EMERGENCE VIA NON-EXISTENCE OF AVERAGES

SHIN KIRIKI, YUSHI NAKANO, AND TERUHIKO SOMA

Abstract. Inspired by a recent work by Berger, we introduce the concept of pointwise emergence. This concept provides with a new quantitative perspective into the study of non-existence of averages for dynamical systems. We show that high pointwise emergence on a large set appears for abundant dynamical systems: Any continuous maps on a compact metric space with the specification property have super-polynomial pointwise emergence on a residual subset of the state space. Furthermore, there is a dense subset of any Newhouse open set each element of which has super-polynomial pointwise emergence on a positive Lebesgue measure subset of the state space.

1. Introduction

The study of infinitude or non-existence of averages for dynamical systems has a long history, despite being beyond the Smale-Palis program [39, 38, 15], which has been a guiding principle in modern dynamical systems theory. It is only 70’s that Newhouse showed in [35] that there is a residual subset of any Newhouse open set (see Section 2 for precise definition) each element of which has infinitely many sinks. Furthermore, it is Bowen who first studied dynamics without time averages on a positive Lebesgue measure set (although it was never published by himself, see [45]). On the other hand, it is recent that Berger [9] proved that typical dynamics (in the sense of Kolmogorov) in any Newhouse open set have infinitely many sinks, and that the first and third authors [32] showed that there is a dense subset of any Newhouse set each element of which has a positive Lebesgue measure set where time averages do not exist. We refer to [15, 9, 32] for detailed history.

Recently, Berger [10] introduced a quantitative viewpoint into the study of infinitude of averages, and further developed it in [13] with Bochi. In the paper [10], a “global” \( \epsilon \)-approximation of empirical measures (i.e. measures representing averages) of a dynamical system is called emergence at scale \( \epsilon > 0 \) (we note that “emergence” is one of the most important concepts in complexity science [34], but had no rigorous formulation before [10] appeared), and it is shown that the growth rate of emergence in the limit \( \epsilon \to 0 \) captures the complexity of a dynamical system with infinitude of averages. Our purpose in this paper is to investigate “local” emergence (called pointwise emergence, Definition 1.1). We will see that pointwise emergence well adapts to the study of complexity of non-existence of averages, resulting in a strong contrast between pointwise emergence and Berger’s emergence (see Section 1.3). Furthermore, we prove that high pointwise emergence on a large set appears for abundant dynamical systems (Proposition 1.2 and Theorem A).
1.1. **Emergences.** We first briefly recall the definition of Berger’s emergence. Let \( X \) be a compact metric space and \( f : X \to X \) a continuous map. We study empirical measures \( \{ \delta^n_x \}_{n \geq 1} \) given by

\[
\delta^n_x \equiv \delta^n_x(f) = \frac{1}{n} \sum_{j=0}^{n-1} \delta_{f^j(x)} \quad (x \in X, \ n \geq 1),
\]

where \( \delta_y \) is the Dirac measure at \( y \in X \). Note that \( \int \varphi d\delta^n_x = 1/n \sum_{j=0}^{n-1} \varphi(f^j(x)) \) is the (partial) time average of a continuous function \( \varphi \) (along the orbit of \( x \) by \( f \)) at \( n \geq 1 \), so that the study of asymptotic behavior of \( \{ \delta^n_x \}_{n \geq 1} \) in weak topology would be most fundamental in ergodic theory. We metrize the weak topology of the space \( \mathcal{P}(X) \) of probability measures on \( X \) by the first Wasserstein metric \( d \equiv W_1 \) (see Section 2 for the definition of \( W_1 \): recall that convergence with respect to \( W_1 \) is equivalent to the weak convergence (refer to e.g. [51] Theorems 6.9); other classical metrics to metrize the weak topology of \( \mathcal{P}(X) \), such as Lévy-Prokhorov metric, were also considered in [13]). By virtue of (a straightforward modification of) Proposition 1.10 of [10], \( x \mapsto d(\delta^n_x, \nu) \) is continuous for any \( n \geq 1 \) and \( \nu \in \mathcal{P}(X) \).

Let \( M \) be a compact manifold and \( f \) a continuous map on \( M \). In [10], Berger defined the emergence \( \delta_{Leb}(\epsilon) \equiv \delta_{Leb}(\epsilon, f) \) of \( f \) at scale \( \epsilon > 0 \) by

\[
(1.1) \quad \delta_{Leb}(\epsilon) = \min \left\{ N \in \mathbb{N} \mid \text{there exists } \{ \mu_j \}_{j=1}^N \subset \mathcal{P}(M) \text{ such that} \right. \\
\left. \lim\sup_{n \to \infty} \int_M \min_{1 \leq j \leq N} d(\delta^n_x, \mu_j) d\text{Leb}(x) \leq \epsilon \right\},
\]

where \( \text{Leb} \) is the normalized Lebesgue measure on \( M \). This was called metric emergence in [13], because they needed to distinguish it with another emergence (called topological emergence, see [17] for definition). We also use their terminology, that is, \( \delta_{Leb}(\epsilon) \) will be called metric emergence at scale \( \epsilon > 0 \). (To be more precise, in [13] they also studied metric emergences \( \delta_{\mu}(\epsilon) \) for any probability measure \( \mu \) on a compact metric space \( X \) (not necessarily a manifold) defined by (1.1) with \( \mu \) instead of \( \text{Leb} \), and obtained a variational principle for metric and topological emergences.)

The inequality in (1.1) means that \( \{ \mu_j \}_{j=1}^N \) approximates the statistics of \( f \) in the \( \epsilon \) scale. Hence, once one fixes \( \epsilon \), the complexity of statistics of \( f \) “emerges” as \( \delta_{Leb}(\epsilon) \). Interesting examples are as follows: It is shown in [10] Section 1.2] that if \( f \) has finitely many ergodic probability measures such that the union of basins of the measures covers \( M \) up to a zero Lebesgue measure set, then \( \delta_{Leb}(\epsilon) \) is bounded by the number of the measures for any \( \epsilon \). On the other hand, \( \lim_{\epsilon \to 0} \delta_{Leb}(\epsilon) = \infty \) if \( f \) has infinitely many sinks ([10 Claim 1.13] or if \( f \) is a conservative system on the annulus \( S^1 \times [0, 1] \) which preserves each circle \( S^1 \times \{ \rho \} \) with \( \rho \in [0, 1] \) and satisfies a very mild condition ([13] Proposition 4.1)). In this sense, we may say that metric emergence well captures infinitude of averages.

Under the background of naive and massive uses of computer approximation of statistics in many branches of sciences, Berger started a program to prove that for each typical dynamics (in the sense of Kolmogorov) in an open set of the space of diffeomorphisms, the metric emergence is super-polynomial, that is, \( \limsup_{\epsilon \to 0} \log \delta_{Leb}(\epsilon)/(-\log \epsilon) = \infty \) (or equivalently, \( \limsup_{\epsilon \to 0} \epsilon^\alpha \delta_{Leb}(\epsilon) = \infty \) for any \( \alpha \geq 0 \), see [10] Problem 1.14). Among computer scientists, an algorithm of super-polynomial complexity is thought to be not feasible in practice by a computer
so that the accomplishment of the program may give an alarm to the aforementioned optimistic trend. A great contribution to the program was recently made in \[13\]. We also remember that another quantitative study of generic non-hyperbolic dynamics by Kaloshin \[29\] (i.e. super-exponential growth of number of periodic orbits for generic dynamics in Newhouse open sets, in which infinitely many sinks exist for generic dynamics) opened up a fruitful research field.

A feature of the metric emergence is the integration in (1.1) resulting in a grasp of “global” statistical information of the dynamics. In this paper, we consider “local” emergence as follows. Let \(X\) be a compact metric space (not necessarily a manifold).

**Definition 1.1.** Given \(\epsilon > 0\) and \(x \in X\), the pointwise emergence \(E_x(\epsilon) \equiv E_x(\epsilon, f)\) of \(f\) at scale \(\epsilon\) at \(x\) is defined by

\[
E_x(\epsilon) = \min \left\{ N \in \mathbb{N} \mid \text{there exists} \{\mu_j\}_{j=1}^N \subset \mathcal{P}(X) \text{ such that} \limsup_{n \to \infty} \min_{1 \leq j \leq N} d(\delta_n^x, \mu_j) \leq \epsilon \right\}.
\]

The pointwise emergence at \(x \in X\) is called super-polynomial if

\[
\limsup_{\epsilon \to 0} \frac{\log E_x(\epsilon)}{-\log \epsilon} = \infty.
\]

**1.2. Historic behavior.** We can see that the pointwise emergence gives a quantitative perspective into non-existence of averages, or historic behavior. Recall that a point \(x \in X\) (or its forward orbit) is said to have historic behavior if the time average \(\lim_{n \to \infty} \delta_n^x\) does not exist. (This terminology originates from Ruelle \[43\]; see also \[46\].) Although the set of points with historic behavior is a \(\mu\)-zero measure set for any invariant measure \(\mu\) due to Birkhoff’s ergodic theorem (so that the set is called the irregular set or the non-typical set in the context of thermodynamic formalism \[5, 48\]), the set is known to be remarkably large for many dynamical systems.

Known dynamical systems with historic behavior on a measure-theoretically large set are as follows. It is a famous folklore that Bowen knew that a surface flow with heteroclinically connected two dissipative saddle points has a positive Lebesgue measure set consisting of points with historic behavior (see \[26, 45\] for precise proof). We emphasize that for Bowen’s example, there are many “abnormal” results other than historic behavior, refer to e.g. \[1, 2, 3, 37, 3\]. However, Bowen’s example is easily broken by small perturbations, and thus Takens asked in \[46\] whether there is a persistent class of diffeomorphisms for which the set of points with historic behavior is of positive Lebesgue measure (called Takens’ Last Problem). The first and third authors \[32\] affirmatively answered it by showing that there is a dense subset of any Newhouse open set in the set of \(C^r\) surface diffeomorphisms (\(2 \leq r < \infty\)) such that any element of the dense set has a wandering domain consisting of points with historic behavior, by employing the best technology developed by Colli-Vargas \[21\] for wandering domains near homoclinic tangency. Very recently, Berger and Biebler extended it to the \(C^\infty\) and analytic cases \([12]\).

One can also find other interesting examples with a positive Lebesgue measure set consisting of points with historic behavior for some quadratic maps in \[28\], for flows generated by 3-dimensional vector fields (in a locally dense set) with heteroclinic
cycles between periodic solutions in [33], and for some partially hyperbolic dynamics which is a compactification of an \(\mathbb{R}\)-extension of an Anosov diffeomorphism sharing properties with the Brownian motion on \(\mathbb{R}\) in [22].

From topological viewpoint, we can find more examples with historic behavior on a large set. Sigmund essentially showed in [44] that any continuous map on a compact metric space with the specification property has a residual subset of the state space consisting of points with historic behavior (this is explicitly stated and proven by himself in [23, Proposition 21.18]). See Subsection 2.1 for the definition of the specification property. We here just remember that any topologically mixing subshift of finite type satisfies the specification property ([23, Proposition 21.2]). This result was extended to shifts with weak specification in [6], geometric Lorenz flows in [31], sectionally hyperbolic flows in [3], and \(C^1\)-generic diffeomorphisms with non-hyperbolic homoclinic classes in [8]. In the context of thermodynamic formalism, a large contribution to historic behavior was also made by several authors. A very incomplete list of them is [42, 5, 18, 48, 16, 17, 14, 7]. We here merely mention that Pesin and Pitskel’ [42] showed that full shifts curries full topological entropy and full Hausdorff dimension on the set of points with historic behavior.

A fundamental relation between historic behavior and pointwise emergence is as follows. Let \(A_x \equiv A_x(f)\) be the set of accumulation points of \(\{\delta^n_x(f)\}_{n \geq 1}\) with respect to \(d\). Notice that \(x\) has historic behavior if and only if \(\#A_x(f) > 1\). For a subset \(\mathcal{Y}\) of a compact metric space \(X\), let \(N(\epsilon, \mathcal{Y})\) be the \(\epsilon\)-covering number of \(\mathcal{Y}\) by closed balls, and denote the upper and lower box-counting dimension of \(\mathcal{Y}\) by \(\dim(\mathcal{Y})\) and \(\dim(\mathcal{Y})\) respectively, that is,

\[
\dim(\mathcal{Y}) = \limsup_{\epsilon \to 0} \frac{\log N(\epsilon, \mathcal{Y})}{-\log \epsilon}, \quad \dim(\mathcal{Y}) = \liminf_{\epsilon \to 0} \frac{\log N(\epsilon, \mathcal{Y})}{-\log \epsilon}.
\]

When they coincide, we simply write it as \(\dim(\mathcal{Y})\). Then, it is straightforward to observe that

\[
\mathcal{E}_x(\epsilon, f) = N(\epsilon, A_x(f)),
\]

and thus, by the well-known fact that \(A_x(f)\) is a connected set in \(\mathcal{P}(X)\) (cf. [23, Proposition 3.8]), we conclude that \(x\) has historic behavior if and only if

\[
\liminf_{\epsilon \to 0} \frac{\log \mathcal{E}_x(\epsilon, f)}{-\log \epsilon} = \dim(A_x(f)) \geq 1
\]

(in particular, \(\mathcal{E}_x(\epsilon, f)\) diverges as \(\epsilon \to 0\)).

From this dimensional perspective, we can easily get the following useful criterion for super-polynomial pointwise emergence (recall that each ergodic measure in \(\mathcal{P}_f(X)\) is an extremal point; cf. [23, Proposition 5.6]): If \(A_x(f)\) includes the convex hull of infinitely many distinct ergodic invariant probability measures, then

\[
\limsup_{\epsilon \to 0} \frac{\log \mathcal{E}_x(\epsilon, f)}{-\log \epsilon} = \dim(A_x(f)) = \infty,
\]

that is, the pointwise emergence at \(x\) diverges super-exponentially fast. In the previously mentioned paper by Sigmund [44], he in fact showed that if \(f : X \to X\) satisfies the specification property, then there exists a residual subset \(R\) of \(X\) such that \(A_x(f) = \mathcal{P}_f(X)\) for any \(x \in R\), where \(\mathcal{P}_f(X)\) is the set of \(f\)-invariant probability measures on \(X\). Furthermore, for any continuous map with the specification property, the set of periodic points is dense in \(X\) (see [44, Proposition 2]), in particular, if \(X\) is an infinite set, then \(\mathcal{P}_f(X)\) includes infinitely many distinct ergodic
invariant probability measures (note that any dense subset of an infinite metric space is an infinite set). By combining these results with (1.5), we immediately get the following conclusion for super-polynomial pointwise emergence.

**Proposition 1.2.** Let $X$ be an infinite compact metric space and $f : X \to X$ a continuous map with the specification property. Then, there is a residual subset $R$ of $X$ such that

$$\lim_{\epsilon \to 0} \frac{\log \delta_x(\epsilon, f)}{-\log \epsilon} = \infty \text{ for all } x \in R. \tag{1.6}$$

**Remark.** It seems that super-polynomial pointwise emergence in (1.6) holds for more general classes of dynamical systems without the specification property, such as dynamics with historic behavior in [6, 31, 5, 8]. Moreover, it is of great interest to see whether one can develop thermodynamic formalism on the set of points with super-polynomial pointwise emergence, refer to e.g. [42, 5, 48].

1.3. **Metric and topological emergences.** A formula similar to (1.3) is seen in [13, Proposition 3.14] for metric emergence: if $f$ is a conservative map, then

$$\delta_{\text{Leb}}(\epsilon, f) \leq N(\epsilon, \mathcal{P}_{\text{erg}}(f)), \tag{1.7}$$

where $\mathcal{P}_{\text{erg}}(f)$ is the set of ergodic probability measures of $f$. In [13], the quantity of the right-hand side of (1.7) is called the topological emergence of $f$ at scale $\epsilon > 0$, and its complexity and connection with metric emergence were deeply investigated. Conformal expanding repellers and hyperbolic sets of conservative surface diffeomorphisms are important examples for which the inequality in (1.7) is strict ([13, Theorem A]). Note that there is no inclusion relationship between $A_x$ and $\mathcal{P}_{\text{erg}}(f)$ in general (see examples below), and so is between pointwise and topological emergences. A basic property for pointwise and metric emergences is the following (its proof will be given in Section 3).

**Lemma 1.3.** Let $f : M \to M$ be a continuous map on a compact manifold $M$. For any $\epsilon > 0$ and Borel set $D \subset M$ of positive Lebesgue measure,

$$\min_{x \in D} \delta_x(\epsilon, f) \leq \delta_{\text{Leb}}(\text{Leb}(D)\epsilon, f).$$

We summarize the differences between metric/topological and pointwise emergences: unlike the similarity in definition, the properties of these emergences are rather in strong contrast. Firstly, recall that if $f$ has infinitely many sinks, then its metric emergence $\delta_{\text{Leb}}(\epsilon)$ diverges in the limit $\epsilon \to 0$ ([10 Claim 1.13]). On the other hand, it is obvious that for such dynamics $f$, the pointwise emergence is trivial (i.e. $\delta_x(\epsilon) = 1$ for any $\epsilon > 0$) on the basin of the sinks. Similarly, in [13 Section 4], many conservative systems $f$ on the annulus with super-polynomially diverging metric emergence were constructed, while, since the constructed dynamics decomposes the annulus into $f$-invariant circles, one can easily see that pointwise emergence of the conservative system is minimal everywhere. That is, the inequality in Lemma 1.3 is strict for such dynamical systems. Conversely, by virtue of Proposition 1.2 for expanding maps or Anosov diffeomorphisms on a compact manifold $M$, the pointwise emergence diverges super-polynomially fast on a residual subset of $M$, while the metric emergence is bounded because there exist finitely many SRB measures whose basins cover $M$ up to a zero Lebesgue measure set (cf. [15]).

Furthermore, Bowen’s example has only finitely many ergodic probability measures under an appropriate setting (cf. [27]), so that its topological emergence is
bounded, while the pointwise emergence diverges with polynomial order of degree at least 1 on a positive Lebesgue measure set because it has historic behavior and [1.4] holds. Conversely, for expanding maps or Anosov diffeomorphisms, the pointwise emergence is bounded almost everywhere, while the topological emergence may diverge super-polynomially fast due to [13, Theorem A].

1.4. Main result. By examples in Subsection 1.2 together with (1.4), one can find many dynamical systems whose pointwise emergence diverges on a topologically or measure-theoretically large set. In fact, we saw in Proposition 1.2 that there are abundant dynamical systems with high pointwise emergence on a residual set, due to the established theory for $\mathcal{A}_x$. However, in measure-theoretic context, to the best of our knowledge, the box-counting dimension of $\mathcal{A}_x$ (i.e. the degree of polynomial growth of pointwise emergence) for any known dynamics with historic behavior is only 1, except some quadratic maps investigated by Keller and Hofbauer [28]. Furthermore, as mentioned in Subsection 1.3, any known dynamical system with super-polynomial metric emergence is not helpful to construct high pointwise emergence. However, we can show that there are abundant dynamical systems with super-polynomial pointwise emergences on a positive Lebesgue measure set, which is our main result.

Let $\text{Diff}^r(M)$ be the space of $C^r$ diffeomorphisms on a closed surface $M$. In this paper, we mean by a basic set a compact hyperbolic and locally maximal invariant set which is transitive and contains a dense subset of periodic orbits. We tacitly assume throughout this paper that any basic set is not a single orbit. Recall that a non-empty connected open set $D$ is called a wandering domain of $f$ if $f^i(D) \cap f^j(D) = \emptyset$ for all nonnegative integers $i, j$ with $i \neq j$. Theorem A. There exists a dense subset $D$ of any Newhouse open set (definition given in Section 2) of $\text{Diff}^r(M)$ with $2 \leq r < \infty$ such that for each $f \in D$, one can find a wandering domain (in particular, a positive Lebesgue measure set) $D \subset M$ such that the union of $\omega$-limit set of each point in $D$ includes a basic set $\Lambda$ and

$$(1.8) \quad \mathcal{A}_x \supset \{(1 - \zeta)\delta_{\hat{p}} + \zeta \mu \mid \mu \in \mathcal{P}_f(\Lambda)\} \quad \text{for all } x \in D$$

with some $\zeta \in [0, 1]$ and a saddle fixed point $\hat{p} \in \Lambda$ (refer to Theorem 2.3), where $\mathcal{P}_f(\Lambda)$ is the set of $f$-invariant probability measures whose supports are included in $\Lambda$. Furthermore,

$$(1.9) \quad \lim_{\epsilon \to 0} \frac{\log E_x(\epsilon, f)}{-\log \epsilon} = \infty \quad \text{for all } x \in D.$$

Notice that (1.9) is an immediate consequence of (1.8) due to (1.5). By Lemma 1.3, we can contribute to the previously mentioned Berger program for metric emergence [11, Problem 1.14] as follows:

Corollary 1.4. There exists a dense subset $D$ of any Newhouse open set of $\text{Diff}^r(M)$ with a closed surface $M$ and $2 \leq r < \infty$ such that for each $f \in D$,

$$\limsup_{\epsilon \to 0} \frac{\log E_{\text{Leb}}(\epsilon, f)}{-\log \epsilon} = \infty.$$

We emphasize that Theorem A would be substantially stronger than Corollary 1.4 because metric emergence quantifies infinitude of averages while pointwise emergence quantifies non-existence of averages (historic behavior) as explained in...
Section 1.3 Therefore, we would rather say that Theorem A is a result about a quantitative version of Takens’ Last Problem in the spirit of Berger program.

Remark. Berger and Bochi proved in [13, Theorem D] that there exists a residual subset of any Newhouse open set of $\text{Diff}^r(M)$ with $\dim(M) = 2$ and $1 \leq r \leq \infty$ such that any element of the subset enjoys super-polynomial metric emergence (although the result is proven by reducing it to a conservative surface diffeomorphism whose metric emergence is super-polynomial but pointwise emergence is everywhere minimal). That is, Corollary 1.4 is an alternative proof of a part of their result, and so, it is natural to ask whether Theorem A holds with a residual set instead of the dense set $D$.

Remark. After we completed the proof of super-polynomial pointwise emergence in Theorem A, we learned from Pierre Berger and Sébastien Biebler that they independently obtained a similar result, although their proof is quite different from ours (for example, they use a geometric model for parameter families of surface real mappings while we do not). We also refer to [11, 47] for other recent results about emergences.

1.5. Stretched exponential emergences. In this subsection, we give a supplementary result related with stretched exponential growths of pointwise emergences: we separated the results in this subsection from Section 1.4 because the proofs essentially use the results of Berger-Bochi [13] and Berger-Biebler [12], while the proof of Theorem A is self-contained.

Let $f : X \to X$ be a continuous map on a metric space $X$. Let $\mathcal{P}_f(X)$ be the set of $f$-invariant probability measures on $X$ equipped with the first Wasserstein metric $d$. Then, it is not difficult to see that $\mathcal{A}_x \subset \mathcal{P}_f(X)$ for all $x \in X$ (cf. [23, Proposition 3.8]). On the other hand, it follows from [13, Theorem 1.3] that

$$\limsup_{\epsilon \to 0} \frac{\log \log N(\epsilon, \mathcal{P}_f(X))}{-\log \epsilon} \leq \dim(X)$$

(see Subsection 1.2 for the definition of $\dim(X)$). So, it follows from [13] that

$$\limsup_{\epsilon \to 0} \frac{\log \log \mathcal{E}_x(\epsilon, f)}{-\log \epsilon} \leq \dim(X) \quad \text{for any } x \in X.$$

Hence, it is natural to ask when the above inequality is an equality (i.e. when stretched exponential pointwise emergences with maximal exponent are observed). We refer to [13] for a systematic study of stretched exponential growths for topological and metric emergences.

We first give an answer for a special case of dynamics with the specification property, that is, subshifts with the specification property, as follows.

Proposition 1.5. Let $X \subset \{1, 2, \ldots, m\}^\mathbb{N}$ be a subshift ($m \geq 2$), endowed with a standard metric $d_X(x, y) = \sum_{j=0}^{\infty} \frac{|x_j - y_j|}{\beta^j}$ for $x = (x_0, x_1, \ldots), y = (y_0, y_1, \ldots) \in X$ with some $\beta > 1$. Let $f : X \to X$ be the left shift operator. Assume that $f$ satisfies the specification property. Then, there is a residual subset $R$ of $X$ such that

$$\lim_{\epsilon \to 0} \frac{\log \log \mathcal{E}_x(\epsilon, f)}{-\log \epsilon} = \dim(X) \quad \text{for any } x \in R.$$
We recall Furstenberg’s formula \( \dim(X) = \frac{\htop(f)}{\log \beta} \) for subshifts \((25)\), where \( \htop(f) \) is the topological entropy of \( f \). Using this formula, we will show that

\[
\liminf_{\epsilon \to 0} \frac{\log \log N(\epsilon, P_f(X))}{-\log \epsilon} \geq \dim(X)
\]

for \( f \) in Proposition 1.3 under the help of Berger-Bochi’s key estimate, see Appendix A. Hence, Proposition 1.5 immediately follows from (1.3), (1.11) and the aforementioned Sigmund’s theorem for continuous maps with the specification property. Note that (1.10) and (1.12) also imply the stretched exponential topological emergence with the maximal exponent \( \dim(X) \) because \( \mathcal{P}_{\text{erg}}(f) \) is dense in \( \mathcal{P}_f(X) \) for any continuous map on a compact metric space with the specification property (cf. \((23)\)). Furthermore, one may get a two-sided subshift version of Proposition 1.5 with a small effort, and (the version of) Proposition 1.5 may be directly applicable to basic sets of conservative surface diffeomorphisms or conformal repellers, for which maximal topological emergence is known in \((13, \text{Theorem A})\).

Next we consider dynamics in Newhouse open sets. We have no idea of whether

\[
\liminf_{\epsilon \to 0} \frac{\log \log N(\epsilon, P_f(\Lambda))}{-\log \epsilon} \geq \dim(\Lambda)
\]

holds (compare with \((13, \text{Theorem 2.4})\) in which \( f \) is a conservative surface diffeomorphism having a basic set). On the other hand, we have the following result by Berger and Biebler.

**Theorem 1.6.** \((12, \text{Theorem 0.4})\) \(\) Let \( f : M \to M \) be a \( C^r \) surface diffeomorphism \((1 < r \leq \infty)\) on a closed surface \( M \) having a basic set \( \Lambda \) of saddle type. Then,

\[
\liminf_{\epsilon \to 0} \frac{\log \log N(\epsilon, P_f(\Lambda))}{-\log \epsilon} \geq \dim_u(\Lambda).
\]

where \( \dim_u(\Lambda) > 0 \) is the box-counting dimension of \( \Lambda \cap W^u_{\text{loc}}(x) \) for any \( x \in \Lambda \).

Under the establishment of Theorem 1.6, Theorem A together with (1.3) implies stretched exponential pointwise emergences with a positive exponent:

**Corollary 1.7.** Let \( f : M \to M \) be the surface diffeomorphism and \( D \subset M \) the wandering domain given in Theorem A. Then,

\[
\liminf_{\epsilon \to 0} \frac{\log \log \mathcal{E}_u(\epsilon, f)}{-\log \epsilon} \geq \dim_u(\Lambda) \quad \text{for any } x \in D.
\]

### 2. Key definitions and outline of proof

In this section, we provide key definitions used in the proof of main theorem, and briefly explain outline of the proof.

---

1. Here we mean by saddle type that the tangent bundle \( TM \) over \( \Lambda \) can be decomposed into one-dimensional stable and unstable bundles. Theorem 0.4 of \((12)\) only dealt with the case when \( \Lambda \) is a horseshoe. However, the proof was given by projecting the saddle-type basic set of the surface to a repeller of the real line and applying \((13, \text{Theorem A})\), which is proven for any conformal repeller. Thus, the proof of \((12, \text{Theorem 0.4})\) can be applied literally to obtain Theorem 1.6.

2. Note that any saddle-type basic set of a surface is \( u \)-conformal in the sense of \((41, \text{Section 22})\), so it follows from \((41, \text{Theorem 22.1})\) that \( \dim(W^u_{\text{loc}}(x) \cap \Lambda) \) is independent of \( x \in \Lambda \) and positive.
2.1. Preliminary definitions. We first give precise definitions to undefined terminologies in Section 1. Let $X$ be a compact metric space endowed with a metric $d_X$. A continuous map $f : X \to X$ on $X$ is said to satisfy the specification property if for any $\varepsilon > 0$, there exists a constant $\tau(\varepsilon) \geq 0$ such that for any integer $K \geq 1$, any points $p_1, \ldots, p_K \in X$, any integers $n_1, \ldots, n_K$ and $m_1, \ldots, m_K$ satisfying $m_k - n_k \geq \tau(\varepsilon)$ for any $1 \leq k \leq K$, there exists a periodic point $x \in X$ of period $N_K$ such that, with $N_k = \sum_{j=1}^k m_j$ and $N_0 = 0$, for every $1 \leq k \leq K$,

$$f^n(x) \in B_r(f^n(p_k)) \quad \text{if } n \in I_k := [N_{k-1}, N_{k-1} + n_k - 1]$$

where $B_r(y)$ is the ball with radius $r > 0$ and centered at $y \in X$. Refer to [23].

For $j = 1, 2$, let $p_j : X \times X \to X$ be the canonical projection to the $j$-th coordinate, and $(p_j)_\# \pi$ the pushforward measure of a probability measure $\pi$ on $X \times X$ by $p_j$. Let $\Pi(\mu, \nu)$ be the set of probability measures $\pi$ on $X \times X$ such that $(p_1)_\# \pi = \mu$ and $(p_2)_\# \pi = \nu$. (Such a measure $\pi$ is called a transport plan or coupling from $\mu$ to $\nu$.) The first Wasserstein metric $W_1$ is defined as

$$W_1(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} \int_{X \times X} d_X(x, y) d\pi(x, y) \quad \text{for } \mu, \nu \in \mathcal{P}(X).$$

(The integral in this formula is called the cost of the transport plan $\pi$ with respect to the cost function $d_X$.) A standard reference for Wasserstein metric is [49, 50]. What we need in this paper is the following Kantorovich-Rubinstein dual representation of the first Wasserstein metric:

$$(2.1) \quad W_1(\mu, \nu) = \sup_{\varphi \in \text{Lip}^1(X, \mathbb{R})} \left| \int_X \varphi(x) d\mu(x) - \int_X \varphi(x) d\nu(x) \right|,$$

where $\text{Lip}^1(X, \mathbb{R})$ is the space of all Lipschitz continuous real-valued functions $\varphi$ on $X$ whose Lipschitz constant of $\varphi$ is bounded by 1. Recall that we denoted $W_1$ by $d$ in Section 1.

Next let us define a Newhouse open set. Let $M$ be a closed surface. It was shown by Newhouse that, for any $g \in \text{Diff}^r(M)$ ($r \geq 2$) with a homoclinic tangency of a dissipative saddle fixed point $\hat{p}$, there is an open set $\mathcal{O} \subset \text{Diff}^r(M)$ whose closure contains $g$ and such that any element of $\mathcal{O}$ is arbitrarily $C^r$-approximated by a diffeomorphism $f$ with a homoclinic tangency associated with a dissipative saddle fixed point $\hat{p}_f$ which is the continuation of $\hat{p}$, and moreover $f$ has a $C^r$-persistent tangency associated with some basic sets $\Lambda_f$ containing $p_f$ (i.e. there is a $C^r$ neighborhood of $f$ any element of which has a homoclinic tangency for the continuation of $\Lambda_f$). Such an open set $\mathcal{O}$ is called a Newhouse open set (associated with $g$), and call $(\hat{p}_f, \Lambda_f)$ the associated pair of $f$. See [30].

2.2. Infinite dimensional simplex. As seen in Subsection 1.2 in order to show super-polynomial pointwise emergence, it suffices to prove that $\mathcal{A}_r$ includes an infinite dimensional simplex. We here prepare some notation to explore the idea in detail. For each sequence $\mathcal{J} = \{\mu(\ell)\}_{\ell \geq 0}$ of probability measures on $X$, we define $\Delta(\mathcal{J})$ by

$$\Delta(\mathcal{J}) = \bigcup_{L \geq 1} \Delta_L(\mathcal{J}), \quad \Delta_L(\mathcal{J}) = \{ \mu(\ell, \mathcal{J}) \mid \ell \in A_L \},$$

where

$$(2.2) \quad A_L = \left\{(t_0, t_1, \ldots, t_L) \in [0, 1]^{L+1} \mid \sum_{\ell=0}^L t_\ell = 1 \right\}.$$
endowed with the Euclidean norm induced from \([0,1]^{L+1}\), and

\[
\mu_t \equiv \mu_t(J) = \sum_{\ell=0}^{L} t_{\ell} \mu^{(\ell)}
\]

for \(t = (t_0, t_1, \ldots, t_L) \in A_L\).

We also define \(E(J, f)\) by

\[
E(J, f) = \{ x \in X \mid \Delta(J) \subset A_x(f) \}.
\]

2.3. Homogeneous coding on a wandering domain. In the next subsection, we will construct just one nice code such that the associated point is in \(E(J, f)\) with a sequence \(J = \{\mu^{(\ell)}\}_{\ell \geq 0}\) of infinitely many ergodic invariant probability measures. So we need to “enlarge” the point to a positive Lebesgue measure set.

Let \(f : M \to M\) be a \(C^r\) diffeomorphism (\(r \geq 1\)) on a compact manifold \(M\). The following is the key definition in the proof of Theorem [A] which is reminiscent of the specification property

**Definition 2.1.** Let \(\hat{p}\) be a fixed point and \(\{p^{(\ell)}\}_{\ell \geq 0}\) a sequence of periodic points. Let \(\{m_k\}_{k \geq 1}\) be a sequence of positive integers and \(\{\ell_k\}_{k \geq 1}\) a sequence of nonnegative integers. We say that a wandering domain \(D \subset M\) of \(f\) is coded by \(\{\ell_k\}_{k \geq 1}\) over the base order \(\{m_k\}_{k \geq 1}\) if there exist sequences \(\{\hat{I}_k\}_{k \geq 1}\) and \(\{I_k\}_{k \geq 1}\) of disjoint discrete intervals and a sequence \(\{\epsilon_k\}_{k \geq 1}\) of positive numbers with \(\lim_{k \to \infty} \epsilon_k = 0\) satisfying the following conditions.

(C1) For any \(k \in \mathbb{N}\),

\[
\hat{I}_k \cup I_k \subset [N_{k-1}, N_k - 1],
\]

where \(N_k = \sum_{j=1}^{k} m_j\) and \(N_0 = 0\). Furthermore,

\[
\lim_{k \to \infty} \frac{\# \hat{I}_k + \# I_k}{m_k} = 1.
\]

(C2) For any sufficiently large \(k \in \mathbb{N}\),

\[
f^n(D) \subset B_{\epsilon_k}(\hat{p}) \quad \text{if} \quad n \in \hat{I}_k.
\]

(C3) For any sufficiently large \(k \in \mathbb{N}\),

\[
f^n(D) \subset B_{\epsilon_k}\left(f^n(p^{(\ell_k)})\right) \quad \text{if} \quad n \in I_k,
\]

and \(#I_k\) is a multiple of \(\text{per}(p^{(\ell_k)})\). Furthermore,

\[
\zeta = \lim_{k \to \infty} \frac{\# I_k}{m_k}
\]

exists as a strictly positive number.

See Figure 1.

**Figure 1.** Travel of \(f^n(D)\) for \(n \in [N_{k-1}, N_k)\).
We say that a sequence of positive integers \( \{m_k\}_{k \geq 1} \) is moderate if
\[
\lim_{k \to \infty} \frac{m_k}{N_k} = 0.
\]
The following theorem is a key generalization of the idea behind a number of estimates in \([22]\) to the context of pointwise emergences and will be used in the proof of Theorem [A].

**Theorem 2.2.** Let \( M \) be a closed surface and \( 2 \leq r < \infty \). For any Newhouse open set \( \mathcal{O} \subset \text{Diff}^r(M) \), any element \( \hat{f} \in \mathcal{O} \) having an associated pair \((\hat{p}, \Lambda_{\hat{f}})\), any neighborhood \( \mathcal{U}(\hat{f}) \) of \( \hat{f} \) in \( \mathcal{O} \), and any sequence of periodic points \( \{p^{(\ell)}_f\}_{\ell \geq 0} \subset \Lambda_{\hat{f}} \), there exists a moderate sequence of positive integers \( \{m_k\}_{k \geq 1} \) such that for any sequence of nonnegative integers \( \{\ell_k\}_{k \geq 1} \), one can find an element \( f \) of \( \mathcal{U}(\hat{f}) \) having a wandering domain coded by \( \{\ell_k\}_{k \geq 1} \) for \( (\hat{p}, \{p^{(\ell)}_f\}_{\ell \geq 0}) \) over \( \{m_k\}_{k \geq 1} \), where \( \hat{p} \) and \( p^{(\ell)}_f \) are the continuations of \( p^{(\ell)}_f \) and \( p^{(\ell)}_f \) for \( \ell \geq 0 \), respectively.

**2.4. Adapted code.** The final step of the proof of Theorem [A] is to construct an adapted code in the following sense. Denote by \( \text{per}(p) \) the period of a periodic point \( p \).

**Theorem 2.3.** Let \( f : M \to M \) be a \( C^r \) diffeomorphism on a compact manifold \( M \) with \( r \geq 1 \). For any moderate sequence of positive integers \( \{m_k\}_{k \geq 1} \), any fixed point \( \hat{p} \) and any sequence of periodic points \( \{p^{(\ell)}_f\}_{\ell \geq 0} \), there is a sequence of nonnegative integers \( \{\ell_k\}_{k \geq 1} \) such that the following holds: Suppose that \( f \) has a wandering domain \( D \) coded by \( \{\ell_k\}_{k \geq 1} \) for \( (\hat{p}, \{p^{(\ell)}_f\}_{\ell \geq 0}) \) over \( \{m_k\}_{k \geq 1} \). Then we have
\[
D \subset E(\{\mu^{(\ell)}\}_{\ell \geq 0}, f) \quad \text{with} \quad \mu^{(\ell)} = (1 - \zeta)\delta_{\hat{p}} + \zeta \delta_{\text{per}(p^{(\ell)})}, \ \ell \geq 0,
\]
where \( \zeta \) is the positive number given in Definition [2.1].

**2.5. Plan of the proof.** Observe that
\[
N(\epsilon, \Delta(\{(1 - \zeta)\hat{p} + \zeta \mu^{(\ell)}\}_{\ell \geq 0})) = N(\zeta^{-1} \epsilon, \Delta(\{\mu^{(\ell)}\}_{\ell \geq 0}))
\]
for any \( \epsilon > 0, \zeta \in (0, 1], \hat{p} \in \mathcal{P}(X) \) and \( \{\mu^{(\ell)}\}_{\ell \geq 0} \subset \mathcal{P}(X) \). Hence, (1.8) immediately follows from Theorems 2.2 and 2.3 by taking \( \{p^{(\ell)}_f\}_{\ell \geq 0} \) as
\[
\text{Per}(f|_{\Lambda_f}) = \{\hat{p}_f\} \cup \left( \bigcup_{\ell \geq 0} O(p^{(\ell)}_f) \right),
\]
where \( O(p^{(\ell)}_f) \) is the forward orbit of \( p^{(\ell)}_f \) and \( \text{Per}(f|_{\Lambda_f}) \) is the set of periodic points of \( f \) on \( \Lambda_f \). Furthermore, as mentioned, (1.9) is a consequence of (1.5) and (1.8). Therefore, we get the conclusion of Theorem [A]. We give the proof of Theorem 2.2 in Section 4 and the proof of Theorem 2.3 in Section 5.

3. Preliminary

In this section we prove Lemma [1.3] together with some basic properties of \( \delta_x^c \) with respect to \( d \) that will be used in the following sections. Let \( f : X \to X \) be a continuous map on a compact metric space \( X \) equipped with a metric \( d_X \). Recall that \( d \) is the first Wasserstein metric on \( \mathcal{P}(X) \). Since \( X \) is compact, we can assume
Lemma 3.4. For each $\mu, \nu \in \mathcal{M}(X)$, the Kantorovich-Rubinstein dual representation \((2.1)\) implies that for each probability measures $\mu, \nu$ on $X$,

\[
d(\mu, \nu) = \sup_{\varphi \in \text{Lip}^1([0,1])} \left| \int_X \varphi(x) d\mu(x) - \int_X \varphi(x) d\nu(x) \right|
\]

where $\text{Lip}^1(X, [0,1])$ is the space of functions $\varphi$ on $X$ with values in $[0,1]$ such that the Lipschitz constant of $\varphi$ is bounded by 1 (notice that for any $\varphi \in \text{Lip}^1(X, \mathbb{R})$, one can find a constant $a$ such that $\varphi + a \in \text{Lip}^1(X, [0,1])$).

Lemma 3.1. For any $m > n \geq 1$ and $x \in X$, we have

\[
d\left(\delta_x^m, \delta_x^{n-m} \right) \leq \frac{2n}{m}.
\]

Proof. For any continuous function $\varphi : X \to \mathbb{R}$ with value in $[0,1]$,

\[
\left| \int_X \varphi \delta_x^m - \int_X \varphi \delta_x^{n-m} \right| \leq \left| \int_X \varphi \left( \frac{1}{m} - \frac{1}{m-n} \right) \sum_{j=m}^{n-1} \varphi(f^j(x)) \right| + \left| \int_X \varphi \frac{1}{m} \sum_{j=m}^{n-1} \varphi(f^j(x)) \right|
\]

which is bounded by $\frac{2n}{m}$, so we get the conclusion due to \(3.1\).

The next lemma follows from a similar argument.

Lemma 3.2. For any $n \geq 1$ and $x \in X$, we have

\[
d(\delta_x^n, \delta_x^{n+1}) \leq \frac{2}{n+1}.
\]

We also recall the following basic fact, refer to e.g. \[49, \text{Section 7.2}\].

Lemma 3.3. For each $x, y \in X$,

\[
d(\delta_x, \delta_y) = d_X(x, y).
\]

Finally, we will use the following lemma.

Lemma 3.4. For each $L \geq 1$ and $t, s \in A_L$, we have

\[
d(\mu_t, \mu_s) \leq (L+1)|t - s|,
\]

where $\mu_t$ and $A_L$ are given in \([2.2]\) and \([2.3]\).

Proof. For any $\varphi \in \text{Lip}^1(X, [0,1])$ and $t = (t_0, \ldots, t_L)$, $s = (s_0, \ldots, s_L)$ in $A_L$ with $L \geq 1$,

\[
\left| \int_X \varphi d\mu_t - \int_X \varphi d\mu_s \right| \leq \sum_{\ell=0}^{L} |t_\ell - s_\ell| \int_X |\varphi| d\mu(\ell)
\]

\[
\leq (L+1) \max_{0 \leq \ell \leq L} |t_\ell - s_\ell| \leq (L+1)|t - s|,
\]

which implies the conclusion due to \(3.1\).

Proof of Lemma \(3.3\) Fix $\epsilon > 0$ and a positive Lebesgue measure set $D$. Let $N_0$ be the maximal integer $N$ such that for all probability measures $\{\mu_j\}_{j=1}^N$ on $M$, the inequality in \(1.2\) does not hold for any $x \in D$. Note that $\min_{x \in D} d(\epsilon, f) = N_0 + 1$.

Given probability measures $\{\mu_j\}_{j=1}^{N_0}$ on $M$, by definition of $N_0$, we get

\[
\lim_{n \to \infty} \min_{1 \leq j \leq N_0} d(\delta_x^n, \mu_j) > \epsilon
\]

for any $x \in D$. 


Thus, the inequality in (1.1) with $\text{Leb}(D)\epsilon$ instead of $\epsilon$ is not satisfied by the probability measures $\{\mu_j\}_{j=1}^{N_0}$, implying that $\delta_{\text{Leb}}(\text{Leb}(D)\epsilon) \geq N_0 + 1$. This completes the proof. 

4. PROOF OF THEOREM \[2.2\]

4.1. Preliminary. The proof of Theorem \[2.2\] is based on the argument in our previous work [32, Theorem A]: the most important part is a modification of Critical Chain Lemma (32, Lemma 7.1), but the modified version of Critical Chain Lemma can be proven as the original version. In the rest of this section, we first briefly recall necessary definitions and dynamics in [32], and precisely describe how we should modify Critical Chain Lemma, together with a short explanation for the reason why the modification does not affect the proof of Critical Chain Lemma. Finally, we will complete the proof of Theorem 2.2 by translating the argument in the proof of Theorem A of [32] after the establishment of Critical Chain Lemma into our context.

Notation. In this section, we will use the notation $N_0, N_1, N_2$, which are positive integers borrowed from [32], but different from $N_k$ defined in Definition (2.1). To avoid notational confusion, we use $\tilde{N}_k$ for an integer playing the role of $N_k$ in Definition (2.1) (refer to (4.1)).

Let $M$ be a closed surface and $2 \leq r < \infty$. Let $\tilde{f}$ be an element of a Newhouse open set $\mathcal{O} \subset \text{Diff}^r(M)$. By definition of Newhouse open sets, $\tilde{f}$ has a dissipative saddle fixed point $\tilde{p}_f$ and a basic set $\Lambda_f$ such that $\tilde{p}_f \in \Lambda_f$ and $\tilde{f}$ has a persistent homoclinic tangency associated with $\Lambda_f$. In fact, $\tilde{f}^k$ has a basic set $\Lambda$ on which $f^k$ is conjugate to a two-sided full shift of two symbols $\{1,2\}^\mathbb{Z}$ and $\Lambda = \bigcup_{j=0}^{k-1} \tilde{f}^j(\tilde{\Lambda})$ with some $k \in \mathbb{N}$. For simplicity, we assume that $f|\Lambda$ is conjugate to the two-sided full shift of two symbols. We also fix a small neighborhood $\mathcal{U}(\tilde{f})$ of $\tilde{f}$ in $\mathcal{O}$. Then, one can find an element $f$ of $\mathcal{U}(\tilde{f})$ which has the continuations $\tilde{p}_f$ of $\tilde{p}_f$ and $\Lambda_f$ of $\Lambda_f$ such that

(S-i) $\Lambda_f$ contains $\tilde{p}_f$;
(S-ii) $f$ has a quadratic tangency $q_f$ associated with $\tilde{p}_f$;
(S-iii) $f$ is linear in $U(\tilde{p}_f) \cap f^{-1}(U(\tilde{p}_f))$ with a small neighborhood $U(\tilde{p}_f)$ of $\tilde{p}_f$.

We refer to e.g. [40] (compare also with Section 3 in [32]).

We suppress $\tilde{f}$ from the notations $\tilde{p}_f, \Lambda_f$ and $q_f$. By replacing the basic set $\Lambda$ by a smaller one if necessary, we can choose the linearizing coordinate in (S-iii) such that $\Lambda \subset S \subset S'$ where $S = [0,2] \times [0,2]$ and $S' = [-2,2] \times [-2,2]$ with $\tilde{p} = (0,0)$. Set $W^s_{\text{loc}}(\tilde{p}) = [-2,2] \times \{0\}$ and $W^u_{\text{loc}}(\tilde{p}) = \{0\} \times [-2,2]$. Let $\mathcal{F}_{\text{loc}}^s(\Lambda)$ and $\mathcal{F}_{\text{loc}}^u(\Lambda)$ be a local stable foliation and a local unstable foliation on $S$ compatible with $W^s_{\text{loc}}(\Lambda)$ and $W^u_{\text{loc}}(\Lambda)$, respectively. For $\sigma = s, u$, consider the projection $\pi^\sigma : S \to W^\sigma_{\text{loc}}(\tilde{p})$ along the leaves of $\mathcal{F}_{\text{loc}}^\sigma(\Lambda)$ where $\pi = u$ and $\pi = s$, and the Cantor set $K^\sigma_\Lambda = \pi^\sigma(\Lambda)$ in $W^\sigma_{\text{loc}}(\tilde{p})$. 
Let $\sigma = s$ or $u$. We denote by $B^\sigma(0)$ the smallest interval in $W^\sigma_{loc}(\hat{p})$ containing $K^u_{\Lambda}$. There exists a Markov partition of $B^\sigma(0)$ for $K^u_{\Lambda}$ which consists of sub-intervals $B^\sigma(1; 1)$, $B^\sigma(1; 2)$ of $B^\sigma(0)$ with $\partial B^\sigma(0) \cap B^\sigma(1; 1) = \{\hat{p}\}$. Let $\Psi^\sigma : B^\sigma(1; 1) \cup B^\sigma(1; 2) \to B^\sigma(0)$ be the map defined by $\Psi' = \pi^\sigma \circ f^{-1}$ and $\Psi^u = \pi^u \circ f$, which is $C^{1+\alpha}$ for some $0 < \alpha < 1$. For each integer $\ell \geq 1$ and $w_i \in \{1, 2\}$ for $1 \leq i \leq \ell$, we define the interval $B^\sigma(\ell; w_1 \cdots w_\ell)$, called a \textit{\sigma-bridge} of the $\ell$-th generation, by

$$B^\sigma(\ell; w_1 \cdots w_\ell) = \{x \in B^\sigma(0) \mid (\Psi^\sigma)^{i-1}(x) \in B^\sigma(1; w_i), i = 1, \ldots, \ell\}.$$ 

Here we say that the word $(w_1 \cdots w_\ell)$ is the \textit{itinerary} for the $\sigma$-bridge. From the definition, we have

$$(4.1) \quad \Psi^\sigma(B^\sigma(\ell; w_1 w_2 \cdots w_\ell)) = B^\sigma(\ell - 1; w_2 \cdots w_\ell).$$

Let $B^\sigma$ be a $\sigma$-bridge with $\sigma = s$ or $u$. The closure of a connected component of $B^\sigma \setminus K^u_{\Lambda}$ is called a \textit{gap} of $K^u_{\Lambda}$ in $B^\sigma$. Finally, we call $B^\sigma = (\pi^\sigma)^{-1}(B^\sigma)$ the \textit{bridge strip} of $B^\sigma$, and $G^u = (\pi^u)^{-1}(G^u)$ the \textit{gap strip} of $G^u$ (see Subsection 4.2 of [32] for details).

4.2. The dynamics in [32]. We used in [32] two preliminary perturbations (Section 3, 4, 5), and two main perturbations (Section 5, 7). One of the preliminary perturbations is given for the perturbed dynamics to satisfy the conditions (S-i), (S-ii), (S-iii). The other is given to satisfy (S-iv), (S-v), (S-vi), (S-vii) of [32] Section 3. We now let $f$ be an element of $U(\hat{f})$ satisfying all the conditions from (S-i) to (S-vii) (note that we here use the notation $f$ for $f_{\mu_n}$ in (S-iv), (S-v), (S-vi), (S-vii) of [32] with large integer $n = n_*$ given in [32] §5.2)). We merely remember that (S-iv), (S-v) and (S-vi) lead to the existence of another basic set $\Gamma$ of $f$ near the homoclinic tangency $q$ such that

- $\Lambda$ and $\Gamma$ are homoclinically related: both $W^u(\Lambda) \cap W^s(\Gamma)$ and $W^s(\Lambda) \cap W^u(\Gamma)$ contain non-trivial transverse intersections,
- there exists a heteroclinic tangency curve $L$ between $\Lambda$ and $\Gamma$: there are a smooth arc $L$, a local stable foliation $\mathcal{F}^s_{loc}(\Gamma)$ of $\Gamma$ which is compatible with a local stable manifold of $\Gamma$ on a compact region $E$ containing $\Gamma$, and positive integers $N_0, N_2$ such that $L \cap f^{-N_0}(\mathcal{F}^s_{loc}(\Gamma)) = L$ and $L \cap f^{-N_2}(\mathcal{F}^u_{loc}(\Lambda))$ is a sub-arc of $L$ each element of which is a quadratic tangency of $f^{-N_2}(\mathcal{F}^u_{loc}(\Lambda))$ and $f^{-N_0}(\mathcal{F}^s_{loc}(\Gamma))$.

Refer to [32] §5.1] and see Figure 2. (We note that the basic set $\Gamma$ is written as $\Gamma_m$ in [32], where $m$ is the period of a periodic point included in $\Gamma_m$.)

The key dynamics in [32] is the return map on $(L)$ and a neighborhood $U(L)$ of $L$, which is the composition of 3 dynamics:

- the transient dynamics from $U(L)$ to $\Lambda$ through the homoclinic relation between $\Lambda$ and $\Gamma$,
- the hyperbolic dynamics on $\Lambda$,
- the transient dynamics from $\Lambda$ to $U(L)$ through the unstable foliation $f^{N_2}(\mathcal{F}^u_{loc}(\Lambda))$.

To be precise, we borrow more notations from [32]. Let $z_0$ be the positive integer satisfying (8.5) in [32] Subsection 8.1] and $\{z_k\}_{k \geq 1}$ arbitrary sequence of integers such that $z_k \in \{z_0, z_0 + 1\}$ for each $k \geq 1$. Let $\hat{w}_k$ be the itinerary given in [32] Lemma 7.1] for $k \geq 1$. The itinerary originates from Linear Growth Lemma ([32] Lemma 6.1]), which implies that if we denote the length of $\hat{w}_k$ by $\widehat{n}_k$, then
there is a constant $\alpha > 0$ such that the $\alpha|B^s(\tilde{n}_k; \tilde{w}_{k+1})|$-neighborhood of $B^s(\tilde{n}_k; \tilde{w}_k)$ and $\alpha|B^s(\tilde{n}_{k+1}; \tilde{w}_{k+1})|$-neighborhood of $B^s(\tilde{n}_{k+1}; \tilde{w}_{k+1})$ are disjoint,

- $\tilde{n}_k$ is of order $k$ (so the lemma is called Linear Growth Lemma).

Moreover, for arbitrary sequence $\{\tilde{w}_{k+1}\}_{k \geq 1}$ of itineraries with $\tilde{w}_{k+1} \in \{1, 2\}^k$, we consider the bridge $B^u_k = B^u_k(z_k; \tilde{w}_{k+1})$ of $K^u_\Lambda$ and the bridge $B^{s*}_{k+1} = B^{s*}_{k+1}(z_k, \tilde{w}_{k+1})$ of $K^s_\Lambda$ given by

$$B^u_k = B^u(z_k k^2 + k^2 + k + \tilde{n}_{k+1}; \tilde{w}_{k+1}, \tilde{w}_{k+1}^{-1}) \tilde{w}_{k+1}[\tilde{w}_{k+1}^{-1}]^{-1},$$

$$B^{s*}_{k+1} = B^{s*}(z_k k^2 + k^2 + k + \tilde{n}_{k+1}; \tilde{w}_{k+1}, \tilde{w}_{k+1}[\tilde{w}_{k+1}^{-1}]^{-1} k^2),$$

where $1^{(\ell)}$ (resp. $2^{(\ell)}$) is the itinerary consisting of only 1 (resp. 2) with length $\ell$ and $[\tilde{w}]^{-1} = (w_{\ell} \cdots w_2 w_1)$ for each $\tilde{w} = (w_1 w_2 \cdots w_\ell) \in \{1, 2\}^\ell$. It follows from Subsection 5.2 of [32] (refer also to Lemma 7.1 (2) of [32]) that there are a connected component $S_k \subset S$, integers $N_1$ and $\hat{n}_k$ of order $k$ for each $k \geq 1$ such that $f^{1(N_k+N_1)} S_k \cap ([0,1] \times \{t\})$ is the intersection of $E$ and a leaf of $\mathcal{F}_{loc}(\Gamma)$ for each $t \in [0,1]$. (In terms and notations of [32], $S_k$ is the sub-strip of $S$ such that $f^{1(N_k+N_1)} S_k$ is the bridge stripe of a $u$-bridge $\hat{A}^u_k$ of $\Gamma$ along $\mathcal{F}_{loc}(\Gamma)$, and denoted by $S(\hat{A}^u_k)$.) We let the arc $L$ transversely return to itself by backward iterations of $f$ as

$$\tilde{L} = f^{-N_2}(L \cap f^{-N_2}(F^u_{loc}(\Lambda))),$$

$$\tilde{L}_k = f^{-N_2}(z_k k^2 + k^2 + k + \tilde{n}_{k+1}; \tilde{w}_{k+1}) \tilde{w}_{k+1}[\tilde{w}_{k+1}^{-1}]^{-1},$$

so that $L$ and $L_k$ have a transverse intersection $x_k$ for any large $k$ (cf. [32 §7.2]). Then, noting that $f^{2z_k k^2 + k^2 + k + \tilde{n}_{k+1}; \tilde{w}_{k+1}^{-1}} = \mathbb{B}^s_{k+1}$ by construction, we get

(a) $\tilde{x}_k = f^{-N_0+\hat{n}_k+N_1}(x_k) \in \mathbb{B}^u_k$, mapped from $L$ to $\Lambda$,

(b) $\tilde{x}_k = f^{-N_0+\hat{n}_k+N_1}(x_k) \in \mathbb{B}^s_{k+1}$, mapped on $\Lambda$,

(c) $f^{N_2}(\tilde{x}_k) \in L$, mapped from $\Lambda$ to $L$.

The second main perturbation (i.e. the perturbation in Subsection 7.2 of [32]) are made, with the notation $f$ again for the perturbed dynamics, to get the relation

$$f^{m_k}(x_k) = x_{k+1}, \quad m_k = N_2 + (z_k k^2 + k^2 + k + \tilde{n}_{k+1}) + (N_0 + \hat{n}_k + N_1)$$

for all large $k$, and Critical Chain Lemma ensures that the perturbation can be arbitrary small (see also the next subsection).

Finally, let $R_k$ be the rectangle given in Subsection 8.2 of [32], where the center of $R_k$ is $x_k$ and $R_k \cap R_{k'} = \emptyset$ for each $k \neq k'$. We notice that the distance between $x_k$ and $x_{k+1}$ is large in the sense of the first item of the above properties of $\tilde{w}_k$ (and similar property for the $u$-bridges of $\Gamma$ in [32 Lemma 6.1]). Indeed, Rectangle Lemma ([32 Lemma 8.2]) states that $f^{m_k}(R_k) \subset R_{k+1}$ for sufficiently large $k$, and thus $D = R_k$ with a large $k$ is a wandering domain.

Furthermore, both $N_0 + \hat{n}_k + N_1$ and $N_2$ as well as the length of $\tilde{w}_{k+1}[\tilde{w}_{k+1}^{-1}]^{-1}$ are at most of order $k$, while the lengths of $1(z_k k^2)$ and $2(k^2)$ are of order $k^2$ (so we called the dynamics (a) and (c) transient). Therefore, we can find sequences of discrete intervals $\{\tilde{I}_k\}_{k \geq 1}$ and $\{I_k\}_{k \geq 1}$, and a sequence of positive measure $\{\epsilon_k\}_{k \geq 1}$
with \( \lim_{k \to \infty} \epsilon_k = 0 \) such that

\[
\hat{I}_k \cup I_k \subset [\bar{N}_{k-1}, \bar{N}_k - 1]
\]
where \( \bar{N}_k = \sum_{j=1}^{k} m_j \),

\[
\lim_{k \to \infty} \frac{\#\hat{I}_k + \#I_k}{m_k} = 1, \quad \lim_{k \to \infty} \frac{\#\hat{I}_k - \hat{z}_k}{\hat{z}_k + 1} = 0,
\]

and for any sufficiently large \( k \),

\[
f^n(D) \subset B_{\epsilon_1}(\hat{p}) \quad \text{if } n \in \hat{I}_k \quad \text{and} \quad f^n(D) \subset B_{\epsilon_1}(p') \quad \text{if } n \in I_k,
\]

where \( p' \) is the another saddle fixed point of \( f \) in \( \Lambda \). Refer to Subsection 4.3 for the calculation, and compare with Definition 4.4

4.3. Modification in Critical Chain Lemma. Let \( \pi^u : B^u(0) \to L \) be the projection along the leaves of \( f^{N_0}(F^u_{\text{loc}}(\Lambda)) \) and \( \pi^u_k : B^u(0) \to L \) the projection given by

\[
\pi^u_k = \pi^u \circ f^{-N_0(N_1)} \circ \pi_{S_k},
\]

where \( \pi^u : E \to L \) is the projection along the leaves of \( f^{-N_0}(F^u_{\text{loc}}(\Gamma)) \) and \( \pi_{S_k} \) is the projection from \( B^u(0) \) to a component of the boundary of \( S_k \) along the leaves of \( f^{(N_1+N_k)}(F^u_{\text{loc}}(\Gamma)) \). Let \( B_{k,L}^u = \pi^u_k(B_k^u) \) and \( B_{k,L}^{w*} = \pi^w(B_k^{w*}) \) (cf. [32 §5.1]).

Critical Chain Lemma states that there are constants \( \epsilon_0 > 0, r > 1 \) and an interval \( J_k \subset (-\epsilon_0 r^{-k}, \epsilon_0 r^k) \) such that

\[
(B_{k+1,L}^w + t) \cap B_{k,L}^w \neq \emptyset \quad \text{if and only if } t \in J_{k+1}.
\]

The second main perturbation is of the form

\[
\tilde{z}_k + u_k = f^{-N_2(x_{k+1})} \quad \text{on } \tilde{L}
\]

to obtain (4.3), and (4.7) ensures that \(|u_k|\) is of order \( r_1^{-k} \) with some \( r_1 > 1 \).

Let \( \tilde{w}_k \) be arbitrary sequence of itineraries with \( \tilde{w}_k \in \{1, 2\}^{k+3k_1+1} \). We modify Critical Chain Lemma by replacing \( B_k^u \) and \( B_{k+1}^{w*} \) in (4.2) with \( B_k^u = B_k^u(z_0, \tilde{w}_k) \) and \( B_{k+1}^{w*} = B_{k+1}^{w*}(z_0, \tilde{w}_k) \) given by

\[
B_k^u = B^u(z_0 k^2 + k^2 + 3k + 1 + \hat{n}_k + 1; \sum_{j=1}^{k} \bar{w}_k^{-1} \hat{w}_k^{-1} (\hat{w}_k - 1) (z_0 k^2)),
\]

\[
B_{k+1}^{w*} = B^w(z_0 k^2 + k^2 + 3k + 1 + \hat{n}_k + 1; \sum_{j=k+1}^{2k+1} \bar{w}_k^{-1} \hat{w}_k^{-1} (\hat{w}_k - 1) (z_0 k^2)).
\]

In [32], the itinerary \( 2(k^2) \) together with the integer \( z_k \) in (4.2) is chosen as the wandering domain \( D \) consists of points with historic behavior (recall (4.1), (4.5) and (4.6)), and the itinerary \( \tilde{w}_{k+1} \) is used just to show that the \( \omega \)-limit set of the forward orbit contains \( \Lambda \). In Theorem 2.2 such properties are not required. So, all \( z_k \) are unified to \( z_0 \), and \( 2(k^2) \) and \( \tilde{w}_{k+1} \) are deleted. However, for the proof of the existence of a wandering domain, it is crucial that the orbit stays long time in a small neighborhood of \( \hat{p} \), and that the distance between \( x_k \) and \( x_{k+1} \) are sufficiently large. So the roles of the itineraries \( 1(z_0 k^2) \) and \( \tilde{w}_{k+1} \) are indispensable. On the other hand, the itinerary \( 2(k^2) \) can be replaced by any itinerary of length \( k^2 + O(k) \). Hence one can use any itinerary \( \tilde{w}_k \) with \(|\tilde{w}_k| = k^2 + 3k + 1 \) instead of \( 2(k^2) \).
4.4. The end of the proof of Theorem 2.2 Here we set \( \hat{j}_k = N_0 + \hat{i}_k + N_1 \) and
\[
(4.10) \quad m_k = N_2 + (z_0 k^2 + k^2 + 3k + 1 + \hat{n}_{k+1}) + \hat{j}_k, \quad \tilde{N}_k = \sum_{j=1}^{k} m_j
\]
(instead of (4.3) and (4.4)). As in the proof of Theorem A in [32], there exists an element \( f \) of \( \mathcal{U}(\hat{f}) \) which has a contracting wandering domain \( D \) such that \( f^\hat{j}_k + \tilde{N}_k - 1 (D) \) is contained in the gap strip \( G^u_k \) for all sufficiently large \( k \). Since the second perturbation is made only in the interior of \( G^u(0) \) (in fact \( \hat{L} \) is included in \( G^u(0) \), recall also (4.5)), \( \Lambda \) and \( \hat{p} \) do not change by the perturbation. See Figure 2.

According to Lemma 7.1 in [32], \( \tilde{N}_{k+1} \) is the itinerary of length \( \tilde{n}_{k+1} = O(k) \) which is arranged such that \( f^{\tilde{N}_k}(D) \subset B^{s}_{k+1} \cap G^u(0) \) is sent into \( G^u_k \subset B^{u}_{k+1} \) by \( f^\tilde{j}_k + 1 \). Since \( m_k = O(k^2) \) and \( \tilde{N}_k = O(k^3) \), we have \( \lim_{k \to \infty} m_k / \tilde{N}_k = 0 \). This means that \( \{m_k\}_{k \geq 1} \) is moderate.

Now we take a sequence \( \{\ell_k\}_{k \geq 1} \) of non-negative integers arbitrarily. For any non-negative integer \( a \) with \( \{k \mid \ell_k = a\} \neq \emptyset \), set \( \kappa(a) = \min \{k \mid \ell_k = a\} \). Let \( p(\ell_k) \) be the periodic point of \( \Lambda \) corresponding to the bi-infinite itinerary \( (\underline{y}_k)^2 \), where
\[
\underline{y}_k = 1 \cdot 2^{(\kappa(a))}, \quad \text{for} \quad a = \ell_k. \tag{4.11}
\]
If \( a \neq a' \), then \( \{k \mid \ell_k = a\} \cap \{k' \mid \ell_{k'} = a'\} = \emptyset \) and hence \( \kappa(a) \neq \kappa(a') \). It follows that the orbit of \( p(\ell_k) \) is disjoint from that of \( p(\ell_{k'}) \) if \( \ell_k \neq \ell_{k'} \). By (4.11), there exists an integer \( s(k) \) with \( k^2 + s(k) \leq k^2 + 1 \) which is a multiple of \( \kappa(a) \). Then \( q(k) = s(k) / \kappa(a) + 1 \leq k + 1 \).

Consider itineraries \( \underline{y}_k \) and \( \underline{b}_k \) with \(|\underline{a}_k| = k, |\underline{b}_k| = k^2 + 2k + 1 - s(k) \) and such that \( \underline{a}_k = \underline{y}_k (q(k)) \underline{b}_k \) is a sub-itinerary of \( (\underline{y}_k)^2 \). From our definition, we have \( k \leq |\underline{b}_k| \leq 2k + 1 \) and \( |\underline{a}_k| = k^2 + 3k + 1 \), the latter of which is one of our required conditions. Consider the discrete intervals \( \hat{I}_k = [\tilde{N}_{k-1} + \hat{j}_k + \hat{n}_{k+1}] \) and \( I_k = [\tilde{N}_k + \hat{n}_{k+1}] \), where \( \tilde{N}_k = z_0 k^2 + \hat{j}_k + \tilde{N}_k - 1 \). See Figure 3.

Figure 2. Travels of \( D \) by \( f \).
Moreover, since (5.1) $\bar{k}$ of length $k^2$. Furthermore, for each $M \in L$, $a$ closed surface $B$. By (4.1), for any $j \in I$, we will prove Theorem 2.3. Let $f$ be a $C^r$ diffeomorphism on a closed surface $M$. We fix a moderate sequence of positive integers $\{m_k\}_{k \geq 1}$, a fixed point $\hat{x}$ and a sequence of periodic points $\{p(\ell)\}_{\ell \geq 0}$ throughout the rest of this section.

5. Proof of Theorem 2.3

In this section, we will prove Theorem 2.3. Let $f$ be a $C^r$ diffeomorphism on a closed surface $M$. We fix a moderate sequence of positive integers $\{m_k\}_{k \geq 1}$, a fixed point $\hat{x}$ and a sequence of periodic points $\{p(\ell)\}_{\ell \geq 0}$ throughout the rest of this section.

5.1. Reduction to $A_L$. We start the proof of Theorem 2.3 by approximating the empirical measures $\{\delta^n_k\}_{n \geq 0}$ along the orbit of $x$ in a wandering domain with a code by measures $\mu_k$ with parameters $\ell \in A_L$ induced by the code.

**Definition 5.1.** For each finite increasing sequence $k = \{k(\ell)\}_{\ell = 1}^L$ of positive integers (i.e. $k(\ell - 1) < k(\ell)$ for $\ell \in [0, L]$), we say that a sequence of nonnegative integers $\{\ell_k\}_{k \geq 1}$ is associated with $k$ if

$$\ell_k = \ell \quad \text{for all } \ell \in [0, L] \text{ and } k \in [k(\ell - 1) + 1, k(\ell)].$$

See Figure 4 for the travel of $f^n(D)$ for $D$ coded by $\{k(\ell)\}_{k \geq 1}$ associated with $\{k(\ell)\}_{\ell = 1}^L$ (compare with Definition 2.1 and Figure 1).

For each finite increasing sequence of positive integers $k = \{k(\ell)\}_{\ell = 1}^L$, we define $\bar{M}(k) \in \mathbb{N}^{L+1}$ by

$$\bar{M}(k)(\ell) = \sum_{k(k(\ell - 1) + 1)}^{k(\ell)} m_k \quad \text{for } \ell \in [0, L].$$

Furthermore, for each $M = (M_0, \ldots, M_L) \in \mathbb{N}^{L+1}$, we define $\bar{t}(M) \in A_L$ by

$$\bar{t}(M) = \left( \frac{M_0}{S_L}, \frac{M_1}{S_L}, \ldots, \frac{M_L}{S_L} \right) \quad \text{with } S_\ell = M_0 + \cdots + M_\ell.$$

---

**Figure 3.** Locations of $f^j(D)$. Then we have

$$\# I_k = s(k) = q(k)\text{per}(p(\ell_k)).$$

Moreover, since $m_k = (z_0 + 1)k^2 + O(k)$ and $\# I_k = k^2 + O(k),$

$$\zeta := \lim_{k \to \infty} \frac{\# I_k}{m_k} = \frac{1}{z_0 + 1} > 0.$$

By (4.1), for any $j \in \hat{I}_k$, $f^j(D) \subset \mathbb{B}^u(k; \frac{1}{(k)}) \cap \mathbb{B}^s(k; \frac{1}{(k)})$. Similarly, for any $j \in I_k$, $f^j(D) \subset \mathbb{B}^u(k; y_{k,j}) \cap \mathbb{B}^s(k; [y_{k,j}]^{-1})$ for some sub-itinerary $y_{k,j}$ of $(y_k)^{(q)}$ of length $k$. Since the diameters of $\mathbb{B}^u(k; \frac{1}{(k)}) \cap \mathbb{B}^s(k; \frac{1}{(k)})$ and $\mathbb{B}^u(k; y_{k,j}) \cap \mathbb{B}^s(k; [y_{k,j}]^{-1})$ uniformly converge to zero as $k \to \infty$, $f$ satisfies the property (2.1) of Definition 2.1. This completes the proof of Theorem 2.2.
Remark 5.3. For a wandering domain $D$ coded by a sequence of nonnegative integers over $(\ell_k)_{k \geq 1}$, let $k_D$ be the minimal integer such that (C2) and (C3) in Definition 2.1 hold for all $k \geq k_D$. Then, for any $x \in D$, let $f$ be a $C^r$ diffeomorphism having a wandering domain $D$ coded by $(\ell_k)_{k \geq 1}$ for $(p, \{p^{(\ell)}\}_{\ell \geq 0})$ over $(m_k)_{k \geq 1}$. Suppose that $k(-1) \geq k_D$. Then, for any $x \in D$, let $N_k = \sum_{j=1}^{k} m_j$.

Lemma 5.2. Let $L \geq 1$ and $\{\ell_k\}_{k \geq 1}$ be a sequence of nonnegative integers associated with $k$. Let $f$ be a $C^r$ diffeomorphism having a wandering domain $D$ coded by $(\ell_k)_{k \geq 1}$ for $(\hat{p}, \{p^{(\ell)}\}_{\ell \geq 0})$. Suppose that $k(-1) \geq k_D$. Then, for any $x \in D$,

\[
d \left( N_{k(t)}(f), N_{k(t)}(f) \right) \leq \frac{2N_{k(-1)} + 2}{N_{k(t)}} + 2 \max_{0 \leq \ell \leq L} \left[ \sum_{k=k(\ell-1)+1}^{k(\ell)} \#I_k \right] \cdot \frac{\max_{0 \leq \ell \leq L} \left[ \sum_{k=k(\ell-1)+1}^{k(\ell)} m_k \right]}{2} + \frac{\max_{0 \leq \ell \leq L} \left[ \sum_{k=k(\ell-1)+1}^{k(\ell)} \#I_k \right]}{1 - \frac{\max_{0 \leq \ell \leq L} \left[ \sum_{k=k(\ell-1)+1}^{k(\ell)} m_k \right]}{2}}.
\]

where $\mu = \sum_{\ell=0}^{L} t_{\mu}^{(\ell)}$ for $t = (t_0, \ldots, t_L)$ and $p^{(\ell)}$ is given in Theorem 2.3.

Remark 5.3. It is easy to see that by (C1) and (C3), both of

\[
\sup_{k_2 > k_1 + 1} \left| \frac{\sum_{k=k_1+1}^{k_2} \#I_k \cdot m_k}{\sum_{k=k_1+1}^{k_2} m_k} - \zeta \right| \quad \text{and} \quad \sup_{k_2 > k_1 + 1} \left| 1 - \frac{\sum_{k=k_1+1}^{k_2} \#I_k + \#I_{k_1}}{\sum_{k=k_1+1}^{k_2} m_k} \right|
\]

go to 0 as $k_1 \to \infty$.

Proof of Lemma 5.2. Fix $x \in D$. For each $k \geq 1$, let

\[
s_{1,k} = \sum_{n \in I_k} \delta f^n(x), \quad s_{2,k} = \sum_{n \in I_k} \delta f^n(x), \quad s_{3,k} = \sum_{n \in [N_{k-1}, N_{k-1} - I_k - I_k]} \delta f^n(x).
\]
Let $s_j(\ell) = \sum_{k=k(\ell-1)+1}^{k(\ell)} s_{j,k}$ for each $j = 1, 2, 3$ and $\ell \geq 0$. Then, we have a decomposition

$$
\sum_{j=N_k(-1)}^{N_k(L)-1} \delta_{fj}(x) = \sum_{\ell=0}^{L} (s_1(\ell) + s_2(\ell) + s_3(\ell)).
$$

(5.2)

Note also that, for each $j = 1, 2, 3$,

$$
\sum_{\ell=0}^{L} s_j(\ell) = \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k}{N_k(L) - N_k(-1)} \times t_{\ell},
$$

(5.3)

where $t_{\ell} = \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k}{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k}$, and that $\overline{7}(\overline{M}(k)) = (t_0, \ldots, t_L)$.

Fix $\varphi \in \text{Lip}^{1}(M, [0, 1])$. Since $f^n(x) \in B_{\epsilon_k}(\hat{p})$ for all $k \geq 1$ and $n \in \hat{I}_k$, it follows from Lemma 3.3 that

$$
\left| \int_{X} \varphi ds_1(\ell) \sum_{k=k(\ell-1)+1}^{k(\ell)} m_k - \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} \# I_k}{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k} \int_{X} \varphi d\delta_{\hat{p}} \right| \leq \max_{k(\ell-1)+1 \leq k \leq k(\ell)} \epsilon_k
$$

for every $\ell \in [0, L]$. So we have

$$
\left| \int_{X} \varphi ds_1(\ell) \sum_{k=k(\ell-1)+1}^{k(\ell)} m_k - (1 - \zeta) \int_{X} \varphi d\delta_{\hat{p}} \right| \leq \max_{k(\ell-1)+1 \leq k \leq k(\ell)} \epsilon_k + \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} \# I_k}{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k} - \zeta + 1 - \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} (\# I_k + \# I_k)}{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k}.
$$

(5.4)

Similarly, since $f^n(x) \in B_{\epsilon_k}(f^n(\ell_k))$ for all $k \geq 1$ and $n \in I_k$, and $\# I_k$ is a multiple of $\text{per}(\ell_k)$, we get that

$$
\left| \int_{X} \varphi ds_{2,k} - \# I_k \int_{X} \varphi d\mu(\ell_k) \right| \leq \# I_k \epsilon_k \quad \text{for every } k \geq 1.
$$

Therefore, by the assumption that $\{\ell_k\}_{k \geq 1}$ is associated with $\ell_k$,

$$
\left| \int_{X} \varphi ds_2(\ell) \sum_{k=k(\ell-1)+1}^{k(\ell)} m_k - \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} \# I_k}{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k} \int_{X} \varphi d\mu(\ell) \right| \leq \max_{k(\ell-1)+1 \leq k \leq k(\ell)} \epsilon_k
$$

for every $\ell \in [0, L]$, and we have

$$
\left| \int_{X} \varphi ds_2(\ell) \sum_{k=k(\ell-1)+1}^{k(\ell)} m_k - \zeta \int_{X} \varphi d\mu(\ell) \right| \leq \max_{k(\ell-1)+1 \leq k \leq k(\ell)} \epsilon_k + \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} \# I_k}{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k} - \zeta.
$$

(5.5)

Furthermore, it is easy to check that

$$
\left| \int_{X} \varphi ds_3(\ell) \sum_{k=k(\ell-1)+1}^{k(\ell)} m_k \right| \leq 1 - \frac{\sum_{k=k(\ell-1)+1}^{k(\ell)} (\# I_k + \# I_k)}{\sum_{k=k(\ell-1)+1}^{k(\ell)} m_k}.
$$

(5.6)

By (5.2), (5.3), (5.4), (5.5) and (5.6), together with Lemma 3.3, we immediately get the conclusion. □
5.2. Filling of $A_L$. The following lemma is elementary but crucial. Notice that
the choice of $k$ is independent of both $f$ and $(\mu, \{\mu^t\}_{t \geq 0})$.

Lemma 5.4. For any positive integer $L$, nonnegative number $\tilde{c}$, positive number
$\epsilon$ and $t \in A_L$, there is an increasing sequence of positive integers $k = \{k(\ell)\}_{\ell=1}^L$
with $k(0) > \tilde{c}$ such that

$$|\mathcal{T}(\mathcal{M}(k)) - t| \leq \epsilon.$$  

Lemma 5.4 easily follows from the following lemma. For $\mathcal{M} = (M_0, \ldots, M_L) \in \mathbb{N}^{L+1}$, we define
\[
\mathcal{T}(\mathcal{M}) = \left(\frac{M_1}{S_1}, \frac{M_2}{S_2}, \ldots, \frac{M_L}{S_L}\right).
\]
(Recall (5.1) for $S_L$.)

Lemma 5.5. For any positive integer $L$, nonnegative number $\tilde{c}$, positive number $\epsilon$ and $T \in [0, 1]^L$, there is an increasing sequence of positive integers $k = \{k(\ell)\}_{\ell=1}^L$
with $k(0) > \tilde{c}$ such that

$$|\mathcal{T}(\mathcal{M}(k)) - T| \leq \epsilon.$$  

Proof. We use the notation $N_{k', k} = \sum_{j=k'+1}^{k} m_j$ for $k' < k$, so that we have

\[
\left(\mathcal{T}(\mathcal{M}(k))\right)_\ell = \frac{N_{k(\ell-1), k(\ell)}}{N_{k(-1), k(\ell)}} \text{ for } \ell \in [1, L].
\]

Fix $L \geq 1$, $\tilde{c} \geq 0$, $\epsilon > 0$ and $T = (T_1, T_2, \ldots, T_L) \in [0, 1]^L$. By the assumption
(5.4) for the moderate sequence \{m_k\}_{k \geq 1}, we can take an integer $k(0) > \tilde{c}$ such that

\[
(5.7) \quad \frac{m_{k'}}{N_{k(-1), k'}} \leq \frac{\epsilon}{\sqrt{L}} \text{ for any } k' > k(0).
\]

Therefore,

\[
\frac{N_{k(0), k'}}{N_{k(-1), k'}} - \frac{N_{k(0), k'-1}}{N_{k(-1), k'-1}} = \frac{N_{k(-1), k}m_{k'}}{N_{k(-1), k'}N_{k(-1), k'-1}} \leq \frac{\epsilon}{\sqrt{L}}.
\]

Moreover, $N \ni k'' \mapsto \frac{N_{k(0), k(0)+k''}}{N_{k(-1), (k(0)+k'')}}$ is monotonically increasing with value in $(0, 1)$.

So, there is a positive integer $k(1) > k(0)$ such that

\[
\left|\frac{N_{k(0), k(1)}}{N_{k(-1), k(1)}} - T_1\right| \leq \frac{\epsilon}{\sqrt{L}}.
\]

Assume that one can find $k(\ell) > \ldots > k(0)$ satisfying
\[
\left|\frac{N_{k(-1), k(j)}}{N_{k(-1), k(j)}^+} - T_j\right| \leq \frac{\epsilon}{\sqrt{L}}
\]
for every $j \in [1, \ell]$. Then, by virtue of (5.7),
\[
\frac{N_{k(\ell), k'}}{N_{k(-1), k'}} - \frac{N_{k(\ell), k'-1}}{N_{k(-1), k'-1}} \leq \frac{\epsilon}{\sqrt{L}},
\]
implicating that one can find $k(\ell + 1) > k(\ell)$ such that
\[
\left|\frac{N_{k(\ell), k(\ell+1)}}{N_{k(-1), k(\ell+1)}} - T_{\ell+1}\right| \leq \frac{\epsilon}{\sqrt{L}}.
\]
From this, the conclusion immediately follows.
5.3. Construction of an adapted code. Let \( \{\ell_{L}\}_{L \geq 0} \) be a sequence of positive numbers such that \( \lim_{L \to \infty} \ell_{L} = 0 \). For each \( L \geq 1 \), let \( \{t_{L,j}\}_{j=1}^{J(L)} \) be a finite subset of \( A_{L} \) such that \( \{B_{\ell_{L}/(L+1)}(t_{L,j})\}_{j=1}^{J(L)} \) covers \( A_{L} \). (Recall that \( B_{\ell}(t) \) is the ball of radius \( \ell \) and center \( t \).) If we write \( \mu_{L,j} \) for \( \mu_{\ell} \) with \( t = t_{L,j} \), then it follows from Lemma 5.4 that \( \{B_{\ell_{L}}(\mu_{L,j})\}_{j=1}^{J(L)} \) covers \( \Delta_{L}(J) \). We consider a lexicographic order in \( A = \{(L,j)\}_{L \geq 1, 1 \leq j \leq J(L)} \) by

\[
(L',j') \lesssim (L,j) \quad \text{if} \quad L' < L, \text{ or } L' = L \text{ and } j' < j.
\]

We define finite increasing sequences of positive integers \( k_{L,j} = \{k_{L,j}(\ell)\}_{\ell=1}^{L} \) inductively with respect to \( (L,j) \in A \). Let \( k_{1,1} = \{k_{1,1}(\ell)\}_{\ell=1}^{1} \) be a finite increasing sequence of positive integers such that

\[
\left| \bar{\mu}(k_{1,1}) - t_{1,1} \right| \leq \epsilon_{1}.
\]

We can take such \( k_{1,1} \) by virtue of Lemma 5.4. Let \( (L,j) \in A \), and assume that \( k_{L',j'} = \{k_{L',j'}(\ell)\}_{\ell=1}^{L'} \) is defined for any \( (L',j') \in A \) satisfying \( (L',j') \lesssim (L,j) \). Then we take \( k_{L,j} = \{k_{L,j}(\ell)\}_{\ell=1}^{L} \) as a finite increasing sequence of positive integers such that, if we write \( (L',j') \) for the predecessor of \( (L,j) \) (i.e. \( L' = L \) and \( j' = j - 1 \), or \( L' = L - 1 \) and \( j' = J(L'), j = 1 \)), then

\[
k_{L,j}(-1) = k_{L',j'}(L') \quad \text{and} \quad \frac{2N_{k_{L,j}(-1)}}{N_{k_{L,j}(L)}} < \bar{\varepsilon}_{L},
\]

and that

\[
\left| \bar{\mu}(k_{L,j}) - t_{L,j} \right| \leq \frac{\bar{\varepsilon}_{L}}{L + 1}.
\]

Again, we can take such \( k_{L,j} \) due to Lemma 5.4. Finally, let \( \{\ell_{k}\}_{k \geq 1} \) be a sequence of nonnegative integers associated with \( k_{L,j} \) for all \( (L,j) \in A \), and \( f \) a \( C^{r} \) diffeomorphism with a wandering domain \( D \) coded by \( \{\ell_{k}\}_{k \geq 1} \) over \( \{m_{k}\}_{k \geq 1} \).

We now complete the proof of Theorem 2.3. Fix \( \check{L} \geq 1, t \in \Delta_{L} \) and \( \epsilon > 0 \). Let \( L \geq \check{L} \) be an integer such that \( \bar{\varepsilon}_{L} < \epsilon / 4 \) and

\[
\ell_{k} + \frac{\sum_{k=k_{1}}^{k_{2}} I_{k}}{\sum_{k=k_{1}}^{k_{2}} m_{k}} - \zeta + \left| 1 - \frac{\sum_{k=k_{1}}^{k_{2}} \# I_{k} + \# \hat{I}_{k}}{\sum_{k=k_{1}}^{k_{2}} m_{k}} \right| < \frac{\epsilon}{8}
\]

for all \( k_{2} > k_{1} \geq k_{L,1}(-1) \) (see Remark 5.3). Let \( 1 \leq j \leq J(L) \) be an integer such that \( d(\mu_{k}, \mu_{L,j}) \leq \bar{\varepsilon}_{L} \) (one can find such \( j \) by the construction of \( \{t_{L,j}\}_{j=1}^{J(L)} \)) and the fact \( \Delta_{\check{L}}(J) \subset \Delta_{L}(J) \), together with Lemma 3.4. Then, with the notation \( \check{k} = k_{L,j} \) and \( \check{N} = N_{k_{L,j}(L)} \), it follows from Lemma 5.2, 5.8, and 5.10 that

\[
d\left( \bar{\delta}^{\check{N}}_{x}, \mu_{\bar{\mu}(\check{k})} \right) \leq \bar{\varepsilon}_{L} + \frac{\epsilon}{4} \quad \text{for all} \quad x \in D,
\]

and from 5.9 and Lemma 5.4 that

\[
d\left( \mu_{\bar{\mu}(\check{k})}, \mu_{L,j} \right) \leq \bar{\varepsilon}_{L}.
\]

Therefore, we get \( d \left( \bar{\delta}^{\check{N}}_{x}, \mu_{t} \right) < \epsilon \) for all \( x \in D \). Since \( \check{L} \geq 1, t \in \Delta_{L} \) and \( \epsilon > 0 \) are arbitrary, we conclude that \( D \subset E(J,f) \). This completes the proof of Theorem 2.3.
Appendix A. Proof of (1.12)

In this appendix we give the proof of (1.12), by following [13] Theorem 2.1. Let $X \subset \{1, 2, \ldots, m\}$ be a subshift with the specification property with $m \geq 2$, endowed with the metric $d_X$ given by $d_X(x, y) = \sum_{j=0}^{\infty} \frac{|x_j - y_j|}{\beta^j}$ for $x = (x_0, x_1, \ldots), y = (y_0, y_1, \ldots) \in X$ with $\beta > 1$. Let $f : X \to X$ be the left shift operator. For each $\epsilon > 0$, we let $T(\epsilon, X)$ be the maximal cardinality of a set $\tilde{F}$ consisting of periodic orbits of $f$ such that $d_X(x, y) \geq \epsilon$ for any $O, O' \in \tilde{F}$ with $O \neq O'$ and $x \in O, y \in O'$. Then, it easily follows from [13] Theorem 1.6 that
\[
\liminf_{\epsilon \to 0} \frac{\log \log N(\epsilon, \mathcal{P}_f(X))}{-\log \epsilon} \geq \liminf_{\epsilon \to 0} \frac{\log T(\epsilon, X)}{-\log \epsilon}.
\]
Therefore, (1.12) follows from the following claim:
\[
(A.1) \quad \liminf_{\epsilon \to 0} \frac{\log T(\epsilon, X)}{-\log \epsilon} \geq \dim(X).
\]
We will prove (A.1). Given $\epsilon > 0$, let $N(\epsilon)$ be a positive integer such that $\beta^{-N(\epsilon)} \leq \epsilon < \beta^{-N(\epsilon)+1}$. Then, for any periodic orbitss $O, O'$ of period $N(\epsilon)$ with $O \neq O'$, we have $\min\{d_X(x, y) \mid x \in O, y \in O'\} \geq \beta^{-N(\epsilon)+1} > \epsilon$. On the other hand, since any subshift is expansive (23 Section 16), by applying [23] Theorem 22.7 for continuous maps with expansiveness and the specification property, we get that
\[
h_{\text{top}}(f) = \lim_{n \to \infty} \frac{1}{n} \log \text{Per}_n(f),
\]
where $h_{\text{top}}(f)$ is the topological entropy of $f$ and $\text{Per}_n(f)$ is the number of periodic points of period $n$. Therefore, for any $\delta > 0$, by taking $\epsilon$ sufficiently small, we get
\[
T(\epsilon, X) \geq \frac{\text{Per}_N(\epsilon)}{N(\epsilon)} \geq e^{h_{\text{top}}(f)(1-\delta)N(\epsilon)} \geq e^{-h_{\text{top}}(f)(1-\delta) \log \epsilon / \log \beta}.
\]
So, it follows from Furstenberg’s formula $\dim(X) = h_{\text{top}}(f) / \log \beta$ for subshifts (24) that
\[
\liminf_{\epsilon \to 0} \frac{\log T(\epsilon, X)}{-\log \epsilon} \geq \frac{(1-\delta)h_{\text{top}}(f)}{\log \beta} = (1-\delta)\dim(X).
\]
Since $\delta > 0$ is arbitrary, we get the claim.

Acknowledgments

We would like to express our gratitude to P. Barrientos, Y. Cao, Y. Chung, T. Persson, A. Raibekas, H. Takahasi, K. Yamamoto and D. Yang for many fruitful discussions and valuable comments. Furthermore, we are grateful to P. Berger and S. Biebler for their valuable comments on stretched exponential emergences, in particular, their pointing out of a mistake in an earlier version of the proof of Corollary 1.7. Moreover, we express our great appreciation to the anonymous reviewers for many suggestions, all of which substantially improved the paper, especially on the presentation of the paper and the exposition of pointwise emergence for dynamics with the specification property. The first and second authors are grateful to the members of Soochow University for their kind hospitality when they visited there. The second author is also grateful to the members of Lund university for their warm hospitality. This work was partially supported by JSPS KAKENHI Grant Numbers 21K03332, 19K14575 and 18K03376.
References

[1] V. Araújo, Attractors and time averages for random maps, Annales de l’Institut Henri Poincaré (C) Non Linear Analysis 17 (2000), 307–369.
[2] V. Araújo, Infinitely many stochastically stable attractors, Nonlinearity 14 (2001), 583–596.
[3] V. Araújo, V. Pinheiro, Abundance of wild historic behavior, Bulletin of the Brazilian Mathematical Society, New Series (2019), 1–36.
[4] V. Baladi, C. Bonatti, B. Schmitt, Abnormal escape rates from nonuniformly hyperbolic sets, Ergodic Theory and Dynamical Systems 19 (1999), 1111–1125.
[5] L. Barreira, J. Schmeling, Sets of “non-typical” points have full topological entropy and full Hausdorff dimension, Israel Journal of Mathematics 116 (2000), 29–70.
[6] L. Barreira, J. Li, C. Valls, Irregular sets are residual, Tohoku Mathematical Journal 66 (2014), 471–489.
[7] L. Barreira, J. Li, C. Valls, Topological entropy of irregular sets, Revista Matemática Iberoamericana 34 (2018), 853–878.
[8] P. Barrientos, S. Kiriki, Y. Nakano, A. Raibekas, T. Soma, Historic behavior in non-hyperbolic homoclinic classes, Proceedings of the American Mathematical Society 148 (2020), 1195–1206.
[9] P. Barrientos, S. Biebler, Emergence of wandering stable components, arXiv preprint arXiv:2001.08649 (2020).
[10] P. Berger, On Emergence and Complexity of Ergodic Decompositions, Advances in Mathematics 390 (2021), 107904.
[11] T. Bomfim, P. Varandas, Multifractal analysis for weak Gibbs measures: from large deviations to irregular sets, Ergodic Theory and Dynamical Systems 37 (2017), 79–102.
[12] C. Bonatti, L. Díaz, M. Viana, Dynamics Beyond Uniform Hyperbolicity: A Global Geometric and Probabilistic Perspective, Encyclopaedia of Mathematical Science 102, Springer-Verlag, 2004.
[13] Y. Cao, L. Zhang, Y. Zhao, The asymptotically additive topological pressure on the irregular set for asymptotically additive potentials, Nonlinear Analysis: Theory, Methods & Applications 74 (2011), 5015–5022.
[14] E. Catsigeras, X. Tian, E. Vargas, Topological entropy on points without physical-like behaviour, Mathematische Zeitschrift (2015), 1–13.
[15] Y. Chen, K. Küpper, L. Shi, Topological entropy for divergence points, Ergodic Theory and Dynamical Systems (2005), 1173–1208.
[16] Y. M. Chung, J. Rivera-Letelier, H. Takahasi, Large deviation principle in one-dimensional dynamics, Inventions mathematicae 218 (2019), 853–888.
[17] A. Cobham, The intrinsic computational difficulty of functions (1965).
[18] E. Colli, E. Vargas, Non-trivial wandering domains and homoclinic bifurcations, Ergodic Theory and Dynamical Systems 21 (2001), 1657–1681.
[19] S. Crovisier, D. Yang, J. Zhang, Empirical measures of partially hyperbolic attractors, Communications in Mathematical Physics (2020), 1–40.
[20] M. Denker, C. Grillenberger, K. Sigmund, Ergodic theory on compact spaces, Lecture Notes in Mathematics 527, Springer-Verlag, 1976.
[21] Y. Dowker, The mean and transitive points of homeomorphisms, Annals of Mathematics (1953), 123–133.
[22] H. Fürstenberg, Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation, Mathematical systems theory 1 (1967), 1–49.
[23] A. Gaunersdorfer, Time averages for heteroclinic attractors, SIAM Journal on Applied Mathematics 52 (1992), 1476–1489.
[24] P. Guarino, P. A. Guihenéuf, B. Santiago, Dirac physical measures on saddle-type fixed points, Journal of Dynamics and Differential Equations (2020), 1–66.
[28] F. Hofbauer, G. Keller, *Quadratic maps without asymptotic measure*, Communications in mathematical physics **127** (1990), 319–337.

[29] V. Y. Kaloshin, *Generic diffeomorphisms with superexponential growth of number of periodic orbits*, Communications in Mathematical Physics **211** (2000), 253–271.

[30] A. Katok, *Lyapunov exponents, entropy and periodic orbits for diffeomorphisms*, Publications Mathématiques de l’IHÉS **51** (1980), 137–173.

[31] S. Kiriki, M. Li, T. Soma, *Geometric Lorenz flows with historic behavior*, Discrete & Continuous Dynamical Systems-A **36** (2016), 7021–7028.

[32] S. Kiriki, T. Soma, *Takens’ last problem and existence of non-trivial wandering domains*, Advances in Mathematics **306** (2017), 524–588.

[33] I. S. Labouriau, A. A. P. Rodrigues, *On Takens’ Last Problem: tangencies and time averages near heteroclinic networks*, Nonlinearity **30** (2016), 1876.

[34] R. MacKay, *Nonlinearity in complexity science*, Nonlinearity **21** (2008), 273–281.

[35] S. Newhouse, *Diffeomorphisms with infinitely many sinks*, Topology **13** (1974), 9–18.

[36] S. Newhouse, *The abundance of wild hyperbolic sets and non-smooth stable sets for diffeomorphisms*, Publications Mathématiques de l’IHÉS **50** (1979), 101–151.

[37] W. Ott, J. Yorke, *When Lyapunov exponents fail to exist*, Physical Review E **78** (2008), 056203–056203.

[38] J. Palis, *A global view of dynamics and a conjecture on the denseness of finitude of attractors*, Astérisque **261** (2000), 335–347.

[39] J. Palis, S. Smale, *Structural stability theorems*, Global Analysis (eds. S.-S. Chern and S. Smale), American Mathematical Society **14** (1970), 223–231.

[40] J. Palis, F. Takens, *Hyperbolicity and sensitive chaotic dynamics at homoclinic bifurcations: Fractal dimensions and infinitely many attractors in dynamics*, Cambridge University Press, 1995.

[41] Y. Pesin, *Dimension theory in dynamical systems: contemporary views and applications*, University of Chicago Press, 2008.

[42] Y. Pesin, B. Pitskel’, *Topological pressure and the variational principle for noncompact sets*, Functional Analysis and its Applications **18** (1984), 307–318.

[43] D. Ruelle, *Historical behaviour in smooth dynamical systems*, Global Analysis of Dynamical Systems (eds. H. W. Broer et al), Institute of Physics Publishing (2001), 63–66.

[44] K. Sigmund, *On dynamical systems with the specification property*, Transactions of the American Mathematical Society **190** (1974), 285–299.

[45] F. Takens, *Heteroclinic attractors: time averages and moduli of topological conjugacy*, Boletim da Sociedade Brasileira de Matemática-Bulletin/Brazilian Mathematical Society **25** (1994), 107–120.

[46] F. Takens, *Orbits with historic behaviour, or non-existence of averages*, Nonlinearity **21** (2008), 33–36.

[47] A. Talebi, *Non-statistical rational maps*, arXiv preprint arXiv:2003.02185 (2020).

[48] D. Thompson, *Irregular sets, the β-transformation and the almost specification property*, Transactions of the American Mathematical Society **364** (2012), 5395–5414.

[49] C. Villani, *Topics in optimal transportation*, American Mathematical Soc., 2003.

[50] C. Villani, *Optimal transport: old and new*, Springer Science & Business Media, 2008.