



A projector-based convergence proof of the Ginelli algorithm for covariant Lyapunov vectors

Florian Noethen

Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany

HIGHLIGHTS

- First proof that treats the algorithm as a whole.
- Handles even degenerate Lyapunov spectra.
- Allows for different amounts of past and future data.
- Treats discrete and continuous data.
- Relates the speed of convergence to Lyapunov exponents.

ARTICLE INFO

Article history:

Received 18 May 2018
Received in revised form 20 November 2018
Accepted 25 February 2019
Available online 5 March 2019
Communicated by G. Froyland

Keywords:

Ginelli algorithm
Convergence proof
Covariant Lyapunov vectors
Lyapunov exponents

ABSTRACT

Linear perturbations of solutions of dynamical systems exhibit different asymptotic growth rates, which are naturally characterized by so-called covariant Lyapunov vectors (CLVs). Due to an increased interest of CLVs in applications, several algorithms were developed to compute them. The Ginelli algorithm is among the most commonly used. Although several properties of the algorithm have been analyzed, there exists no mathematically rigorous convergence proof yet.

In this article we extend existing approaches in order to construct a projector-based convergence proof of Ginelli's algorithm. One of the main ingredients will be an asymptotic characterization of CLVs via the Multiplicative Ergodic Theorem. In the proof, we keep a rather general setting allowing even for degenerate Lyapunov spectra.

© 2019 Elsevier B.V. All rights reserved.

1. Introduction

In this paper, we provide a projector-based convergence proof of the Ginelli algorithm [1], which computes the so-called CLVs. They form an intrinsic basis of the tangent space along a given reference trajectory and, thus, describe its local model. In fact, CLVs can be seen as a generalization to eigenvectors describing the local model of a steady-state.

By the Hartman–Grobman theorem the local model of a steady-state is linked with the original system. Eigenspaces correspond to invariant manifolds of the flow and eigenvalues indicate exponential growth/decay rates of perturbations of the equilibrium. Similar relations can be established for periodic orbits via Floquet theory.

In 1968 Oseledets managed to find a suitable generalization that goes beyond the analysis of steady-states and periodic orbits. In his celebrated *Multiplicative Ergodic Theorem* (MET) [2] the

long-term behavior of linear perturbations of arbitrary trajectories is explained. Similar to the case of steady-states, the tangent space is split into invariant subspaces that capture directions of different asymptotic growth rates. Instead of a splitting into eigenspaces, we obtain the *Oseledets splitting* with its corresponding *Lyapunov spectrum* consisting of *Lyapunov exponents* (LEs) as opposed to eigenvalues. In the nondegenerate case, i.e. if the Lyapunov spectrum is simple, the Oseledets spaces are one-dimensional and, hence, can be identified with a basis of vectors for each point of the trajectory. Those vectors are called *covariant Lyapunov vectors* (CLVs).

Despite their prominent role, it was not until a few years ago that first algorithms to compute CLVs were developed. Following Ginelli's algorithm [1] in 2007 several other approaches emerged [3–5], some of which are explained only for nondegenerate scenarios. With computational tools like Ginelli's algorithm at hand, CLVs became a frequent interest in applications. Amongst others, CLVs reveal structures in turbulent flows [6,7] and are used to analyze hard-disk systems [8–11] and climate

E-mail address: florian.noethen@uni-hamburg.de.

models [12–14]. Moreover, they constitute a hyperbolic decoupling of the tangent space of dissipative systems that extracts the physically relevant modes [15]. Furthermore, the angle between CLVs is used as an indicator for critical transitions in long-term behavior of solutions [16,17] and as a degree of hyperbolicity [6,18–20] in dynamical systems. However, despite the existence of numerous applications, many theoretical aspects of CLV-algorithms are still unexplored. This paper is a step to reducing the gap between theory and applications. Our goal is to verify convergence of Ginelli's algorithm by correcting and extending previous results.

In 1998 Ershov and Potapov investigated what could be called the first phase of Ginelli's algorithm, where past states of a reference point are explored to compute the fastest growing directions [21]. 15 years later Ginelli et al. built upon the work of Ershov and Potapov to formulate a convergence proof of their full algorithm [22]. They focused on a second phase, where future states are probed to obtain the fastest decaying directions. By a certain relation between both phases it is possible to extract the CLVs.

While [21] and [22] present fundamental ideas on convergence of Ginelli's algorithm, we find it necessary to be more precise in some arguments. In particular, [21] shows that almost all initial vectors propagated from present to future will align with left singular vectors of the propagator asymptotically. Then, an estimate for propagation from past to present is obtained by shifting the estimate for propagation from present to future. This argument requires a more detailed analysis of the $O(1)$ -terms as the condition on exceptional vectors that will not yield convergence depends on the starting point, which is neglected in [21]. In particular, the set of admissible initial vectors can be different for each starting point that is associated with a chosen runtime. Additionally, we find that both phases of Ginelli's algorithm should be treated as connected. Whereas, until now perfect convergence of the first phase was assumed to simplify the analysis of the second phase. However, despite the criticism, both papers are significant steps to better understand the connection between Oseledets MET and the Ginelli algorithm. In fact, they inspire many ideas presented here.

Our new convergence proof fills missing details and even extends the existing results. Unlike in a nondegenerate scenario, we do not pose any restrictions on the Lyapunov spectrum. Arbitrary dimensions for Oseledets spaces are allowed. Moreover, we distinguish between a discrete and continuous time version of the algorithm. It turns out that both versions converge, however the precise notion of convergence is different. Namely, the discrete time version of Ginelli's algorithm converges for almost all configurations of initial vectors, whereas the continuous time version only converges in measure. Furthermore, by incorporating the Lyapunov index notation we find an estimate for the speed of convergence. As already predicted and observed [4,14,21,22], the speed of convergence is exponential with a rate determined by the minimal distance of LEs.

The main part of our article is divided into three sections. Section 2 sets the notation and constructs tools needed for the convergence proof later on. A special interest lies in the evolution of vectors/subspaces in terms of distances and angles. In particular, the relation of propagated vectors to singular vectors is of importance, since singular vectors form directions of optimal growth rates for finite time. In Section 3 we present Ginelli's algorithm and state a deterministic version of Oseledets MET. By having a fairly general setting, we try to include as many scenarios as possible, though, we assume finite dimensional dynamics. With all preparations finished, we are in a position to precisely formulate and prove convergence of Ginelli's algorithm. The main work of Section 4 consists in assembling the tools obtained in Section 2, while the MET from Section 3 serves as an interface between evolution of singular vectors and CLVs.

2. Notation and tools

This section is primarily concerned with the evolution of vectors and subspaces. In order to keep track of the speed of convergence, we define the notion of a Lyapunov index. Next, we set up necessary notation to describe distances and angles of subspaces. In particular, we are interested in how those quantities change after applying a propagation map and after orthogonalization, e.g. the Gram–Schmidt procedure. An estimate of the rate of change is given based on a relation to singular vectors of the propagating linear map. As it turns out, there are configurations of vectors and subspaces that perform better than others. A distinction between them will be made by introducing a so-called admissibility parameter. Later on we will use the admissibility parameter to describe how well a configuration behaves in Ginelli's algorithm.

2.1. 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 so-called 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 an exponential time scale, which can be captured by the Lyapunov index notation.

Definition 2.1. The Lyapunov index $\lambda(f) \in \mathbb{R} \cup \{\pm\infty\}$ of a function $f : \mathbb{T}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is defined as the limit

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

Roughly speaking, the function f behaves similar to $e^{t\lambda(f)}$ on an exponential scale. 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 for the Lyapunov index, which can be found in Arnold's book [23] and are easily verified:

Proposition 2.2. Let $f, g : \mathbb{T}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$. 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. $\lambda(f + g) \leq \max(\lambda(f), \lambda(g))$,
7. $\lambda(fg) \leq \lambda(f) + \lambda(g)$ (if the right-hand side makes sense).

As the algorithm consists of two subsequent phases, the Lyapunov index is not enough for discussing Ginelli's 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

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

that allows two different runtimes. For this purpose, we extend the notion of a Lyapunov index to a formulation depending on two parameters.

Definition 2.3. The extended Lyapunov index $\bar{\lambda}(f) \in \mathbb{R} \cup \{\pm\infty\}$ of a function $f : \mathbb{T}_{\geq 0} \times \mathbb{T}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is defined as the limit

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

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

Proposition 2.4. Rules 1–7 of Proposition 2.2 hold true with λ replaced by $\bar{\lambda}$. Furthermore, if we extend a function $f : \mathbb{T}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ to $\bar{f} : \mathbb{T}_{\geq 0} \times \mathbb{T}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ by setting $\bar{f}(t_1, t_2) := f(t_1)$, then

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

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

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

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

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

Now, let \bar{f} be the extension of some function $f : \mathbb{T}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ as above. The relation $\lambda(f) \leq \bar{\lambda}(\bar{f})$ 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 right-hand side converging to $\lambda(f)$ for $t \rightarrow \infty$. \square

We demonstrate two exceptional cases where the function is growing/decaying either too slow or too fast to be captured by the notation.

Example 2.5. 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

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

and

$$\bar{\lambda}(g) = \bar{\lambda}\left(\frac{1}{\alpha}(\alpha^f)^2\right) = 2\bar{\lambda}(\alpha^f) \leq 2\bar{\lambda}(\alpha^{\min(t_1, t_2)^2}) = -\infty.$$

2.2. Orthogonal projections

We present some essential results about orthogonal projection. For most facts, we specifically refer to chapter 1.6 of Kato's book [24], the chapter on projections in Galántai's book [25] and the chapter by Deutsch [26].

Amongst others, orthogonal projections are a tool to describe geometric properties of subspaces. We associate a subspace $M \subset \mathbb{R}^d$ and its corresponding orthogonal projection P_M using the standard inner product. Through this identification we can define

distances and angles between subspaces, or even speak of converging sequences of subspaces. Since we focus on the euclidean norm $\|\cdot\|_2$, let us drop the subscript and simply write $\|\cdot\|$.³

Definition 2.6. 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 2.7. 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. $d(M, N) = d(M^\perp, N^\perp)$,
3. $d(M, N) < 1 \implies \dim(M) = \dim(N)$.

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

4. $d(M, N) = \|P_M P_{N^\perp}\|$,
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 2.8. 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 map A . For subspaces of different dimension, the inequality is trivially satisfied. So, let M and N be of the same dimension. We compute:

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

Here, A^* denotes the adjoint map of A with respect to the standard inner product. \square

The next concept needed is the (minimal) angle between two subspaces. A lot on this topic can be found in [26].

Definition 2.9. The cosine of the angle between M and N is given by

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

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

³ As all norms on \mathbb{R}^d are equivalent, quantities that are defined on an exponential scale remain the same. In particular, LEs and CLVs are independent of the chosen norm. Moreover, estimates on the exponential speed of convergence in our main theorems in Section 4 remain unchanged.

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 2.10. *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. $c_0(M, N) = \|P_M P_N\|$,
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 [27].

Theorem 2.11. *For each two subspaces M and N , the method of alternating projections converges:*

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

A discussion on the speed of convergence can be found in [26]. The following estimate will be enough for our purposes.

Proposition 2.12. *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 2.13. *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

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

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 $k \in \mathbb{N}$:

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

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

$$c(M', N') = c((M')^\perp, (N')^\perp)$$

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

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

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

For the last summand, we remark

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

Combining the above yields the desired estimate. \square

Now, assume we are given two converging sequences of subspaces $(M_t)_{t \in \mathbb{T}}$ and $(N_t)_{t \in \mathbb{T}}$ with transversal⁴ limits M and N . As an immediate consequence of Proposition 2.13 with the right choice of $k = k(t)$, we see that the sequence of intersections $(M_t \cap N_t)_{t \in \mathbb{T}}$ converges to the intersection of the limits $M \cap N$. Moreover, we show that the speed of convergence on an exponential scale can be preserved in a uniform manner.

Corollary 2.14. *Let $M, N \subset \mathbb{R}^d$ be two transversal subspaces. Moreover, assume $(\mathcal{M}_t)_{t \in \mathbb{T}}$ and $(\mathcal{N}_t)_{t \in \mathbb{T}}$ are two sequences of collections of subspaces that converge to M , resp. N , exponentially fast:*

$$\lambda_M := \lambda \left(\sup_{M' \in \mathcal{M}_t} d(M', M) \right) < 0$$

and

$$\lambda_N := \lambda \left(\sup_{N' \in \mathcal{N}_t} d(N', N) \right) < 0.$$

Then,

$$\bar{\lambda} \left(\sup_{M' \in \mathcal{M}_{t_1}} \sup_{N' \in \mathcal{N}_{t_2}} d(M' \cap N', M \cap N) \right) \leq \max(\lambda_M, \lambda_N).$$

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

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

for all $t_1, t_2 \geq T$. Invoking Proposition 2.13, we get

$$\sup_{M' \in \mathcal{M}_{t_1}} \sup_{N' \in \mathcal{N}_{t_2}} d(M' \cap N', M \cap N)$$

⁴ 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 $c_0(M^\perp, N^\perp) < 1$.

$$\leq \delta^{2k-1} + \left(\frac{1+\delta}{2}\right)^{2k-1} + k \left(\sup_{M' \in \mathcal{M}_{t_1}} d(M', M) + \sup_{N' \in \mathcal{N}_{t_2}} d(N', N) \right)$$

with 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 2.4 and Example 2.5 we compute

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

2.3. Singular value decomposition

We assume degeneracies $d_1 + \dots + d_p = d$ with $d_i \geq 1$ to be given. The case $p = d$ is called *nondegenerate*. Moreover, the standard basis of \mathbb{R}^d is denoted by

$$(e) := (e_{1_1}, e_{1_2}, \dots, e_{1_{d_1}}, e_{2_1}, \dots, e_{2_{d_2}}, \dots, e_{p_1}, \dots, e_{p_{d_p}}).$$

In the nondegenerate case, we drop the subindex, i.e. $(e) = (e_1, \dots, e_d)$. Both cases can be translated into each other via $e_{i_k} = e_{d_1 + \dots + d_{i-1} + k}$. To further shorten notation, we write (Ae) for the d -tuple of vectors we get from applying a linear map A to each vector of (e) .

Definition 2.15. Let $A \in \mathbb{R}^{d \times d}$. The singular value decomposition (SVD) of A is given by

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

where

$$\Sigma = \text{diag}(\sigma_1, \dots, \sigma_{p_{dp}})$$

is the diagonal matrix of *singular values* $\sigma_{i_k} \geq 0$ 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*.

A connection between left and right singular vectors is established via

$$Av_{i_k} = \sigma_{i_k} u_{i_k}.$$

In general, the SVD is not unique. Given $A \in \text{Gl}(d, \mathbb{R})$ we settle for the following ordering:

$$\sigma_{1_1} \geq \dots \geq \sigma_{p_{dp}} > 0. \quad (1)$$

Later on, every group of singular values will correspond to a different LE. Hence, the inequalities between $\sigma_{i_{d_i}}$ and $\sigma_{(i+1)_1}$ will eventually be strict. In that case, the spaces spanned by singular

vectors of one group, i.e. $\text{span}(u_{i_1}, \dots, u_{i_{d_i}})$ and $\text{span}(v_{i_1}, \dots, v_{i_{d_i}})$, are uniquely determined independent of our choice of SVD with Eq. (1).

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

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

$$\sigma_i^{\min} := \min_{k=1, \dots, d_i} \sigma_{i_k} \quad \text{and} \quad \sigma_i^{\max} := \max_{k=1, \dots, d_i} \sigma_{i_k}.$$

2.4. Gram–Schmidt procedure

We define the Gram–Schmidt procedure for subspaces. To this end, let $U_1 \oplus \dots \oplus U_p$ be a decomposition of \mathbb{R}^d into subspaces of dimension $\dim U_i = d_i$. Inductively, set

$$F_i := \bigoplus_{j=1}^i U_j \cap \left(\bigoplus_{j=1}^{i-1} U_j \right)^\perp$$

for $i = 1, \dots, p$. Then, $\mathbb{R}^d = F_1 \oplus \dots \oplus F_p$ is a decomposition with $\dim F_i = d_i$, $F_i \perp F_j$ for $i \neq j$, and with

$$\bigoplus_{j=1}^i U_j = \bigoplus_{j=1}^i F_j$$

for all i . Actually, the outcome only depends on the filtration

$$\{0\} \subset \bar{U}_1 \subset \bar{U}_2 \subset \dots \subset \bar{U}_p = \mathbb{R}^d$$

given by

$$\bar{U}_i := \bigoplus_{j=1}^i U_j.$$

In later scenarios the above spaces are spanned by groups of vectors. Thus, for a given basis (b) , set $U_i^{(b)}$ as the span of $b_{i_1}, \dots, b_{i_{d_i}}$. From $U_i^{(b)}$ we get $\bar{U}_i^{(b)}$ and $F_i^{(b)}$. The associated orthogonal projection onto $F_i^{(b)}$ will be denoted by

$$P_i^{(b)} := P_{F_i^{(b)}}.$$

It follows that

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

is the orthogonal projection onto $\bar{U}_i^{(b)}$. Another consequence of our notation is the relation

$$A(\bar{U}_i^{(b)}) = \bar{U}_i^{(Ab)}$$

for an invertible linear map A .

2.5. Admissibility

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 application of linear maps and the Gram–Schmidt procedure, we relate it to singular vectors. That relation is represented by a single parameter δ . It describes how strong the corresponding filtrations are correlated. Here, a value of 0

means no correlation and a value of 1 implies equality. Thus, we call tuples that have a certain level of correlation admissible. A special task will be to understand how many tuples are at least δ -admissible. For this purpose, we denote by μ the Lebesgue-measure for the respective dimension.

Definition 2.16. Let $0 < \delta \leq 1$ and a basis (c) of \mathbb{R}^d be given. A d -tuple (b) is called δ -admissible with respect to (c) if it is linearly independent and

$$\forall i < p : d(\bar{U}_i^{(b)}, \bar{U}_i^{(c)})^2 \leq 1 - \delta^2.$$

We denote the set of all δ -admissible tuples by $\mathcal{Ad}^{(c)}(\delta)$ and the set of all tuples that are admissible for some $\delta > 0$ by $\mathcal{Ad}^{(c)}$.

As admissibility is described by distances of filtration spaces, we are allowed to interchange the involved tuples with their Gram–Schmidt bases. So, let us assume (c) to be an ONB from now on. Moreover, the invariance of distances under orthogonal transformations implies that δ -admissibility of (b) w.r.t. (c) is equivalent to δ -admissibility of (Vb) w.r.t. (Vc) for all $V \in O(d, \mathbb{R})$. Hence, $V^d(\mathcal{Ad}^{(c)}(\delta))$ and $\mathcal{Ad}^{(Vc)}(\delta)$ coincide.

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

Lemma 2.17. A basis (b) is δ -admissible w.r.t. (c) if, and only if, for all $i < p$ and $x \in \bar{U}_i^{(b)}$ with $\|x\| = 1$, we have

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

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

$$\begin{aligned} \|(I - \bar{P}_i^{(c)}) \bar{P}_i^{(b)}\|^2 &= \max_{\|x\|=1} \|(I - \bar{P}_i^{(c)}) \bar{P}_i^{(b)} x\|^2 \\ &= \max_{\substack{x \in \bar{U}_i^{(b)} \\ \|x\|=1}} \|(I - \bar{P}_i^{(c)}) x\|^2 \\ &= 1 - \min_{\substack{x \in \bar{U}_i^{(b)} \\ \|x\|=1}} \|\bar{P}_i^{(c)} x\|^2 \\ &= 1 - \min_{\substack{x \in \bar{U}_i^{(b)} \\ \|x\|=1}} \sum_{j=1}^i \sum_k |\langle x, c_{j_k} \rangle|^2. \quad \square \end{aligned}$$

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 2.18. Let $A = U \Sigma V^T$ be invertible and $0 < \delta \leq 1$. For all $(b) \in \mathcal{Ad}^{(v)}(\delta)$, it holds

$$\forall i : d(\bar{U}_i^{(Ab)}, \bar{U}_i^{(u)}) \leq \frac{1}{\delta} \frac{\sigma_{i+1}^{\max}}{\sigma_{i+1}^{\min}}.$$

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

$$Ax = \sum_{j_k} \langle x, v_{j_k} \rangle \sigma_{j_k} u_{j_k} \Rightarrow \|Ax\|^2 = \sum_{j_k} |\langle x, v_{j_k} \rangle|^2 \sigma_{j_k}^2.$$

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

$$\begin{aligned} \|Ax\|^2 &\geq \sum_{j=1}^i \sum_k |\langle x, v_{j_k} \rangle|^2 \sigma_{j_k}^2 \\ &\geq (\sigma_i^{\min})^2 \sum_{j=1}^i \sum_k |\langle x, v_{j_k} \rangle|^2 \geq \delta^2 (\sigma_i^{\min})^2 \end{aligned}$$

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

$$\|(I - \bar{P}_i^{(u)}) Ax\|^2 = \sum_{j>i} \sum_k |\langle x, v_{j_k} \rangle|^2 \sigma_{j_k}^2 \leq (\sigma_{i+1}^{\max})^2.$$

Now, we compute:

$$\begin{aligned} d(\bar{U}_i^{(Ab)}, \bar{U}_i^{(u)}) &= \|(I - \bar{P}_i^{(u)}) \bar{P}_i^{(Ab)}\| \\ &= \max_{y \in \bar{U}_i^{(Ab)} \setminus \{0\}} \frac{\|(I - \bar{P}_i^{(u)}) y\|}{\|y\|} \\ &= \max_{x \in \bar{U}_i^{(b)} \setminus \{0\}} \frac{\|(I - \bar{P}_i^{(u)}) Ax\|}{\|Ax\|} \\ &= \max_{\substack{x \in \bar{U}_i^{(b)} \\ \|x\|=1}} \frac{\|(I - \bar{P}_i^{(u)}) Ax\|}{\|Ax\|} \\ &\leq \frac{1}{\delta} \frac{\sigma_{i+1}^{\max}}{\sigma_i^{\min}}. \quad \square \end{aligned}$$

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

$$\mathcal{Ad}^{(c)} = \{(b) \text{ basis} \mid \forall i : d(\bar{U}_i^{(b)}, \bar{U}_i^{(c)}) < 1\} \subset (\mathbb{R}^d)^d$$

must be a set of measure zero. Using Proposition 2.7, we can rewrite the condition as follows:

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

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

$$d(\bar{U}_i^{(b)}, \bar{U}_i^{(c)}) < 1 \iff \det(b_{1_1}, \dots, b_{i_d}, c_{(i+1)_1}, \dots, c_{p_{dp}}) \neq 0.$$

This form easily reveals the following:

Proposition 2.19. The set of nonadmissible tuples $(\mathbb{R}^d)^d \setminus \mathcal{Ad}^{(c)}$ 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 all $k \geq 1$. \square

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

Corollary 2.20. For each subset $\mathcal{F} \subset (\mathbb{R}^d)^d$ of finite Lebesgue-measure, it holds

$$\lim_{\delta \searrow 0} \mu(\mathcal{F} \setminus \mathcal{Ad}^{(c)}(\delta)) = 0.$$

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

$$\begin{aligned} \lim_{\delta \searrow 0} \mu(\mathcal{F} \setminus \mathcal{Ad}^{(c)}(\delta)) &= \mu\left(\bigcap_{0 < \delta \leq 1} \mathcal{F} \setminus \mathcal{Ad}^{(c)}(\delta)\right) \\ &= \mu(\mathcal{F} \setminus \mathcal{Ad}^{(c)}) = 0. \quad \square \end{aligned}$$

In the second part of Ginelli's algorithm, we need a special domain for initial tuples (b). Namely, we look at

$$b_{i_1}, \dots, b_{i_{d_i}} \in \text{span}(c_{i_1}, \dots, c_{p_{d_p}}) = U_i^{(c)} \oplus \dots \oplus U_p^{(c)}.$$

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

$$(*, \dots, *, b_{i_1}, \dots, b_{i_{d_i}}, *, \dots, *) \in \mathcal{Ad}^{(c)}(\delta)$$

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

As before, one readily checks that $V^d(\mathcal{Ad}_{\text{ext}}^{(c)}(\delta)) = \mathcal{Ad}_{\text{ext}}^{(Vc)}(\delta)$ for $V \in O(d, \mathbb{R})$. Moreover, we again conclude that almost all tuples satisfy extendable admissibility.

Proposition 2.21. *The set*

$$\begin{aligned} &\left(\left(U_1^{(c)} \oplus \dots \oplus U_p^{(c)} \right)^{d_1} \times \left(U_2^{(c)} \oplus \dots \oplus U_p^{(c)} \right)^{d_2} \times \dots \right. \\ &\quad \left. \times \left(U_p^{(c)} \right)^{d_p} \right) \setminus \mathcal{Ad}_{\text{ext}}^{(c)} \end{aligned}$$

has Lebesgue-measure zero.

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

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

not satisfying the extension property has Lebesgue-measure zero.

The idea is to apply Proposition 2.19 to a reduced setting for fixed i . To this end, look at $\mathbb{R}^{d'}$ with degeneracies $d' = d_1' + \dots + d_{p'}'$, given by $d_j' := d_{i-1+j}$ for all $j = 1, \dots, p' := p + 1 - i$, and let (e') be its standard basis. We get

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

In particular, this implies

$$\mu\left(\left(\mathbb{R}^{d'}\right)^{d'} \setminus \left\{ \left(b'_{i_1}, \dots, b'_{i_{d'_i}} \right) \text{ has admissible extension} \right\} \right) = 0.$$

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

$$d\left(\text{span}(b_{i_1}, \dots, b_{i_{d_i}}), U_i^{(c)}\right)^2 \leq 1 - \delta^2$$

and

$$\begin{aligned} \forall j > i: d\left(\text{span}(b_{i_1}, \dots, b_{i_{d_i}}, g_{(i+1)_1}, \dots, g_{j_{d_j}}), \right. \\ \left. U_i^{(c)} \oplus \dots \oplus U_j^{(c)}\right)^2 \leq 1 - \delta^2. \end{aligned}$$

We can extend such a tuple

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

to a δ -admissible tuple (g) by setting $g_{j_k} := c_{j_k}$ for $j < i$. This concludes the proof. \square

As a consequence, we get the following corollary:

Corollary 2.22. *Given a subset $\mathcal{F} \subset \left(U_1^{(c)} \oplus \dots \oplus U_p^{(c)} \right)^{d_1} \times \dots \times \left(U_p^{(c)} \right)^{d_p}$ of finite Lebesgue-measure, it holds*

$$\lim_{\delta \searrow 0} \mu\left(\mathcal{F} \setminus \mathcal{Ad}_{\text{ext}}^{(c)}(\delta)\right) = 0.$$

In the discrete time convergence proof of Ginelli's algorithm, a more precise measure-estimate on non- δ -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 2.5 will be devoted to a rather technical derivation of explicit estimates needed only for the proof of Theorem 4.2.

Proposition 2.23. *Let $d > 1$. There is a constant $\eta = \eta(d, M) > 0$ such that*

$$\mu\left(B_d(0, M)^d \setminus \mathcal{Ad}^{(c)}(\delta)\right) \leq \eta \delta^{\frac{1}{d-1}}.$$

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

Lemma 2.24. *Let (f) be an ONB of \mathbb{R}^d . Fix $1 < i < d$ and $0 < \delta_1, \delta_2 \leq 1$. If*

$$\|P_{\text{span}(f_1, \dots, f_{i-1}, c_{i+1}, \dots, c_d)} f_i\|^2 \leq 1 - \delta_1^2$$

and

$$\|\bar{P}_{i-1}^{(f)}(I - \bar{P}_i^{(c)})\|^2 \leq 1 - \delta_2^2,$$

then

$$d\left(\bar{U}_i^{(f)}, \bar{U}_i^{(c)}\right)^2 \leq 1 - (\delta_1 \delta_2)^2.$$

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

$$\begin{aligned} \|\bar{P}_i^{(f)}(I - \bar{P}_i^{(c)})\|^2 &= \|\bar{P}_i^{(f)} c'_3\|^2 = \|\bar{P}_{i-1}^{(f)} c'_3\|^2 + |\langle f_i, c'_3 \rangle|^2 \\ &= |\langle f'_1, c'_3 \rangle|^2 + |\langle f'_2, c'_3 \rangle|^2 \end{aligned}$$

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

$$\|P_{\text{span}(f'_1, c'_3)} f'_2\|^2 \leq \|P_{\text{span}(f_1, \dots, f_{i-1}, c_{i+1}, \dots, c_d)} f_i\|^2 \leq 1 - \delta_1^2$$

and

$$|\langle f'_1, c'_3 \rangle|^2 \leq \|\bar{P}_{i-1}^{(f)} c'_3\|^2 \leq \|\bar{P}_{i-1}^{(f)}(I - \bar{P}_i^{(c)})\|^2 \leq 1 - \delta_2^2.$$

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

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

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

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

$$\|P_{\text{span}(f'_1, c'_3)} f'_2\|^2 = |\langle f'_1, f'_2 \rangle|^2 + \left\| \frac{c'_3 - \langle f'_1, c'_3 \rangle f'_1}{\|c'_3 - \langle f'_1, c'_3 \rangle f'_1\|}, f'_2 \right\|^2$$

$$\begin{aligned}
&= \frac{|\langle c'_3, f'_2 \rangle|^2}{\|c'_3 - \langle f'_1, c'_3 \rangle f'_1\|^2} \\
&= \frac{|\langle c'_3, f'_2 \rangle|^2}{1 - |\langle f'_1, c'_3 \rangle|^2}.
\end{aligned}$$

Thus, by our assumptions:

$$\begin{aligned}
|\langle f'_2, c'_3 \rangle|^2 &= \|P_{\text{span}(f'_1, c'_3)} f'_2\|^2 (1 - |\langle f'_1, c'_3 \rangle|^2) \\
&\leq (1 - \delta_1^2)(1 - |\langle f'_1, c'_3 \rangle|^2).
\end{aligned}$$

We estimate:

$$\begin{aligned}
|\langle f'_1, c'_3 \rangle|^2 + |\langle f'_2, c'_3 \rangle|^2 &\leq |\langle f'_1, c'_3 \rangle|^2 + (1 - \delta_1^2)(1 - |\langle f'_1, c'_3 \rangle|^2) \\
&= 1 - \delta_1^2 + \delta_1^2 |\langle f'_1, c'_3 \rangle|^2 \\
&\leq 1 - \delta_1^2 + \delta_1^2(1 - \delta_2^2) \\
&= 1 - (\delta_1 \delta_2)^2. \quad \square
\end{aligned}$$

The previous lemma can be used to give a sufficient condition for a tuple to be δ -admissible.

Lemma 2.25. *If a basis (b) satisfies*

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

where $(f) := \mathcal{GS}(b)$, then (b) is δ -admissible.

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

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

For $i = 1$, we have

$$\begin{aligned}
d\left(\overline{U}_1^{(f)}, \overline{U}_1^{(c)}\right)^2 &= \left\| \left(I - \overline{P}_1^{(c)}\right) f_1 \right\|^2 = \|P_{\text{span}(c_2, \dots, c_d)} f_1\|^2 \\
&\leq 1 - \left(\delta^{\frac{1}{d-1}}\right)^2.
\end{aligned}$$

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

$$\begin{aligned}
\left\| \overline{P}_{i-1}^{(f)} \left(I - \overline{P}_i^{(c)}\right) \right\|^2 &\leq \left\| \overline{P}_{i-1}^{(f)} \left(I - \overline{P}_{i-1}^{(c)}\right) \right\|^2 = d\left(\overline{U}_{i-1}^{(f)}, \overline{U}_{i-1}^{(c)}\right)^2 \\
&\leq 1 - \left(\delta^{\frac{1}{d-1}}\right)^2.
\end{aligned}$$

Simply apply Lemma 2.24 to close the induction step. \square

Now, we prove the proposition.

Proof of Proposition 2.23. Set $\tilde{\delta} := \delta^{\frac{1}{d-1}}$ and let

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

be the set of all nonadmissible vector tuples inside $B_d(0, M)^d$. From Proposition 2.19 we know that \mathcal{N} has measure zero. On its complement we define a continuous mapping into the d -fold product of spheres:

$$w: B_d(0, M)^d \setminus \mathcal{N} \rightarrow (S^{d-1})^d$$

with components

$$w_i(b_1, \dots, b_d) := \mathcal{GS}_d(b_1, \dots, b_{i-1}, c_{i+1}, \dots, c_d, c_i),$$

where \mathcal{GS}_d is the last component of the Gram–Schmidt procedure. By construction $w_i = w_i(b_1, \dots, b_d)$ is the unique unit-vector orthogonal to

$$\text{span}(b_1, \dots, b_{i-1}, c_{i+1}, \dots, c_d)$$

with $\langle w_i, c_i \rangle > 0$, and only depends on the first $i - 1$ vectors of (b). w will help us to measure sets of admissible vectors.

The Gram–Schmidt basis of (b) is constructed by setting $f_i := \frac{b'_i}{\|b'_i\|}$ with $b'_i := \left(I - \overline{P}_{i-1}^{(b)}\right) b_i$. Assuming $|\langle w_i, b_i \rangle| \geq M\tilde{\delta}$, we get

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

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

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

then (b) is δ -admissible by Lemma 2.25. In particular, the subset of all non- δ -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:

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

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 $B_d(0, M) \cap ((-M\tilde{\delta}, M\tilde{\delta}) \times \mathbb{R}^{d-1})$ to $(-M\tilde{\delta}, M\tilde{\delta}) \times (-M, M)^{d-1}$ for a simple 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- δ -admissible tuples inside the special domain.

Proposition 2.26. Let $d > 1$. There is a constant $\eta = \eta(d, M) > 0$ such that

$$\mu \left(B(M) \setminus \mathcal{Ad}_{\text{ext}}^{(c)}(\delta) \right) \leq \eta \delta^{\frac{1}{d-1}},$$

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

$$B_d(0, M)^{d_1} \times \cdots \times B_{d_p}(0, M)^{d_p} \subset \left(U_1^{(c)} \oplus \cdots \oplus U_p^{(c)} \right)^{d_1} \times \cdots \times \left(U_p^{(c)} \right)^{d_p}.$$

Proof. The proof is similar to the one of Proposition 2.21. Again, it is enough to find such a bound for the set of all tuples in

$$B_{d_1+\dots+d_p}(0, M)^{d_i} \subset \left(U_i^{(c)} \oplus \cdots \oplus U_p^{(c)} \right)^{d_i}$$

that cannot be extended to a δ -admissible tuple.

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

$$B_{d'}(0, M)^{d'_i} \setminus \left\{ \left(b'_{1_1}, \dots, b'_{1_{d'_i}} \right) \text{ has a } \delta\text{-admissible extension} \right\}.$$

Proposition 2.23 yields η' only depending on d' and M with

$$\mu \left(B_{d'}(0, M)^{d'_i} \setminus \mathcal{Ad}^{(e')}(\delta) \right) \leq \eta' \delta^{\frac{1}{d'-1}}.$$

This implies

$$\begin{aligned} & \mu \left(B_{d'}(0, M)^{d'_i} \setminus \left\{ \left(b'_{1_1}, \dots, b'_{1_{d'_i}} \right) \right. \right. \\ & \quad \left. \left. \text{has a } \delta\text{-admissible extension} \right\} \right) \\ & \leq \left(\frac{1}{\text{vol}(B_{d'}(0, M))} \right)^{d'-d'_i} \eta' \delta^{\frac{1}{d'-1}}. \end{aligned}$$

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

3. Ginelli's algorithm

In this section we define a minimalistic setting suitable for both the MET and Ginelli's algorithm.

3.1. Setting and multiplicative ergodic theorem

Since we want to cover as many applications for Ginelli's algorithm as possible, we do not specify a type of state space or system. Instead, we assume a non-empty set $\Omega = \{\theta_t \omega_0 \mid t \in \mathbb{T}\}$ to be the abstract orbit of our state of interest ω_0 respective to the flow $(\theta_t)_{t \in \mathbb{T}}$. Here, $\theta_t : \Omega \rightarrow \Omega$ represents the time- t -flow on our orbit. The flow should satisfy $\theta_0 = \text{id}_\Omega$ and $\theta_{s+t} = \theta_s \theta_t$. Remaining information of the linear model is encoded in a cocycle $\Phi(t, \omega)$ assigning a timestep t and a state ω to the linear propagator on tangent space from ω to $\theta_t \omega$.

Definition 3.1. A map $\Phi : \mathbb{T} \times \Omega \rightarrow \mathbb{R}^{d \times d}$ is called a (linear) cocycle (over θ) if

1. $\Phi(0, \omega) = \text{id}$,
2. $\Phi(s+t, \omega) = \Phi(s, \theta_t \omega) \Phi(t, \omega)$,

for all $s, t \in \mathbb{T}$ and $\omega \in \Omega$.

Since \mathbb{T} is two-sided, every cocycle is pointwise invertible with inverse

$$\Phi(t, \omega)^{-1} = \Phi(-t, \theta_t \omega).$$

The Multiplicative Ergodic Theorem of Oseledets [2] not only gives us existence of CLVs, but will play a crucial role in our convergence proof. We state a deterministic version found in [23]. It assumes that changes during a short timestep do not matter on an exponential scale and, furthermore, that expansion rates of different volumes are well-defined and do not exceed the exponential scale.

Proposition 3.2 (Deterministic MET). Let Φ be a cocycle satisfying

$$\lambda \left(\sup_{s \in [0, 1] \cap \mathbb{T}} \|\Phi(s, \theta_t \omega_0)^{\pm 1}\| \right) \leq 0$$

and assume that

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|\wedge^i \Phi(t, \omega_0)\| \in \mathbb{R} \cup \{-\infty\}$$

exists for all orders of the wedge product of $\Phi(t, \omega_0)$. Then, there exists a Lyapunov spectrum with a corresponding filtration capturing subspaces of different 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 $d_1 + \cdots + d_p = d$:

$$\forall i : \forall k = 1, \dots, d_i : \lambda_i = \lim_{t \rightarrow \infty} \frac{1}{t} \log \sigma_{i_k}(\Phi(t, \omega_0)).$$

2. There is a filtration

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

given by subspaces

$$V_i := \left\{ x \in \mathbb{R}^d \mid \lim_{t \rightarrow \infty} \frac{1}{t} \log \|\Phi(t, \omega_0)x\| \leq \lambda_i \right\}.$$

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 proposition only requires one-sided time and an invertible cocycle to provide the Lyapunov spectrum and filtration at state ω_0 of the orbit. However, since we assumed two-sided time, we immediately get the existence of these quantities for all states along the orbit.

Corollary 3.3. In the setting of Proposition 3.2 Lyapunov spectrum and filtration are defined for all $\omega \in \Omega$. Furthermore, $p(\omega)$, $\lambda_i(\omega)$ and $d_i(\omega)$ are independent of ω , and the filtration changes in a covariant way:

$$\Phi(t, \omega)V_i(\omega) = V_i(\theta_t \omega).$$

Proof. The first assumption of Proposition 3.2 is trivially satisfied if we replace ω_0 by $\omega = \theta_u \omega_0$. To prove the second assumption, we use the following properties of the wedge product, which can be found in [23]:

1. $\|\wedge^i A\| = \sigma_1(A) \cdots \sigma_i(A)$,
2. $\|\wedge^i(AB)\| \leq \|\wedge^i A\| \|\wedge^i B\|$,

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

$$\lim_{t \rightarrow \infty} \frac{1}{t} \log \|\wedge^i \Phi(t, \theta_u \omega_0)\| < \infty$$

follows due to the cocycle property:

$$\left(\frac{t+u}{t} \right) \left(\frac{1}{t+u} \log \|\wedge^i \Phi(t+u, \omega_0)\| \right) - \frac{1}{t} \log \|\wedge^i \Phi(u, \omega_0)\|$$

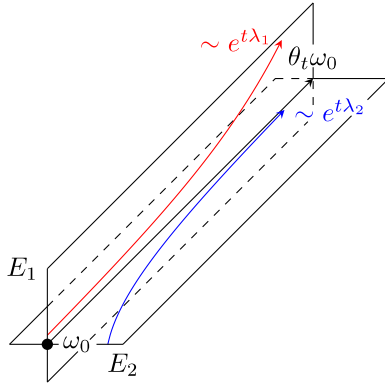


Fig. 1. Diagonal cocycle for dimension 2.

$$\begin{aligned} &\leq \frac{1}{t} \log \|\wedge^i \Phi(t, \theta_u \omega_0)\| \\ &\leq \left(\frac{t+u}{t}\right) \left(\frac{1}{t+u} \log \|\wedge^i \Phi(t+u, \omega_0)\|\right) \\ &\quad + \frac{1}{t} \log \|\wedge^i \Phi(-u, \theta_u \omega_0)\|. \end{aligned}$$

Thus, the proposition gives us the existence of a Lyapunov spectrum and filtration at state $\omega = \theta_u \omega_0$. In particular, the above shows that limits of singular values for ω and ω_0 coincide on an exponential scale. Hence, the Lyapunov exponents and their multiplicities are the same for ω_0 and ω . Finally, the identity for filtrations spaces follows from the definition. \square

Similar statements can be derived for the time-reversed cocycle $\Phi^-(t, \omega) := \Phi(-t, \omega)$ over the time-reversed flow $\theta_t^- := \theta_{-t}$. We denote its Lyapunov spectrum by $(\lambda_i^-, d_i^-)_{i=1, \dots, p^-}$ and the corresponding filtration spaces by $V_i^-(\omega)$.

In order to define a covariant splitting of the tangent space that captures asymptotic growth rates in both forward and backward time, we require additional assumptions on Lyapunov spectra and associated splittings of Φ and Φ^- :

1. $p = p^-$, $d_i^- = d_{p+1-i}$ and $\lambda_i^- = -\lambda_{p+1-i}$,
2. $V_{i+1}(\omega_0) \cap V_{p+1-i}^-(\omega_0) = \{0\}$.

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

Proposition 3.4. *Assuming the above relations between the Lyapunov spectra of Φ and Φ^- , there exists a splitting $\mathbb{R}^d = E_1(\omega) \oplus \dots \oplus E_p(\omega)$ of the tangent space into so-called Oseledets spaces*

$$E_i(\omega) := V_i(\omega) \cap V_{p+1-i}^-(\omega). \quad (2)$$

Furthermore, Oseledets spaces can be characterized via

$$x \in E_i(\omega) \setminus \{0\} \iff \lim_{t \rightarrow \pm\infty} \frac{1}{|t|} \log \|\Phi(t, \omega)x\| = \pm\lambda_i, \quad (3)$$

are covariant

$$\Phi(t, \omega)E_i(\omega) = E_i(\theta_t \omega),$$

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

Proof. The proof is purely algebraic and can be found along the lines of the proof of the MET for two-sided time in [23]. \square

For random dynamical systems satisfying a particular integrability condition, it is shown in [23] that the cocycle along almost all orbits of the system admits an Oseledets splitting. Moreover, in an ergodic setting the Lyapunov spectrum coincides

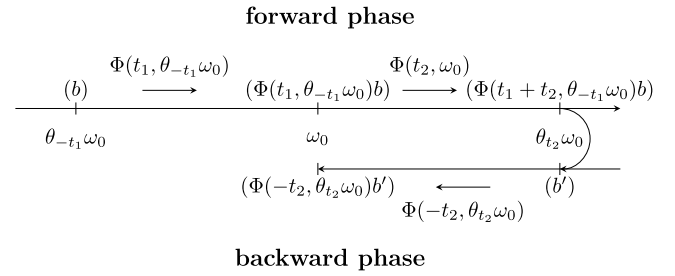


Fig. 2. Schematic picture of the Ginelli algorithm.

for almost all orbits. Therefore, in applications it is often assumed that the underlying system is ergodic at least near an interesting structure.⁵ Via CLVs one hopes to better understand the local flow around that structure.

Definition 3.5. Normalized basis vectors, which are covariant and chosen subject to the Oseledets splitting for each $\omega \in \Omega$, are called *covariant Lyapunov vectors (CLVs)*.

CLVs represent directions of different asymptotic growth rates⁶ by Eq. (3). However, they are uniquely defined (up to sign) only for nondegenerate Lyapunov spectra.

3.2. The algorithm

The Ginelli algorithm [1,22] computes Oseledets spaces (or CLVs) for a given cocycle by using its asymptotic characterization Eq. (3). The main idea is that each vector with a nonzero E_1 -part will approach E_1 asymptotically, since its E_1 -component has the largest exponential growth rate. More abstractly, almost all $(d_1 + \dots + d_i)$ -dimensional subspaces will align with $E_1 \oplus \dots \oplus E_i$, the fastest expanding (or slowest contracting) subspace of the corresponding dimension, in forward time. Reversing time, we are able to extract the slowest expanding (or fastest contracting) subspaces. In particular, almost all d_i -dimensional subspaces of $E_1 \oplus \dots \oplus E_i$ will align with E_i in backward time.

Taking these traits into consideration, the abstract formalism of Ginelli's algorithm is as follows: (also see Fig. 2)

Ginelli Algorithm (analytical kernel)⁷

- 1.1. Randomly choose a basis (b) of the tangent space at a past state $\theta_{-t_1}\omega_0$ and propagate it forward until ω_0 . If the propagation time t_1 is chosen large enough, we expect $\Phi(t_1, \theta_{-t_1}\omega_0)\overline{U}_i^{(b)} = \overline{U}_i^{(\Phi(t_1, \theta_{-t_1}\omega_0)b)}$ to be a good approximation to $E_1(\omega_0) \oplus \dots \oplus E_i(\omega_0)$.
- 1.2. Continue propagating forward until a state $\theta_{t_2}\omega_0$ is reached. This state should be far enough in the future, so that we have a sufficiently good approximation to $E_1 \oplus \dots \oplus E_i$ on a long enough timeframe for the second phase.

⁵ See the concept of SRB-measures for attractors [28].

⁶ Aside from asymptotic growth, the angle between CLVs can be used as a measure of hyperbolicity (see, e.g., [6,18–20]).

⁷ The intended implementation of Ginelli's algorithm has a few more details (see [1,22]). To avoid that all vectors collapse onto the first Oseledets space in forward time, the propagated vectors are frequently orthonormalized via a QR-decomposition. Analytically, however, orthonormalizations do not change the filtration of subspaces. Hence, the resulting subspace approximations are the same independent of how often the vectors were corrected.

During phase 1.2 the R matrices are stored and later reused in phase 2. By expressing (b') as coefficients (α) with respect to forward propagated vectors appearing as columns of the Q matrices, it suffices to apply inverses of the R matrices to the coefficient matrix given by (α) . In-between propagation steps one normalizes the columns of (α) . Similar to the forward phase, propagated subspaces are not changed by the added details. Thus, analytical and numerical approximations coincide if computed with exact precision.

2. For each i , randomly choose d_i vectors $b'_{i_1}, \dots, b'_{i_{d_i}}$ in $\overline{U}_i^{(\phi(t_1+t_2, \theta_{-t_1}\omega_0)b)}$ $\approx E_1(\theta_{t_2}\omega_0) \oplus \dots \oplus E_i(\theta_{t_2}\omega_0)$ and propagate them backward until ω_0 . The evolved subspace, i.e. $U_i^{(\phi(-t_2, \theta_{t_2}\omega_0)b')}$, is our approximation to $E_i(\omega_0)$.

Since we propagate vectors forward, we call steps 1.1 and 1.2 *forward phase*, and by the same reasoning step 2 is called *backward phase*.

The asymptotic expansion rate $(\lambda_1 + \dots + \lambda_i)$ of $E_1 \oplus \dots \oplus E_i$ is usually computed as a byproduct of the forward phase of the algorithm. Using this information, we can derive the Lyapunov spectrum.⁸

In Section 4 we provide a convergence proof of the whole algorithm as $\min(t_1, t_2) \rightarrow \infty$. The speed of convergence turns out to be exponential in relation to the minimum distance of LEs. Furthermore, the kind of convergence differs between discrete and continuous time. The discrete version with $t_1, t_2 \in \mathbb{N}$ converges for almost all initial tuples, whereas the continuous version with $t_1, t_2 \in \mathbb{R}_{>0}$ only converges in measure.

3.2.1. Two examples

Next, we present two exemplary cocycles to make the reader familiar with the subtleties of convergence in Ginelli's algorithm.

Example 3.6 (Diagonal Cocycle, see Fig. 1). Assume $\Omega = \{\omega_0\}$ with trivial flow $\theta_t\omega_0 = \omega_0$. For given $\lambda_1 > \dots > \lambda_p$, define $D := \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_p})$. Then, $\Phi(t, \omega_0) := D^t$ is a cocycle and the CLVs (at ω_0) 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

$$\langle D^t b_1, e_i \rangle = \langle b_1, e_i \rangle e^{t\lambda_i}.$$

Thus, we compute

$$\left\| \frac{D^t b_1}{\|D^t b_1\|}, e_i \right\|^2 = \frac{|\langle b_1, e_i \rangle|^2 e^{2t\lambda_i}}{\sum_j |\langle b_1, e_j \rangle|^2 e^{2t\lambda_j}} = \frac{|\langle b_1, e_i \rangle|^2 e^{-2t|\lambda_1 - \lambda_i|}}{\sum_j |\langle b_1, e_j \rangle|^2 e^{-2t|\lambda_1 - \lambda_j|}}.$$

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:

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

In general, it holds

$$\lambda \left(d \left(\overline{U}_i^{(D^t b)}, \overline{U}_i^{(e)} \right) \right) \leq -|\lambda_i - \lambda_{i+1}|$$

for all tuples (b) that are admissible w.r.t. (e) . A more general statement for arbitrary cocycles will be proved in Section 4.3 when analyzing convergence of the forward phase.

Ginelli's algorithm starts with a random choice of initial vectors to prevent nonadmissible configurations. One such configuration would be the unlikely case where the first vector lies in

the second Oseledets space. As Oseledets spaces are covariant, the first vector will stay inside the second Oseledets space when propagated. Consequently, it will not be a good representation of a vector lying in the first Oseledets space. The next example shows that all vectors might be nonadmissible when initiated at a wrong time in the continuous version of Ginelli's algorithm.

Example 3.7 (Rotating Oseledets Spaces). Let $\Omega := S^1 \cong \mathbb{R}/\mathbb{Z}$ be a periodic orbit with homogeneous flow $\theta_t\omega := \omega + t$. Furthermore, let $R : \mathbb{R} \rightarrow \text{SO}(2)$ be the parametrization of $\text{SO}(2)$ 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 $D := \text{diag}(e^{\lambda_1}, e^{\lambda_2})$ for some $\lambda_1 > \lambda_2$, and define the cocycle to be

$$\Phi(t, \omega) := R(\theta_t\omega) D^t R(-\omega).$$

One readily checks that $\Phi(t, \omega)$ indeed is a cocycle over θ .

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

$$\begin{aligned} \lim_{t \rightarrow \pm\infty} \frac{1}{t} \log \left\| \Phi(t, \omega) R(\omega) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right\| &= \begin{cases} \lambda_1 & x_1 \neq 0 \text{ and } x_2 = 0 \\ \lambda_2 & x_1 = 0 \text{ and } x_2 \neq 0 \end{cases} \\ \Rightarrow E_1(\omega) &= \text{span} \left(R(\omega) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right) \quad \text{and} \\ E_2(\omega) &= \text{span} \left(R(\omega) \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

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 $b_1 \in E_2(\theta_{-t_1}\omega)$. This implies that the continuous version of Ginelli's algorithm does not converge for all fixed choices of b_1 . 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 $\bigcup_{t_1 \in \mathbb{N}} E_2(\theta_{-t_1}\omega)$ 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 all initial tuples in the discrete time case.

Setting $D = \text{diag}(e^{\lambda_1}, e^{\lambda_1})$ in the previous example yields a trivial Oseledets space $E_1(\omega) = \mathbb{R}^2$ with inner rotation. In general, Oseledets spaces can have complicated internal dynamics that prevent single propagated vectors from converging. Additionally, we already remarked that CLVs are not uniquely defined 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.⁹

As a closing remark for this section, we would like to mention that there are several other recently developed algorithms, see [3–5], some of which can be treated in a similar fashion to Ginelli's algorithm with the tools developed here.

4. Convergence of Ginelli's algorithm

Finally, we have gathered enough background knowledge to prove convergence of Ginelli's algorithm. During the proof, we will not distinguish between discrete and continuous time until after we have shown convergence in measure for both cases. Most results will be formulated using the Lyapunov index notation, thus, providing us with a direct link to the exponential speed of convergence.

⁸ This concept was already used in 1980 by Benettin [29,30] to compute the Lyapunov spectrum. Therefore, subsequent applications of the cocycle and orthonormalizations are sometimes called *Benettin steps*.

⁹ In practice, degeneracies can be derived from growth rates of propagated vectors during the forward phase. Moreover, they might be forced by symmetries (e.g. in equivariant systems), whereas, for some classes of systems degenerate scenarios are the exception [31].

4.1. Convergence theorems

Before formulating our new convergence results, we make one simplification that is motivated by the implementation of Ginelli's algorithm and helps us to formulate the theorems in a more compact way. Namely, as the domain of (b') in the backward phase depends on evolved vectors from the forward phase, it will be convenient to identify the backward domain with a time-independent one. To this end, we set $A^{(f)} \in O(d)$ as the orthogonal transformation sending the standard basis (e) to the Gram-Schmidt basis of evolved vectors from the forward phase, i.e. $(f) = \mathcal{GS}(\Phi(t_1 + t_2, \theta_{-t_1} \omega_0)b)$. Note that forward initial vectors (b) need to be linearly independent in order to get a well-defined mapping. By identifying $\mathbb{R}^{d_1 + \dots + d_i}$ with $\mathbb{R}^{d_1 + \dots + d_i} \times \{0\} \subset \mathbb{R}^d$ we may regard the restriction of $A^{(f)}$ as an identification between time-independent coefficients and time-dependent vectors:

$$\mathbb{R}^{d_1 + \dots + d_i} \rightarrow \overline{U}_i^{(\Phi(t_1 + t_2, \theta_{-t_1} \omega_0)b)} \\ \alpha_{i_k} \mapsto b'_{i_k}.$$

Thus, we use

$$(\mathbb{R}^{d_1})^{d_1} \times (\mathbb{R}^{d_1 + d_2})^{d_2} \times \dots \times (\mathbb{R}^{d_1 + \dots + d_{p-1}})^{d_{p-1}} \times (\mathbb{R}^d)^{d_p} \subset (\mathbb{R}^d)^d$$

as the domain for coefficient of the backward phase.

Theorem 4.1 (Convergence in Measure of Ginelli's algorithm). *For each compact subset*

$$\mathcal{K} \subset (\mathbb{R}^d)^d \times \left((\mathbb{R}^{d_1})^{d_1} \times (\mathbb{R}^{d_1 + d_2})^{d_2} \times \dots \times (\mathbb{R}^{d_1 + \dots + d_{p-1}})^{d_{p-1}} \times (\mathbb{R}^d)^{d_p} \right)$$

and $\epsilon > 0$, it holds

$$\lim_{T \rightarrow \infty} \sup_{t_1, t_2 \geq T} \mu \left(\mathcal{K} \setminus \left\{ ((b), (\alpha)) \in \mathcal{K} \mid (b) \right. \right. \\ \left. \left. \text{linearly independent and } \forall i : \right. \right. \\ \left. \frac{1}{\min(t_1, t_2)} \log d \left(U_i^{(\Phi(-t_2, \theta_{t_2} \omega_0)b')}, E_i(\omega_0) \right) \right. \\ \left. \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|) + \epsilon \right. \\ \left. \left. \text{with } (b') = \left(A^{\mathcal{GS}(\Phi(t_1 + t_2, \theta_{-t_1} \omega_0)b)} \alpha \right) \right\} \right) \\ = 0.$$

Compared to the somewhat more involved notation of Theorem 4.1 the convergence theorem for discrete time can be formulated quite nicely using the Lyapunov index notation.

Theorem 4.2 (Convergence a.e. of Ginelli's Algorithm for $\mathbb{T} = \mathbb{Z}$). *For almost all pairs of tuples $((b), (\alpha))$*

$$(\mathbb{R}^d)^d \times \left((\mathbb{R}^{d_1})^{d_1} \times (\mathbb{R}^{d_1 + d_2})^{d_2} \times \dots \times (\mathbb{R}^{d_1 + \dots + d_{p-1}})^{d_{p-1}} \times (\mathbb{R}^d)^{d_p} \right),$$

(b) is linearly independent and the algorithm converges:

$$\bar{\lambda} \left(d \left(U_i^{(\Phi(-n_2, \theta_{n_2} \omega_0)b')}, E_i(\omega_0) \right) \right) \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|)$$

$$\text{with } (b') = \left(A^{\mathcal{GS}(\Phi(n_1 + n_2, \theta_{-n_1} \omega_0)b)} \alpha \right).$$

Theorem 4.2 tells us that, for almost all choices of initial vectors (b) for the forward phase and initial coefficients (α) for the backward phase, the i th output subspace $U_i^{(\Phi(-n_2, \theta_{n_2} \omega_0)b')}$ of Ginelli's algorithm converges to $E_i(\omega_0)$ exponentially fast with a

rate of $\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|)$. In particular, the speed of convergence of the whole algorithm is approximately given by

$$e^{-\min(n_1, n_2) \min_i |\lambda_i - \lambda_{i+1}|}.$$

In applications one usually wants to compute CLVs at more than just one point along a trajectory.¹⁰ In fact, it is feasible to use propagated vectors near ω_0 as approximations to CLVs in Ginelli's algorithm. Thus, it is enough to run the algorithm once. Similar statements on convergence are possible. We only formulate a version for discrete time.

Corollary 4.3 (Convergence a.e. of Ginelli's Algorithm on Interval for $\mathbb{T} = \mathbb{Z}$). *Let $I \subset \mathbb{T}$ be a bounded interval. For almost all pairs of tuples $((b), (\alpha))$ in*

$$(\mathbb{R}^d)^d \times \left((\mathbb{R}^{d_1})^{d_1} \times (\mathbb{R}^{d_1 + d_2})^{d_2} \times \dots \times (\mathbb{R}^{d_1 + \dots + d_{p-1}})^{d_{p-1}} \times (\mathbb{R}^d)^{d_p} \right),$$

(b) is linearly independent and the algorithm converges on I :

$$\bar{\lambda} \left(\sup_{m \in I} d \left(U_i^{(\Phi(-n_2 + m, \theta_{n_2} \omega_0)b')}, E_i(\theta_m \omega_0) \right) \right) \\ \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|)$$

$$\text{with } (b') = \left(A^{\mathcal{GS}(\Phi(n_1 + n_2, \theta_{-n_1} \omega_0)b)} \alpha \right).$$

Proof. Writing

$$d \left(U_i^{(\Phi(-n_2 + m, \theta_{n_2} \omega_0)b')}, E_i(\theta_m \omega_0) \right) \\ = d \left(\Phi(m, \omega_0) U_i^{(\Phi(-n_2, \theta_{n_2} \omega_0)b')}, \Phi(m, \omega_0) E_i(\omega_0) \right),$$

the claim is a direct consequence of Theorem 4.2 and Corollary 2.8. \square

In order to prove both theorems, we derive asymptotic characterizations of each phase of Ginelli's algorithm. However, first, we need to understand how singular vectors and Oseledets spaces are connected by invoking the proof of Proposition 3.2 as it can be found in [23].

4.2. The link between multiplicative ergodic theorem and singular value decomposition

Let

$$\Phi(t, \omega_0) = U(t) \Sigma(t) (V(t))^T$$

be a SVD of the cocycle $\Phi(t, \omega_0)$ for $t \geq 0$, where singular values are ordered as in Eq. (1). Using right singular vectors, Arnold shows that the filtration $V_1(t) \supset \dots \supset V_p(t)$ given by

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

converges exponentially fast to the filtration $V_1(\omega_0) \supset \dots \supset V_p(\omega_0)$. Distances between filtrations are measured in a special metric. Unraveling the notation, we end up with

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

where $P_p + \dots + P_i$ is the orthogonal projection onto $V_i(\omega_0)$ for each i .

¹⁰ It is much harder to predict how the rate of convergence changes when switching to another orbit. For example, in the scenario of random dynamical systems as in [23] Lyapunov spectrum and Oseledets spaces depend only measurably on ω_0 .

Lemma 4.4. *It holds*

$$\forall i : \lambda \left(d \left(\overline{U}_i^{(v(t))}, (V_{i+1}(\omega_0))^\perp \right) \right) \leq -|\lambda_i - \lambda_{i+1}|.$$

Proof. We compute

$$\begin{aligned} \lambda \left(d \left(\overline{U}_i^{(v(t))}, (V_{i+1}(\omega_0))^\perp \right) \right) &= \lambda \left(\left\| \overline{P}_i^{(v(t))} P_{V_{i+1}(\omega_0)} \right\| \right) \\ &\leq \lambda \left(\sum_{\substack{k,j \\ k \leq i < j}} \left\| P_k^{(v(t))} P_j \right\| \right) \\ &\leq \max_{\substack{k,j \\ k \leq i < j}} -|\lambda_k - \lambda_j| \\ &= -|\lambda_i - \lambda_{i+1}|. \quad \square \end{aligned}$$

A similar result holds for the time-reversed cocycle Φ^- with SVD

$$\Phi(-t, \omega_0) = U^-(t) \Sigma^-(t) (V^-(t))^T$$

for $t \geq 0$, where singular values are ordered as in Eq. (1). Note that, for the time-reversed cocycle, we need to consider reversed degeneracies: d_1^-, \dots, d_p^- . To distinguish between both types of degeneracies we equip the notation introduced in Section 2.4 with a minus sign following the subindex, whenever we count with respect to reversed degeneracies.

Lemma 4.5. *It holds*

$$\forall i : \lambda \left(d \left(\overline{U}_{i-}^{(v^-(t))}, (V_{i+1}^-(\omega_0))^\perp \right) \right) \leq -|\lambda_i^- - \lambda_{i+1}^-|.$$

The algorithm of Ginelli starts by propagating vectors from past to present, i.e. we apply $\Phi(t, \theta_{-t}\omega_0) = (\Phi(-t, \omega_0))^{-1}$, and ends with propagating vectors from future to present, i.e. we apply $\Phi(-t, \theta_t\omega_0) = (\Phi(t, \omega_0))^{-1}$. Thus, it is important to keep track of singular vectors for inverted cocycles as well.

Lemma 4.6. *It holds*

$$\forall i : \lambda \left(d \left(\overline{U}_{i-}^{(\hat{u}(t))}, V_{p+1-i}(\omega_0) \right) \right) \leq -|\lambda_{p-i} - \lambda_{p+1-i}|.$$

Proof. This is a consequence of Lemma 4.4, since

$$\begin{aligned} d \left(\overline{U}_{i-}^{(\hat{u}(t))}, V_{p+1-i}(\omega_0) \right) &= d \left(\overline{U}_{i-}^{(v(t))^T}, V_{p+1-i}(\omega_0) \right) \\ &= d \left(\left(\overline{U}_{p-i}^{(v(t))} \right)^\perp, V_{p+1-i}(\omega_0) \right) \\ &= d \left(\overline{U}_{p-i}^{(v(t))}, (V_{p+1-i}(\omega_0))^\perp \right). \end{aligned}$$

Here, we used the identity

$$\overline{U}_{i-}^{(c)^T} = \left(\overline{U}_{p-i}^{(c)} \right)^\perp,$$

which is true for all ONBs (c) . \square

Again, we derive a similar result for reversed time.

Lemma 4.7. *It holds*

$$\forall i : \lambda \left(d \left(\overline{U}_i^{(\hat{u}^-(t))}, V_{p+1-i}^-(\omega_0) \right) \right) \leq -|\lambda_i - \lambda_{i+1}|.$$

4.3. Forward phase

Step 1.1 of Ginelli's algorithm propagates vectors from past to present. It turns out that admissible tuples yield good approximations to $V_{p+1-i}^-(\omega_0) = E_1(\omega_0) \oplus \dots \oplus E_i(\omega_0)$. Moreover, changes of the admissibility parameter on subexponential scales do not influence the exponential speed of convergence of the algorithm.

Lemma 4.8. *Let $0 < \delta(t) < 1$ be a sequence with $\lambda(\frac{1}{\delta}) = 0$. We have*

$$\lambda \left(\sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t))}(\delta(t))} d \left(\overline{U}_i^{(\Phi(t, \theta_{-t}\omega_0)b)}, V_{p+1-i}^-(\omega_0) \right) \right) \leq -|\lambda_i - \lambda_{i+1}|.$$

Proof. First use the triangle inequality, then apply Proposition 2.18 to the map $A = (\Phi(-t, \omega_0))^{-1}$, and finally use Lemma 4.7 to obtain

$$\begin{aligned} &\lambda \left(\sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t))}(\delta(t))} d \left(\overline{U}_i^{(\Phi(t, \theta_{-t}\omega_0)b)}, V_{p+1-i}^-(\omega_0) \right) \right) \\ &\leq \max \left(\lambda \left(\sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t))}(\delta(t))} d \left(\overline{U}_i^{(\Phi(t, \theta_{-t}\omega_0)b)}, \overline{U}_i^{(\hat{u}^-(t))} \right) \right), \right. \\ &\quad \left. \lambda \left(d \left(\overline{U}_i^{(\hat{u}^-(t))}, V_{p+1-i}^-(\omega_0) \right) \right) \right) \\ &\leq \max \left(\lambda \left(\frac{1}{\delta(t)} \frac{(\hat{\sigma}_{i+1}^-(t))^{\max}}{(\hat{\sigma}_i^-(t))^{\min}} \right), -|\lambda_i - \lambda_{i+1}| \right) \\ &\leq \max \left(\lambda \left(\frac{(\sigma_{p+1-i}^-(t))^{\max}}{(\sigma_{p-i}^-(t))^{\min}} \right), -|\lambda_i - \lambda_{i+1}| \right) \\ &= -|\lambda_i - \lambda_{i+1}|. \quad \square \end{aligned}$$

To continue using our tools for step 1.2 we need to retain admissibility for tuples propagated in step 1.1.

Lemma 4.9. *Let $0 < \delta(t) < 1$ with $\lambda(\frac{1}{\delta}) = 0$. There are $0 < \epsilon < 1$ and $T > 0$ such that admissible tuples in step 1.1 get mapped to admissible tuples for step 1.2, i.e.*

$$(\Phi(t_1, \theta_{-t_1}\omega_0))^d \left(\mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1)) \right) \subset \mathcal{Ad}^{(v(t_2))}(\epsilon),$$

for all $t_1, t_2 \geq T$.

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

$$d(V_{p+1-i}^-(\omega_0), (V_{i+1}(\omega_0))^\perp) \leq \sqrt{1 - \epsilon^2} - 2\epsilon.$$

This is possible due to Proposition 2.7, since we assumed $V_{p+1-i}^-(\omega_0) \cap V_{i+1}(\omega_0) = \{0\}$. Now, Lemma 4.8 gives us the existence of $T_1 > 0$ such that for all $t_1 \geq T_1$ and all $(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))$ it holds

$$d \left(\overline{U}_i^{(\Phi(t_1, \theta_{-t_1}\omega_0)b)}, V_{p+1-i}^-(\omega_0) \right) \leq \epsilon.$$

Moreover, Lemma 4.4 yields $T_2 > 0$ with

$$d \left((V_{i+1}(\omega_0))^\perp, \overline{U}_i^{(v(t_2))} \right) \leq \epsilon$$

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

$$\begin{aligned} d \left(\overline{U}_i^{(\Phi(t_1, \theta_{-t_1}\omega_0)b)}, \overline{U}_i^{(v(t_2))} \right) &\leq d \left(\overline{U}_i^{(\Phi(t_1, \theta_{-t_1}\omega_0)b)}, V_{p+1-i}^-(\omega_0) \right) \\ &\quad + d \left(V_{p+1-i}^-(\omega_0), (V_{i+1}(\omega_0))^\perp \right) \\ &\quad + d \left((V_{i+1}(\omega_0))^\perp, \overline{U}_i^{(v(t_2))} \right) \\ &\leq \sqrt{1 - \epsilon^2}. \end{aligned}$$

This concludes the proof. \square

The following lemma combines steps 1.1 and 1.2 into a characterization of the forward phase.

Lemma 4.10. Let $0 < \delta(t) < 1$ with $\lambda(\frac{1}{\delta}) = 0$. There is $T > 0$ such that

$$\lambda \left(\sup_{t_1 \geq T} \sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))} d \left(\bar{U}_i^{(\Phi(t_1+t_2, \theta_{-t_1} \omega_0) b)}, \bar{U}_i^{(u(t_2))} \right) \right) \leq -|\lambda_i - \lambda_{i+1}|$$

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

Proof. Write

$$\Phi(t_1 + t_2, \theta_{-t_1} \omega_0) = \Phi(t_2, \omega_0) \Phi(t_1, \theta_{-t_1} \omega_0).$$

By Lemma 4.9 we find $T > 0$ and $0 < \epsilon < 1$ such that for all $t_1, t_2 \geq T$ and $(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))$ the tuple $(\Phi(t_1, \theta_{-t_1} \omega_0) b)$ is ϵ -admissible w.r.t. $v(t_2)$. Now, apply Proposition 2.18 with $A = \Phi(t_2, \omega_0)$ to see that

$$d \left(\bar{U}_i^{(\Phi(t_1+t_2, \theta_{-t_1} \omega_0) b)}, \bar{U}_i^{(u(t_2))} \right) \leq \frac{1}{\epsilon} \frac{\sigma_{i+1}^{\max}(t_2)}{\sigma_i^{\min}(t_2)}.$$

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

4.4. Backward phase

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

Lemma 4.11. Let $0 < \delta(t) < \frac{1}{\sqrt{2}}$ with $\lambda(\frac{1}{\delta}) = 0$ be given. There is $T > 0$ such that for all $t_1, t_2 \geq T$ and all $(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))$ we have

$$\mathcal{Ad}_-^{(f)^r}(\sqrt{2}\delta(t_2)) \subset \mathcal{Ad}_-^{(\hat{v}(t_2))}(\delta(t_2)),$$

where $(f) := \mathcal{GS}(\Phi(t_1 + t_2, \theta_{-t_1} \omega_0) b)$ and admissibility holds with respect to reversed degeneracies.

Proof. Let $(f) := \mathcal{GS}(\Phi(t_1 + t_2, \theta_{-t_1} \omega_0) b)$ for $(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))$ be given, and let $(g) \in \mathcal{Ad}_-^{(f)^r}(\sqrt{2}\delta(t_2))$ be an admissible tuple. We estimate

$$\begin{aligned} d(\bar{U}_{i-}^{(g)}, \bar{U}_{i-}^{(\hat{v}(t_2))}) &\leq d(\bar{U}_{i-}^{(g)}, \bar{U}_{i-}^{(f)^r}) + d(\bar{U}_{i-}^{(f)^r}, \bar{U}_{i-}^{(\hat{v}(t_2))}) \\ &\leq \sqrt{1 - 2\delta(t_2)^2} + d\left(\left(\bar{U}_{p-i}^{(f)}\right)^\perp, \left(\bar{U}_{p-i}^{(u(t_2))}\right)^\perp\right) \\ &= \sqrt{1 - 2\delta(t_2)^2} \\ &\quad + d\left(\bar{U}_{p-i}^{(\Phi(t_1+t_2, \theta_{-t_1} \omega_0) b)}, \bar{U}_{p-i}^{(u(t_2))}\right). \end{aligned}$$

The last summand is bounded by

$$d(t_2) := \sup_{t_1 \geq T} \sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))} d\left(\bar{U}_{p-i}^{(\Phi(t_1+t_2, \theta_{-t_1} \omega_0) b)}, \bar{U}_{p-i}^{(u(t_2))}\right)$$

for $t_2 \geq T$ with T as in Lemma 4.10. In particular, it holds $\lambda(d(t_2)) < 0$. Now, for (g) to be $\delta(t_2)$ -admissible w.r.t. $(\hat{v}(t_2))$, it suffices to show that

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

for t_2 large enough, which in turn is equivalent to

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

and to

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

The latter is true for t_2 large enough, since we have

$$\begin{aligned} \lambda \left(\frac{d(t_2) \left(2\sqrt{1 - 2\delta(t_2)^2} + d(t_2) \right)}{\delta(t_2)^2} \right) \\ \leq \lambda \left(\frac{d(t_2) (2 + d(t_2))}{\delta(t_2)^2} \right) < 0. \quad \square \end{aligned}$$

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 arguments from the forward phase can be repeated by reversing the cocycle.

Lemma 4.12. Let $0 < \delta(t) < \frac{1}{\sqrt{2}}$ with $\lambda(\frac{1}{\delta}) = 0$ be given. It holds

$$\begin{aligned} \bar{\lambda} \left(\sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))} \sup_{(b') \in \left(\mathcal{Ad}_{\text{ext}_-}^{(f)^r}(\sqrt{2}\delta(t_2)) \right)^r} d \left(U_i^{(\Phi(-t_2, \theta_{t_2} \omega_0) b')}, E_i(\omega_0) \right) \right) \\ \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|), \end{aligned}$$

where $(f) := \mathcal{GS}(\Phi(t_1 + t_2, \theta_{-t_1} \omega_0) b)$.

Proof. Applying Lemma 4.8 to Φ and Φ^- , we get

$$\lambda \left(\sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))} d \left(\bar{U}_i^{(\Phi(t, \theta_{-t} \omega_0) b)}, V_{p+1-i}^-(\omega_0) \right) \right) \leq -|\lambda_i - \lambda_{i+1}|$$

and

$$\lambda \left(\sup_{(g) \in \mathcal{Ad}_-^{(\hat{v}(t))}(\delta(t))} d \left(\bar{U}_{i-}^{(\Phi(-t, \theta_t \omega_0) g)}, V_{p+1-i}(\omega_0) \right) \right) \leq -|\lambda_i^- - \lambda_{i+1}^-|.$$

By switching indices we can rewrite the latter as

$$\lambda \left(\sup_{(g) \in \mathcal{Ad}_-^{(\hat{v}(t))}(\delta(t))} d \left(\bar{U}_{(p+1-i)-}^{(\Phi(-t, \theta_t \omega_0) g)}, V_i(\omega_0) \right) \right) \leq -|\lambda_i - \lambda_{i-1}|.$$

In short, we have exponentially fast converging approximations to $V_{p+1-i}^-(\omega_0)$ and $V_i(\omega_0)$, which are transversal subspaces with intersection $E_i(\omega_0)$ (see Eq. (2)). Thus, we can apply Corollary 2.14 to

$$\mathcal{M}_t := \left\{ \bar{U}_i^{(\Phi(t, \theta_{-t} \omega_0) b)} \mid (b) \in \mathcal{Ad}^{(\hat{v}^-(t))}(\delta(t)) \right\}$$

and

$$\mathcal{N}_t := \left\{ \bar{U}_{(p+1-i)-}^{(\Phi(-t, \theta_t \omega_0) g)} \mid (g) \in \mathcal{Ad}_-^{(\hat{v}(t))}(\delta(t)) \right\}$$

to get a convergence rate estimate for intersections¹¹:

¹¹ Following this statement, one can prove convergence of algorithms that initiate randomly chosen vectors in the past and future, propagate them to the present, and then take intersections of involved subspaces to get an approximation of $E_i(\omega_0)$. Similar convergence theorems for continuous and discrete time can be derived.

$$\begin{aligned} & \bar{\lambda} \left(\sup_{(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))} \sup_{(g) \in \mathcal{Ad}_-^{(\hat{v}(t_2))}(\delta(t_2))} d \left(\overline{U}_i^{(\Phi(t_1, \theta_{-t_1} \omega_0) b)} \cap \overline{U}_{(p+1-i)^-}^{(\Phi(-t_2, \theta_{t_2} \omega_0) g)}, E_i(\omega_0) \right) \right) \\ & \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|). \end{aligned}$$

By Lemma 4.11 we can take the supremum over

$$(g) \in \mathcal{Ad}_-^{(f)^r}(\sqrt{2}\delta(t_2))$$

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

$$(b')^r \in \mathcal{Ad}_{\text{ext}_-}^{(f)^r}(\sqrt{2}\delta(t_2)).$$

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

$$\left((b')_{(p+1-i)_1^-}^r, \dots, (b')_{(p+1-i)_d^-}^r \right) = (b'_{i_d}, \dots, b'_{i_1})$$

satisfies

$$U_i^{(\Phi(-t_2, \theta_{t_2} \omega_0) b')} = \overline{U}_i^{(\Phi(t_1, \theta_{-t_1} \omega_0) b)} \cap \overline{U}_{(p+1-i)^-}^{(\Phi(-t_2, \theta_{t_2} \omega_0) g)}.$$

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

$$U_i^{(\Phi(-t_2, \theta_{t_2} \omega_0) b')} \subset \overline{U}_{(p+1-i)^-}^{(\Phi(-t_2, \theta_{t_2} \omega_0) g)}$$

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

$$\begin{aligned} (b')_{(p+1-i)_1^-}^r, \dots, (b')_{(p+1-i)_d^-}^r & \in U_{(p+1-i)^-}^{(f)^r} \oplus \dots \oplus U_{p^-}^{(f)^r} \\ & = U_i^{(f)} \oplus \dots \oplus U_1^{(f)} \\ & = \overline{U}_i^{(\Phi(t_1+t_2, \theta_{-t_1} \omega_0) b)} \\ & = \Phi(t_2, \omega_0) \overline{U}_i^{(\Phi(t_1, \theta_{-t_1} \omega_0) b)}, \end{aligned}$$

or equivalently, it holds

$$U_i^{(\Phi(-t_2, \theta_{t_2} \omega_0) b')} \subset \overline{U}_i^{(\Phi(t_1, \theta_{-t_1} \omega_0) b)}.$$

Thus, we have

$$U_i^{(\Phi(-t_2, \theta_{t_2} \omega_0) b')} \subset \overline{U}_i^{(\Phi(t_1, \theta_{-t_1} \omega_0) b)} \cap \overline{U}_{(p+1-i)^-}^{(\Phi(-t_2, \theta_{t_2} \omega_0) g)}.$$

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 $E_i(\omega_0)$. Hence, we have equality of subspaces, which concludes the proof. \square

4.5. Proof of theorems

Lemma 4.12 describes how admissible tuples fare in Ginelli's algorithm. The remaining work lies in connecting the lemma to measurement results from Section 2.5.

Proof of Theorem 4.1. Fix $\epsilon > 0$. By compactness of \mathcal{K} we find $M > 0$ with $\mathcal{K} \subset B_d(0, M)^d \times (B(M))^r$. Note that it is enough to prove the claim for the product of balls instead of \mathcal{K} . Furthermore,

set $\delta(t) := \min\left(\frac{1}{t}, \frac{1}{2\sqrt{2}}\right)$, so that $\lambda\left(\frac{1}{\delta}\right) = 0$. Now, we use δ in Lemma 4.12 to get

$$\begin{aligned} & \frac{1}{\min(t_1, t_2)} \log d \left(U_i^{(\Phi(-t_2, \theta_{t_2} \omega_0) b')}, E_i(\omega_0) \right) \\ & \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|) + \epsilon \end{aligned}$$

for all $(b) \in \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1))$ and $(b') \in \left(\mathcal{Ad}_{\text{ext}_-}^{(f)^r}(\sqrt{2}\delta(t_2)) \right)^r$ if t_1 and t_2 are large enough. Using the identification via $A^{(f)}$, we could equivalently assume $(b') = (A^{(f)}\alpha)$ for

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

Hence, it is enough to show that nonadmissible tuples have measure zero in the limit:

$$\begin{aligned} & \mu \left((B_d(0, M)^d \times (B(M))^r) \setminus \left(\mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1)) \right. \right. \\ & \quad \times \left. \left. \mathcal{Ad}_{\text{ext}_-}^{(e)^r}(\sqrt{2}\delta(t_2)) \right)^r \right) \\ & \leq \mu \left(B_d(0, M)^d \setminus \mathcal{Ad}^{(\hat{v}^-(t_1))}(\delta(t_1)) \right) \mu \left((B(M))^r \right) \\ & \quad + \mu \left(B_d(0, M)^d \right) \mu \left((B(M))^r \setminus \left(\mathcal{Ad}_{\text{ext}_-}^{(e)^r}(\sqrt{2}\delta(t_2)) \right)^r \right) \\ & = \mu \left(B_d(0, M)^d \setminus \mathcal{Ad}^{(e)}(\delta(t_1)) \right) \mu \left((B(M))^r \right) \\ & \quad + \mu \left(B_d(0, M)^d \right) \mu \left((B(M)) \setminus \mathcal{Ad}_{\text{ext}_-}^{(e)^r}(\sqrt{2}\delta(t_2)) \right). \end{aligned}$$

Here, we used invariance under orthogonal transformations of $B_d(0, M)$ to switch from $(\hat{v}^-(t_1))$ to (e) . By Corollaries 2.20 and 2.22 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 4.2. Assume discrete time $\mathbb{T} = \mathbb{Z}$ and $d > 1$. We define $\delta_\epsilon(n) := \left(\frac{\epsilon}{\sqrt{2n^2}} \right)^{d-1}$ as our admissibility parameter satisfying $\lambda\left(\frac{1}{\delta_\epsilon}\right) = 0$ for each $0 < \epsilon < 1$. Using δ_ϵ , we invoke Lemma 4.12 to find that

$$\begin{aligned} & \bar{\lambda} \left(d \left(U_i^{(\Phi(-n_2, \theta_{n_2} \omega_0) b')}, E_i(\omega_0) \right) \right) \\ & \leq -\min(|\lambda_i - \lambda_{i-1}|, |\lambda_i - \lambda_{i+1}|) \end{aligned}$$

for $(b') = \left(A^{\mathcal{GS}(\Phi(n_1+n_2, \theta_{-n_1} \omega_0) b)} \alpha \right)$, whenever

$$((b), (\alpha)) \in \bigcap_{n_1, n_2 \in \mathbb{N}} \mathcal{Ad}^{(\hat{v}^-(n_1))}(\delta_\epsilon(n_1)) \times \left(\mathcal{Ad}_{\text{ext}_-}^{(e)^r}(\sqrt{2}\delta_\epsilon(n_2)) \right)^r.$$

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

$$\bigcup_{0 < \epsilon < 1} \bigcap_{n_1, n_2 \in \mathbb{N}} \mathcal{Ad}^{(\hat{v}^-(n_1))}(\delta_\epsilon(n_1)) \times \left(\mathcal{Ad}_{\text{ext}_-}^{(e)^r}(\sqrt{2}\delta_\epsilon(n_2)) \right)^r \quad (4)$$

has measure zero,¹² which can be proved by exhausting the domain of $((b), (\alpha))$ with products of balls: It holds

$$\begin{aligned} & \mu \left((B_d(0, M)^d \times (B(M))^r) \setminus \right. \\ & \quad \left. \bigcup_{0 < \epsilon < 1} \bigcap_{n_1, n_2 \in \mathbb{N}} \mathcal{Ad}^{(\hat{v}^-(n_1))}(\delta_\epsilon(n_1)) \times \left(\mathcal{Ad}_{\text{ext}_-}^{(e)^r}(\sqrt{2}\delta_\epsilon(n_2)) \right)^r \right) \end{aligned}$$

¹² Note that the statement is not true in general for continuous time. In fact, in Example 3.7 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) would be empty.

$$\begin{aligned}
&\leq \inf_{0 < \epsilon < 1} \left(\sum_{n_1 \in \mathbb{N}} \mu(B_d(0, M)^d \setminus \mathcal{A}d^{(\epsilon)}(\delta_\epsilon(n_1))) \mu(B(M)) \right. \\
&\quad \left. + \sum_{n_2 \in \mathbb{N}} \mu(B_d(0, M)^d) \mu(B(M) \setminus \mathcal{A}d_{\text{ext}}^{(\epsilon)^r}(\sqrt{2}\delta_\epsilon(n_2))) \right) \\
&\leq \inf_{0 < \epsilon < 1} \left(\sum_{n_1 \in \mathbb{N}} \eta_1(\delta_\epsilon(n_1))^{\frac{1}{d-1}} \mu(B(M)) \right. \\
&\quad \left. + \sum_{n_2 \in \mathbb{N}} \mu(B_d(0, M)^d) \eta_2(\sqrt{2}\delta_\epsilon(n_2))^{\frac{1}{d-1}} \right) \\
&= \inf_{0 < \epsilon < 1} \epsilon \left(\sum_{n \in \mathbb{N}} \frac{\eta_1 \mu(B(M)) + \eta_2 \sqrt{2}^{\frac{1}{d-1}} \mu(B_d(0, M)^d)}{\sqrt{2}n^2} \right) \\
&= 0
\end{aligned}$$

for all $M > 0$. Here, it was crucial to use Propositions 2.23 and 2.26 to get a more precise measure estimate on nonadmissible tuples. \square

5. Conclusions

We defined Ginelli's algorithm as a means to compute CLVs/Oseledets spaces, which are the most natural choice for directions describing asymptotic expansion and contraction in the tangent linear model along a given trajectory. The existence of those characteristic directions was provided by the MET of Oseledets. Moreover, 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 configurations 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 configuration of initial vectors is called admissible if it is not too far from the optimal initial vectors, i.e. right singular vectors of the propagator. The term “not too far” was made more precise by a parameter $\delta \in [0, 1]$. In our formulation, δ -values close to one imply a good correlation, whereas small values of δ stand for greater distances to the configuration of singular vectors.

In [21] it is shown that configurations with $\delta > 0$ will align with left singular vectors when propagated from the present state to future states. While the condition of admissibility depends on the chosen runtime, according to the MET the configuration of right singular vectors defining admissibility at the present state converges. Using the limit configuration, it is possible to show that almost all initial configurations will yield a good approximation to left singular vectors if propagated long enough from the present state. However, in Ginelli's algorithm initial vectors are first propagated from past states to the present state. In this case, left singular vectors converge to an orthonormalization of CLVs. But, as the admissibility condition depends on right singular vectors at the past state, the set of admissible initial vectors varies with the runtime and in general does not converge to a limit set. In fact, we presented an example where no fixed initial configuration is admissible for all past states simultaneously. Consequently, the continuous time version of Ginelli's algorithm cannot be expected to converge for fixed initial configurations in general. Instead, we have shown convergence in measure of the continuous time version by carefully analyzing the time-sensitivity of propagated vectors. Moreover, due to suitable measure estimates

for sets of admissible vectors, we were able to prove convergence for almost all initial vectors in the discrete time case.

The convergence results for both time cases 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. Interestingly, this was already predicted and observed in applications.

It is important to point out that the Lyapunov index notation neglects system-dependent prefactors for 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 two different initial conditions, can be ignored. Moreover, nonadmissible initial configurations will in general turn admissible due to numerical noise. Hence, the concept of admissibility and the different versions 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 of convergence.

While the proof assumes perfect computations, it is often not known how perturbed data affects LEs and CLVs. In particular, the possibly noncontinuous dependence of the Lyapunov spectrum on the choice of trajectory adds to the uncertainty. In this regard, it would be interesting to know more about how perturbations affect the outcome of Ginelli's algorithm in numerical simulations as well as in analytical computations.

Ultimately, a wide range of applications, some of which are referenced here, underline the importance of CLVs for dynamical systems. Our convergence proof not only verifies the use of Ginelli's algorithm in those applications, but encourages one to apply the concept of CLVs to further scenarios. Moreover, the tools obtained during the proof can be used to investigate other algorithms, such as Wolfe–Samelson's algorithm [3], as well. In general, we expect our rigorous mathematical treatment to enable a more in-depth analysis that will lead to new insights and improvements of CLV-algorithms, which are important instruments to finding structure in the chaos of dynamical systems.

Acknowledgments

This paper is a contribution to the project M1 (Instabilities across scales and statistical mechanics of multi-scale GFD systems) of the Collaborative Research Centre TRR 181 “Energy Transfer in Atmosphere and Ocean” funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 274762653. Special thanks goes to my colleagues from project M1, in particular, R. Lauterbach and I. Gasser for frequent feedback and S. Schubert for motivational discussions on applications of CLVs.

References

- [1] Francesco Ginelli, et al., Characterizing dynamics with covariant lyapunov vectors, *Phys. Rev. Lett.* 99 (13) (2007) 130601, <http://dx.doi.org/10.1103/PhysRevLett.99.130601>.
- [2] Valery I. Oseledets, A multiplicative ergodic theorem: lyapunov characteristic exponents for dynamical systems, *Trudy Mosk. Mat. Obsc.* 19 (1968) 179–210.
- [3] Christopher L. Wolfe, Roger M. Samelson, An efficient method for recovering lyapunov vectors from singular vectors, *Tellus A: Dynamical Meteorology and Oceanography* 59 (3) (2007) 355–366, <http://dx.doi.org/10.1111/j.1600-0870.2007.00234.x>.
- [4] Gary Froyland, et al., Computing covariant lyapunov vectors, oseledets vectors, and dichotomy projectors: a comparative numerical study, *Physica D* 247 (1) (2013) 18–39, <http://dx.doi.org/10.1016/j.physd.2012.12.005>.
- [5] Pavel V. Kuptsov, Ulrich Parlitz, Theory and computation of covariant Lyapunov vectors, *J. Nonlinear Sci.* 22 (5) (2012) 727–762, <http://dx.doi.org/10.1007/s00332-012-9126-5>.

- [6] Giovanni Conti, Gualtiero Badin, Hyperbolic covariant coherent structures in two dimensional flows, *Fluids* 2 (50) (2017) <http://dx.doi.org/10.3390/fluids2040050>, 4.
- [7] Masanobu Inubushi, Shin-ichi Takehiro, Michio Yamada, Regeneration cycle and the covariant Lyapunov vectors in a minimal wall turbulence, *Phys. Rev. E* 92 (2) (2015) 023022, <http://dx.doi.org/10.1103/PhysRevE.92.023022>.
- [8] Hadrien Bosetti, Harald A Posch, Covariant Lyapunov vectors for rigid disk systems, *Chem. Phys.* 375 (2–3) (2010) 296–308, <http://dx.doi.org/10.1016/j.chemphys.2010.06.010>.
- [9] Hadrien Bosetti, Harald A. Posch, Orthogonal versus covariant Lyapunov vectors for rough hard disc systems, *J. Phys. A* 46 (25) (2013) 254011, <http://dx.doi.org/10.1088/1751-8113/46/25/254011>.
- [10] Gary P Morriss, Daniel P Truant, A review of the hydrodynamic Lyapunov modes of hard disk systems, *J. Phys. A* 46 (25) (2013) 254010, <http://dx.doi.org/10.1088/1751-8113/46/25/254010>.
- [11] Daniel P. Truant, Gary P. Morriss, Backward and covariant lyapunov vectors and exponents for hard-disk systems with a steady heat current, *Phys. Rev. E* 90 (5) (2014) 052907, <http://dx.doi.org/10.1103/PhysRevE.90.052907>.
- [12] Sebastian Schubert, Valerio Lucarini, Covariant Lyapunov vectors of a quasi-geostrophic baroclinic model: analysis of instabilities and feedbacks, *Q. J. R. Meteorol. Soc.* 141 (693) (2015) 3040–3055, <http://dx.doi.org/10.1002/qj.2588>.
- [13] Sebastian Schubert, Valerio Lucarini, Dynamical analysis of blocking events: spatial and temporal fluctuations of covariant lyapunov vectors, *Q. J. R. Meteorol. Soc.* 142 (698) (2016) 2143–2158, <http://dx.doi.org/10.1002/qj.2808>.
- [14] Stéphane Vannitsem, Valerio Lucarini, Statistical and dynamical properties of covariant lyapunov vectors in a coupled atmosphere-ocean model-multiscale effects, geometric degeneracy, and error dynamics, *J. Phys. A* 49 (22) (2016) 224001, <http://dx.doi.org/10.1088/1751-8113/49/22/224001>.
- [15] Kazumasa A. Takeuchi, et al., Hyperbolic decoupling of tangent space and effective dimension of dissipative systems, *Phys. Rev. E* 84 (4) (2011) 046214, <http://dx.doi.org/10.1103/PhysRevE.84.046214>.
- [16] Marcus W. Beims, Jason A.C. Gallas, Alignment of Lyapunov vectors: a quantitative criterion to predict catastrophes? *Sci. Rep.* 6 (2016) 37102, <http://dx.doi.org/10.1038/srep37102>.
- [17] Nahal Sharafi, Marc Timme, Sarah Hallerberg, Critical transitions and perturbation growth directions, *Phys. Rev. E* 96 (3) (2017) 032220, <http://dx.doi.org/10.1103/PhysRevE.96.032220>.
- [18] Yoshitaka Saiki, Miki U. Kobayashi, Numerical identification of nonhyperbolicity of the lorenz system through lyapunov vectors, *JSIAM Lett.* 2 (2010) 107–110, <http://dx.doi.org/10.14495/jsiaml.2.107>.
- [19] Mu Xu, M.R. Paul, Covariant Lyapunov vectors of chaotic Rayleigh-Bénard convection, *Phys. Rev. E* 93 (6) (2016) 062208, <http://dx.doi.org/10.1103/PhysRevE.93.062208>.
- [20] Hong-liu Yang, et al., Hyperbolicity and the effective dimension of spatially extended dissipative systems, *Phys. Rev. Lett.* 102 (7) (2009) 074102, <http://dx.doi.org/10.1103/PhysRevLett.102.074102>.
- [21] Sergey V. Ershov, Alexei B. Potapov, On the concept of stationary lyapunov basis, *Physica D* 118 (3–4) (1998) 167–198, [http://dx.doi.org/10.1016/S0167-2789\(98\)00013-X](http://dx.doi.org/10.1016/S0167-2789(98)00013-X).
- [22] Francesco Ginelli, et al., Covariant Lyapunov vectors, *J. Phys. A* 46 (25) (2013) 254005, <http://dx.doi.org/10.1088/1751-8113/46/25/254005>.
- [23] Ludwig Arnold, Random Dynamical Systems, in: Springer Monographs in Mathematics, Springer-Verlag Berlin Heidelberg New York, 1998, <http://dx.doi.org/10.1007/978-3-662-12878-7>.
- [24] Tosio Kato, Perturbation theory for linear operators, in: Classics in Mathematics, vol. 132, Springer-Verlag Berlin Heidelberg, 1995, <http://dx.doi.org/10.1007/978-3-642-66282-9>.
- [25] Aurl Galntai, Projectors and Projection Methods, *Advances in Mathematics*, vol. 6, Springer US, 2004, <http://dx.doi.org/10.1007/978-1-4419-9180-5>.
- [26] Frank Deutsch, The angle between subspaces of a Hilbert space, in: Approximation Theory, Wavelets and Applications, Springer Science+Business Media Dordrecht, 1995, pp. 107–130, http://dx.doi.org/10.1007/978-94-015-8577-4_7.
- [27] John von Neumann, The geometry of orthogonal spaces, *functional operators-vol. II*, *Ann. Math. Stud.* 22 (1950).
- [28] Rufus Bowen, David Ruelle, The ergodic theory of axiom a flows, in: The Theory of Chaotic Attractors, Springer, 1975, pp. 55–76, http://dx.doi.org/10.1007/978-0-387-21830-4_5.
- [29] Giancarlo Benettin, 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) 9–20, <http://dx.doi.org/10.1007/BF02128236>.
- [30] Giancarlo Benettin, 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) 21–30, <http://dx.doi.org/10.1007/BF02128237>.
- [31] Ludwig Arnold, Nguyen D. Cong, Linear cocycles with simple lyapunov spectrum are dense in L^∞ , *Ergodic Theory Dynam. Systems* 19 (6) (1999) 1389–1404, <http://dx.doi.org/10.1017/S014338579915199X>.