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

Florian Noethen*

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 ones. 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.

Key words. Dynamical Systems, Stability, Covariant Lyapunov Vectors, Oseledets Spaces, Lyapunov Exponents, Ginelli Algorithm, Convergence Proof

AMS subject classifications. 37M25, 37H15, 37B25, 37C75

1. Introduction. Dynamical systems play a fundamental role in mathematics. They are of interest not only as a theoretical tool, but they provide methods to study a wide range of interdisciplinary applications. However, due to high dimension and complexity it can be difficult or numerically expensive to completely explore a system. Nevertheless, local structures around simple objects can be understood in terms of linear models. For example, the Hartman-Grobman theorem links the linearization of a hyperbolic steady-state 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 formulated his celebrated *Multiplicative Ergodic Theorem* (MET) [20]. He managed to find a suitable generalization that goes beyond the analysis of steady-states and periodic orbits. In his theorem the long-term behavior of linear perturbations of arbitrary trajectories is explained. Similar to the situation before, the tangent space is split into invariant subspaces that capture different directions of asymptotic growth rates. We refer to this splitting as *Oseledets splitting*. It was shown afterwards [21] that *Oseledets spaces* are linked to invariant manifolds of the original system, making them a valuable tool for understanding dynamics.

Despite their prominent role, it was not until a few years ago that first algorithms to compute Oseledets spaces were developed. Following Ginelli's algorithm [15] in 2007 several other approaches emerged [12, 18, 30], some of which are explained only for nondegenerate scenarios. That is, Oseledets spaces are one-dimensional, and, thus, can be identifies with a basis of vectors for each point of the trajectory. These vectors are called *covariant Lyapunov vectors*

^{*}Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany (florian.noethen@uni-hamburg.de).

(CLVs). Similar to eigenvectors for the linear model of a steady-state, they are preserved by the linear propagator along a trajectory except for rescaling factors. Increasing propagation time, it turns out that those rescaling factors grow exponentially fast. The specific exponents associated with them are called *Lyapunov exponents* (LEs). They can be seen as generalizations to eigenvalues corresponding to CLVs instead of eigenvectors.

With computational tools like Ginelli's algorithm at hand, CLVs became a frequent interest in applications. Amongst others, CLVs reveal structures in turbulent flows [9, 16] and are used to analyze hard-disk systems [6, 7, 19, 27] and climate models [23, 24, 28]. Moreover, they constitute a hyperbolic decoupling of the tangent space of dissipative systems that extracts the physically relevant modes [26]. Furthermore, the angle between CLVs is used as an indicator for critical transitions in long-term behavior of solutions [3, 25] and as a degree of hyperbolicity [9, 22, 31, 32] in dynamical systems.

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. It is our goal 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 [11]. 15 years later Ginelli et al. built upon the work of Ershov and Potapov to formulate a convergence proof of their full algorithm [14]. 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 [11] and [14] present fundamental ideas on convergence of Ginelli's algorithm, we find it necessary to be more precise in some arguments. In particular, [11] shows that almost all initial vectors propagated from present to future will align with a so-called stationary Lyapunov basis asymptotically. Then, an estimate for propagation from past to present is obtained by shifting the estimate for propagation from present to future. This argument not only requires uniformity, but the condition on exceptional vectors that will not yield convergence depends on the starting point, which is neglected in [11]. Hence, 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 second phase.

Despite the criticism, both papers are significant steps to better understand the connection between Oseledets MET and the Ginelli algorithm. Moreover, they inspire many ideas presented here. These ideas fill missing details and even extend the existing results. Namely, 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:

Theorem 1.1. Ginelli's algorithm converges in measure.

Given discrete time, we are able to prove a stronger kind of convergence, which does not hold for continuous time in general. Theorem 1.2. The discrete version of Ginelli's algorithm converges for almost all configurations of initial vectors.

Furthermore, we use the Lyapunov index notation to find an estimate for the speed of convergence. As already predicted and observed [11, 12, 14, 28], the speed of convergence is exponential with a rate determined by the minimal distance of LEs.

The 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 previous tools obtained in section 2, while the MET from section 3 serves as an interface between evolution of singular vectors and CLVs.

A summary and concluding remarks will be provided in section 5.

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 these quantities change after applying a propagation map and after orthogonalization, e.g. the Gram-Schmidt procedure. An estimate of the evolution rate is based on a relation to singular vectors of the propagating linear map. As it turns out, there are configurations of vectors that perform better than others. A distinction between them will be made by introducing a so-called admissibility parameter.

We formulate our statements for discrete time $\mathbb{T} = \mathbb{Z}$ and continuous time $\mathbb{T} = \mathbb{R}$ simultaneously.

2.1. Lyapunov Index. When analyzing an algorithm, one of the main aspects to consider is the speed of convergence. Often this rate is defined as the change of distance between a current and a sough-after state as a parameter, such as time, is increased. Moreover, the nature of the problem or features of the algorithm might prescribe certain timescales to consider. Behavior on an exponential scale can be captured by the Lyapunov index.

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

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

Roughly speaking, the function f behaves similar to $e^{t\lambda(f)}$ on an exponential scale. For exam-

ple, a negative Lyapunov index means exponential decay with rate $\lambda(f)$. However, one should note that variations on smaller scales are not included in this notation¹, but very well may be of importance for limited time scenarios such as numerical computations.

We list some useful properties for the Lyapunov index, which can be found in Arnold's book [1] and are easily verified:

Proposition 2.2. Let $f, g: \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$. The following are true: 1. $\lambda(0) = -\infty$ 2. $\lambda(c) = 0$ for c > 0 constant 3. $\lambda(\alpha f) = \lambda(f)$ for $\alpha > 0$ 4. $\lambda(f^{\alpha}) = \alpha\lambda(f)$ for $\alpha > 0$ 5. $f \leq g \implies \lambda(f) \leq \lambda(g)$ 6. $\lambda(f+g) \leq \max(\lambda(f), \lambda(g))$ 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 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 $\overline{\lambda}(f) \in \mathbb{R} \cup \{\pm \infty\}$ of a function $f : \mathbb{T}_{\geq 0} \times \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is defined as the limit

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

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 $\overline{\lambda}$. Furthermore, if we extend a function $f: \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ to $\overline{f}: \mathbb{T}_{\geq 0} \times \mathbb{T}_{\geq 0} \to \mathbb{R}_{\geq 0}$ by setting $\overline{f}(t_1, t_2) := f(t_1)$, then

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

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

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

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

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

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

The relation $\lambda(f) \leq \overline{\lambda}(\overline{f})$ is always satisfied. To show equality, we remark that $\lambda(f) < 0$ implies the existence of some T > 0 with $\log f(t) < 0$ for all $t \geq T$. In particular, it holds

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

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

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

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

and

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

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 [17], the chapter on projections in Galántai's book [13] and the chapter by Deutsch [10].

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. This lets us 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 $\|\cdot\|_{\cdot}^{3}$

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 [13].

Proposition 2.7. The distance d is a metric on the set of subspaces. Moreover, the following holds for all subspaces $M, N \subset \mathbb{R}^d$:

³As 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 stay unchanged.

 $\begin{array}{ll} 1. & 0 \leq d(M,N) \leq 1 \\ 2. & d(M,N) = d\left(M^{\perp},N^{\perp}\right) \\ 3. & d(M,N) < 1 \Rightarrow \dim(M) = \dim(N) \\ In \ case \ that \ \dim(M) = \dim(N), \ we \ also \ have: \\ 4. & d(M,N) = \|P_M P_{N^{\perp}}\| \\ 5. & d(M,N) = 1 \iff M \cap N^{\perp} \neq \{0\} \\ If \ V \in O(d,\mathbb{R}) \ is \ an \ orthogonal \ transformation, \ then \\ 6. & d(V(M),V(N)) = d(M,N). \end{array}$

Each 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

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

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:

$$d(A(M), A(N)) = \|P_{A(M)}P_{(A(N))^{\perp}}\|$$

$$= \|P_{A(M)}P_{(A^*)^{-1}N^{\perp}}\|$$

$$= \max_{x \in M \setminus \{0\}, y \in N^{\perp} \setminus \{0\}} \frac{|\langle Ax, (A^*)^{-1}y \rangle|}{\|Ax\| \|(A^*)^{-1}y\|}$$

$$= \max_{x \in M \setminus \{0\}, y \in N^{\perp} \setminus \{0\}} \frac{|\langle x, y \rangle|}{\|x\| \|y\|} \frac{\|A^{-1}(Ax)\|}{\|Ax\|} \frac{\|A^*((A^*)^{-1}y)\|}{\|(A^*)^{-1}y\|}$$

$$\leq \max_{x \in M \setminus \{0\}, y \in N^{\perp} \setminus \{0\}} \frac{|\langle x, y \rangle|}{\|x\| \|y\|} \|A^{-1}\| \|A^*\|$$

$$= \|A\| \|A^{-1}\| \|P_M P_{N^{\perp}}\|$$

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

Here, A^* denotes the adjoint map of A with respect to the standard inner product. The next concept needed is the (minimal) angle between two subspaces. A lot on this topic can be found in [10].

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

$$c(M,N) := \max\left\{ |\langle x,y \rangle| : x \in M \cap (M \cap N)^{\perp}, ||x|| \le 1, y \in N \cap (M \cap N)^{\perp}, ||y|| \le 1 \right\}$$

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

$$c_0(M,N) := \max\{|\langle x, y \rangle| : x \in M, ||x|| \le 1, y \in N, ||y|| \le 1\}.$$

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 those quantities.

Proposition 2.10. The following statements are true for all subspaces $M, N \subset \mathbb{R}^d$:

1. $0 \le c(M, N) \le c_0(M, N) \le 1$ 2. c(M, N) < 13. $c_0(M, N) < 1 \iff M \cap N = \{0\}$ 4. $c(M, N) = c(N, M) \text{ and } c_0(M, N) = c_0(N, M)$ 5. $c(M, N) = c(M^{\perp}, N^{\perp})$ 6. $c_0(M, N) = ||P_M P_N||$ 7. $c(M, N) = ||P_M P_N - P_{M \cap N}||$

The minimal angle depends continuously on M and N, whereas the angle between M and N is a bit more involved. It also depends on the intersection $M \cap N$.

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 [29].

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

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

A discussion on the speed of convergence can be found in [10]. The following estimate will be enough for our purposes.

Proposition 2.12. For each two subspaces M and N, it holds

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

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

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

we have

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

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

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

$$||P_{M'\cap N'} - P_{M\cap N}|| \le ||P_{M'\cap N'} - (P_{M'}P_{N'})^{k}|| + ||(P_{M'}P_{N'})^{k} - (P_{M}P_{N})^{k}|| + ||(P_{M}P_{N})^{k} - P_{M\cap N}|| \le c(M', N')^{2k-1} + ||(P_{M'}P_{N'})^{k} - (P_{M}P_{N})^{k}|| + c(M, N)^{2k-1}$$

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

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

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

$$\begin{split} \| (P_{M'}P_{N'})^{k} - (P_{M}P_{N})^{k} \| \\ &\leq \sum_{l=0}^{k-1} \| (P_{M}P_{N})^{l} (P_{M'}P_{N'})^{k-l} - (P_{M}P_{N})^{l} P_{M}P_{N'} (P_{M'}P_{N'})^{k-(l+1)} \| \\ &+ \| (P_{M}P_{N})^{l} P_{M}P_{N'} (P_{M'}P_{N'})^{k-(l+1)} - (P_{M}P_{N})^{l+1} (P_{M'}P_{N'})^{k-(l+1)} \| \\ &\leq \sum_{l=0}^{k-1} \| P_{M'} - P_{M} \| + \| P_{N'} - P_{N} \| \\ &= k \left(\| P_{M'} - P_{M} \| + \| P_{N'} - P_{N} \| \right). \end{split}$$

For the last summand, we remark

$$c(M,N) = c\left(M^{\perp}, N^{\perp}\right) \le c_0\left(M^{\perp}, N^{\perp}\right) = \delta.$$

Combining the above yields the desired estimate.

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, 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:

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

⁴Two 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 $c_0(M^{\perp}, N^{\perp}) < 1$.

Then,

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

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

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

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

with arbitrary $k \in \mathbb{N}$. For our purposes, choose $k = k(t_1, t_2) := \left[\min(t_1, t_2)^2\right]$.

By means of Proposition 2.4 and Example 2.5 we compute

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

$$\leq \max \left(\overline{\lambda} \left(\delta^{2k(t_1, t_2) - 1} \right), \overline{\lambda} \left(\left(\frac{1 + \delta}{2} \right)^{2k(t_1, t_2) - 1} \right),$$

$$\overline{\lambda} (k(t_1, t_2)) + \max \left(\overline{\lambda} \left(\sup_{M' \in \mathcal{M}_{t_1}} d(M', M) \right), \overline{\lambda} \left(\sup_{N' \in \mathcal{N}_{t_2}} d(N', N) \right) \right) \right)$$

$$= \max(\lambda_M, \lambda_N).$$

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

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

In the nondegenerate case, we drop the subindex, i.e. $(e) = (e_1, \ldots, e_d)$. Both cases can be translated into each other via $e_{i_k} = e_{d_1 + \cdots + d_{i-1} + k}$.

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

$$\Sigma = \operatorname{diag}\left(\sigma_{1_1}, \ldots, \sigma_{p_{d_p}}\right)$$

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. One needs to assume distinct nonzero singular values and have them ordered by size for uniqueness of U, V and Σ . For our purposes, it will be enough to have uniqueness of spaces spanned by groups of singular vectors. To this end, let $A \in \text{Gl}(d, \mathbb{R})$ be invertible with SVD $U\Sigma V^T$. We assume that singular values of A are ordered in groups by size, i.e. we assume

(2.1)
$$\sigma_{1_1}, \ldots, \sigma_{1_{d_1}} \ge \sigma_{2_1}, \ldots, \sigma_{2_{d_2}} \ge \cdots \ge \sigma_{p_1}, \ldots, \sigma_{p_{d_p}} > 0.$$

Later on, each group of singular values will correspond to a different LE. Hence, the above inequalities will eventually be strict. In that case, the spaces spanned by singular vectors of one group, i.e. $\operatorname{span}\left(u_{i_1},\ldots,u_{i_{d_i}}\right)$ and $\operatorname{span}\left(v_{i_1},\ldots,v_{i_{d_i}}\right)$, are uniquely determined independent of our choice of SVD with (2.1). Thus, each choice satisfying (2.1) is sufficient for an asymptotic analysis. In particular, we get a SVD $\hat{U}\hat{\Sigma}\hat{V}^T$ for the inverse of A by inverting $A = U\Sigma V^T$ and, heeding (2.1), reversing the order of singular values and vectors. In other words, a SVD for the inverse is given by $(\hat{\sigma}) = \left(\frac{1}{\sigma}\right)^r$, $(\hat{u}) = (v)^r$, and $(\hat{v}) = (u)^r$ with $(.)^r$ being the tuple in reversed order.

We denote the smallest and largest singular value 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 a generalization of the Gram-Schmidt procedure. To this end, let $\mathbb{R}^d = U_1 \oplus \cdots \oplus U_p$ be a decomposition 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, ..., p. Then, $\mathbb{R}^d = F_1 \oplus \cdots \oplus F_p$ 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

$$\{0\} \subset \overline{U}_1 \subset \overline{U}_2 \subset \cdots \subset \overline{U}_p = \mathbb{R}^d$$

given by

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

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 $b_{i_1}, \ldots, b_{i_{d_i}}$. From $U_i^{(b)}$ we get $\overline{U}_i^{(b)}$ 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

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

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

$$A\left(\overline{U}_i^{(b)}\right) = \overline{U}_i^{(Ab)}$$

for an invertible linear map A. Moreover, the resulting orthogonal projections after applying A do not depend on whether (b) is orthogonal or not.

Proposition 2.16. For each $A \in Gl(d, \mathbb{R})$ and each basis (b), we have

$$\forall i: P_i^{(Ab)} = P_i^{(Af)},$$

where $(f) := \mathcal{GS}(b)$ is the Gram-Schmidt basis of (b).

Proof. Since (b) and (f) create the same filtration, the filtrations of (Ab) and (Af) coincide as well.

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. This relation is represented by a single parameter δ . It describes how strong the corresponding filtrations are correlated. A value of 0 means no correlation and a value of 1 implies equality. Thus, we call tuples that have a 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 of respective dimension.

Let $0 < \delta \leq 1$ and a basis (c) of \mathbb{R}^d be given.

Definition 2.17. A d-tuple (b) is called δ -admissible with respect to (c) if it is linearly independent and

$$\forall i$$

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

As admissibility only depends on filtrations, we are allowed to interchange involved tuples with their Gram-Schmidt bases. So, let us assume (c) to be an ONB from now on.

Moreover, invariance of distances between subspaces 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 each $V \in O(d, \mathbb{R})$. Hence, $V^d(\mathcal{A}d^{(c)}(\delta))$ and $\mathcal{A}d^{(Vc)}(\delta)$ coincide.

Let us proceed with an alternative characterization of admissibility.

Lemma 2.18. A basis (b) is δ -admissible w.r.t. (c) if, and only if, for all i < p and $x \in \overline{U}_i^{(b)}$ with ||x|| = 1, we have

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

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

$$\begin{split} \left\| \left(I - \overline{P}_{i}^{(c)}\right) \overline{P}_{i}^{(b)} \right\|^{2} &= \max_{\|x\|=1} \left\| \left(I - \overline{P}_{i}^{(c)}\right) \overline{P}_{i}^{(b)} x \right\|^{2} = \max_{\substack{x \in \overline{\mathcal{U}}_{i}^{(b)} \\ \|x\|=1}} \left\| \left(I - \overline{P}_{i}^{(c)}\right) x \right\|^{2} \\ &= 1 - \min_{\substack{x \in \overline{\mathcal{U}}_{i}^{(b)} \\ \|x\|=1}} \left\| \overline{P}_{i}^{(c)} x \right\|^{2} = 1 - \min_{\substack{x \in \overline{\mathcal{U}}_{i}^{(b)} \\ \|x\|=1}} \sum_{j=1}^{i} \sum_{k} |\langle x, c_{j_{k}} \rangle|^{2} \end{split}$$

Now, we are able to relate the evolution of a tuple under a linear map to singular vectors. It turns out that this important relation is sensitive to the admissibility parameter. In fact, controlling the dependence was a major reason to introduce the concept of admissibility.

Proposition 2.19. Let $A = U\Sigma V^T$ be invertible and $0 < \delta \leq 1$. For all $(b) \in \mathcal{A}d^{(v)}(\delta)$, it holds

$$\forall i: d\left(\overline{U}_i^{(Ab)}, \overline{U}_i^{(u)}\right) \le \frac{i(p-i)}{\delta} \frac{\sigma_{i+1}^{\max}}{\sigma_i^{\min}}.$$

Proof. Assume p > 1. We will derive the estimate from

$$\forall l < m : \left\| P_m^{(u)} P_l^{(Ab)} \right\| \le \frac{1}{\delta} \frac{\sigma_m^{\max}}{\sigma_l^{\min}}.$$

To show the above, use right singular vectors and write $x \in \mathbb{R}^d$ as

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

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

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

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

$$\|Ax\|^{2} \geq \sum_{j=1}^{l} \sum_{k} |\langle x, v_{j_{k}} \rangle|^{2} \sigma_{j_{k}}^{2} \geq \left(\sigma_{l}^{\min}\right)^{2} \sum_{j=1}^{l} \sum_{k} |\langle x, v_{j_{k}} \rangle|^{2} \geq \delta^{2} \left(\sigma_{l}^{\min}\right)^{2}$$

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

$$\left\|P_m^{(u)}Ax\right\|^2 = \sum_k |\langle x, v_{m_k}\rangle|^2 \sigma_{m_k}^2 \le (\sigma_m^{\max})^2$$

Since

$$\begin{split} \left\| P_m^{(u)} P_l^{(Ab)} \right\| &= \max_{y \in \operatorname{im} P_l^{(Ab)} \setminus \{0\}} \frac{\left\| P_m^{(u)} y \right\|}{\|y\|} \le \max_{y \in \overline{U}_l^{(Ab)} \setminus \{0\}} \frac{\left\| P_m^{(u)} y \right\|}{\|y\|} \\ &= \max_{x \in \overline{U}_l^{(b)} \setminus \{0\}} \frac{\left\| P_m^{(u)} Ax \right\|}{\|Ax\|} = \max_{\substack{x \in \overline{U}_l^{(b)} \\ \|x\| = 1}} \frac{\left\| P_m^{(u)} Ax \right\|}{\|Ax\|}, \end{split}$$

it follows that

$$\left\|P_m^{(u)}P_l^{(Ab)}\right\| \le \frac{1}{\delta} \frac{\sigma_m^{\max}}{\sigma_l^{\min}}.$$

Now, we compute for i < p:

$$\begin{split} d\Big(\overline{U}_i^{(Ab)}, \overline{U}_i^{(u)}\Big) &= \Big\|\Big(I - \overline{P}_i^{(u)}\Big)\overline{P}_i^{(Ab)}\Big\| \leq \sum_{\substack{l,m \\ l \leq i < m}} \Big\|P_m^{(u)}P_l^{(Ab)}\Big\| \\ &\leq \frac{1}{\delta} \sum_{\substack{l,m \\ l \leq i < m}} \frac{\sigma_m^{\max}}{\sigma_l^{\min}} \quad \leq \frac{i(p-i)}{\delta} \frac{\sigma_{i+1}^{\max}}{\sigma_i^{\min}}. \end{split}$$

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

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

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

$$d\left(\overline{U}_{i}^{(b)},\overline{U}_{i}^{(c)}\right) < 1 \iff \overline{U}_{i}^{(b)} \oplus \left(\overline{U}_{i}^{(c)}\right)^{\perp} = \mathbb{R}^{d}$$

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

$$d\left(\overline{U}_{i}^{(b)},\overline{U}_{i}^{(c)}\right) < 1 \iff \det\left(b_{1_{1}},\ldots,b_{i_{d_{i}}},c_{(i+1)_{1}},\ldots,c_{p_{d_{p}}}\right) \neq 0$$

This form easily reveals the following:

Proposition 2.20. The set of nonadmissible tuples $(\mathbb{R}^d)^d \setminus \mathcal{A}d^{(c)}$ has Lebesgue-measure zero.

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

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

Corollary 2.21. For any subset $\mathcal{F} \subset (\mathbb{R}^d)^d$ of finite Lebesgue-measure, it holds

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

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

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

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

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

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

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

for each index *i*. The set of all (*b*) satisfying this extension property will be denoted by $\mathcal{A}d_{\text{ext}}^{(c)}(\delta)$. We write $\mathcal{A}d_{\text{ext}}^{(c)}$ for the union of those sets over $0 < \delta \leq 1$. As before, one readily checks that $V^d \left(\mathcal{A}d_{\text{ext}}^{(c)}(\delta) \right) = \mathcal{A}d_{\text{ext}}^{(Vc)}(\delta)$ for $V \in O(d, \mathbb{R})$. Moreover, we again conclude that almost all tuples satisfy extendable admissibility.

Proposition 2.22. The set

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

has Lebesgue-measure zero.

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

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

not satisfying the extension property has Lebesgue-measure zero.

The idea is to apply Proposition 2.20 to a reduced setting for fixed *i*. To this end, look at $\mathbb{R}^{d'}$ with degeneracies $d' = d'_1 + \cdots + d'_{p'}$ given by $d'_j := d_{i-1+j}$ for all $j = 1, \ldots, p' := p+1-i$, and let (e') be its standard basis. We get

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

In particular, this implies

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

Now, we transfer the result from $\mathbb{R}^{d'}$ to $U_i^{(c)} \oplus \cdots \oplus U_p^{(c)}$ by identifying (e') with $(c_{i_1}, \ldots, c_{p_{d_p}})$. As an identification between orthonormal bases, Lebesgue-measure, distance between subspaces, and admissibility are preserved. Hence, for almost all given tuples $(b_{i_1}, \ldots, b_{i_{d_i}}) \in (U_i^{(c)} \oplus \cdots \oplus U_p^{(c)})^{d_i}$, we find $0 < \delta \leq 1$ and $g_{(i+1)_1}, \ldots, g_{p_{d_p}} \in U_i^{(c)} \oplus \cdots \oplus U_p^{(c)}$ such that $d(\operatorname{span}(b_{i_1}, \ldots, b_{i_{d_i}}), U_i^{(c)})^2 \leq 1 - \delta^2$

and

$$\forall j > i : d \Big(\operatorname{span}\Big(b_{i_1}, \dots, b_{i_{d_i}}, g_{(i+1)_1}, \dots, g_{j_{d_j}} \Big), U_i^{(c)} \oplus \dots \oplus U_j^{(c)} \Big)^2 \le 1 - \delta^2.$$

We can extend such a tuple

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

to a δ -admissible tuple (g) by setting $g_{j_k} := c_{j_k}$ for j < i. This concludes the proof. As a consequence, we get the following corollary:

Corollary 2.23. Given a subset $\mathcal{F} \subset \left(U_1^{(c)} \oplus \cdots \oplus U_p^{(c)}\right)^{d_1} \times \cdots \times \left(U_p^{(c)}\right)^{d_p}$ of finite Lebesguemeasure, it holds

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

In the discrete time convergence proof of Ginelli's algorithm, a more precise 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 subsection 2.5 will be devoted to a rather technical derivation of explicit estimates needed only for this proof.

Proposition 2.24. Let
$$d > 1$$
. There is a constant $\eta = \eta(d, M) > 0$ such that

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

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

Lemma 2.25. Let (f) be an ONB of \mathbb{R}^d . Fix 1 < i < d and $0 < \delta_1, \delta_2 \leq 1$. If

$$\left\|P_{\operatorname{span}(f_1,\dots,f_{i-1},c_{i+1},\dots,c_d)}f_i\right\|^2 \le 1 - \delta_1^2 \quad and \quad \left\|\overline{P}_{i-1}^{(f)}\left(I - \overline{P}_i^{(c)}\right)\right\|^2 \le 1 - \delta_2^2,$$

then

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

Proof. First, we reduce the problem to the case i = 2 and d = 3:

There are unit vectors $f'_1 \in \text{span}(f_1, \ldots, f_{i-1})$ and $c'_3 \in \text{span}(c_{i+1}, \ldots, c_d)$ such that

$$\left\|\overline{P}_{i}^{(f)}\left(I-\overline{P}_{i}^{(c)}\right)\right\|^{2} = \left\|\overline{P}_{i}^{(f)}c_{3}'\right\|^{2} = \left\|\overline{P}_{i-1}^{(f)}c_{3}'\right\|^{2} + |\langle f_{i}, c_{3}'\rangle|^{2} = |\langle f_{1}', c_{3}'\rangle|^{2} + |\langle f_{2}', c_{3}'\rangle|^{2}$$

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

$$\left\| P_{\operatorname{span}(f_1',c_3')} f_2' \right\|^2 \le \left\| P_{\operatorname{span}(f_1,\dots,f_{i-1},c_{i+1},\dots,c_d)} f_i \right\|^2 \le 1 - \delta_1^2$$

and

$$|\langle f_1', c_3' \rangle|^2 \le \left\| \overline{P}_{i-1}^{(f)} c_3' \right\|^2 \le \left\| \overline{P}_{i-1}^{(f)} \left(I - \overline{P}_i^{(c)} \right) \right\|^2 \le 1 - \delta_2^2.$$

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

$$d\left(\overline{U}_{2}^{(f')}, \overline{U}_{2}^{(c')}\right)^{2} = \left\|\overline{P}_{2}^{(f')}c_{3}'\right\|^{2} = |\langle f_{1}', c_{3}'\rangle|^{2} + |\langle f_{2}', c_{3}'\rangle|^{2}$$

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

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

$$\begin{split} \left\| P_{\text{span}(f_1',c_3')} f_2' \right\|^2 &= |\langle f_1', f_2' \rangle|^2 + \left| \left(\frac{c_3' - \langle f_1', c_3' \rangle f_1'}{\|c_3' - \langle f_1', c_3' \rangle f_1'\|}, f_2' \right) \right|^2 \\ &= \frac{|\langle c_3', f_2' \rangle|^2}{\|c_3' - \langle f_1', c_3' \rangle f_1'\|^2} \\ &= \frac{|\langle c_3', f_2' \rangle|^2}{1 - |\langle f_1', c_3' \rangle|^2}. \end{split}$$

Thus, by our assumptions:

$$|\langle f'_2, c'_3 \rangle|^2 = \left\| P_{\operatorname{span}(f'_1, c'_3)} f'_2 \right\|^2 (1 - |\langle f'_1, c'_3 \rangle|^2) \le (1 - \delta_1^2) (1 - |\langle f'_1, c'_3 \rangle|^2)$$

We estimate:

$$\begin{split} |\langle f_1', c_3' \rangle|^2 + |\langle f_2', c_3' \rangle|^2 &\leq |\langle f_1', c_3' \rangle|^2 + (1 - \delta_1^2)(1 - |\langle f_1', c_3' \rangle|^2) = 1 - \delta_1^2 + \delta_1^2 \, |\langle f_1', c_3' \rangle|^2 \\ &\leq 1 - \delta_1^2 + \delta_1^2(1 - \delta_2^2) \qquad \qquad = 1 - (\delta_1 \delta_2)^2 \qquad \qquad \blacksquare$$

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

Lemma 2.26. If a basis (b) satisfies

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

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

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

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

For i = 1, we have

$$d\left(\overline{U}_{1}^{(f)}, \overline{U}_{1}^{(c)}\right)^{2} = \left\| \left(I - \overline{P}_{1}^{(c)}\right) f_{1} \right\|^{2} = \left\| P_{\operatorname{span}(c_{2}, \dots, c_{d})} f_{1} \right\|^{2} \le 1 - \left(\delta^{\frac{1}{d-1}}\right)^{2}.$$

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

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

Simply apply Lemma 2.25 to close the induction step. Now, we prove the proposition.

Proof of Proposition 2.24. Set $\tilde{\delta} := \delta^{\frac{1}{d-1}}$ and let

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

be the set of all nonadmissible vector tuples inside $B_d(0, M)^d$. From Proposition 2.20 we know that \mathcal{N} has measure zero. On its complement we define a continuous mapping into the *d*-fold product of spheres:

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

with components

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

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

$$span(b_1, ..., b_{i-1}, c_{i+1}, ..., c_d)$$

with $(w_i, c_i) > 0$, 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 $b'_i := (I - \overline{P}_{i-1}^{(b)})b_i$. Assuming $|\langle w_i, b_i \rangle| \ge M\tilde{\delta}$, we get

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

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

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

then (b) is δ -admissible by Lemma 2.26. In particular, the subset of all non- δ -admissible tuples is contained in the subset of all (b), which do not fulfill the above condition. Therefore, a measure-estimate on those tuples is enough for the claim:

$$\begin{split} \mu\Big(\Big\{(b) \in B_{d}(0,M)^{d} \setminus \mathcal{N} \mid \exists i < d : |\langle w_{i}, b_{i} \rangle| < M\tilde{\delta}\Big\}\Big) \\ &\leq \sum_{i < d} \mu\Big(\Big\{(b) \in B_{d}(0,M)^{d} \setminus \mathcal{N} \mid |\langle w_{i}, b_{i} \rangle| < M\tilde{\delta}\Big\}\Big) \\ &= \sum_{i < d} \mu\Big(\{(b) \in B_{d}(0,M)^{d} \mid \det(b_{1}, \dots, b_{i-1}, c_{i}, \dots, c_{d}) \neq 0 \text{ and } |\langle w_{i}, b_{i} \rangle| < M\tilde{\delta}\Big\}\Big) \\ &= \sum_{i < d} (\mu(B_{d}(0,M)))^{d-i} \int_{\{(b_{1},\dots,b_{i-1}) \in B_{d}(0,M)^{i-1} \mid \det(b_{1},\dots,b_{i-1},c_{i},\dots,c_{d}) \neq 0\}} \\ &\int_{\{b_{i} \in B_{d}(0,M) : |\langle w_{i}, b_{i} \rangle| < M\tilde{\delta}\Big\}} 1 \ db_{i} \ d(b_{1},\dots,b_{i-1}) \\ &= \sum_{i < d} (\mu(B_{d}(0,M)))^{d-i} \int_{\{(b_{1},\dots,b_{i-1}) \in B_{d}(0,M)^{i-1} \mid \det(b_{1},\dots,b_{i-1},c_{i},\dots,c_{d}) \neq 0\}} \\ &\int_{\{b_{i} \in B_{d}(0,M) : |\langle e_{1}, b_{i} \rangle| < M\tilde{\delta}\Big\}} 1 \ db_{i} \ d(b_{1},\dots,b_{i-1}) \\ &= \sum_{i < d} (\mu(B_{d}(0,M)))^{d-i} \mu\Big(B_{d}(0,M) \cap \Big(\Big(-M\tilde{\delta},M\tilde{\delta}\Big) \times \mathbb{R}^{d-1}\Big)\Big) \\ &\leq (d-1)(\mu(B_{d}(0,M)))^{d-1}(2M)^{d}\tilde{\delta} \end{split}$$

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 $B_d(0, M) \cap \left(\left(-M\tilde{\delta}, M\tilde{\delta}\right) \times \mathbb{R}^{d-1}\right)$ to $\left(-M\tilde{\delta}, M\tilde{\delta}\right) \times (-M, M)^{d-1}$ for a simple estimate.

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

A similar estimate will be necessary for non- δ -admissible tuples inside the special domain.

Proposition 2.27. Let d > 1. There is a constant $\eta = \eta(d, M) > 0$ such that

$$\mu\Big(B(M)\setminus \mathcal{A}d_{\mathrm{ext}}^{(c)}(\delta)\Big) \leq \eta\delta^{\frac{1}{d-1}},$$

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 2.22. Again, it is enough to find such a bound for the set of all tuples in

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

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

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

Proposition 2.24 yields η' only depending on d' and M with

$$\mu\left(B_{d'}\left(0,\sqrt{2}M\right)^{d'}\setminus\mathcal{A}d^{(e')}(\delta)\right)\leq\eta'\delta^{\frac{1}{d'-1}}.$$

Since $B_{d'}(0, M)^{d'_1} \times B_{d'}(0, M)^{d'-d'_1} \subset B_{d'}(0, \sqrt{2}M)^{d'}$, we have

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

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

3. Ginelli's Algorithm. Understanding long-term behavior of solutions can prove to be a challenging task due to nonlinearity of the underlying system. However, locally, around a reference solution the problem becomes a lot easier. Instead of the full nonlinear dynamics, it may be sufficient to only regard the tangent linear model along the reference solution for a qualitative analysis.

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

3.1. Setting. 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, $\theta_t : \Omega \to \Omega$ represents the time-t-flow on our orbit. The flow should satisfy $\theta_0 = \mathrm{id}_{\Omega}$ 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 $\Phi : \mathbb{T} \times \Omega \to \mathbb{R}^{d \times d}$ is called a (linear) cocycle (over θ) if 1. $\Phi(0, \omega) = \mathrm{id}$ 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, the cocycle is pointwise invertible with inverse

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

The Multiplicative Ergodic Theorem of Oseledets [20] is not only necessary to describe existence of CLVs, but will play a crucial role in our convergence proof. We state a deterministic version found in [1]. 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

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

and assume that

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

exists for all orders of the wedge product of $\Phi(t, \omega_0)$. Then, there is a Lyapunov spectrum with a corresponding filtration capturing subspaces of different growth rates:

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

$$-\infty \leq \lambda_p < \cdots < \lambda_1 < \infty,$$

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

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

2. There is a filtration

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

given by subspaces

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

Limits in the definition of V_i exist for each $x \in \mathbb{R}^d$ and take values in $\{\lambda_1, \ldots, \lambda_p\}$. 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 those 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 [1]:

1. $\|\wedge^{i} A\| = \sigma_{1}(A) \dots \sigma_{i}(A)$ 2. $\|\wedge^{i} (AB)\| \leq \|\wedge^{i} A\| \|\wedge^{i} B\|$

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

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

follows due to the cocycle property:

$$\begin{aligned} \left(\frac{t+u}{t}\right) \left(\frac{1}{t+u} \log \|\wedge^{i} \Phi(t+u,\omega_{0})\|\right) &- \frac{1}{t} \log \|\wedge^{i} \Phi(u,\omega_{0})\| \\ &\leq \frac{1}{t} \log \|\wedge^{i} \Phi(t,\theta_{u}\omega_{0})\| \\ &\leq \left(\frac{t+u}{t}\right) \left(\frac{1}{t+u} \log \|\wedge^{i} \Phi(t+u,\omega_{0})\|\right) + \frac{1}{t} \log \|\wedge^{i} \Phi(-u,\theta_{u}\omega_{0})\| \end{aligned}$$

Thus, the proposition gives us the existence of a Lyapunov spectrum and filtration at state ω . In particular, limits of singular values for ω and ω_0 do not differ on an exponential scale. Hence, the remaining statements of Corollary 3.3 follow.

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 that LEs are finite. For convenience sake, we set $\lambda_0 := \infty$ and $\lambda_{p+1} := -\infty$.

Assuming the above, we get the existence of Oseledets spaces characterizing asymptotic dynamics.

Proposition 3.4. There is a splitting $\mathbb{R}^d = E_1(\omega) \oplus \cdots \oplus E_p(\omega)$ of the tangent space into so-called Oseledets spaces

(3.1)
$$E_i(\omega) := V_i(\omega) \cap V_{p+1-i}^-(\omega).$$

Furthermore, Oseledets spaces can be characterized via

(3.2)
$$x \in E_i(\omega) \setminus \{0\} \iff \lim_{t \to \pm \infty} \frac{1}{|t|} \log \|\Phi(t,\omega)x\| = \pm \lambda_i,$$

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 [1].



Figure 1. diagonal cocycle for dimension 2

For a random dynamical system satisfying some integrability condition, it is shown in [1] that the cocycle along almost all orbits of the system admits an Oseledets splitting. Moreover, in an ergodic setting the Lyapunov spectrum for almost all orbits coincides. Therefore, in applications it is often conveniently 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 this 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 (3.2). However, they are uniquely defined (up to sign) only for nondegenerate Lyapunov spectra.

3.2. The Algorithm. The Ginelli algorithm [14, 15] computes Oseledets spaces (or CLVs) for a given cocycle by using its asymptotic characterization (3.2). The main idea is that every 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 $(d_1 + \cdots + d_i)$ -dimensional subspaces will align with $E_1 \oplus \cdots \oplus E_i$, 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 $E_1 \oplus \cdots \oplus E_i$ will align with E_i

⁵See the concept of SRB-measures for attractors [8].

⁶Aside from asymptotic growth, the angle between CLVs can be used as a measure of hyperbolicity (see, e.g., [9, 22, 31, 32]).

in backward time.

Taking these traits into consideration, the abstract formalism of Ginelli's algorithm is as follows:

Ginelli Algorithm

- 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 $E_1(\omega_0) \oplus \cdots \oplus E_i(\omega_0)$.
- 1.2. Continue forward propagation 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 $E_1 \oplus \cdots \oplus E_i$ on a long enough timeframe for the second phase.
 - 2. For each *i*, randomly choose d_i vectors $b'_{i_1}, \ldots, b'_{i_{d_i}}$ in $\overline{U}_i^{(\Phi(t_1+t_2,\theta_{-t_1}\omega_0)b)} \approx E_1(\theta_{t_2}\omega_0) \oplus \cdots \oplus E_i(\theta_{t_2}\omega_0)$ 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 phase⁷, and by the same reasoning step 2 is called *backward phase*⁸.

forward phase

backward phase

Figure 2. schematic picture of the Ginelli algorithm

The asymptotic expansion rate $(\lambda_1 + \cdots + \lambda_i)$ of $E_1 \oplus \cdots \oplus E_i$ is usually computed as a byproduct in the forward phase of the algorithm. Using this information, we can derive the Lyapunov spectrum.⁹

⁷Numerically, it makes sense to apply the cocycle in small timesteps and orthonormalize the basis inbetween. Otherwise, computations become singular as all vectors will collapse onto the first Oseledets space. However, analytically, Proposition 2.16 tells us that there is no difference in applying the Gram-Schmidt procedure between every step or just once at the end.

⁸Here, assuming the Lyapunov spectrum is known, it appears numerically advantageous to orthonormalize $b'_{i_1}, \ldots, b'_{i_{d_i}}$ between propagation steps.

⁹This concept was already used in 1980 by Benettin [4, 5] to compute the Lyapunov spectrum. Therefore, subsequent applications of the cocycle and orthonormalizations are sometimes called *Benettin steps*.

We provide a convergence proof of the whole algorithm as $\min(t_1, t_2) \to \infty$ in section 4. 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.

Next, we present what might happen in the forward phase for some exemplary cocycles. We start with a very simple example that demonstrates how almost all vectors should evolve in a nondegenerate setting.

Example 3.6 (diagonal cocycle). Assume $\Omega = \{\omega_0\}$ with trivial flow $\theta_t \omega_0 = \omega_0$. For given $\lambda_1 > \cdots > \lambda_p$, define $D := \text{diag}(e^{\lambda_1}, \ldots, e^{\lambda_p})$. 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

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

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

$$\begin{split} \lambda \Big(d\Big(\overline{U}_1^{(D^tb)}, \overline{U}_1^{(e)}\Big) \Big) &= \frac{1}{2} \lambda \Big(\left\| \Big(I - \overline{P}_1^{(e)}\Big) \overline{P}_1^{(D^tb)} \right\|^2 \Big) \\ &= \frac{1}{2} \lambda \left(\left\| \left\langle \frac{D^tb_1}{\|D^tb_1\|}, e_i \right\rangle \right\|^2 \right) \\ &\leq \max_{i \neq 1} \frac{1}{2} \lambda \left(\left\| \left\langle \frac{D^tb_1}{\|D^tb_1\|}, e_i \right\rangle \right\|^2 \right) \\ &= -|\lambda_1 - \lambda_2| \end{split}$$

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 similar statement for arbitrary cocycles will be provided in subsection 4.2, 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 representative of a direction 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 normalized flow $\theta_t \omega := \omega + t$. Furthermore, let $R : \mathbb{R} \to SO(2)$ be the parametrization of SO(2) by 2×2 rotation matrices

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

so that R(0) = R(1) = I and R(s+t) = R(s)R(t). Moreover, we set $D := \text{diag}(e^{\lambda_1}, e^{\lambda_2})$ 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 .

We use the characterization of Oseledets spaces via asymptotic growth rates:

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

In particular, both Oseledets spaces are rotating uniformly with varying ω . Hence, for each 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 any fixed choice of b_1 . Instead, it is shown later on that the continuous case converges in measure.

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 convergence for almost all initial tuples is still possible.

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 presence of degeneracies. Therefore, objects of interest should not be the propagated vectors themselves, but rather the spaces spanned by them subject to degeneracies.¹⁰

As a closing remark for this section, we would like to mention that there are several other recently developed algorithms, see [12, 18, 30], some of which can be treated in a similar fashion to Ginelli's algorithm with tools developed here.

¹⁰Practically speaking, degeneracies can be derived from growth rates of propagated vectors during the forward phase. Moreover, they might be forced by symmetries (e.g. equivariant systems), whereas, for some classes of systems degenerate scenarios are the exception [2].

4. Convergence of Ginelli's Algorithm. Finally, we have gathered enough notation and tools to formulate and 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. All results will be formulated using the Lyapunov index notation, providing us with a direct link to the exponential speed of convergence.

Since the domain for (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:

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

Thus, we use

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

as the domain for coefficient of the backward phase.

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

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

and $\epsilon > 0$, it holds

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

Where the latter result requires more involved notation, the convergence theorem for discrete time can be formulated quite compactly using the extended Lyapunov index notation.

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

$$\left(\mathbb{R}^{d}\right)^{d} \times \left(\left(\mathbb{R}^{d_{1}}\right)^{d_{1}} \times \left(\mathbb{R}^{d_{1}+d_{2}}\right)^{d_{2}} \times \cdots \times \left(\mathbb{R}^{d_{1}+\cdots+d_{p-1}}\right)^{d_{p-1}} \times \left(\mathbb{R}^{d}\right)^{d_{p}}\right),$$

(b) is linearly independent and the algorithm converges:

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

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

In applications one usually wants to compute CLVs at more than just one point along a trajectory.¹¹ 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 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

$$\left(\mathbb{R}^{d}\right)^{d} \times \left(\left(\mathbb{R}^{d_{1}}\right)^{d_{1}} \times \left(\mathbb{R}^{d_{1}+d_{2}}\right)^{d_{2}} \times \cdots \times \left(\mathbb{R}^{d_{1}+\cdots+d_{p-1}}\right)^{d_{p-1}} \times \left(\mathbb{R}^{d}\right)^{d_{p}}\right),$$

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

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

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

Proof. Writing

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

this is a direct consequence of Theorem 4.2 and Corollary 2.8.

In order to prove both theorems we derive asymptotic characterizations of Ginelli's algorithm step by step. However, first, we need to understand how singular vectors and Oseledets spaces are connected by invoking the proof of Proposition 3.2 as found in [1].

4.1. 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 $t \ge 0$, where singular values are ordered as in (2.1). Using right singular vectors, Arnold shows that the filtration $V_p(t) \subset \cdots \subset V_1(t)$ given by

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

¹¹It is much harder to predict how the convergence rate changes, when switching to another orbit. For example, in the scenario of random dynamical systems as in [1] Lyapunov spectrum and Oseledets spaces depend only measurably on ω_0 .

converges exponentially fast to the filtration $V_p(\omega_0) \subset \cdots \subset V_1(\omega_0)$. 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 $P_p + \cdots + P_i$ is the orthogonal projection onto $V_i(\omega_0)$ for each *i*.

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

$$\lambda\left(d\left(\overline{U}_{i}^{(v(t))}, (V_{i+1}(\omega_{0}))^{\perp}\right)\right) = \lambda\left(\left\|\overline{P}_{i}^{(v(t))}P_{V_{i+1}(\omega_{0})}\right\|\right) \le \lambda\left(\sum_{\substack{k,j\\k\leq i< j}}\left\|P_{k}^{(v(t))}P_{j}\right\|\right)$$
$$\le \max_{\substack{k,j\\k\leq i< j}} -|\lambda_{k} - \lambda_{j}| = -|\lambda_{i} - \lambda_{i+1}|.$$

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

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

for $t \ge 0$, where singular values are ordered as in (2.1). Note that, for the time-reversed cocycle, we need to consider reversed degeneracies: d_1^-, \ldots, d_p^- . To distinguish between both types of degeneracies we equip the notation introduced in subsection 2.4 with a minus sign following the subindex, whenever we count with respect to reversed degeneracies.

Lemma 4.5. It holds

$$\forall i: \lambda \left(d \left(\overline{U}_{i^{-}}^{(v^{-}(t))}, \left(V_{i+1}^{-}(\omega_{0}) \right)^{\perp} \right) \right) \leq -|\lambda_{i}^{-} - \lambda_{i+1}^{-}|.$$

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 inversed 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

$$d\left(\overline{U}_{i^{-}}^{(\hat{u}(t))}, V_{p+1-i}(\omega_{0})\right) = d\left(\overline{U}_{i^{-}}^{(v(t))^{r}}, V_{p+1-i}(\omega_{0})\right)$$
$$= d\left(\left(\overline{U}_{p-i}^{(v(t))}\right)^{\perp}, V_{p+1-i}(\omega_{0})\right)$$
$$= d\left(\overline{U}_{p-i}^{(v(t))}, (V_{p+1-i}(\omega_{0}))^{\perp}\right).$$

Here, we used the identity

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

which is true for all ONBs (c).

Again, we derive a similar result for reversed time.

Lemma 4.7. It holds

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

4.2. 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 $V_{p+1-i}^{-}(\omega_0) = E_1(\omega_0) \oplus \cdots \oplus E_i(\omega_0)$. 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

$$\lambda\left(\sup_{(b)\in\mathcal{A}d^{(\hat{v}^{-}(t))}(\delta(t))}d\left(\overline{U}_{i}^{(\Phi(t,\theta_{-t}\omega_{0})b)},V_{p+1-i}^{-}(\omega_{0})\right)\right)\leq-|\lambda_{i}-\lambda_{i+1}|.$$

Proof. Use the triangle inequality, apply Proposition 2.19 to the map $A = (\Phi(-t, \omega_0))^{-1}$, and use Lemma 4.7 to obtain

$$\begin{split} \lambda \left(\sup_{(b) \in \mathcal{A}d^{(\hat{v}^{-}(t))}(\delta(t))} d\left(\overline{U}_{i}^{(\Phi(t,\theta_{-t}\omega_{0})b)}, V_{p+1-i}^{-}(\omega_{0})\right) \right) \\ &\leq \max \left(\lambda \left(\sup_{(b) \in \mathcal{A}d^{(\hat{v}^{-}(t))}(\delta(t))} d\left(\overline{U}_{i}^{(\Phi(t,\theta_{-t}\omega_{0})b)}, \overline{U}_{i}^{(\hat{u}^{-}(t))}\right) \right), \lambda \left(d\left((\overline{U}_{i}^{(\hat{u}^{-}(t))}, V_{p+1-i}^{-}(\omega_{0})\right) \right) \right) \right) \\ &\leq \max \left(\lambda \left(\frac{i(p-i)}{\delta(t)} \frac{(\hat{\sigma}_{i+1}^{-}(t))^{\max}}{(\hat{\sigma}_{i}^{-}(t))^{\min}} \right), -|\lambda_{i} - \lambda_{i+1}| \right) \\ &\leq \max \left(\lambda \left(\frac{(\sigma_{p+1-i}^{-}(t))^{\max}}{(\sigma_{p-i}^{-}(t))^{\min}} \right), -|\lambda_{i} - \lambda_{i+1}| \right) \\ &= -|\lambda_{i} - \lambda_{i+1}|. \end{split}$$

To further use our tools we need to retain admissibility for tuples propagated in step 1.1.

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

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

for all $t_1, t_2 \geq T$.

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

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

This is possible due to Proposition 2.7, since we assumed $V_{p+1-i}^{-}(\omega_0) \cap V_{i+1}(\omega_0) = \{0\}$.

Lemma 4.8 gives us the existence of $T_1 > 0$ such that for all $t_1 \ge T_1$ and all $(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\delta(t_1))$ it holds

$$d\left(\overline{U}_{i}^{\left(\Phi(t_{1},\theta-t_{1}\omega_{0})b\right)},V_{p+1-i}^{-}(\omega_{0})\right)\leq\epsilon.$$

Moreover, Lemma 4.4 yields $T_2 > 0$ with

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

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

$$d\left(\overline{U}_{i}^{\left(\Phi(t_{1},\theta_{-t_{1}}\omega_{0})b\right)},\overline{U}_{i}^{\left(v(t_{2})\right)}\right) \leq d\left(\overline{U}_{i}^{\left(\Phi(t_{1},\theta_{-t_{1}}\omega_{0})b\right)},V_{p+1-i}^{-}(\omega_{0})\right) + d\left(V_{p+1-i}^{-}(\omega_{0}),(V_{i+1}(\omega_{0}))^{\perp}\right) + d\left((V_{i+1}(\omega_{0}))^{\perp},\overline{U}_{i}^{\left(v(t_{2})\right)}\right) \leq \sqrt{1-\epsilon^{2}}.$$

This concludes the proof.

The following lemma combines step 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

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

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 such that for all $t_1, t_2 \ge T$ and $(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\delta(t_1))$ the tuple $(\Phi(t_1, \theta_{-t_1}\omega_0)b)$ is ϵ -admissible w.r.t. $v(t_2)$. Now, apply Proposition 2.19 with $A = \Phi(t_2, \omega_0)$ to see that

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

Since singular values converge to LEs, the claim is proved.

4.3. 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 $t_1, t_2 \ge T$ and all $(b) \in \mathcal{A}d^{(\hat{v}^-(t_1))}(\delta(t_1))$ we have

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

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

$$\begin{split} d\Big(\overline{U}_{i^{-}}^{(g)}, \overline{U}_{i^{-}}^{(\hat{v}(t_{2}))}\Big) &\leq d\Big(\overline{U}_{i^{-}}^{(g)}, \overline{U}_{i^{-}}^{(f)^{r}}\Big) + d\Big(\overline{U}_{i^{-}}^{(f)^{r}}, \overline{U}_{i^{-}}^{(\hat{v}(t_{2}))}\Big) \\ &\leq \sqrt{1 - 2\delta(t_{2})^{2}} + d\Big(\Big(\overline{U}_{p-i}^{(f)}\Big)^{\perp}, \Big(\overline{U}_{p-i}^{(u(t_{2}))}\Big)^{\perp}\Big) \\ &= \sqrt{1 - 2\delta(t_{2})^{2}} + d\Big(\overline{U}_{p-i}^{(\Phi(t_{1}+t_{2},\theta_{-t_{1}}\omega_{0})b)}, \overline{U}_{p-i}^{(u(t_{2}))}\Big). \end{split}$$

As in the proof of Lemma 4.10, the last summand can be estimated by

$$\frac{(p-i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)}$$

with $0 < \epsilon < 1$ for t_1, t_2 large enough. Now, for (g) to be $\delta(t_2)$ -admissible w.r.t. $(\hat{v}(t_2))$, it suffices to show that

$$\sqrt{1 - 2\delta(t_2)^2} + \frac{(p - i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)} \le \sqrt{1 - \delta(t_2)^2}$$

for t_2 large enough, which in turn is equivalent to

$$1 - 2\delta(t_2)^2 + 2\sqrt{1 - 2\delta(t_2)^2} \left(\frac{(p-i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)}\right) + \left(\frac{(p-i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)}\right)^2 \le 1 - \delta(t_2)^2$$

and to

$$\frac{2\sqrt{1-2\delta(t_2)^2} \left(\frac{(p-i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)}\right) + \left(\frac{(p-i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)}\right)^2}{\delta(t_2)} \le 1.$$

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

$$\lambda \left(\frac{2\sqrt{1-2\delta(t_2)^2} \left(\frac{(p-i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)}\right) + \left(\frac{(p-i)i}{\epsilon} \frac{\sigma_{p+1-i}^{\max}(t_2)}{\sigma_{p-i}^{\min}(t_2)}\right)^2}{\delta(t_2)} \right) \le -|\lambda_{p+1-i} - \lambda_{p-i}| < 0.$$

Next, we add backward propagation to the characterization of the forward phase. During the backward phase, it is enough to restrict ourselves to tuples that have admissible extensions. A few argument from the forward phase can be repeated by reversing the cocycle.

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

$$\overline{\lambda}\left(\sup_{\substack{(b)\in\mathcal{A}d^{(\hat{v}^{-}(t_{1}))}(\delta(t_{1})) \ (b')\in\left(\mathcal{A}d_{\mathrm{ext}_{-}}^{(f)r}(\sqrt{2}\delta(t_{2}))\right)^{r}}d\left(U_{i}^{\left(\Phi(-t_{2},\theta_{t_{2}}\omega_{0})b'\right)},E_{i}(\omega_{0})\right)\right)$$
$$\leq-\min(|\lambda_{i}-\lambda_{i-1}|,|\lambda_{i}-\lambda_{i+1}|),$$

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

$$\lambda\left(\sup_{(b)\in\mathcal{A}d^{(\hat{v}^{-}(t))}(\delta(t))}d\left(\overline{U}_{i}^{(\Phi(t,\theta_{-t}\omega_{0})b)},V_{p+1-i}^{-}(\omega_{0})\right)\right)\leq-|\lambda_{i}-\lambda_{i+1}|$$

and

$$\lambda\left(\sup_{(g)\in\mathcal{A}d_{-}^{(\hat{v}(t))}(\delta(t))}d\left(\overline{U}_{i^{-}}^{(\Phi(-t,\theta_{t}\omega_{0})g)},V_{p+1-i}(\omega_{0})\right)\right)\leq-|\lambda_{i}^{-}-\lambda_{i+1}^{-}|.$$

By switching indices we can rewrite the latter as

$$\lambda\left(\sup_{(g)\in\mathcal{A}d_{-}^{(\delta(t))}(\delta(t))}d\left(\overline{U}_{(p+1-i)^{-}}^{(\Phi(-t,\theta_{t}\omega_{0})g)},V_{i}(\omega_{0})\right)\right)\leq-|\lambda_{i}-\lambda_{i-1}|.$$

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 equation (3.1)). Thus, we can apply Corollary 2.14 to

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

and

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

to get a convergence rate estimate for intersections¹²:

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

By Lemma 4.11 we can take the supremum over

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

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

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

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

$$\left((b')_{(p+1-i)_{1}^{-}}^{r},\ldots,(b')_{(p+1-i)_{d_{p+1-i}^{-}}^{-}}^{r}\right) = \left(b'_{i_{d_{i}}},\ldots,b'_{i_{1}}\right)$$

satisfies

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

We clearly have

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

and hence

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

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

$$(b')_{(p+1-i)_{1}^{-}}^{r}, \dots, (b')_{(p+1-i)_{d_{p+1-i}^{-}}^{-}}^{r} \in U_{(p+1-i)^{-}}^{(f)r} \oplus \dots \oplus U_{p^{-}}^{(f)r} = U_{i}^{(f)} \oplus \dots \oplus U_{1}^{(f)}$$
$$= \overline{U}_{i}^{(\Phi(t_{1}+t_{2},\theta_{-t_{1}}\omega_{0})b)} = \Phi(t_{2},\omega_{0})\overline{U}_{i}^{(\Phi(t_{1},\theta_{-t_{1}}\omega_{0})b)}$$

or equivalently, it holds

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

¹²Following 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. Thus, we have

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

Since admissible tuples are linearly independent, the left-hand side has dimension d_i . The right-hand side must have the same dimension for t_1, t_2 large enough, because the intersection converges to $E_i(\omega_0)$. Hence, we have equality of subspaces, which concludes the proof.

4.4. Proof of Theorems. Lemma 4.12 describes how admissible tuples fare in Ginelli's algorithm. All that remains is to connect the lemma to measurement results from subsection 2.5.

Proof of Theorem 4.1. Fix $\epsilon > 0$. It is enough to assume that \mathcal{K} is a product of balls, i.e. $\mathcal{K} = B_d(0, M)^d \times (B(M))^r$ for some M > 0. Furthermore, we set $\delta(t) := \min\left(\frac{1}{t}, \frac{1}{2\sqrt{2}}\right)$, so that $\lambda(\frac{1}{\delta}) = 0$.

By Lemma 4.12 we have

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

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

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

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

$$\begin{split} \mu\Big(\Big(B_d(0,M)^d \times (B(M))^r\Big) \setminus \Big(\mathcal{A}d^{(\hat{v}^-(t_1))}(\delta(t_1)) \times \Big(\mathcal{A}d^{(e)r}_{\mathrm{ext}_-}\Big(\sqrt{2}\delta(t_2)\Big)\Big)^r\Big)\Big) \\ &\leq \mu\Big(B_d(0,M)^d \setminus \mathcal{A}d^{(\hat{v}^-(t_1))}(\delta(t_1))\Big)\mu((B(M))^r) \\ &\quad + \mu\Big(B_d(0,M)^d\Big)\mu\Big((B(M))^r \setminus \Big(\mathcal{A}d^{(e)r}_{\mathrm{ext}_-}\Big(\sqrt{2}\delta(t_2)\Big)\Big)^r\Big) \\ &= \mu\Big(B_d(0,M)^d \setminus \mathcal{A}d^{(e)}(\delta(t_1))\Big)\mu(B(M)) + \mu\Big(B_d(0,M)^d\Big)\mu\Big(B(M) \setminus \mathcal{A}d^{(e)r}_{\mathrm{ext}_-}\Big(\sqrt{2}\delta(t_2)\Big)\Big) \end{split}$$

Here, we used invariance under orthogonal transformations of $B_d(0, M)$ to switch from $(\hat{v}^-(t_1))$ to (e).

By Corollary 2.21 and Corollary 2.23 the estimate converges to zero as $\min(t_1, t_2)$ is increased. Hence, we get the desired convergence result.

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 $\delta_{\epsilon}(n) := \left(\frac{\epsilon}{\sqrt{2}n^2}\right)^{d-1}$ as our admissibility parameter satisfying $\lambda\left(\frac{1}{\delta_{\epsilon}}\right) = 0$ for each $0 < \epsilon < 1$.

Similar to before, we invoke Lemma 4.12 to find that

$$\overline{\lambda}\left(d\left(U_{i}^{\left(\Phi(-n_{2},\theta_{n_{2}}\omega_{0})b'\right)},E_{i}(\omega_{0})\right)\right) \leq -\min(|\lambda_{i}-\lambda_{i-1}|,|\lambda_{i}-\lambda_{i+1}|)$$

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

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

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

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

has measure zero¹³, which can be proved by exhausting the domain of $((b), (\alpha))$ with products of balls: It holds

$$\begin{split} \mu \bigg(\Big(B_d(0, M)^d \times (B(M))^r \Big) \setminus \\ & \bigcup_{0 < \epsilon < 1} \bigcap_{n_1, n_2 \in \mathbb{N}} \mathcal{A}d^{(\hat{v}^-(n_1))}(\delta_\epsilon(n_1)) \times \Big(\mathcal{A}d^{(e)r}_{ext_-} \Big(\sqrt{2}\delta_\epsilon(n_2) \Big) \Big)^r \Big) \\ & \leq \inf_{0 < \epsilon < 1} \left(\sum_{n_1 \in \mathbb{N}} \mu \Big(B_d(0, M)^d \setminus \mathcal{A}d^{(e)}(\delta_\epsilon(n_1)) \Big) \mu(B(M)) \right) \\ & + \sum_{n_2 \in \mathbb{N}} \mu \Big(B_d(0, M)^d \Big) \mu \Big(B(M) \setminus \mathcal{A}d^{(e)r}_{ext_-} \Big(\sqrt{2}\delta_\epsilon(n_2) \Big) \Big) \Big) \\ & \leq \inf_{0 < \epsilon < 1} \left(\sum_{n_1 \in \mathbb{N}} \eta_1(\delta_\epsilon(n_1))^{\frac{1}{d-1}} \mu(B(M)) + \sum_{n_2 \in \mathbb{N}} \mu \Big(B_d(0, M)^d \Big) \eta_2(\delta_\epsilon(n_2))^{\frac{1}{d-1}} \right) \\ & = \inf_{0 < \epsilon < 1} \epsilon \left(\sum_{n \in \mathbb{N}} \frac{\eta_1 \mu(B(M)) + \eta_2 \mu \big(B_d(0, M)^d \big)}{\sqrt{2}n^2} \right) \\ & = 0 \end{split}$$

for each M > 0. Here, it was crucial to use Proposition 2.24 and Proposition 2.27 to get a more precise measure estimate on nonadmissible tuples.

¹³Note 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 (4.1) is empty.

CONVERGENCE PROOF OF THE GINELLI ALGORITHM FOR CLVS

5. Conclusions. We defined Ginelli's algorithm as a mean 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. It rates a configuration based on how well it performs in the finite time scenario. A parameter close to 1 would imply fast convergence and a value near 0 indicates slow convergence. On the one hand we want to exclude nonadmissible tuples, but on the other hand we would like to cover as many configurations as possible. Hence, it was necessary to choose a time-dependent admissibility-parameter that converges to 0, while not influencing the speed of convergence on an exponential timescale. This way we were able to prove convergence in measure of Ginelli's algorithm and, moreover, convergence for almost all initial tuples in the discrete case.

Due to our notation it was possible to establish a direct connection between the speed of convergence and LEs. Namely, the algorithm converges exponentially fast with an exponent given by the minimal distance between LEs. Interestingly, this was already predicted and observed in applications [11, 12, 14, 28].

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. As it stands, one needs to choose a runtime prior to executing Ginelli's algorithm. Hence, without clever adjustments, it is necessary to repeat the whole algorithm after increasing the runtime. Especially, when probing the past or future, a fitting runtime might not be known due to insufficient data.

Numerical applications often do not provide exact integrations of the underlying system. Consequently, one is confronted with perturbed information of orbit and cocycle that might lead to false results in sensitive regions of the system. In particular, the possibly noncontinuous dependence of the Lyapunov spectrum on the choice of trajectory adds to the uncertainty. In this sense it would be interesting to know how perturbations affect the outcome of Ginelli's algorithm in numerical simulations as well as in analytical computations.

Ultimately, a wide range of applications [3, 6, 7, 9, 12, 14, 15, 16, 19, 22, 23, 24, 25, 26, 27, 28, 31, 32] underline the importance of CLVs for dynamical systems. Our convergence proof not only verifies the use of Ginelli's algorithm in these applications, but encourages 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 [30], 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 an important instrument 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 German Research Foundation. A special thanks goes to my colleagues, in particular, R. Lauterbach and I. Gasser for frequent feedback and S. Schubert for motivational discussions on applications of CLVs.

REFERENCES

- L. ARNOLD, Random Dynamical Systems, Springer Monographs in Mathematics, Springer-Verlag Berlin Heidelberg New York, 1998, https://doi.org/10.1007/978-3-662-12878-7.
- [2] L. ARNOLD AND N. D. CONG, Linear cocycles with simple lyapunov spectrum are dense in l[∞], Ergodic Theory and Dynamical Systems, 19 (1999), pp. 1389–1404.
- [3] M. W. BEIMS AND J. A. C. GALLAS, Alignment of lyapunov vectors: A quantitative criterion to predict catastrophes?, Scientific reports, 6 (2016), p. 37102, https://doi.org/10.1038/srep37102.
- [4] G. BENETTIN, L. GALGANI, A. GIORGILLI, AND J.-M. STRELCYN, Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. part 1: Theory, Meccanica, 15 (1980), pp. 9–20, https://doi.org/10.1007/BF02128236.
- [5] G. BENETTIN, L. GALGANI, A. GIORGILLI, AND J.-M. STRELCYN, Lyapunov characteristic exponents for smooth dynamical systems and for hamiltonian systems; a method for computing all of them. part 2: Numerical application, Meccanica, 15 (1980), pp. 21–30, https://doi.org/10.1007/BF02128237.
- [6] H. BOSETTI AND H. A. POSCH, Covariant lyapunov vectors for rigid disk systems, Chemical physics, 375 (2010), pp. 296–308, https://doi.org/10.1016/j.chemphys.2010.06.010.
- H. BOSETTI AND H. A. POSCH, Orthogonal versus covariant lyapunov vectors for rough hard disc systems, Journal of Physics A: Mathematical and Theoretical, 46 (2013), p. 254011, https://doi.org/10.1088/ 1751-8113/46/25/254011.
- [8] R. BOWEN AND D. RUELLE, The ergodic theory of axiom a flows, in The Theory of Chaotic Attractors, Springer, 1975, pp. 55–76, https://doi.org/10.1007/978-0-387-21830-4_5.
- [9] G. CONTI AND G. BADIN, Hyperbolic covariant coherent structures in two dimensional flows, Fluids, 2 (2017), https://doi.org/10.3390/fluids2040050.
- [10] F. DEUTSCH, The Angle Between Subspaces of a Hilbert Space, Springer Science+Business Media Dordrecht, 1995, pp. 107–130, https://doi.org/10.1007/978-94-015-8577-4_7.
- [11] S. V. ERSHOV AND A. B. POTAPOV, On the concept of stationary lyapunov basis, Physica D: Nonlinear Phenomena, (1998), pp. 167–198, https://doi.org/10.1016/S0167-2789(98)00013-X.
- [12] G. FROYLAND, T. HÜLS, G. P. MORRISS, AND T. M. WATSON, Computing covariant lyapunov vectors, oseledets vectors, and dichotomy projectors: A comparative numerical study, Physica D: Nonlinear Phenomena, 247 (2013), pp. 18–39, https://doi.org/10.1016/j.physd.2012.12.005.
- [13] A. GALÁNTAI, Projectors and Projection Methods, no. 6 in Advances in Mathematics, Springer US, 2004, https://doi.org/10.1007/978-1-4419-9180-5.
- [14] F. GINELLI, H. CHATÉ, R. LIVI, AND A. POLITI, Covariant lyapunov vectors, Journal of Physics A: Mathematical and Theoretical, 46 (2013), p. 254005, https://doi.org/10.1088/1751-8113/46/25/254005.
- [15] F. GINELLI, P. POGGI, A. TURCHI, H. CHATÉ, R. LIVI, AND A. POLITI, Characterizing dynamics with covariant lyapunov vectors, Physical Review Letters, 99 (2007), p. 130601, https://doi.org/10.1103/ PhysRevLett.99.130601.
- [16] M. INUBUSHI, S.-I. TAKEHIRO, AND M. YAMADA, Regeneration cycle and the covariant lyapunov vectors in a minimal wall turbulence, Physical Review E, 92 (2015), p. 023022, https://doi.org/10.1103/ PhysRevE.92.023022.
- T. KATO, Perturbation theory for linear operators, no. 132 in Classics in Mathematics, Springer-Verlag Berlin Heidelberg, 1995, https://doi.org/10.1007/978-3-642-66282-9.
- [18] P. V. KUPTSOV AND U. PARLITZ, Theory and computation of covariant lyapunov vectors, Journal of Nonlinear Science, 22 (2012), pp. 727–762, https://doi.org/10.1007/s00332-012-9126-5.
- [19] G. P. MORRISS AND D. P. TRUANT, A review of the hydrodynamic lyapunov modes of hard disk systems,

CONVERGENCE PROOF OF THE GINELLI ALGORITHM FOR CLVS

Journal of Physics A: Mathematical and Theoretical, 46 (2013), p. 254010, https://doi.org/10.1088/1751-8113/46/25/254010.

- [20] V. I. OSELEDETS, A multiplicative ergodic theorem: Lyapunov characteristic exponents for dynamical systems, Trudy Mosk. Mat. Obsc., 19 (1968), pp. 179 – 210.
- [21] D. RUELLE, Ergodic theory of differentiable dynamical systems, Publications mathématiques de l'IHÉS, 50 (1979), pp. 27–58, https://doi.org/10.1007/BF02684768.
- [22] Y. SAIKI AND M. U. KOBAYASHI, Numerical identification of nonhyperbolicity of the lorenz system through lyapunov vectors, JSIAM Letters, 2 (2010), pp. 107–110, https://doi.org/10.14495/jsiaml.2.107.
- [23] S. SCHUBERT AND V. LUCARINI, Covariant lyapunov vectors of a quasi-geostrophic baroclinic model: analysis of instabilities and feedbacks, Quarterly Journal of the Royal Meteorological Society, 141 (2015), pp. 3040–3055, https://doi.org/10.1002/qj.2588.
- [24] S. SCHUBERT AND V. LUCARINI, Dynamical analysis of blocking events: spatial and temporal fluctuations of covariant lyapunov vectors, Quarterly Journal of the Royal Meteorological Society, 142 (2016), pp. 2143–2158, https://doi.org/10.1002/qj.2808.
- [25] N. SHARAFI, M. TIMME, AND S. HALLERBERG, Critical transitions and perturbation growth directions, Physical Review E, 96 (2017), p. 032220, https://doi.org/10.1103/PhysRevE.96.032220.
- [26] K. A. TAKEUCHI, H.-L. YANG, F. GINELLI, G. RADONS, AND H. CHATÉ, Hyperbolic decoupling of tangent space and effective dimension of dissipative systems, Physical Review E, 84 (2011), p. 046214, https:// doi.org/10.1103/PhysRevE.84.046214.
- [27] D. P. TRUANT AND G. P. MORRISS, Backward and covariant lyapunov vectors and exponents for harddisk systems with a steady heat current, Physical Review E, 90 (2014), p. 052907, https://doi.org/10. 1103/PhysRevE.90.052907.
- [28] S. VANNITSEM AND V. LUCARINI, Statistical and dynamical properties of covariant lyapunov vectors in a coupled atmosphere-ocean model—multiscale effects, geometric degeneracy, and error dynamics, Journal of Physics A: Mathematical and Theoretical, 49 (2016), p. 224001, https://doi.org/10.1088/ 1751-8113/49/22/224001.
- [29] J. VON NEUMANN, The geometry of orthogonal spaces, functional operators-vol. ii, Annals of Math. Studies, 22 (1950).
- [30] C. L. WOLFE AND R. M. SAMELSON, An efficient method for recovering lyapunov vectors from singular vectors, Tellus A: Dynamical Meteorology and Oceanography, 59 (2007), pp. 355–366, https://doi. org/10.1111/j.1600-0870.2007.00234.x.
- [31] M. XU AND M. R. PAUL, Covariant lyapunov vectors of chaotic rayleigh-bénard convection, Physical Review E, 93 (2016), p. 062208, https://doi.org/10.1103/PhysRevE.93.062208.
- [32] H.-L. YANG, K. A. TAKEUCHI, F. GINELLI, H. CHATÉ, AND G. RADONS, Hyperbolicity and the effective dimension of spatially extended dissipative systems, Physical review letters, 102 (2009), p. 074102, https://doi.org/10.1103/PhysRevLett.102.074102.