, we had to restrict the propagator to certain subspaces, since it may not be globally invertible in a semi-invertible setting. Last but not least, we combined our estimates to form the convergence proof.

Throughout the proof, we connected estimates to the LEs. Thus, we were able to relate LEs to the speed of convergence. As for the finite-dimensional case, Ginelli's algorithm converges exponentially fast with a rate given by the spectral gap between associated LEs.

While we successfully generalized and proved Ginelli's algorithm for infinite dimensions, it is primarily an analytical tool. The numerical computation of CLVs brings its own set of challenges. Indeed, our results may be seen as a help to understand limit cases of applications of Ginelli's algorithm for systems of increasingly higher resolutions. The transition between finite and infinite dimensions is still an open question and leads to the concept of stability of CLVs. Additionally, numerical inaccuracies in computing the linear propagator can result in a different output of Ginelli's algorithm. We stress again that CLVs may depend only measurably on the trajectory.

Despite the remaining challenges, we made a big step towards computing CLVs in infinite dimensions. Through the connection to semi-invertible METs, our research applies to recent developments in the context of CLVs and paves the way for new advancements of both analytical and numerical aspects of CLV-algorithms.

6

Conclusions

During the last years *covariant Lyapunov vectors (CLVs)* received increased attention in applications. They constitute modes that describe long-term linear perturbations along a given nonlinear background trajectory of a dynamical system. Even though their existence has been known since 1968 due to the celebrated *multiplicative ergodic theorem (MET)* by Oseledets, the interest in applications for CLVs sparkled only recently due to new efficient algorithms to compute CLVs. Among the most famous is *Ginelli's algorithm*, which uses a purely dynamical approach to compute CLVs. While there are numerous applications of the algorithm, mathematical results are quite rare. In this thesis we performed a mathematically rigorous convergence analysis of Ginelli's algorithm. Our main contributions are convergence theorems for finite and for infinite dimensions.

Since we require at least existence of CLVs to compute them, we restricted our analysis to settings of the MET. On the one hand, we stated different versions of the MET for finite-dimensional random dynamical systems (including a fully invertible deterministic version), on the other, we presented a semi-invertible version for random dynamical systems acting on separable Banach spaces. The semi-invertible and fully-invertible versions yield *Oseledets splittings* consisting of *Oseledets spaces* with respect to which the CLVs are defined. In finite dimensions there is a direct link between CLVs and singular vectors of the propagator, whereas the MET on Banach spaces requires different tools to derive the existence of CLVs.

Using the asymptotic characterization of CLVs (or Oseledets spaces) via *Lyapunov exponents (LEs)* as given by the MET, we explained the basic idea behind Ginelli's algorithm. The algorithm propagates two sets of initial perturbations along the trajectory to approximate CLVs at the present state. The first set of perturbations is initiated in the far past and propagated along the whole trajectory to get an approximation of the linear span of the first CLVs in forward-time. The second set of perturbations is expressed in terms of coefficients of forward propagated vectors of the first set and is propagated backwards from the far future to give an approximation of CLVs at the present state.

Prior to the convergence analysis we implemented Ginelli's algorithm and computed CLVs for the Lorenz attractor of the Lorenz 63 model. The computed vectors appeared to converge when the amount of past and future data was increased. Furthermore, the speed of convergence seemed to be exponential with a rate given by the spectral gap between associated LEs. However, since we do not have analytic results for CLVs of the Lorenz attractor, these observations are merely a motivation for the mathematical convergence analysis.

The analysis was split into two parts. The first investigated convergence in a fully-invertible finite-dimensional setting, whereas the second was concerned with convergence in a semi-invertible setting on Hilbert spaces. Both approaches used projection operators to adhere for possibly degenerate Lyapunov spectra. Thus, convergence was analyzed with respect to subspaces rather than single output vectors. Indeed, in the presence of degeneracies isolated output vectors might not converge due to internal dynamics of the Oseledets spaces. Therefore, we proved convergence at the level of subspaces using different notions of distances and angles for finite and for infinite dimensions.

An important part of the analysis for finite dimensions was the link between CLVs and singular vectors of the propagator. Singular vectors are orthogonal directions of optimal growth for finite time. To describe the relation between input vectors and singular vectors, we derived the notion of *admissibility*. A tuple of vectors is said to be admissible with respect to another tuple if the associated filtration spaces are close enough. By letting the admissibility parameter depend on time, we were able to tell if a tuple of initial vectors would align with singular vectors (in terms of filtrations) when propagated long enough. Since tuples of input vectors are generically admissible, it was sufficient to analyze convergence of Ginelli's algorithm in terms of admissible tuples.

Ultimately, we proved convergence of Ginelli's algorithm in measure for continuous time and convergence for almost every input for discrete time. A distinction between both cases of time was necessary since there are examples where the continuous-time version diverges for every fixed choice of initial vectors. However, since initial vectors are usually chosen anew when increasing the runtime, the difference between both notions of convergence is negligible for applications.

Our convergence analysis for infinite dimensions required different techniques. Instead of admissibility we used so-called *well-separating common complements* to find subspaces which stay far enough from the spaces of the Oseledets filtration at different initial times. We proved that almost every tuple of vectors of a Hilbert space spans such a complement. Here, "almost every" is understood in the sense of *prevalence* which generalizes the notion of "Lebesgue almost every" to infinite-dimensional vector spaces. In our convergence analysis we connected properties of well-separating common complements to estimates for forward and for backward propagation. Finally, the obtained tools were combined to arrive at a convergence theorem of Ginelli's algorithm similar to the finite-dimensional version with discrete time: convergence for almost every input.

In addition to the pure convergence statements, we gave an upper bound of the speed of convergence. We proved that Ginelli's algorithm converges at an exponential rate given by the spectral gap between associated LEs. Factors on subexponential scales, e.g., coming from the nonlinearity of the background trajectory or from choosing different inputs do not matter asymptotically. Hence, they did not play a role in our convergence analysis. Nevertheless, they may be important for applications where only limited amounts of past and future data is available. The proportion in which past and future data are needed depends on the specific system and background trajectory. We allowed for different proportions by incorporating two separate runtime variables for past and future data into our analysis. Both need to be increased to reach convergence.

In conclusion we analyzed multiple aspects of convergence such as the difference between continuous and discrete time, invertible and semi-invertible settings, the connection to singular vectors, a version of Ginelli's algorithm on Hilbert spaces, and the speed of convergence. Even though there are still some open questions, e.g., concerning the generalization to Banach spaces and the transition from finite to infinite dimensions, we managed to derive a wide range of tools that help to understand the fundamental ideas behind Ginelli's algorithm and can potentially be applied to other algorithms or scenarios as well. Ultimately, our analysis provides new mathematical insight into CLV-algorithms which are important instruments for finding structure in the chaos of dynamical systems.

APPENDIX A

Derivation of Uniform Bounds

In this chapter we derive *uniform bounds* for asymptotics of the Oseledets filtration and of the Oseledets splitting which appear in Section 2.3. We need those bounds for our convergence proof in Chapter 5. Since [29] provides hints for their derivation and uses them in the proof of Theorem 2.3.2, we think this is a good opportunity to provide the necessary details.

While uniform bounds for systems on $\mathbb{R}^d$ follow from properties of the SVD, systems on Banach spaces require different approaches. We first derive bounds for the Oseledets filtration by invoking [17] and [39]. Then, we extract bounds for the Oseledets splitting via [22] and [29].

A.1 Bounds for Oseledets filtration

In his dissertation Doan proves the existence of an Oseledets filtration for one-sided random dynamical systems on separable Banach spaces by embedding systems into larger systems that have injective cocycles [17]. For those enlarged systems, the MET by Lian and Lu provides existence of an Oseledets filtration with uniform bounds [39]. We state the necessary arguments to derive uniform bounds for the Oseledets filtration of Doan's MET from the MET by Lian and Lu and refer to the selected references for full statements of the results and techniques.

Theorem A.1.1 (MET by Lian and Lu [39])
Let $\mathcal{R} = (\Omega, \mathcal{F}, \mathbb{P}, \sigma, X, \mathcal{L})$ be a separable strongly measurable random dynamical system over an ergodic invertible base such that $\log^+ \|\mathcal{L}\| \in L^1(\Omega, \mathcal{F}, \mathbb{P})$. Furthermore, assume that $\mathcal{L} : \Omega \to L(X)$ is injective $\mathbb{P}$-almost everywhere and $\mathcal{R}$ is quasi-compact.

There exist $1 \leq p \leq \infty$ exceptional LEs $\lambda^* = \lambda_1 > \cdots > \lambda_p > \kappa^*$ (or if $p = \infty$: $\lambda_1 > \lambda_2 > \cdots > \kappa^*$ and $\lim_{i\to\infty} \lambda_i = \kappa^*$), multiplicities $d_1, \ldots, d_p \in \mathbb{N}$, and a unique measurable filtration of X into closed subspaces

$$X = V_1(\omega) \supset \cdots \supset V_p(\omega) \supset V(\omega) \supset \{0\}$$

defined on a σ-invariant subset $\Omega' \subset \Omega$ of full $\mathbb{P}$-measure such that the following hold for $\omega \in \Omega'$:

1. *the splitting is equivariant, i.e.,* $\mathcal{L}(\omega)V_i(\omega) \subset V_i(\sigma\omega)$ *and* $\mathcal{L}(\omega)V(\omega) \subset V(\sigma\omega)$,
2. $\mathrm{codim}V_{i+1}(\omega) = d_1 + \cdots + d_i$,
3. $\lim_{n\to\infty}(1/n)\log\|\mathcal{L}_\omega^{(n)}v\| = \lambda_i$ *for* $v \in V_i(\omega) \setminus V_{i+1}(\omega)$,
4. $\lim_{n\to\infty}(1/n)\log\|\mathcal{L}_\omega^{(n)}|_{V_i(\omega)}\| = \lambda_i$, *and*
5. $\limsup_{n\to\infty}(1/n)\log\|\mathcal{L}_\omega^{(n)}|_{V(\omega)}\| \leq \kappa^*$,

where we set $V_{p+1}(\omega) := V(\omega)$.

The full MET by Lian and Lu even provides an Oseledets splitting that is related to the Oseledets filtration via Eq. (2.5). For our purposes the filtration is enough.

Now, Doan's theorem states that the above is still true if we drop the assumption of $\mathcal{L}$ being injective $\mathbb{P}$-almost everywhere. To be precise, in Doan's formulation the fourth and fifth properties about uniform bounds are left out. Since we need these properties, we derive them as products of Doan's proof.

Given $\gamma > 0$, Doan enlarges the separable Banach space X to the space of sequences of elements of X:

$$X_\gamma := \left\{ \mathbf{x} := (x_n)_{n\in\mathbb{N}_0} \;\middle|\; \lim_{n\to\infty} e^{-\gamma n} x_n \text{ exists} \right\}.$$

He shows that X_γ equipped with the norm

$$\|\mathbf{x}\|_\gamma := \sup_{n\in\mathbb{N}_0} e^{-\gamma n}\|x_n\|$$

is a separable Banach space. On this Banach space an extended cocycle can be defined using the generator

$$\tilde{\mathcal{L}}_\omega \mathbf{x} := (\mathcal{L}_\omega x_0, \alpha_0 x_0, \alpha_1 x_1, \dots),$$

where $(\alpha_n)_{n\in\mathbb{N}_0}$ is a descending sequence of positive scalars satisfying certain growth conditions. As Doan suggests, we set $\alpha_n = e^{-(2n+1)}$. The generated cocycle has the form

$$\tilde{\mathcal{L}}_\omega^{(n)}\mathbf{x} = \left(\mathcal{L}_\omega^{(n)}x_0, \alpha_0\mathcal{L}_\omega^{(n-1)}x_0, \dots, \alpha_{n-1}\dots\alpha_0 x_0, \alpha_n\dots\alpha_1 x_1, \dots\right).$$

If the original system $\mathcal{R}$ is strongly measurable, has an ergodic invertible base, satisfies $\log^+\|\mathcal{L}\| \in L^1$, and is quasi compact, then the enlarged random dynamical system $\tilde{\mathcal{R}}$ fulfills the assumptions of Theorem A.1.1. Thus, we have LEs $\lambda_1 > \cdots > \lambda_p > \kappa^*$ and an Oseledets filtration $X_\gamma = \tilde{V}_1(\omega) \supset \cdots \supset \tilde{V}_p(\omega) \supset \tilde{V}(\omega) \supset \{0\}$ of $\tilde{\mathcal{R}}$. Doan proves that $\pi\mathbf{x} = x_0$ projects the Oseledets filtration of $\tilde{\mathcal{R}}$ onto a filtration of $\mathcal{R}$ via $V_i(\omega) = \pi\tilde{V}_i(\omega)$ with almost the same properties (all except 4. and 5. of Theorem A.1.1). Hence, we call the projected filtration Oseledets filtration. The exceptional LEs of the original and of the enlarged system coincide. We now argue why properties 4. and 5. still hold for the projected filtration:

Lemma A.1.2
In Doan's MET [17] for one-sided random dynamical systems on Banach spaces Eqs. (2.6) *and* (2.7) *hold for* $\mathbb{P}$*-almost every* ω, *i.e., we have*

$$\lim_{n\to\infty}\frac{1}{n}\log\|\mathcal{L}_\omega^{(n)}|_{V_i(\omega)}\| = \lambda_i$$

for $1 \le i \le p$ *and*

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}_\omega^{(n)}|_{V(\omega)}\| \le \kappa^*.$$

In particular, these bounds hold in the setting of Theorem 2.3.2.

Proof. Since the third property of Theorem A.1.1 also holds in Doan's MET, we get λ_i as a lower bound of the limit in Eq. (2.6). It remains to prove that λ_i is an upper bound.

The main idea is to use the uniform bound coming from the enlarged system. For $\mathbf{x} \in \tilde{V}_i(\omega)$ and $x_0 = \pi\mathbf{x}$, we have

$$\|\mathcal{L}_\omega^{(n)} x_0\| = \|\mathcal{L}_\omega^{(n)} \pi \mathbf{x}\| = \|\pi \tilde{\mathcal{L}}_\omega^{(n)} \mathbf{x}\| \le \|\pi\| \, \|\tilde{\mathcal{L}}_\omega^{(n)}|_{\tilde{V}_i(\omega)}\| \, \|\mathbf{x}\|_\gamma.$$

Since $\|\pi\|$ is a constant factor, it vanishes on exponential scales. The second factor on the right can be bounded according to the MET by Lian and Lu. To get rid of the last factor, we show that $\iota V_i(\omega) \subset \tilde{V}_i(\omega)$, where $\iota x_0 = (x_0, 0, 0, \dots)$. If this is true, then $V_i(\omega) = \pi\iota V_i(\omega)$ and

$$\|\mathcal{L}_\omega^{(n)}|_{V_i(\omega)}\| \le \sup_{x_0 \in V_i(\omega)\cap B} \|\mathcal{L}_\omega^{(n)} x_0\| \le \|\pi\| \, \|\tilde{\mathcal{L}}_\omega^{(n)}|_{\tilde{V}_i(\omega)}\|$$

since $\|\iota x_0\|_\gamma = \|x_0\|$. In particular, this would prove the claim.[1]

Let $\mathbf{x} \in \tilde{V}_i(\omega)$. To show that $\iota x_0 \in \tilde{V}_i(\omega)$, it suffices to investigate its growth rate, since $\tilde{V}_i(\omega)$ is the set of all elements whose exponential growth rate is at most λ_i. Thus, if ιx_0 does not have a faster growth rate than $\mathbf{x}$, it is an element of $\tilde{V}_i(\omega)$. We have

$$\|\tilde{\mathcal{L}}_\omega^{(n)} \mathbf{x}\|_\gamma = \max\left(\max_{0\le k\le n} e^{-\gamma k} \|\mathcal{L}_\omega^{(n-k)} x_0\| \prod_{j=0}^{k-1} \alpha_j, \, \sup_{k\in\mathbb{N}} e^{-\gamma(n+k)} \|x_k\| \prod_{j=k}^{k+n-1} \alpha_j \right).$$

The second part decays superexponentially fast:

$$-\log \prod_{j=k}^{k+n-1} \alpha_j = (2k+1) + (2k+1+2) + \cdots + (2k+1+2(n-1)) = 2kn + n^2$$

and

$$\begin{aligned}\frac{1}{n} \log\left(\sup_{k\in\mathbb{N}} e^{-\gamma(n+k)} \|x_k\| \prod_{j=k}^{k+n-1} \alpha_j \right) &\le \frac{1}{n} \log\left(\sup_{k\in\mathbb{N}_0} e^{-\gamma k} \|x_k\| e^{-\gamma n} e^{-2kn-n^2} \right) \\ &\le \frac{\log \|\mathbf{x}\|_\gamma}{n} - \gamma - n \to -\infty.\end{aligned}$$

Hence, the exponential growth rate only depends on x_0. We get $\iota V_i(\omega) \subset \tilde{V}_i(\omega)$. □

The uniform estimates for backward-time, i.e., Eqs. (2.9) and (2.10), follow from a result by Froyland and others:

Lemma A.1.3 ([22])
Let $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ *be an ergodic metric dynamical system over* $\mathbb{Z}$. *If* $(f_n)_{n\in\mathbb{N}}$ *is a* subadditive *sequence of functions* $\Omega \to \mathbb{R} \cup \{\pm\infty\}$, *i.e.,* $f_{m+n}(\omega) \le f_m(\sigma_n\omega) + f_n(\omega)$ *for every* $n, m \in \mathbb{N}$ *and* $\omega \in \Omega$, *and* $f_1^+ \in L^1$, *then there is a constant* $c \in \mathbb{R} \cup \{-\infty\}$ *such that* $f_n(\omega)/n \to c$ *and* $f_n(\sigma_{-n}\omega)/n \to c$.

[1] The bound for $V(\omega)$ follows analogously.

Lemma A.1.4
In the setting of Theorem 2.3.2 Eqs. (2.9) *and* (2.10) *hold for* $\mathbb{P}$*-almost every* ω*, i.e., we have*

$$\lim_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}^{(n)}_{\sigma_{-n}\omega}|_{V_i(\sigma_{-n}\omega)}\| = \lambda_i$$

for $1 \leq i \leq p$ *and*

$$\limsup_{n\to\infty} \frac{1}{n} \log \|\mathcal{L}^{(n)}_{\sigma_{-n}\omega}|_{V(\sigma_{-n}\omega)}\| \leq \kappa^*.$$

Proof. We apply Lemma A.1.3 to the sequences $f_n := \log \|\mathcal{L}^{(n)}_\omega|_{V_i(\omega)}\|$ and $g_n := \log \|\mathcal{L}^{(n)}_\omega|_{V(\omega)}\|$. Indeed, subadditivity follows from the cocycle property and from equivariance of the Oseledets filtration. f_1^+, g_1^+ can both be bounded by the integrable function $\log^+ \|\mathcal{L}\|$. Now, according to Lemma A.1.3 it holds $f_n(\omega)/n \to c_f$, $f_n(\sigma_{-n}\omega)/n \to c_f$, $g_n(\omega)/n \to c_g$, and $g_n(\sigma_{-n}\omega)/n \to c_g$ for $\mathbb{P}$-almost every ω. By Lemma A.1.2 we must have $c_f = \lambda_i$ and $c_g \leq \kappa^*$. This proves the claim. □

A.2 Bounds for Oseledets splitting

In their proof of the MET in [29] Gonzàlez-Tokman and Quas require uniform lower bounds of growth rates inside Oseledets spaces. Even though they argue why those bounds hold for forward-time [29, lemma 2.14], the details are left for the reader to complete. Here, we provide the missing details. Moreover, we show that uniform lower bounds hold for the sum of the first Oseledets spaces instead of single Oseledets spaces. The main idea is to find a basis of the first Oseledets spaces in which we reduce the cocycle to a cocycle on $\mathbb{R}^d$. Then, the finite-dimensional theory applies and gives us uniform bounds via the SVD.

Lemma A.2.1
Let X *be a separable Banach space and* $Y(\omega) \subset X$ *a measurable subspaces of dimension* k*. There is a measurable map* $A : \Omega \times \mathbb{R}^k \to X$ *such that* $A(\omega) : \mathbb{R}^k \to Y(\omega)$ *is a linear isomorphism with*

$$\frac{1}{2^{k+1}-2}\|a\|_2 \leq \|A(\omega)a\| \leq \sqrt{k}\|a\|_2$$

for every $a \in \mathbb{R}^k$*.*

We prove the lemma by finding an ϵ-nice basis of $Y(\omega)$. Using this basis, we may identify $Y(\omega)$ with $\mathbb{R}^k$.

Definition A.2.2 ([29])
Let Y *be a Banach space of dimension* k*. A basis* $(y_1, \ldots, y_k)$ *is called* ϵ-nice *if* $1-\epsilon < \|y_i\| < 1+\epsilon$ *and* $d(y_i, \mathrm{span}(y_1, \ldots, y_{i-1})) > 1-\epsilon$ *for each* $i > 1$*.*

We need [29, lemma B.4] for the proof of Lemma A.2.1:

Lemma A.2.3 ([29])
If $(y_1, \ldots, y_k)$ *is an* ϵ*-nice basis with* $\epsilon < 2^{-k-2}$*, then*

$$\left\|\sum_{i=1}^k a_i y_i\right\| \leq 1 \quad \Longrightarrow \quad |a_i| \leq 2^{k+1-i} \text{ for each } i.$$

proof of Lemma A.2.1. We inductively prove existence of a measurable ϵ-nice basis $(y_1(\omega), \dots, y_k(\omega))$ of $Y(\omega)$ with $\epsilon < 2^{-k-2}$. Additionally, we assume that $\|y_i\| = 1$. Fix a countable dense subset $(x_j)_{j\in\mathbb{N}}$ of the unit sphere $S \subset X$. Assume we already have the first $i-1$ basis vectors for some $i = 1, \dots, k$. We show existence of the i^{th} vector. Define

$$r_1(\omega) := \min\left\{ j \in \mathbb{N} \;\middle|\; d(x_j, \operatorname{span}(y_1(\omega), \dots, y_{i-1}(\omega))) > 1 - \frac{\epsilon}{2} \text{ and } d(x_j, Y(\omega)) < \frac{\epsilon}{2} \right\}$$

and inductively set

$$r_s(\omega) := \min\left\{ j \in \mathbb{N} \;\middle|\; d(x_j, x_{r_{s-1}(\omega)}) < \frac{\epsilon}{2^s} \text{ and } d(x_j, Y(\omega)) < \frac{\epsilon}{2^s} \right\}.$$

The sequence of measurable functions $(x_{r_s(\omega)})_{s\in\mathbb{N}}$ converges pointwise to a measurable function $y_i(\omega)$, which satisfies the required properties.

Now, let $A(\omega)a := \sum a_i y_i(\omega)$. We have

$$\|A(\omega)a\| \leq \sum_{i=1}^{k} |a_i| \, \|y_i(\omega)\| = \|a\|_1 \leq \sqrt{k}\|a\|_2$$

and by Lemma A.2.3

$$\frac{\left\|\frac{a}{c}\right\|_2}{2^{k+1}-2} \leq \frac{\left\|\frac{a}{c}\right\|_1}{2^{k+1}-2} \leq \frac{\sum_{i=1}^{k} 2^{k+1-i}}{2^{k+1}-2} = 1 = \left\|A(\omega)\frac{a}{c}\right\|,$$

where $c := \|A(\omega)a\|$. Since c is a scalar, the above chain of (in-)equalities can be scaled to eliminate c. The claim follows. □

Let $\mathcal{R}$ be a random dynamical system as in Theorem 2.3.2. Using Lemma A.2.1, we get an identification of the sum of the first Oseledets spaces $Y_1(\omega) \oplus \dots \oplus Y_i(\omega)$ with $\mathbb{R}^k$, where $k = d_1 + \dots + d_i$. Moreover, the identification provides a new one-sided cocycle on $\mathbb{R}^k$ via $\tilde{\mathcal{L}}_\omega^{(n)} := A(\omega)^{-1}\mathcal{L}_\omega^{(n)}A(\omega)$. Here, $A(\omega)$ should be understood as an isomorphism $\mathbb{R}^k \to Y_1(\omega) \oplus \dots \oplus Y_i(\omega)$.

By [29, lemma A.5] the composition of strongly measurable maps is again strongly measurable. However, the section on strong measurability in [29] is only formulated for operators acting on a single separable Banach space X. To show that $\tilde{\mathcal{L}}$ is strongly measurable one needs to generalize the whole section to operators between potentially different separable Banach spaces. We leave this to the reader. Special care should be given to $A(\omega)^{-1}$. In fact, it suffices to show strong measurability of $A(\omega)^{-1}\Pi_{Y_1(\omega)\oplus\dots\oplus Y_i(\omega)||V_{i+1}(\omega)}$, which is a well-defined map $\Omega \times X \to \mathbb{R}^k$. This can be done in a similar fashion to the proof of Lemma A.2.1. Fix a countable dense subset $(b_j)_{j\in\mathbb{N}}$ of $\mathbb{R}^k$. For $x \in X$ and $s \in \mathbb{N}$, define

$$r_s(\omega) := \min\left\{ j \in \mathbb{N} \;:\; \|A(\omega)b_j - \Pi_{Y_1(\omega)\oplus\dots\oplus Y_i(\omega)||V_{i+1}(\omega)}x\| < \frac{1}{s} \right\}.$$

Writing $y = \Pi_{Y_1(\omega)\oplus\dots\oplus Y_i(\omega)||V_{i+1}(\omega)}x$, Lemma A.2.1 implies that

$$\frac{1}{s} > \|A(\omega)b_{r_s(\omega)} - y\| = \|A(\omega)(b_{r_s(\omega)} - A(\omega)^{-1}y)\| \geq \frac{\|b_{r_s(\omega)} - A(\omega)^{-1}y\|_2}{2^{k+1}-2}.$$

In particular, $b_{r_s(\omega)} \to A(\omega)^{-1}y$ pointwise in ω for $s \to \infty$. Since the projection onto the sum of the first Oseledets spaces is measurable, we have strong measurability of $A(\omega)^{-1}\Pi_{Y_1(\omega)\oplus\cdots\oplus Y_i(\omega)||V_{i+1}(\omega)}$. Thus, $\tilde{\mathcal{L}}$ is strongly measurable. Finally, the word "strongly" can be omitted, because the strong operator topology and the norm topology on $\mathbb{R}^{k\times k}$ coincide.

Besides the measurability of $\tilde{\mathcal{L}}$, the other cocycle properties are inherited from the original cocycle. Moreover, the norm estimates in Lemma A.2.1 imply that $\|\mathcal{L}_\omega^{(n)}|_{Y_1(\omega)\oplus\cdots\oplus Y_i(\omega)}\|$ and $\|\tilde{\mathcal{L}}_\omega^{(n)}\|$ differ by at most a positive constant depending only on k. In particular, $\log^+\|\tilde{\mathcal{L}}\| \in L^1$ is integrable and we may apply case (A) of Theorem 2.2.7. The LEs of the reduced cocycle coincide with the first i exceptional LEs of the original cocycle. Hence, the lowest singular value $\delta_i^{\min} = \delta_i^{\min}(\omega)$ of the reduced cocycle grows exponentially with a rate given by λ_i. Since $\|\tilde{\mathcal{L}}_\omega^{(n)} a\|_2 \geq \delta_i^{\min}$ for all $\|a\|_2 = 1$, we get a uniform lower bound which can be transferred back to the original cocycle.

Lemma A.2.4
In the setting of Theorem 2.3.2 Eq. (2.8) *holds for* $\mathbb{P}$*-almost every* ω*, i.e., we have*

$$\liminf_{n\to\infty} \inf_{y\in Y_1(\omega)\oplus\cdots\oplus Y_i(\omega)\cap S} \frac{1}{n}\log\|\mathcal{L}_\omega^{(n)} y\| = \lambda_i$$

for $1 \leq i \leq p$.

The last uniform bound needed in Chapter 5 follows immediately from [22, lemma 8.3] which relates singular values of $\tilde{\mathcal{L}}_{\sigma_{-n}\omega}^{(n)}$ to those of $\tilde{\mathcal{L}}_\omega^{(n)}$:

Lemma A.2.5 ([22])
Let $\mathcal{R} = (\Omega, \mathcal{F}, \mathbb{P}, \sigma, \mathcal{L})$ *be an ergodic semi-invertible random dynamical system on* $\mathbb{R}^k$ *with* $\log^+\|\mathcal{L}\| \in L^1$. *By Theorem 2.2.7 the system admits a Lyapunov spectrum with exponents* λ_i. *The singular values of* $\mathcal{L}_{\sigma_{-n}\omega}^{(n)}$ *(sorted in decaying order) grow exponentially according to the LEs:*

$$\forall i_j : \lim_{n\to\infty} \frac{1}{n}\log\delta_{i_j}\left(\mathcal{L}_{\sigma_{-n}\omega}^{(n)}\right) = \lambda_i.$$

Lemma A.2.6
In the setting of Theorem 2.3.2 Eq. (2.11) *holds for* $\mathbb{P}$*-almost every* ω*, i.e., we have*

$$\liminf_{n\to\infty} \inf_{y\in Y_1(\sigma_{-n}\omega)\oplus\cdots\oplus Y_i(\sigma_{-n}\omega)\cap S} \frac{1}{n}\log\|\mathcal{L}_{\sigma_{-n}\omega}^{(n)} y\| = \lambda_i$$

for $1 \leq i \leq p$.

APPENDIX B

Sample Code

We list the code used for computing CLVs in the Lorenz 63 model with MATLAB R2019a. Ginelli's algorithm is implemented in a quite general form, so that it can be adjusted to settings other than the Lorenz model.

Listing B.1: Lorenz63CLVs.m

```
close all;
clc;


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% input
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

system_name = 'Lorenz 63 (rho = 28, sigma = 10, beta = 8/3)';
system_nonlinear_filename = 'Lorenz63';
system_linear_filename = 'Lorenz63_linear';

dim = 3;
x0_past = [-1.01;3.01;2.01]; % initial state of trajectory
tau = 0.0005; % size of one timestep
N = 100000;
N_past = N; % number of timesteps for past cocycle
N_future = N; % number of timesteps for future cocycle
numLEsToCompute = dim; % number of LEs/CLVs to compute

% figure for input data
figure_input = figure;
str_input = sprintf(...
    [...
        '--- Input data ---\n'...
        '\n'...
        'System:\n'...
        '%s\n'...
```

```
        '\n'...
        'Number of timesteps for past data:\n'...
        '%d\n'...
        '\n'...
        'Number of timesteps for future data:\n'...
        '%d\n'...
        '\n'...
        'Stepsize:\n'...
        '%g\n'...
        '\n'...
        'Number of LEs to compute:\n'...
        '%d\n'...
    ],...
    system_name,N_past,N_future,tau,numLEsToCompute);
str1 = sprintf([...
        '\n'...
        'Initial state of past data orbit:\n'...
        '[ '
    ]);
str2 = sprintf('%.5f ',x0_past);
str3 = sprintf(']\n');
str_input = [str_input str1 str2 str3];
annotation('textbox', [0, 0, 1, 1], 'String', str_input);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% computation of LEs and CLVs
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% combine nonlinear with linear system
system_nonlinear = str2func(system_nonlinear_filename);
system_linear = str2func(system_linear_filename);
system_combined = combineNonlinearLinear(system_nonlinear,system_linear,dim);

% prepare plot of trajectory
figure_orbit = figure;
hold on

% compute cocycle and plot trajectory
[x0_present,cocycle_past] = ODEtoCocycle(system_combined,x0_past,N_past,tau);
[~,cocycle_future] = ODEtoCocycle(system_combined,x0_present,N_future,tau);

% finish plot of trajectory
xlabel('x1-axis')
ylabel('x2-axis')
zlabel('x3-axis')
plot3(x0_present(1),x0_present(2),x0_present(3),'ok','linewidth',3);
legend('past data','future data','present state','Location','southwest')
figure_orbit.CurrentAxes.View = [50 20];
```

```
hold off

% compute LEs and CLVs from cocycle data
[LEs,CLVs] = Ginelli(cocycle_past,cocycle_future,numLEsToCompute,tau);

% figure for output data
figure_output = figure;
str_output = sprintf(...
    [...
        '--- Output data ---\n'...
        '\n'...
        'Lyapunov exponents:\n'...
    ]);
str1 = sprintf('%.5g\n',LEs);
str2 = sprintf(...
    [...
        '\n'...
        'Covariant Lyapunov vectors:\n'...
    ]);
str3 = '';
for i = 1:numLEsToCompute
    str3a = sprintf('[ ');
    str3b = sprintf('%.5f ',CLVs(:,i));
    str3c = sprintf(']\n');
    str3 = [str3 str3a str3b str3c];
end
str_output = [str_output str1 str2 str3];
str1 = sprintf([...
        '\n'...
        'System state:\n'...
        '[ '
        ]);
str2 = sprintf('%.5f ',x0_present);
str3 = sprintf(']\n');
str_output = [str_output str1 str2 str3];
annotation('textbox', [0, 0, 1, 1], 'String', str_output);


%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% error evolution of CLVs (for N_past = N_future)
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% number of tests (test i uses i/m times the number of timesteps, which were used to compute the reference CLVs)
m = 100;
CLVs_reference = CLVs;

% test runs of Ginelli's algorithm with error computation
CLV_error = zeros(numLEsToCompute,m);
for i = 1:m
    n = max(floor(N*(i/m)),1);
```

```

    [~,CLVs] = Ginelli(cocycle_past(N-n+1:N,1),cocycle_future(1:n,1),numLEsToCompute,tau);

    for j = 1:numLEsToCompute
        CLV_error(j,i) = min(norm(CLVs(:,j)-CLVs_reference(:,j)),norm(CLVs(:,j)+CLVs_reference(:,j)));
    end
end

% figures for CLV errors
figure_errors = cell(numLEsToCompute,1);
x = (1:m)'*max(floor(N*(1/m)),1)*tau;
for i = 1:numLEsToCompute
    figure_errors{i,1} = figure;
    hold on;

    xlabel('t-axis')
    ylabel('y-axis')
    ylim([-50 10])

    error = log(CLV_error(i,:));
    plot(x,error,'-.or');
    str_error = sprintf('y = log ||CLV%d(t) - CLV%d_r||_2',i,i);

    if i == 1
        if i == numLEsToCompute
            legend(str_error,'Location','southwest');
        else
            y = -abs(LEs(1)-LEs(2))*x;
            plot(x,y);
            str_estimate = 'y = -t * |LE1_r - LE2_r|';
            legend(str_error,str_estimate,'Location','southwest');
        end
    else
        if i == numLEsToCompute
            if numLEsToCompute == dim
                y = -abs(LEs(dim-1)-LEs(dim))*x;
                plot(x,y);
                str_estimate = sprintf('y = -t * |LE%d_r - LE%d_r|',dim,dim-1);
                legend(str_error,str_estimate,'Location','southwest');
            else
                legend(str_error,'Location','southwest');
            end
        else
            y = -min(abs(LEs(i)-LEs(i-1)),abs(LEs(i)-LEs(i+1)))*x;
            str_estimate = sprintf(['y = -t * min(|LE%d_r - LE%d_r|, |LE%d_r - LE%d_r|)'],i,i-1,i,i+1);
            plot(x,y);
            legend(str_error,str_estimate,'Location','southwest');
```

```
        end
    end

    hold off;
end
```

Listing B.2: Lorenz63.m

```
function [ dx ] = Lorenz63( ~, x )
%LORENZ63 Returns vector field of Lorenz 63 system at x.

 rho = 28;
 sigma = 10;
 beta = 8/3;

 dx = [sigma*(x(2)-x(1));...
       x(1)*(rho-x(3))-x(2);...
       x(1)*x(2)-beta*x(3)];

end
```

Listing B.3: Lorenz63_linear.m

```
function [ A ] = Lorenz63_linear( ~, x )
%LORENZ63_LINEAR Returns linearization of Lorenz 63 system at x.

 rho = 28;
 sigma = 10;
 beta = 8/3;

 A = [-sigma, sigma, 0;...
       rho-x(3), -1, -x(1);...
       x(2), x(1), -beta];

end
```

Listing B.4: combineNonlinearLinear.m

```
function [ system_combined ] = combineNonlinearLinear( system_nonlinear,
    system_linear, dim )
%COMBINENONLINEARLINEAR Combines nonlinear and linear autonomous
%systems of dimension dim.

    function [ dz ] = f( ~, z )

        x = z(1:dim);
        dx = system_nonlinear(1,x);

        y = z(dim+1:end);
        for i = 1:dim
            dy(1+(i-1)*dim:i*dim,1) = system_linear(1,x)*y(1+(i-1)*dim:i*dim
                );
```

```
        end

        dz = [dx;dy];
    end

    system_combined = @f;
end
```

Listing B.5: ODEtoCocycle.m

```
function [ x1, cocycle ] = ODEtoCocycle( system_combined, x0, n, tau )
%ODETOCOCYCLE Produces cocycle from combined system.
%
%   The nonlinear system is integrated from x0 for n timesteps of
%   length tau, and the linear propagator is computed for each
%   timestep. Moreover, the trajectory is plotted to the current
%   plot.
%
%   input:
%
%       system - function handle for an ODE combined with its
%           linearization
%       x0 - initial state for the ODE as column vector
%       n - number of timesteps
%       tau - size of one timestep
%
%   output:
%
%       x1 - final state after n timesteps of size tau from x0
%       cocycle - cell array of size [n,1], where cocycle{i,1} is the
%           linear propagator on tangent space between timesteps
%


[dim,~] = size(x0);

z_plot = zeros(dim,n+1);
z_plot(:,1) = x0;

cocycle = cell(n,1);
I = eye(dim);
I_vector = I(:);
x1 = x0;

for i = 1:n
    z0 = [x1;I_vector];
    [~,z1] = rk4(system_combined,0,z0,tau);
    x1 = z1(1:dim);
    y1 = z1(1+dim:end,1);

    z_plot(:,i+1) = x1;

```

```
    cocycle{i,1} = reshape(y1,[dim,dim]);
end

plot3(z_plot(1,:),z_plot(2,:),z_plot(3,:));

end
```

Listing B.6: rk4.m

```
function [ t1, z1 ] = rk4( f, t0, z0, h )
%RK4 One step of Runge-Kutta 4th order method.

    t1 = t0 + h;

    h2 = h/2;
    h6 = h/6;

    k1 = f(t0,z0);
    k2 = f(t0+h2,z0+h2*k1);
    k3 = f(t0+h2,z0+h2*k2);
    k4 = f(t1,z0+h*k3);

    z1 = z0+h6*(k1+2*k2+2*k3+k4);

end
```

Listing B.7: Ginelli.m

```
function [ LEs, CLVs ] = Ginelli( cocycle_past, cocycle_future,
    numLEsToCompute, tau )
%GINELLI Computes LEs and CLVs from cocycle data with size of
%timestep tau.

[dim,~] = size(cocycle_past{1,1});

% forward propagation: past to present
V = rand([dim,numLEsToCompute]); % initial vectors for the forward phase of
    Ginelli's algorithm
[~,V,~] = forward_propagation(cocycle_past,V,tau,0,0);

% forward propagation: present to future
[LEs,~,R] = forward_propagation(cocycle_future,V,tau,1,1);

% backward propagation: future to present
W = triu(rand(numLEsToCompute)); % initial coefficients for the backward
    phase of Ginelli's algorithm
CLVs = backward_propagation(V,R,W);

end
```

Listing B.8: forward_propagation.m

```
function [ LEs, V, R ] = forward_propagation( cocycle, V, tau, compute_LEs,
    store_R )
%FORWARD_PROPAGATION Forward propagation and reorthonormalization
%of initial vectors given by the columns of V.
%
%   The initial vectors given by the columns of V are propagated
%   using the cocycle. After every timestep of length tau the
%   propagated vectors are orthonormalized using a QR-decomposition.
%   The R matrices can be stored for output. The final propagated
%   vectors are returned as V.
%
%   input:
%
%       cocycle - cell array containing the propagation matrices
%       V - invertible matrix whose columns are initial vectors for
%           the propagation
%       tau - size of one timestep
%       compute_LEs - enables computation of Lyapunov exponents if
%           set to 1
%       store_R - stores R matrices from QR-decompositions if set to
%           1
%
%   output:
%
%       LEs - column vector of Lyapunov exponents (set to NaN if
%           compute_LEs is set to 0)
%       V - matrix of orthonormal final propagated vectors
%       R - cell array with upper triangular matrices from
%           QR-decompositions (set to NaN if store_R is set to 0)
%

[n,~] = size(cocycle);
[~,numLEsToCompute] = size(V);

LEs = NaN;
if compute_LEs == 1
    LE_sums = zeros(numLEsToCompute,1);
end

R = NaN;
if store_R == 1
    R = cell(n,1);
end

for i = 1:n
    [V,S] = qr(cocycle{i,1}*V);
    V = V(:,1:numLEsToCompute);

    if store_R == 1
        R{i,1} = S(1:numLEsToCompute,:);
```

```
    end

    if compute_LEs == 1
        D = diag(S);
        for j = 1:numLEsToCompute
            r = prod(D(1:j));
            LE_sums(j) = LE_sums(j)*((i-1)/i) + (1/tau*log(abs(r)))/i;
        end
    end
end

if compute_LEs == 1
    LEs = LE_sums(1:numLEsToCompute) - [0;LE_sums(1:numLEsToCompute-1)];
end

end
```

Listing B.9: backward_propagation.m

```
function [ CLVs ] = backward_propagation( V, R, W )
%BACKWARD_PROPAGATION Backward propagation and renormalization of
%coefficient vectors given by columns of W.
%
%   The coefficient vectors given by the columns of W are propagated
%   backwards using R. After every step the propagated vectors are
%   renormalized to have euclidean norm 1. The final propagated
%   vectors are expressed in terms of columns of V.
%
%   input:
%
%       V - matrix in whose columns to express the final propagated
%           coefficient vectors
%       R - cell array of upper triangular transition matrices with
%           respect to which the coefficient vectors will be
%           propagated backwards; inverses of transition matrices
%           are used in reversed order for propagation
%       W - upper triangular matrix containing initial coefficients
%
%   output:
%
%       CLVs - matrix of final propagated vectors expressed in terms
%           of V
%

[n,~] = size(R);
[~,numLEsToCompute] = size(V);

for i = 1:n
    W = R{n+1-i,1}\W;
    for j = 1:numLEsToCompute
        W(:,j) = W(:,j)/norm(W(:,j));
    end
```

```
end

CLVs = V*W;

end
```

Bibliography

[1] ARNOLD, L. *Random Dynamical Systems.* Springer monographs in mathematics. Berlin and Heidelberg: Springer, 1998. DOI: 10.1007/978-3-662-12878-7.

[2] ARNOLD, L. and NGUYEN, D. C. "Linear cocycles with simple Lyapunov spectrum are dense in L^∞." *Ergodic Theory and Dynamical Systems* 19.6 (1999), pp. 1389–1404. DOI: 10.1017/S014338579915199X.

[3] BARREIRA, L. and SILVA, C. "Lyapunov exponents for continuous transformations and dimension theory." *Discrete and Continuous Dynamical Systems* 13.2 (2005), pp. 469–490. DOI: 10.3934/dcds.2005.13.469.

[4] BEIMS, M. W. and GALLAS, J. A. C. "Alignment of Lyapunov Vectors: A Quantitative Criterion to Predict Catastrophes?" *Scientific reports* 6 (2016), p. 37102. DOI: 10.1038/srep37102.

[5] BENETTIN, G. et al. "Lyapunov Characteristic Exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. Part 1: Theory." *Meccanica* 15.1 (1980), pp. 9–20. DOI: 10.1007/BF02128236.

[6] BENETTIN, G. et al. "Lyapunov Characteristic Exponents for smooth dynamical systems and for hamiltonian systems; A method for computing all of them. Part 2: Numerical application." *Meccanica* 15.1 (1980), pp. 21–30. DOI: 10.1007/BF02128237.

[7] BLUMENTHAL, A. "A volume-based approach to the multiplicative ergodic theorem on Banach spaces." *Discrete and Continuous Dynamical Systems* 36.5 (2016), pp. 2377–2403. DOI: 10.3934/dcds.2016.36.2377.

[8] BOSETTI, H. and POSCH, H. A. "Orthogonal versus covariant Lyapunov vectors for rough hard disc systems." *Journal of Physics A: Mathematical and Theoretical* 46.25 (2013), p. 254011. DOI: 10.1088/1751-8113/46/25/254011.

[9] BOSETTI, H. and POSCH, H. A. "Covariant Lyapunov vectors for rigid disk systems." *Chemical Physics* 375.2-3 (2010), pp. 296–308. DOI: 10.1016/j.chemphys.2010.06.010.

[10] BOWEN, R. and RUELLE, D. "The Ergodic Theory of Axiom A Flows." *Inventiones mathematicae* 29.3 (1975), pp. 181–202. DOI: 10.1007/BF01389848.

[11] BURGER, T., GRITZMANN, P., and KLEE, V. "Polytope Projection and Projection Polytopes." *The American Mathematical Monthly* 103.9 (1996), pp. 742–755. DOI: 10.2307/2974444.

[12] CARLU, M. et al. "Lyapunov analysis of multiscale dynamics: the slow bundle of the two-scale Lorenz 96 model." *Nonlinear Processes in Geophysics* 26.2 (2019), pp. 73–89. DOI: 10.5194/npg-26-73-2019.

[13] CONTI, G. and BADIN, G. "Hyperbolic Covariant Coherent Structures in Two Dimensional Flows." *Fluids* 2.4 (2017), p. 50. DOI: 10.3390/fluids2040050.

[14] CRIMMINS, H. "Stability of hyperbolic Oseledets splittings for quasi-compact operator cocycles" (2019). arXiv: 1912.03008v1.

[15] Cruz, L. D. et al. “Exploring the Lyapunov instability properties of high-dimensional atmospheric and climate models.” *Nonlinear Processes in Geophysics* 25.2 (2018), pp. 387–412. doi: 10.5194/npg-25-387-2018.

[16] Deutsch, F. “The Angle Between Subspaces of a Hilbert Space.” In: *Approximation Theory, Wavelets and Applications*. Ed. by S. P. Singh. Dordrecht: Springer Netherlands, 1995, pp. 107–130. doi: 10.1007/978-94-015-8577-4_7.

[17] Doan, T. S. “Lyapunov Exponents for Random Dynamical Systems.” PhD thesis. Dresden: Fakultät Mathematik und Naturwissenschaften der Technischen Universität Dresden, 2010.

[18] Drivaliaris, D. and Yannakakis, N. “Subspaces with a common complement in a Banach space.” *Studia Mathematica* 182.2 (2007), pp. 141–164. doi: 10.4064/sm182-2-4.

[19] Drivaliaris, D. and Yannakakis, N. “Subspaces with a Common Complement in a Separable Hilbert Space.” *Integral Equations and Operator Theory* 62.2 (2008), pp. 159–167. doi: 10.1007/s00020-008-1622-5.

[20] Ershov, S. V. and Potapov, A. B. “On the concept of stationary Lyapunov basis.” *Physica D: Nonlinear Phenomena* 118.3-4 (1998), pp. 167–198. doi: 10.1016/S0167-2789(98)00013-X.

[21] Froyland, G., González-Tokman, C., and Quas, A. “Hilbert space Lyapunov exponent stability.” *Transactions of the American Mathematical Society* 372 (2019), pp. 2357–2388. doi: 10.1090/tran/7521.

[22] Froyland, G., Lloyd, S., and Quas, A. “Coherent structures and isolated spectrum for Perron–Frobenius cocycles.” *Ergodic Theory and Dynamical Systems* 30.3 (2010), pp. 729–756. doi: 10.1017/S0143385709000339.

[23] Froyland, G., Lloyd, S., and Quas, A. “A semi-invertible Oseledets Theorem with applications to transfer operator cocycles.” *Discrete and Continuous Dynamical Systems* 33.9 (2013), pp. 3835–3860. doi: 10.3934/dcds.2013.33.3835.

[24] Froyland, G. et al. “Computing covariant Lyapunov vectors, Oseledets vectors, and dichotomy projectors: A comparative numerical study.” *Physica D: Nonlinear Phenomena* 247.1 (2013), pp. 18–39. doi: 10.1016/j.physd.2012.12.005.

[25] Galántai, A. *Projectors and Projection Methods*. Springer Science+Business Media New York, 2004. doi: 10.1007/978-1-4419-9180-5.

[26] Ginelli, F. et al. “Characterizing Dynamics with Covariant Lyapunov Vectors.” *Physical review letters* 99.13 (2007), p. 130601. doi: 10.1103/PhysRevLett.99.130601.

[27] Ginelli, F. et al. “Covariant Lyapunov vectors.” *Journal of Physics A: Mathematical and Theoretical* 46.25 (2013), p. 254005. doi: 10.1088/1751-8113/46/25/254005.

[28] González-Tokman, C. “Multiplicative ergodic theorems for transfer operators: Towards the identification and analysis of coherent structures in non-autonomous dynamical systems.” In: *Contributions of Mexican mathematicians abroad in pure and applied mathematics*. Ed. by F. Galaz-García, J. C. P. Millán, and P. Solórzano. Vol. 709. Contemporary Mathematics. Guanajuato, Mexico: American Mathematical Society, 2018, pp. 31–52. doi: 10.1090/conm/709/14290.

[29] González-Tokman, C. and Quas, A. “A semi-invertible operator Oseledets theorem.” *Ergodic Theory and Dynamical Systems* 34.4 (2014), pp. 1230–1272. doi: 10.1017/etds.2012.189.

[30] González-Tokman, C. and Quas, A. “A concise proof of the multiplicative ergodic theorem on Banach spaces.” *Journal of Modern Dynamics* 9 (2015), pp. 237–255. doi: 10.3934/jmd.2015.9.237.

[31] González-Tokman, C. and Quas, A. “Stability and Collapse of the Lyapunov spectrum for Perron-Frobenius Operator cocycles” (2018). arXiv: 1806.08873v2.

[32] Inubushi, M., Takehiro, S.-i., and Yamada, M. “Regeneration cycle and the covariant Lyapunov vectors in a minimal wall turbulence.” *Physical Review E* 92.2 (2015), p. 023022. doi: 10.1103/PhysRevE.92.023022.

[33] John, F. "Extremum Problems with Inequalities as Subsidiary Conditions." In: *Traces and Emergence of Nonlinear Programming*. Ed. by G. Giorgi and T. H. Kjeldsen. Basel: Birkhäuser, 2014, pp. 197–215. doi: 10.1007/978-3-0348-0439-4_9.

[34] Kato, T. *Perturbation Theory for Linear Operators*. Vol. 132. Classics in mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg, 1995. doi: 10.1007/978-3-642-66282-9.

[35] Kryloff, N. and Bogoliouboff, N. "La Theorie Generale De La Mesure Dans Son Application A L'Etude Des Systemes Dynamiques De la Mecanique Non Lineaire." *The Annals of Mathematics* 38.1 (1937), pp. 65–113. doi: 10.2307/1968511.

[36] Kuptsov, P. V. and Parlitz, U. "Theory and Computation of Covariant Lyapunov Vectors." *Journal of Nonlinear Science* 22.5 (2012), pp. 727–762. doi: 10.1007/s00332-012-9126-5.

[37] Lauzon, M. and Treil, S. "Common complements of two subspaces of a Hilbert space." *Journal of Functional Analysis* 212.2 (2004), pp. 500–512. doi: 10.1016/S0022-1236(03)00253-2.

[38] Leonov, G. A. and Kuznetsov, N. V. "Time-varying Linearization and the Perron effects." *International Journal of Bifurcation and Chaos* 17.04 (2007), pp. 1079–1107. doi: 10.1142/S0218127407017732.

[39] Lian, Z. and Lu, K. "Lyapunov exponents and invariant manifolds for random dynamical systems in a Banach space." *Memoirs of the American Mathematical Society* 206.967 (2010). doi: 10.1090/S0065-9266-10-00574-0.

[40] Lorenz, E. N. "Deterministic Nonperiodic Flow." *Journal of the Atmospheric Sciences* 20 (1963), pp. 130–141. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

[41] Lorenz, E. N. "Does the Flap of a Butterfly's wings in Brazil Set Off a Tornado in Texas?" *Conference held at the 139th Meeting of the American Association for the Advancement of Science* (1972).

[42] Lorenz, E. N. *The Essence of Chaos*. 1. paperback ed. The Jessie and John Danz lectures. Seattle, Wash.: University of Washington Press, 1995.

[43] Luzzatto, S., Melbourne, I., and Paccaut, F. "The Lorenz Attractor is Mixing." *Communications in Mathematical Physics* 260.2 (2005), pp. 393–401. doi: 10.1007/s00220-005-1411-9.

[44] Mañé, R. "Lyapounov exponents and stable manifolds for compact transformations." In: *Geometric dynamics*. Ed. by J. E. Palis. Vol. 1007. Lecture Notes in Mathematics. Berlin: Springer-Verlag, 1983, pp. 522–577. doi: 10.1007/BFb0061433.

[45] Metzger, R. J. "Sinai-Ruelle-Bowen measures for contracting Lorenz maps and flows." *Annales de l'I.H.P. Analyse non linéaire* 17.2 (2000), pp. 247–276.

[46] Morriss, G. P. and Truant, D. P. "A review of the hydrodynamic Lyapunov modes of hard disk systems." *Journal of Physics A: Mathematical and Theoretical* 46.25 (2013), p. 254010. doi: 10.1088/1751-8113/46/25/254010.

[47] Neumann, J. V. *Functional Operators (AM-22), Volume 2: The Geometry of Orthogonal Spaces. (AM-22)*. Vol. v.2. Annals of mathematics studies. Princeton: Princeton University Press, 1951.

[48] Noethen, F. "A projector-based convergence proof of the Ginelli algorithm for covariant Lyapunov vectors." *Physica D: Nonlinear Phenomena* 396 (2019), pp. 18–34. doi: 10.1016/j.physd.2019.02.012.

[49] Noethen, F. "Computing covariant Lyapunov vectors in Hilbert spaces" (2019). arXiv: 1907.12458v1.

[50] Noethen, F. "Well-separating common complements of a sequence of subspaces of the same codimension in a Hilbert space are generic" (2019). arXiv: 1906.08514v1.

[51] Ochs, G. "Stability of Oseledets spaces is equivalent to stability of Lyapunov exponents." *Dynamics and Stability of Systems* 14.2 (1999), pp. 183–201. doi: 10.1080/026811199282056.

[52] Oseledets, V. I. “A multiplicative ergodic theorem. Characteristic Ljapunov, exponents of dynamical systems.” *Transactions of the Moscow Mathematical Society* 19 (1968), pp. 197–231.

[53] Ott, W. and Yorke, J. A. “Prevalence.” *Bulletin of the American Mathematical Society* 42.3 (2005), pp. 263–290. doi: 10.1090/S0273-0979-05-01060-8.

[54] Raghunathan, M. S. “A proof of Oseledec’s multiplicative ergodic theorem.” *Israel Journal of Mathematics* 32.4 (1979), pp. 356–362. doi: 10.1007/BF02760464.

[55] Ruelle, D. “Characteristic Exponents and Invariant Manifolds in Hilbert Space.” *The Annals of Mathematics* 115.2 (1982), pp. 243–290. doi: 10.2307/1971392.

[56] Saiki, Y. and Kobayashi, M. U. “Numerical identification of nonhyperbolicity of the Lorenz system through Lyapunov vectors.” *JSIAM Letters* 2 (2010), pp. 107–110. doi: 10.14495/jsiaml.2.107.

[57] Schubert, S. and Lucarini, V. “Covariant Lyapunov vectors of a quasi-geostrophic baroclinic model: analysis of instabilities and feedbacks.” *Quarterly Journal of the Royal Meteorological Society* 141.693 (2015), pp. 3040–3055. doi: 10.1002/qj.2588.

[58] Sharafi, N., Timme, M., and Hallerberg, S. “Critical transitions and perturbation growth directions.” *Physical Review E* 96.3 (2017), p. 032220. doi: 10.1103/PhysRevE.96.032220.

[59] Shinbrot, T. et al. “Chaos in a double pendulum.” *American Journal of Physics* 60.6 (1992), pp. 491–499. doi: 10.1119/1.16860.

[60] Stickler, B. A. and Schachinger, E. “The Double Pendulum.” In: *Basic Concepts in Computational Physics*. Ed. by B. A. Stickler and E. Schachinger. Cham: Springer, 2016, pp. 85–101. doi: 10.1007/978-3-319-27265-8_6.

[61] Takeuchi, K. A. et al. “Hyperbolic decoupling of tangent space and effective dimension of dissipative systems.” *Physical Review E* 84.4 (2011), p. 046214. doi: 10.1103/PhysRevE.84.046214.

[62] Thieullen, P. “Fibrés dynamiques asymptotiquement compacts exposants de Lyapounov. Entropie. Dimension.” *Annales de l’I.H.P. Analyse non linéaire* 4.1 (1987), pp. 49–97.

[63] Todd, A. R. “Covers by Linear Subspaces.” *Mathematics Magazine* 63.5 (1990), pp. 339–342. doi: 10.2307/2690910.

[64] Truant, D. P. and Morriss, G. P. “Backward and covariant Lyapunov vectors and exponents for hard-disk systems with a steady heat current.” *Physical Review E* 90.5 (2014), p. 052907. doi: 10.1103/PhysRevE.90.052907.

[65] Tucker, W. “A Rigorous ODE Solver and Smale’s 14th Problem.” *Foundations of Computational Mathematics* 2.1 (2002), pp. 53–117. doi: 10.1007/s002080010018.

[66] Vannitsem, S. and Lucarini, V. “Statistical and dynamical properties of covariant lyapunov vectors in a coupled atmosphere-ocean model—multiscale effects, geometric degeneracy, and error dynamics.” *Journal of Physics A: Mathematical and Theoretical* 49.22 (2016), p. 224001. doi: 10.1088/1751-8113/49/22/224001.

[67] Wolfe, C. L. and Samelson, R. M. “An efficient method for recovering Lyapunov vectors from singular vectors.” *Tellus A: Dynamic Meteorology and Oceanography* 59.3 (2007), pp. 355–366. doi: 10.1111/j.1600-0870.2007.00234.x.

[68] Xu, M. and Paul, M. R. “Covariant Lyapunov vectors of chaotic Rayleigh-Bénard convection.” *Physical Review E* 93.6 (2016), p. 062208. doi: 10.1103/PhysRevE.93.062208.

[69] Yang, H.-l. et al. “Hyperbolicity and the Effective Dimension of Spatially Extended Dissipative Systems.” *Physical Review Letters* 102.7 (2009), p. 074102. doi: 10.1103/PhysRevLett.102.074102.

List of Figures

List of Acronyms

CLV covariant Lyapunov vector
LE Lyapunov exponent
MET multiplicative ergodic theorem
ODE ordinary differential equation
SVD singular value decomposition

Index

Acknowledgments

I would like to thank my supervisor Reiner Lauterbach and co-supervisor Ingenuin Gasser for their kind support and for constant feedback. While giving me the freedom to develop and to follow my own research, they were always open for questions and discussions. Moreover, I want to thank my panel chair Jens Rademacher for his interest in my research and for our friendly conversations.

Let me also thank Sebastian Schubert for introducing me to CLVs and to Ginelli's algorithm. His applied work has always been a great motivation for me and bridged the gap between my purely mathematical background and the project TRR 181 "Energy Transfers in Atmosphere and Ocean". The project provided an interdisciplinary platform to share research with the common goal of improving climate models. Moreover, I am grateful for the financial support received from the DFG through this project.

Special thanks goes to Gary Froyland, who hosted my research stay at UNSW Sydney. It was very inspiring to meet one of the leading experts working in my field. I was happy to hear his thought on my research and left Australia with new motivation and ideas.

I am thankful to my office colleague Sören Schwenker for letting me interrupt him and for listening to my spontaneous ideas and thoughts during the last three years. Moreover, I want to thank my colleagues Alexander Lohse, Ivan Ovsyannikov, Mara Sommerfeld, Maxim Kirsebom, and Philipp Kunde for helpful discussions and conversations during our seminar and during lunch breaks.

Let me also thank my friends for providing the necessary counterweight to work. Last but not least, I wholeheartedly thank my family for all their support, loving care, and trust in me.

List of Publications and Preprints

[48] Noethen, F. "A projector-based convergence proof of the Ginelli algorithm for covariant Lyapunov vectors." *Physica D: Nonlinear Phenomena* 396 (2019), pp. 18–34. doi: `10.1016/j.physd.2019.02.012`.

[49] Noethen, F. "Computing covariant Lyapunov vectors in Hilbert spaces" (2019). arXiv: `1907.12458v1`.

[50] Noethen, F. "Well-separating common complements of a sequence of subspaces of the same codimension in a Hilbert space are generic" (2019). arXiv: `1906.08514v1`.

Eidesstattliche Versicherung
Declaration on Oath

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Weiterhin versichere ich, dass die Dissertation in keinem anderen Prüfungsverfahren eingereicht wurde.

I hereby declare, on oath, that I have written the presented dissertation by myself and have not used other than the acknowledged recourses and aids. Moreover, I declare that the dissertation has not been submitted in any other doctoral examination procedure.

Hamburg, den