Synchronization and averaging in partially hyperbolic systems with fast and slow variables

Federico Bonetto[1] **and Guido Gentile**[2]

[1] School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332, USA

[2] Dipartimento di Matematica e Fisica, Università Roma Tre, Roma, 00146, Italy

e-mail: fb49@gatech.edu (corresponding author), guido.gentile@uniroma3.it

Abstract

We consider a family of dynamical systems obtained by coupling an Anosov map on the two-dimensional torus – the *chaotic system* – with the identity map on the one-dimensional torus – the *neutral system* – through a dissipative interaction. Firstly, we show that the two systems synchronize: the trajectories evolve toward an attracting invariant manifold, with exponential rates which are slower the smaller the interaction, and the full dynamics is conjugated to its linearization around the invariant manifold. As a byproduct, we obtain that there exists a unique exponentially mixing physical measure. Then, we focus on the instance of very weak interactions: in such a case the variable which describes the neutral system, since its evolution is very close to the identity, appears as a *slow* variable with respect to the variable which describes the chaotic system, and which is wherefore named the *fast* variable. We demonstrate that, seen on a suitably long time scale, the slow variable effectively follows the solution of a deterministic differential equation obtained by averaging over the fast variable. More precisely, we prove that the invariant manifold is in probability close to the fixed point of the averaged dynamics and that the difference between the exact evolution of the slow variable, seen from the invariant manifold, and the solution to the averaged system is in probability exponentially decreasing in time with a uniform rate.

1 Introduction

In this paper, we study the long time evolution of a dynamical system obtained by coupling a chaotic system on the two-dimensional torus with a one-dimensional neutral system[1] through a dissipative interaction, with special emphasis on the case of weak interaction, which corresponds to partially hyperbolic systems included in the class studied in [26]. Our goal is to show the occurrence of two related phenomena of interest.

The first one consists in the emergence, regardless of the size of the interaction as far as it is dissipative, of an attracting invariant manifold, on which the dynamics is essentially a copy of the evolution of the chaotic system. This is what is usually referred to as *synchronization*. In this respect, we extend the results in [30] to a more general class of systems. In addition, we show that in a large neighborhood of the invariant manifold there exists a *conjugation*, that is a function establishing a bijection between the actual dynamics and its linearization around the invariant manifold.

The second phenomenon can be described as follows. When the interaction between the two systems is very small, the evolution of the neutral variable becomes independent of the chaotic variable and is governed only by the average of the interaction over the chaotic variable.[2] More precisely, there exists a set $\mathcal{A}$ of the two-dimensional torus such that (1) for all initial conditions of the chaotic variable in $\mathcal{A}$ the evolution of the neutral variable, on a suitable time scale τ, is described by a differential equation that does not involve the chaotic variable, and (2) in the limit of vanishing interaction, the measure of the set

[1]By 'neutral system' we mean a system which in the absence of the interaction is at rest, so that the Lyapunov exponent of the central direction is zero – and remains close to zero when the interaction is switched on and is small.

[2]The 'chaotic variable' and the 'neutral variable' are the variables that, in the absence of interaction, describe the chaotic system and the neutral system, respectively.

$\mathcal{A}$ is asymptotically full and the time scale τ diverges. This phenomenon is known as *averaging* and it is studied for systems related to ours in [22, 23], among others. In the class of systems we consider, thanks to the hypothesis of uniform dissipation, we get more control on the dynamics: we show that, excluding a set of initial conditions for the chaotic variable whose measure vanishes with the size of the interaction, the averaged differential equation describes the evolution of the neutral variables toward the invariant manifold up to an error which remains uniformly bounded in time for all times and vanishes with the interaction. Moreover, when the dynamics is seen from the invariant manifold, the averaged differential equation describes the evolution of the neutral variables with an error exponentially decreasing in time, still after excluding a set of initial conditions for the chaotic variable which vanishes with the interaction.

We prove the existence of both the invariant manifold and the conjugation via an explicit construction relying on the hyperbolic properties of the chaotic system. To obtain averaging we prove several, progressively more complex correlation inequalities for the chaotic evolution that, combined with the explicit construction, allow us to control effectively the fluctuations of the full evolution around the averaged one.

We now give a more detailed description of the results briefly outlined above, without dwelling on the technical aspects. Formal statements are provided in Section 2, after introducing all the appropriate notations and definitions.

1.1 Deterministic description: synchronization

Synchronization in quasi-integrable systems is well known to occur in the presence of dissipation; a typical example is the orbital resonance in celestial mechanics [50]. By contrast, in chaotic systems, where trajectories starting at close initial conditions tend to diverge from each other, synchronization may appear as an unlikely phenomenon, and was considered so for a long time, up to the end of the last century [5] and the publication of seminal works such as Pecora and Carroll's paper [52]. In particular, the presence of negative Lypaunov exponents due to a dissipative coupling may still produce synchronization.

One of the simplest models one can think of is the partially hyperbolic system obtained by coupling a chaotic system, for instance an Anosov automorphism on the two-dimensional torus $\mathbb{T}^2$, such as the Arnold's cat map, with a one-dimensional neutral system on $\mathbb{T}$ through a dissipative perturbation which is unidirectional, that is which affects only the motion of the neutral system. Such a model has been explicitly considered in [30], by fully exploiting the perturbative nature of the coupling and under suitable assumptions which simplify the analysis to a great extent. More complicated and realistic models can be easily envisaged (see for instance the discussion in Section 1.3, in particular about the heat equation, and refer to [53, 56, 8, 35, 2] for further examples), however the main advantage of the simple model studied in [30] is that its solution can be explicitly worked out and studied in great detail, without resorting to numerical simulations or heuristic arguments. What is found in [30] is that a two-dimensional invariant manifold appears on which the dynamics is conjugated to that of the unperturbed automorphism: as a consequence, the two systems synchronize asymptotically, in the sense that they tend to realize a drive-response configuration, with the originally neutral system slaved to follow the dynamics of the chaotic system (see [8] and references therein for an introduction to the topics). The invariant manifold is no more than Hölder continuous, but, with the hypotheses considered in [30], its oscillations are small, that is of the same order as the perturbation – a property which is not expected to hold in general.

In the present paper we study a class of dynamical systems that include those considered in [30], and show that an invariant manifold exists in a more general setting, and it is still the graph of a Hölder continuous function over the two-dimensional torus (see Theorem 1). Moreover, we extend the analysis beyond the perturbative regime, in the sense that the coupling is required only to be dissipative in a finite region and is not necessarily small. The oscillations of the invariant manifold may be rather large in general, even in the perturbative regime, albeit large oscillations are rare in the latter case, given that the variance of the function whose graph describes the invariant manifold is of the order of the perturbation (see Theorem 4). We also provide a detailed description of the dynamics away from the invariant manifold, by demonstrating that it is conjugated to its linearization around the invariant manifold (see Theorem 2). In particular, the invariant manifold is proved to be an attractor. The conjugation too, in general, is no more than Hölder continuous. The variance of the conjugating function, seen as a function on $\mathbb{T}^2$ for fixed value of the neutral variable, is found to be of the order of the coupling, despite the fact that deviations of order 1 from its average are possible even in the case of small coupling (see Theorem 5). As a byproduct, we obtain that any system in the class we consider admits a unique physical measure given by the lift to the invariant manifold of the normalised Lebesgue measure of $\mathbb{T}^2$. Thanks to the uniformly dissipative nature of the interaction, the physical measure is exponentially mixing (see Theorem 3), with mixing rate of the order of the perturbation due to the low regularity of the invariant manifold (see Corollary 2.23). This can be compared with the results in [22, 46], in the case of expanding maps, where logarithmic corrections appear in the decay rate, due the non-uniform dissipation.

1.2 Probabilistic description: averaging

When the perturbation is very small, two time scales naturally appear in the evolution: the *fast* time scale of the chaotic dynamics on $\mathbb{T}^2$ as opposed to the *slow* time scale of the neutral system, whose evolution is driven only by the perturbation. The study of systems with fast and slow variables is well established for quasi-integrable systems where the motion of the slow variable is close to periodic, and probably originated with Lagrange's analysis of the secular variations of the orbital elements of planets [45]. In such systems, the slow variable, for a very long time, only feels the average over one period of its interaction with the fast variables. Thus, to study the drift of the slow variable, one applies the so-called *averaging method* [37, 4, 57]: once the oscillations of the fast variables have been integrated out, one finds an approximate solution which provides, in general, a reliable description of the dynamics up to a time inversely proportional to the slow time scale. However, this is only the first step in the study of the dynamics of the system, and one has to control the corrections in a deeper way in order to make the analysis rigorous and extend it up to longer – even infinite – time scales.

In a similar spirit, we investigate the dynamics of the systems we are considering when the size ρ of the perturbation is very small. We find that, in this situation, in order to describe the evolution of the slow variable, the fast variable, despite being deterministic, can somehow be treated as a random variable. More precisely, the evolution of the slow variable becomes essentially independent of the dynamics on the torus and is effectively described by an *averaged dynamics* in which the function describing the influence of the fast variable on the slow variable is replaced by its average on the torus. We show that, taking a random initial condition for the fast variable, the probability of seeing a sizable deviation from the deterministic averaged evolution is of the order of the perturbation (see Theorem 6, complemented by Lemma 2.28, and the ensuing Corollary 2.33). Furthermore, the deviations can be controlled and proved to decay exponentially with rate of order ρ (see Theorem 7), consistently with the the bounds on the correlations implied by Theorem 3 (see Corollary 2.23). We can interpret the above results in the form of a *scaling limit*, in the sense that we can fix a finite time t and study, when the size ρ of the perturbation is very small, what happens after $\lfloor t/\rho \rfloor$ iterations of the map. In this regime, henceforth called the *scaling regime*, the averaged dynamics is described by a suitable ordinary differential equation. This implies that the evolution of the slow variable, seen as a continuous time stochastic process in the rescaled time t, converges exponentially to the solution of an ordinary differential equation in probability in the topology of the uniform convergence (see Theorem 8).

Properties as those discussed above are exhibited by the class of models studied in [22, 24, 23, 46], where the one-dimensional system described by the slow variable is coupled with an expanding circle map and the coupling is dissipative only in average. When the system is not uniformly dissipative, the dynamics may become much more complicated. This is especially true when the averaged system admits more than one attractor. If this happens, starting close to one of them does not prevent the dynamics to move toward some other attractor, and, as a consequence, a further time scale naturally appears, on which the dynamics looks like as a random walk between metastable states. A result of this kind has been worked out in the aforementioned models based on expanding maps [22], where the rate of decay of correlations, in the case of more coexisting sinks for the one-dimensional averaged system, is proved to be much slower than in the case of only one attractor. An analogous behaviour may be expected to occur also in the case of the Anosov automorphisms, when the interaction is dissipative only in average. On the other hand, in the case of fully dissipative systems, where synchronization is achieved at exponential rate, the class of models we consider is not covered by the analysis in [22], because of the presence of both stable and unstable directions for the fast chaotic dynamics on $\mathbb{T}^2$; furthermore, considering Anosov systems instead of expanding maps has the advantage that the systems one deals with are reversible and exhibit symplectic structure [46]. Noteworthily, the stronger assumption on the dissipation allows us to identify the stable state with the metastable state (in particular, the last time scale listed in the introduction of [24] does not arise) and to obtain stronger bounds for the decay of the correlations, with no logarithmic corrections. For a more thorough comparison with the literature we refer to Section 2.5.

1.3 More general settings: a brief overview of possible applications

Systems of the kind described above, usually in more complex and articulated variants, appear in many applications. In fact, the class of systems we study in this paper can be considered as a prototype – or at least a simplified version – of physically more structured models, in the same spirit as in [30, 23].

One of the first problems to be studied, where fast and slow variables are coupled to each other, were the planetary motions in celestial mechanics. A well-known example are the effects of the revolution of the Moon around the Earth (the fast motion) on the revolution of the Earth around the Sun (the slow motion). Krylov-Bogolyubov theory provides a useful tool to deal with such a kind of problems and, more generally, to study the behaviour of oscillating systems where at least two very different time scales are involved: an averaged equation for the slow variables is obtained after integrating out the motions of the fast variables [42, 9, 38]. The theory has been successfully applied to a wide class of dynamical systems, which range from very simple two-dimensional systems, such as the Van der Pol equation or the inverted pendulum, to much more complicated ones, such as the stability of the Solar system, where, because of the complexity of the equations, numerical analysis plays a dominant role.

Recently the averaging method has been applied to study the stability of the H_2^+ ion within the framework of classical mechanics [15]. For certain initial conditions of the electron coordinates, the protons are found numerically to be captured in an oscillatory state. On the other hand, the numerical simulations also show that for other initial conditions the motion of the electron becomes chaotic. It would be interesting to investigate further the chaotic regime, in the light of the increasing results in the literature showing that synchronization may still occur when the dynamics of the fast variable moves from regular to chaotic; for instance, a behaviour of this kind is observed numerically in electromechanical systems with flexible arms (see [43] and references quoted therein). Another recent application of the theory of averaging is provided by the derivation of Haff's law for the system formed by two hard disks on a two-dimensional torus that dissipate energy due to inelastic collisions [36]: by rewriting the dynamics in terms of a dispersing billiard, in the limit of vanishing dissipation, the speed of the particle is approximated uniformly by the solution of a suitable ordinary differential equation.

Undoubtedly, also in the light of the relation of the problem of climate change with society and life on our planet, one of the most significant examples of chaotic systems where fast and slow variables interact with each other is the Earth's climate system: while weather processes, such as the atmospheric and ocean dynamics, behave as fast motions, what one is really interested in is the slow evolution of the climate [40]; moreover, one has to take into account also intentional and unintentional human-induced perturbations, such as the global warming due to human activities. Because of the wide range of external forces involved in the climate system, the mathematical models which are used to treat the problem in full generality are inevitably complicated, and the corresponding differential equations are mainly studied numerically. However, analytically more accessible models have been studied, not only because there are problems which admit a simpler description, but also on the grounds that obtaining analytical results allows us to improve our general understanding of the problem. A class of simple climate models are the *energy balance models*, where only a few variables appear. An example is provided by a two-dimensional model where the evolution of the mean surface temperature and of the mean deep ocean temperature is governed by a system of two stochastic differential equations: the climate system response is characterised by two time scales, with the deep ocean temperature reacting much more slowly [59]. The use of mathematical models in order to deal with the climate change has intensified in the last few years, thanks to the recent developments in dynamical systems theory as well as in statistical mechanics and probability (see [25, 47, 33] for reviews on the topics). The tools we use in the present paper provide a possible path to follow in order to address the analysis of climate models. Studying coupled Anosov systems, which in principle could appear as mathematical abstractions, is justified in consideration of Gallavotti-Cohen chaotic hypothesis [31]. In this regard, we stress that the results obtained by relying on Ruelle response theory [48, 33] exploit the very same assumption.

Another well established – and more ambitious – line of research aims at deriving rigorously the macroscopic law of transport of energy in a crystal, that is the *heat equation*, starting from the deterministic microscopic dynamics (see [11] for a review). Since one is interested in the long-time dynamics, in order to avoid the difficulties inherent in perturbations of integrable Hamiltonian systems one may either add some randomness, which leads to stochastic equations, or turn to a different approach and move toward strongly chaotic dynamics. In this perspective, as pointed out in [46], one would like to investigate the case of a large number N of chaotic systems, locally coupled with an equal number of neutral systems and weakly interacting with each other, and look for results which are uniform in N. One is ultimately interested in taking the *hydrodynamical limit*, that is considering the system obtained by a scaling limit in which both N and the (discrete) time k go to infinity in a suitable way (see [60] for a review). A model of this kind, with a different approach with respect to ours, is studied in [13], where the local systems are chaotic maps, the neutral variables play the role of local energies and a further conservative small interaction is introduced between the local systems; then, for initial conditions with the energies confined in a very small region, a diffusion equation is proved to be satisfied by the local energies at finite time.

Stated in its full generality, the problem is too hard for our present knowledge. As discussed in [22, 46], a possible strategy to pursue is to split the study into two separate steps:

1. First one studies a single local system weakly interacting with a neutral variable in the limit in which the size ρ of the interaction vanishes (the *scaling regime*). In order to obtain a non-trivial evolution one studies the behaviour of the system for times k diverging with a law which depends on ρ. In this way one obtains a differential equation (the *mesoscopic equation*) describing the dynamics of the neutral variable.
2. Next, one couples a large number N of such systems and takes the limit in which both N and the rescaled time go to infinity according to the hydrodynamic limit. The heat equation should emerge as the partial differential equation describing the evolution of the local energy concentration, that is the average of the neutral variable in a small region.

As far as the first step is concerned, one wishes the dynamics to be well understood in the absence of interaction. Since integrable systems have to be excluded because they are too special and are expected to display a non-typical behaviour, the local systems are usually assumed to be chaotic. A further simplifying hypothesis is to consider an interaction which makes non-conservative the evolution of the neutral variable, so that local attractors appear and the dynamics is controlled over long times. In such a situation the scaling limit requires $k \sim \rho^{-1}$ and the mesoscopic equation is an ordinary differential equation. In the more difficult conservative case, a different scaling law $k \sim \rho^{-2}$ is looked for and the mesoscopic equation is expected to be a stochastic differential equation.

The present paper deals with the first step of the strategy outlined above, in the fully dissipative case. Of course, the problem one has in mind is the general case of partially hyperbolic systems, where all manifolds – unstable, central and stable – exist, in the presence of a dissipative interaction along the central direction. Thus, the fact that the neutral system does not influence the chaotic evolution on the torus, although it still remains highly non-trivial, makes the analysis easier. Notwithstanding this simplification, we think that the model we study here contains most of the relevant features we need to control the statistical properties over long times of models with more general couplings (see Section 2.6). At the same time, it has the advantage of being well suited for explicit computations and eliminating details which would introduce technical complications without really adding anything to the underlying physics.[3] Therefore, in our opinion, results obtained for such a model represent a first step toward a full mathematical understanding of the problem of synchronization and averaging in chaotic systems, before considering more realistic situations.

2 Model and Results

In this section, first, we introduce the basic ingredients that will be used in the rest of the paper. Then, we give the formal definition of the model we study and present our main results, referring to Sections 3 to 5 – and the Appendices – for the proofs. A more detailed plan of the paper is given in Subsection 2.7.

2.1 Basic Ingredients

Set $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ and consider an Anosov automorphism A_0 on $\mathbb{T}^2$ [1, 14], such as Arnold's cat map [3]. Call $\lambda_\pm$ the eigenvalues and $v_\pm$ the eigenvectors of A_0, with $|v_\pm| = 1$, $|\lambda_+| > 1 > |\lambda_-|$ and $|\lambda_-\lambda_+| = 1$, and set $\lambda := |\lambda_+|$.

Consider $\Omega := \mathcal{U} \times \mathbb{T}^2$, where $\mathcal{U}$ is a non-empty closed interval of either $\mathbb{T}$ or $\mathbb{R}$, and, for any function $f\colon \Omega \to \mathbb{R}$ and any $\varphi \in \mathcal{U}$, set

$$\langle f\rangle(\varphi) := \langle f(\varphi,\cdot)\rangle := \int_{\mathbb{T}^2} f(\varphi,\psi)\, m_0(\mathrm{d}\psi), \qquad \tilde{f}(\varphi,\psi) := f(\varphi,\psi) - \langle f\rangle(\varphi), \tag{2.1}$$

where $m_0(\mathrm{d}\psi) := \mathrm{d}\psi/(2\pi)^2$ is the (normalized) Lebesgue measure on $\mathbb{T}^2$. Observe that m_0 is the unique *physical measure* for A_0, that is m_0 is invariant under A_0 and for every continuous function $f : \mathbb{T}^2 \to \mathbb{R}$, one has that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(A^n\psi) = \langle f\rangle \tag{2.2}$$

for almost every ψ with respect to the Lebesgue measure.

Let $\|f\|_\infty := \sup_{(\varphi,\psi)\in\Omega} |f(\varphi,\psi)|$ be the supremum norm on Ω and let $\mathscr{B}(\Omega,\mathbb{R})$ denote the Banach space of the bounded continuous real-valued functions f defined on Ω equipped with the norm $\|f\|_\infty$. For $\alpha \in (0,1]$, introduce the Hölder seminorm

$$|f|_\alpha := \sup_{(\varphi,\psi),(\varphi,\psi')\in\Omega} \frac{|f(\varphi,\psi') - f(\varphi,\psi)|}{|\psi'-\psi|^\alpha}, \tag{2.3}$$

the two directional Hölder seminorms

$$|f|_\alpha^\pm := \sup_{(\varphi,\psi)\in\Omega} \sup_{x\in\mathbb{R}} \frac{|f(\varphi,\psi+xv_\pm) - f(\varphi,\psi)|}{|x|^\alpha} \tag{2.4}$$

and the norms

$$\|f\|_{\alpha_-,\alpha_+} := \|f\|_\infty + \alpha_-|f|_{\alpha_-}^- + \alpha_+|f|_{\alpha_+}^+, \qquad \|f\|_\alpha^- := \|f\|_{\alpha,0}, \qquad \|f\|_\alpha^+ := \|f\|_{0,\alpha}, \tag{2.5}$$

and let $\mathscr{B}^*_{\alpha_-,\alpha_+}(\Omega,\mathbb{R})$, $\mathscr{B}^+_\alpha(\Omega,\mathbb{R})$ and $\mathscr{B}^-_\alpha(\Omega,\mathbb{R})$ denote the anisotropic Banach spaces of the Hölder continuous functions defined on Ω equipped with the norms $\|\cdot\|_{\alpha_-,\alpha_+}$, $\|\cdot\|_\alpha^+$ and $\|\cdot\|_\alpha^-$, respectively.[4] If $\alpha_- = \alpha_+ = \alpha$ the norm $\|\cdot\|_{\alpha,\alpha}$ is equivalent to the norm $\|\cdot\|_\alpha := \|\cdot\|_\infty + \alpha|\cdot|_\alpha$ of the α-Hölder continuous functions, so that $\mathscr{B}^*_{\alpha,\alpha}(\Omega,\mathbb{R}) = \mathscr{B}_\alpha(\Omega,\mathbb{R})$, where $\mathscr{B}_\alpha(\Omega,\mathbb{R})$ is the Banach space defined by the norm $\|\cdot\|_\alpha$; on the other hand, using anisotropic Banach spaces allows to treat differently the stable and unstable manifolds [7, 20], and this will be exploited in what follows. Finally, observe that $\mathscr{B}_0(\Omega,\mathbb{R}) = \mathscr{B}(\Omega,\mathbb{R})$.

[3]Similar considerations are at the basis of both the models studied in [30, 18] and the skew products given in [22, 23] as examples for the general theory in the case of expanding maps.

[4] With respect to the cone framework widely used in the recent literature (see for instance [26, 27, 22, 34, 16]) we use a more direct geometric approach that exploit the explicitly known stable and unstable directions.

Remark 2.1. It is easy to see that one has

$$|f|_\alpha^\pm \le |f|_\alpha \le |f|_\alpha^+ + |f|_\alpha^-, \qquad \|f\|_\infty \le c_\alpha |f|_\alpha + \|\langle f\rangle\|_\infty, \tag{2.6a}$$
$$|fg|_\alpha \le \|f\|_\infty |g|_\alpha + |f|_\alpha \|g\|_\infty, \qquad |fg|_\alpha^\pm \le \|f\|_\infty |g|_\alpha^\pm + |f|_\alpha^\pm \|g\|_\infty, \tag{2.6b}$$
$$\|fg\|_\alpha \le \|f\|_\alpha \|g\|_\alpha, \qquad \|fg\|_{\alpha_-,\alpha_+} \le \|f\|_{\alpha_-,\alpha_+} \|g\|_{\alpha_-,\alpha_+}, \tag{2.6c}$$

with $c_\alpha := (\pi\sqrt{2})^\alpha$. Indeed, writing $\psi' = \psi + x_+ v_+ + x_- v_-$ in (2.3), we obtain the inequality $|f|_\alpha^\pm \le |f|_\alpha$ in (2.6a), while, splitting

$$f(\varphi,\psi') - f(\varphi,\psi) = f(\varphi,\psi + x_+v_+ + x_-v_-) - f(\varphi,\psi + x_-v_-) + f(\varphi,\psi + x_-v_-) - f(\varphi,\psi)$$

and, bounding $|x|^\alpha \ge \max\{|x_+|^\alpha, |x_-|^\alpha\}$, we find

$$\frac{f(\varphi,\psi') - f(\varphi,\psi)}{|x|^\alpha} \le \frac{f(\varphi,\psi + x_+v_+ + x_-v_-) - f(\varphi,\psi + x_-v_-)}{|x_+|^\alpha} + \frac{f(\varphi,\psi + x_-v_-) - f(\varphi,\psi)}{|x_-|^\alpha},$$

which yields the inequality $|f|_\alpha \le |f|_\alpha^+ + |f|_\alpha^-$ in (2.6a). The second inequality in (2.6a) follows writing

$$f(\varphi,\psi) = \int_{\mathbb{T}^2} f(\varphi,\psi)\, m_0(\mathrm{d}\psi') = \int_{\mathbb{T}^2} \big((f(\varphi,\psi) - f(\varphi,\psi')) + f(\varphi,\psi') \big)\, m_0(\mathrm{d}\psi')$$

and bounding $|\psi - \psi'| \le \pi\sqrt{2}$ on $\mathbb{T}^2$. The inequalities (2.6b) are obtained by writing

$$f(\varphi,\psi')g(\varphi,\psi') - f(\varphi,\psi)g(\varphi,\psi) = \big(f(\varphi,\psi') - f(\varphi,\psi)\big)g(\varphi,\psi') + f(\varphi,\psi)\big(g(\varphi,\psi') - g(\varphi,\psi)\big).$$

Finally, the inequalities (2.6c) follow from (2.6b) by noting that the right hand sides contain additional terms with respect to the left hand sides.

For $k \ge 1$, define also $\mathscr{B}_{\alpha,k}(\Omega,\mathbb{R})$ as the Banach space of the functions $f\colon \Omega \to \mathbb{R}$ which are k-times continuously differentiable in the first variable, that is in the variable $\varphi \in \mathcal{U}$, and such that the first $k-1$ derivatives in the first variable are α-Hölder continuous in the second one, that is in the variable $\psi \in \mathbb{T}^2$, equipped with the norm

$$|||f|||_{\alpha,k} := \sum_{n=0}^{k-1} \|\partial_\varphi^n f\|_\alpha + \|\partial_\varphi^k f\|_\infty = \sum_{n=0}^{k} \|\partial_\varphi^n f\|_\infty + \alpha \sum_{n=0}^{k-1} |\partial_\varphi^n f|_\alpha\,, \tag{2.7}$$

where ∂_φ denotes the derivative with respect to the first variable. Similarly, define the norms $|||\cdot|||_{\alpha,k}^+$ and $|||\cdot|||_{\alpha,k}^-$, as in (2.7) with $|\cdot|_\alpha$ replaced with $|\cdot|_\alpha^+$ and $|\cdot|_\alpha^-$, respectively, and call the respective Banach spaces $\mathscr{B}_{\alpha,k}^+(\Omega,\mathbb{R})$ and $\mathscr{B}_{\alpha,k}^-(\Omega,\mathbb{R})$. As in (2.6c) we get

$$|||fg|||_{\alpha,k} \le |||f|||_{\alpha,k} |||g|||_{\alpha,k}, \qquad |||fg|||_{\alpha,k}^\pm \le |||f|||_{\alpha,k}^\pm |||g|||_{\alpha,k}^\pm. \tag{2.8}$$

We call $\mathfrak{B}^*_{\alpha_+,\alpha_-}(\mathbb{T}^2,\mathbb{R})$ the subspace of $\mathscr{B}^*_{\alpha_+,\alpha_-}(\Omega,\mathbb{R})$ whose functions do not depend on φ, and similarly for $\mathfrak{B}_\alpha^\pm(\mathbb{T}^2,\mathbb{R})$, $\mathfrak{B}_\alpha(\mathbb{T}^2,\mathbb{R})$ and $\mathfrak{B}(\mathbb{T}^2,\mathbb{R})$. For such functions, the seminorms $|\cdot|_\alpha^\pm$ and the norms $\|\cdot\|_\infty$, $\|\cdot\|_{\alpha,\alpha_+}$ and $\|\cdot\|_\alpha^\pm$ are defined as previously, with the supremum taken over $\psi \in \mathbb{T}^2$ only.

Remark 2.2. We do not consider spaces of functions more regular than Hölder continuous in the variable ψ because Hölder continuity is the highest regularity we obtain in general for both the invariant manifold and the conjugation, independently of the regularity of the dynamical systems we consider (see Section 2.2).

The behavior of the seminorms (2.3) for $f \in \mathfrak{B}_\alpha(\mathbb{T}^2,\mathbb{R})$ under the action of A_0^n, for $n \in \mathbb{Z}$, is given by

$$|f \circ A_0^n|_\alpha \le \lambda^{\alpha|n|} |f|_\alpha\,, \qquad |f \circ A_0^n|_\alpha^+ \le \lambda^{\alpha n} |f|_\alpha^+\,, \qquad |f \circ A_0^n|_\alpha^- \le \lambda^{-\alpha n} |f|_\alpha^-\,. \tag{2.9}$$

If the functions $g_0, \ldots, g_{n-1}$ are in $\mathfrak{B}_\alpha^\pm(\mathbb{T}^2,\mathbb{R})$, for some $\alpha \in (0,1]$, then we have

$$\Big| \prod_{i=0}^{n-1} g_i \circ A_0^{\mp i} \Big|_\alpha^\pm \le \sum_{i=0}^{n-1} \lambda^{-\alpha i} |g_i|_\alpha^\pm \prod_{\substack{j=0 \\ j\neq i}}^{n-1} \|g_i\|_\infty. \tag{2.10}$$

Remark 2.3. In the following we also consider sets of the form $\mathcal{A} = \{(\varphi, \psi) : a_-(\psi) \le \varphi \le a_+(\psi)\}$ in $\mathbb{R} \times \mathbb{T}^2$, where $a_\pm : \mathbb{T}^2 \to \mathbb{R}$ are Hölder continuous functions. All the definitions of the norms given above extend naturally if the set Ω is replaced with any other closed subset of $\mathcal{A}$ as above. We only need to take the supremum over $(\varphi, \psi) \in \mathcal{A}$ and replace (2.3) with

$$|f|_\alpha := \sup_{(\varphi,\psi),(\varphi,\psi') \in \mathcal{A}} \frac{|f(\varphi,\psi') - f(\varphi,\psi)|}{|\psi' - \psi|^\alpha}, \qquad |f|^\pm_\alpha := \sup_{(\varphi,\psi)\in\mathcal{A}} \sup_{\substack{x\in\mathbb{R} \\ (\varphi,\psi+xv_\pm)\in\mathcal{A}}} \frac{|f(\varphi,\psi+xv_\pm) - f(\varphi,\psi)|}{|x|^\alpha}.$$

Then, we may define the corresponding Banach spaces in the same way as before, with $\mathcal{A}$ instead of Ω.

In order not to introduce further symbols, we use the notation $|||f|||_{0,k}$ also to denote the C^k-norm of any function f depending only on the first variable φ. We thus identify $C^k(\mathcal{D}, \mathbb{R})$, for any given subset $\mathcal{D} \subseteq \mathbb{R}$, with the subspace of $\mathscr{B}_{0,k}(\mathcal{D} \times \mathbb{T}^2, \mathbb{R})$ of the functions independent of ψ.

A crucial role in our analysis is played by the correlation functions and their decay due to the hyperbolicity of A_0. The following estimate is rather standard (see Appendix A.2 for details).

Proposition 2.4. *Let the functions g_+ and g_- be, respectively, in $\mathfrak{B}^+_\alpha(\mathbb{T}^2, \mathbb{R})$ and in $\mathfrak{B}^-_\alpha(\mathbb{T}^2, \mathbb{R})$, for some $\alpha \in (0, 1]$. Then for all $n \in \mathbb{N}$ one has*

$$|\langle g_+ \, g_- \circ A_0^n \rangle - \langle g_+ \rangle \langle g_- \rangle| \le C_0 (1 + \alpha n) \lambda^{-\alpha n} \|\tilde{g}_+\|^+_\alpha \|\tilde{g}_-\|^-_\alpha,$$

for a suitable positive constant C_0 independent of n, α, g_- and g_+.

Remark 2.5. Proposition 2.4 implies that, under the same assumptions, for every $\alpha' < \alpha$ one has $|\langle g_+ \, g_- \circ A_0^n \rangle - \langle g_+ \rangle \langle g_- \rangle| \le C_0' \lambda^{-\alpha' n} \|\tilde{g}_+\|^+_\alpha \|\tilde{g}_-\|^-_\alpha$, with the positive constant C_0' depending only on α'. Often, the correlation inequality is stated in this form (see for instance [55]): the sharper inequality in Proposition 2.4 motivates the discussion provided in Appendix A.2.

2.2 The model

We consider the dynamical system defined by the map $\mathscr{S}$ on $\mathbb{T} \times \mathbb{T}^2$ given by

$$\mathscr{S}(\varphi, \psi) = (\mathscr{S}_\varphi(\varphi, \psi), \mathscr{S}_\psi(\varphi, \psi)) = (G(\varphi, \psi), A_0 \psi), \qquad G(\varphi, \psi) := \varphi + F(\varphi, \psi), \tag{2.11}$$

with $F \in \mathscr{B}_{\alpha_0,6}(\mathbb{T} \times \mathbb{T}^2, \mathbb{R})$, for some $\alpha_0 \in (0, 1]$. We do not require more than Hölder continuity in ψ inasmuch as requiring stronger regularity does not provide in general stronger regularity of the invariant manifold and the conjugation (see also Remark 2.2).

Remark 2.6. Throughout the paper, for any map $\mathscr{S}$, the notation $\mathscr{S}^n$ means $\mathscr{S} \circ \mathscr{S}^{n-1} = \mathscr{S} \circ \ldots \circ \mathscr{S}$, that is the composition of $\mathscr{S}$ with itself n times.

2.2.1 Hypotheses on the map

We assume $\mathscr{S}$ in (2.11) to satisfy the following further hypotheses.

Hypothesis 1. *There exists a non-empty closed interval $\mathcal{U} \subsetneq \mathbb{T}$ such that $\partial_\varphi \mathscr{S}_\varphi(\varphi, \psi) > 0$ for all $(\varphi, \psi) \in \Omega := \mathcal{U} \times \mathbb{T}^2$. If, after a suitable parametrization of $\mathbb{T}$, one writes $\mathcal{U} = [\phi_m, \phi_M]$, with $\phi_m < \phi_M$, then $\mathscr{S}_\varphi(\phi_m, \psi) > \phi_m$ and $\mathscr{S}_\varphi(\phi_M, \psi) < \phi_M$ for all $\psi \in \mathbb{T}^2$.*

Hypothesis 2. *For every $\psi \in \mathbb{T}^2$ there is a unique point $S(\psi) \in \mathrm{int}\,\mathcal{U}$ such that $F(S(\psi), \psi) = 0$.*

Set also

$$S_m := \inf_{\psi\in\mathbb{T}^2} S(\psi), \quad S_M := \sup_{\psi\in\mathbb{T}^2} S(\psi), \quad \Lambda := [S_m, S_M] \times \mathbb{T}^2, \quad \Gamma := -\sup_{(\varphi,\psi)\in\Lambda} \partial_\varphi F(\varphi, \psi). \tag{2.12}$$

Hypothesis 3. *One has $\Gamma > 0$.*

A map $\mathscr{S}$ satisfying Hypotheses 1–3 defines a partially hyperbolic system [26]. By Hypothesis 1, $\mathscr{S}$ is injective on Ω and $\Gamma < 1$. By Hypotheses 1 and 2, $\mathscr{S}(\Omega) \subsetneq \operatorname{int}\Omega$ and $S(\psi) \in (\phi_m, \phi_M)$ $\forall \psi \in \mathbb{T}^2$, and hence $[S_m, S_M] \subsetneq (\phi_m, \phi_M)$. Hypothesis 3 yields a stronger assumption on the dissipation when compared with [22, 24, 23, 46], because it ensures that the dissipation is uniform.

The following lemma lists a few immediate consequences of Hypotheses 1–3.

Lemma 2.7. *If $\mathscr{S}$ satisfies Hypotheses 1–3, then the following properties hold:*

1. *one has $F(\varphi,\psi) > 0$ for $\varphi < S_m$, while $F(\varphi,\psi) < 0$ for $\varphi > S_M$;*
2. *one has $\partial_\varphi F(\varphi,\psi) > -1$ for $(\varphi,\psi) \in \Omega$ and hence $-1 < \partial_\varphi F(\varphi,\psi) \le -\Gamma$ for $(\varphi,\psi) \in \Lambda$;*
3. *for any $r > 0$ there exists $N_r \in \mathbb{N}$ such that $\mathscr{S}^{N_r}(\Omega) \subset \Lambda_r := [S_m - r, S_M + r] \times \mathbb{T}^2$, with*
$$N_r \le \frac{\max\{\phi_M - S_M, S_m - \phi_m\}}{\inf\limits_{\Omega\setminus\Lambda_r} |F(\varphi,\psi)|}\,; \tag{2.13}$$
4. *for any $\Gamma' \in (0,\Gamma)$ there exists $r = r(\Gamma')$ such that $\partial_\varphi F(\varphi,\psi) \le -\Gamma'$ for all $(\varphi,\psi) \in \Lambda_r$;*
5. *the set Λ is positively invariant under $\mathscr{S}$ and is attracting for $\mathscr{S}$ on Ω;*
6. *there exists a unique $\overline{\varphi} \in (S_m, S_M)$ such that $\langle F(\overline{\varphi},\cdot)\rangle = 0$.*

Remark 2.8. Thanks to property 6 in Lemma 2.7, without loss of generality we choose the parametrization in Hypothesis 2 in such a way that $\overline{\varphi} = 0$.

Remark 2.9. By Remark 2.8, we can write $F(\varphi,\psi) = \beta(\psi) - \nu(\psi)\,\varphi + \delta(\varphi,\psi)\,\varphi^2$, with $\langle\beta\rangle = 0$. Conversely, if we replace F with the function $F_0(\varphi,\psi) := \beta(\psi) - \nu(\psi)\,\varphi + \delta(\psi)\,\varphi^2$, with $\langle\beta\rangle = 0$ and $4\|\delta\|_\infty\|\beta\|_\infty < \nu_0^2$, where we have set $\nu_0 := \inf_{\psi\in\mathbb{T}^2}|\nu(\psi)|$, then it is easy to see that (2.11) defined by F_0 instead of F satisfies Hypotheses 1–3 with $\phi_M = \|\beta\|_\infty/2\nu_0$ and $\phi_m = -\phi_M$.

Remark 2.10. Although $\mathscr{S}$ is defined as a map on $\mathbb{T}\times\mathbb{T}^2$, given that $\mathscr{S}(\Omega) \subsetneq \Omega$ what we are really interested in only the action of $\mathscr{S}$ on Ω, that we identify as a subset of $\mathbb{R}\times\mathbb{T}^2$. For technical reasons, we also need to extend $\mathscr{S}|_\Omega$ to a map $\mathscr{S}_{\rm ext}$ on $\Omega_{\rm ext} = \mathcal{U}_{\rm ext}\times\mathbb{T}^2$, with $\mathcal{U}_{\rm ext} \supset \mathcal{U}$ a closed interval, such that $\mathscr{S}_{\rm ext}$ still verifies Hypotheses 1–3. We provide a concrete construction of $\mathscr{S}_{\rm ext}$, as simple as possible, in Subsection 4.1. We stress here that the construction of the invariant manifold $\mathcal{W}$ and of the conjugation $\mathscr{H}$ dicussed in Theorems 1 and 2 below does not require any extension of $\mathscr{S}$. Thus the statement of Theorem 3 and the estimates in Theorems 4 and 5 do not depend on the particular choice of the extension $\mathscr{S}_{\rm ext}$ that only appears in the intermediate steps of the proof of Theorems 4 and 5 (see Remarks 4.2, 4.6 and 4.13 for further details).

We call φ the *slow variable* and ψ the *chaotic* or *fast variable* – such a terminology is motivated by the fact that we are mainly interested in the limit of small Γ, where the neutral variable φ moves slowly with respect to the chaotic variable ψ describing the hyperbolic system.

2.2.2 Iterated products

We introduce here some further notation that will be used widely throughout the rest of the paper. Let $\mathscr{S}$ be any map on $\mathcal{U}\times\mathbb{T}^2$ of the form
$$\mathscr{S}(\varphi,\psi) = (\mathscr{S}_\varphi(\varphi,\psi), \mathscr{S}_\psi(\varphi,\psi)) := (G(\varphi,\psi), A_0\psi)\,, \tag{2.14}$$
and let $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ be functions defined in $\mathcal{U}\times\mathbb{T}^2$. Define, for $k = 1,\dots,n$ and $i = 0,\dots,n-k$,
$$\mathfrak{p}^{(k)}_{[i]}(\mathscr{S};\varphi,\psi) := \prod_{j=0}^{k-1} \mathfrak{p}_{i+j}(\mathscr{S}^j(\varphi,\psi)), \qquad \mathfrak{p}^{(k)}(\mathscr{S};\varphi,\psi) := \mathfrak{p}^{(k)}_{[0]}(\mathscr{S};\varphi,\psi). \tag{2.15}$$

Remark 2.11. Throughout the paper we use the convention that a product over an empty set of indices is 1, while a sum over an empty set of indices is 0. This means that the definition (2.15) extends to $k = 0$ and $k = -1$ and yields $\mathfrak{p}^{(-1)}_{[i]}(\mathscr{S};\varphi,\psi) = \mathfrak{p}^{(0)}_{[i]}(\mathscr{S};\varphi,\psi) = 1$.

Remark 2.12. The map $\mathscr{S}$ in (2.14) is not necessarily the map (2.11) which defines our model. In particular, in what follows, we shall use the notation (2.15) for several maps, including the translated map $\mathscr{S}_1$ and the auxiliary map $\mathscr{S}_2$ which will be introduced in Subsection 4.7.1.

If the function $\mathscr{S}_\varphi(\varphi,\psi)$ does not depend on ψ and the functions $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ are independent of ψ as well, instead of (2.15) we may consider

$$\mathfrak{p}^{(k)}_{[i]}(\mathscr{S};\varphi) := \prod_{j=0}^{k-1}\mathfrak{p}_{i+j}(G^j(\varphi)), \qquad \mathfrak{p}^{(k)}(\mathscr{S};\varphi) := \mathfrak{p}^{(k)}_{[0]}(\mathscr{S};\varphi), \tag{2.16}$$

with $G(\varphi) := \mathscr{S}_\varphi(\varphi,\psi)$ and G^j denoting is the composition of G with itself j times. In particular, if $\overline{\mathscr{S}}(\varphi,\psi) = (\overline{G}(\varphi), A_0\psi)$ is the averaged map (2.26) and $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ are functions on $\mathcal{U}\times\mathbb{T}^2$, we have

$$\langle\mathfrak{p}\rangle^{(k)}_{[i]}(\overline{\mathscr{S}};\varphi) = \prod_{j=0}^{k-1}\langle\mathfrak{p}_{i+j}\rangle(\overline{G}^j(\varphi)), \qquad \langle\mathfrak{p}\rangle^{(k)}(\overline{\mathscr{S}};\varphi) := \langle\mathfrak{p}\rangle^{(k)}_{[0]}(\overline{\mathscr{S}};\varphi). \tag{2.17}$$

Write $\mathfrak{p}^{(n)}(\varphi,\psi) = \mathfrak{p}^{(k)}(\mathscr{S};\varphi,\psi)$ in (2.15) and set $\mathfrak{P}_{ij}(\varphi,\psi) := (\partial^i_\varphi\mathfrak{p}_j)(\mathscr{S}^j(\varphi,\psi)) = (\partial^i_\varphi\mathfrak{p}_j\circ\mathscr{S}^j)(\varphi,\psi)$ and $\mathfrak{S}_{ij}(\varphi,\psi) := \partial^i_\varphi(\mathscr{S}^j)_\varphi(\varphi,\psi)$. The following lemma is easily checked by direct computation.

Lemma 2.13. *Let $\mathscr{S}$ be a map on $\mathcal{U}\times\mathbb{T}^2$ of the form (2.14), with $\mathscr{S}_\varphi(\varphi,\psi)$ of class C^3 in φ, and let $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ be any functions defined in $\mathcal{U}\times\mathbb{T}^2$ and of class C^3 in the first variable φ. One has*

$$\partial_\varphi\mathfrak{p}^{(n)} = \sum_{i=0}^{n-1}\mathfrak{P}_{1i}\mathfrak{S}_{1i}\prod_{\substack{j=0\\ j\neq i}}^{n-1}\mathfrak{P}_{0j} \tag{2.18a}$$

$$\partial^2_\varphi\mathfrak{p}^{(n)} = \sum_{i=0}^{n-1}\Big(\mathfrak{P}_{2i}\mathfrak{S}_{1i}^2 + \mathfrak{P}_{1i}\mathfrak{S}_{2i}\Big)\prod_{\substack{j=0\\ j\neq i}}^{n-1}\mathfrak{P}_{0j} + \sum_{\substack{i,j=0\\ i\neq j}}^{n-1}\mathfrak{P}_{1i}\mathfrak{S}_{1i}\mathfrak{P}_{1j}\mathfrak{S}_{1j}\prod_{\substack{k=0\\ k\neq i,j}}^{n-1}\mathfrak{P}_{0k} \tag{2.18b}$$

$$\begin{aligned}\partial^3_\varphi\mathfrak{p}^{(n)} = {}& \sum_{i=0}^{n-1}\Big(\mathfrak{P}_{3i}\mathfrak{S}_{1i}^3 + 3\mathfrak{P}_{2i}\mathfrak{S}_{1i}\mathfrak{S}_{2i} + \mathfrak{P}_{1i}\mathfrak{S}_{3i}\Big)\prod_{\substack{j=0\\ j\neq i}}^{n-1}\mathfrak{P}_{0j} \\ & + 2\sum_{\substack{i,j=0\\ i\neq j}}^{n-1}\Big(\mathfrak{P}_{2i}\mathfrak{S}_{1i}^2 + \mathfrak{P}_{1i}\mathfrak{S}_{2i}\Big)\left(\partial_\varphi\mathfrak{p}_j\circ\mathscr{S}^j\right)\mathfrak{P}_{1j}\mathfrak{S}_{1j}\prod_{\substack{k=0\\ k\neq i,j}}^{n-1}\mathfrak{P}_{0k} \\ & + \sum_{\substack{i,j,k=0\\ i\neq j\neq k\neq i}}^{n-1}\mathfrak{P}_{1i}\mathfrak{S}_{1i}\mathfrak{P}_{1j}\mathfrak{S}_{1j}\mathfrak{P}_{1k}\mathfrak{S}_{1k}\prod_{\substack{h=0\\ h\neq i,j,k}}^{n-1}\mathfrak{P}_{0h}.\end{aligned} \tag{2.18c}$$

Remark 2.14. If both $\mathscr{S}_\varphi(\varphi,\psi)$ and the functions $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ do not depend on ψ, then setting $\mathfrak{p}^{(n)}(\varphi) = \mathfrak{p}^{(n)}(\mathscr{S};\varphi)$, equations (2.18) still hold, with $\mathscr{S}(\varphi,\psi)$ and $\mathfrak{p}_i(\mathscr{S}^i(\varphi,\psi))$ replaced, respectively, with $G(\varphi)$ and $\mathfrak{p}_i(G^i(\varphi))$, and with $\mathfrak{P}_{ij}(\varphi,\psi)$ and $\mathfrak{S}_{ij}(\varphi,\psi)$ defined accordingly.

2.3 Synchronization

2.3.1 The invariant manifold

If in (2.11) we replace A_0 with the identity $\mathbb{1}$, that is if we consider the dynamical system defined by the map $\mathscr{S}_{\mathbb{1}}(\varphi,\psi) = (\varphi + F(\varphi,\psi),\psi)$, with the function F still satisfying Hypotheses 1–3, then, setting $\mathcal{W}_{\mathbb{1}} = \{(S(\psi),\psi) : \psi\in\mathbb{T}^2\}$, we have $\mathscr{S}_{\mathbb{1}}(\mathcal{W}_{\mathbb{1}}) = \mathcal{W}_{\mathbb{1}}$. It is natural to ask whether a similar property remains true for $\mathscr{S}$ notwithstanding the chaotic nature of the evolution generated by A_0 on $\mathbb{T}^2$. We say that $\mathcal{W} := \{(W(\psi),\psi) : \psi\in\mathbb{T}^2\}$ is an *invariant manifold* for $\mathscr{S}$ in (2.11) if we have, for all $\psi\in\mathbb{T}^2$,

$$\mathscr{S}(W(\psi),\psi) = (W(A_0\psi), A_0\psi). \tag{2.19}$$

This means that on $\mathcal{W}$ the dynamics generated by $\mathscr{S}$ is conjugated to the dynamics generated by A_0.

The following result is proved in Subsection 3.1.

Theorem 1 (Synchronization). *Consider the dynamical system described by $\mathscr{S}$ in* (2.11) *satisfying Hypotheses 1–3. There exists a unique invariant manifold $\mathcal{W} = \{(W(\psi), \psi) : \psi \in \mathbb{T}^2\} \subset \Lambda$ for $\mathscr{S}$, with $W \in \mathfrak{B}^*_{\alpha_-,\alpha_+}(\mathbb{T}^2, \mathbb{R})$ for suitable $\alpha_-, \alpha_+ \in (0, \alpha_0]$.*

Remark 2.15. In general, $\mathcal{W}$ is only Hölder continuous in ψ even if F is smooth in ψ and W is of order 1 in $\|F\|_\infty$ even if F is very close to 0, its size depending mainly on $S(\psi)$ (see also Remark 3.1).

2.3.2 The linearized map and the conjugation

To analyze the evolution generated by $\mathscr{S}$ outside $\mathcal{W}$ we can try to conjugate it with its linearization around $\mathcal{W}$, that is with the simpler system $\mathscr{S}_0$ given by

$$\mathscr{S}_0(\eta, \psi) = (\kappa(\psi)\eta, A_0\psi), \qquad \kappa(\psi) := 1 + \partial_\varphi F(W(\psi), \psi), \tag{2.20}$$

where, by Hypothesis 3, one has $\kappa(\psi) \in (0, 1 - \Gamma)$ for all $\psi \in \mathbb{T}^2$. In other words, we look for a function $\mathscr{H}$ – called the *conjugating function* (or simply *conjugation*) – such that

$$\mathscr{H} \circ \mathscr{S} = \mathscr{S}_0 \circ \mathscr{H}. \tag{2.21}$$

Remark 2.16. Since $\mathscr{S}_0$ is linear in η, it is easy to see that if $\mathscr{H}$ is a solution to (2.21) then, for any $a \neq 0$, also $\mathscr{H}(a\varphi, A_0\psi)$ is a solution. Thus, we say that $\mathscr{H}$ is *the* conjugation if it solves (2.21) and can be written as $\mathscr{H}(\varphi, \psi) = (\mathcal{H}(\varphi, \psi), \psi)$, with $\partial_\varphi \mathcal{H}(0, \psi) = 1$ for every $\psi \in \mathbb{T}^2$.

The following result on the conjugation is proved in Subsection 3.2.

Theorem 2 (Conjugation). *Consider the dynamical system described by $\mathscr{S}$ in* (2.11) *satisfying Hypotheses 1–3. There exist a set $\Omega_0 \subset \mathbb{R} \times \mathbb{T}^2$ and a function $\mathscr{H} : \Omega \to \Omega_0$ that conjugates $\mathscr{S}$ via* (2.21) *to $\mathscr{S}_0$ given by* (2.20)*. Moreover $\mathscr{H}$ can be written as in Remark 2.16 with $\mathcal{H} \in \mathscr{B}_{\alpha_*,2}(\Omega, \mathbb{R})$, for a suitable $\alpha_* \in (0, \min\{\alpha_-, \alpha_+\}]$ and α_-, α_+ as in Theorem 1.*

Remark 2.17. Theorem 2 implies that, for any initial datum in Ω, the evolution generated by $\mathscr{S}$ leads towards $\mathcal{W}$. Therefore, $\mathcal{W}$ is a global attractor for the dynamical system $(\Omega, \mathscr{S})$.

From the proof of Theorem 2 it is easy to see that the function $\mathscr{H}$ is invertible; indeed, the following result is proved in Subsection 3.2.4.

Corollary 2.18. *The map $\mathscr{H}$ in Theorem 2 is invertible, and its inverse $\mathscr{H}^{-1} : \Omega_0 \to \Omega$, which satisfies $\mathscr{H}^{-1} \circ \mathscr{S}_0 = \mathscr{S} \circ \mathscr{H}^{-1}$, can be written as $\mathscr{H}^{-1}(\eta, \psi) = (\mathcal{L}(\eta, \psi), \psi)$, with $\mathcal{L} \in \mathscr{B}_{\alpha_*,1}(\Omega_0, \mathbb{R})$.*

As for W, the conjugation and its inverse are no more than Hölder continuous even if F is very smooth in ψ. The proofs of Theorem 2 and of Corollary 2.18 show that $\mathcal{H}$ and $\mathcal{L}$ can be written as

$$\mathcal{H}(\varphi, \psi) = \varphi - W(\psi) + (\varphi - W(\psi))^2 h(\varphi, \psi), \qquad \mathcal{L}(\eta, \psi) = W(\psi) + \eta + \eta^2 l(\eta, \psi), \tag{2.22}$$

with $h \in \mathscr{B}_{\alpha_*,1}(\Omega, \mathbb{R})$ and $l \in \mathscr{B}_{\alpha_*,1}(\Omega_0, \mathbb{R})$. These representations are used in Subsection 2.4.3, when the deviations of the conjugation are studied.

2.3.3 The physical measure

Let $m_0(\mathrm{d}\psi)$ be the Lebesgue measure on $\mathbb{T}^2$ (see Section 2.1) and let $\nu_0(\mathrm{d}\varphi\mathrm{d}\psi) := m_0(\mathrm{d}\psi)\mathrm{d}\varphi/(\phi_M - \phi_m)$ be the Lebesgue measure on $\Omega = \mathcal{U} \times \mathbb{T}^2$. By Theorems 1 and 2, given any observable $\mathcal{O} \in \mathscr{B}_{\alpha,0}(\Omega, \mathbb{R})$, one has

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mathcal{O}(\mathscr{S}^i(\varphi_0, \psi_0)) = \int \mathcal{O}(W(\psi), \psi)\, m_0(\mathrm{d}\psi) =: \nu_W(\mathcal{O}) \tag{2.23}$$

for almost every $(\varphi_0, \psi_0) \in \Omega$ with respect to the measure ν_0. Indeed, for all $n \geq 0$ and for any initial condition $(\varphi_0, \psi_0) \in \Omega$, setting $\eta_0 = \mathcal{H}(\varphi_0, \psi_0)$ and defining

$$\kappa^{(n)}(\psi) := \prod_{i=0}^{n-1} \kappa(A_0^i \psi)\,, \tag{2.24}$$

with $\kappa(\psi)$ as in (2.20), because of (2.22) and the regularity of $\mathcal{L}$ in η (see Corollary 2.18) we may write

$$\begin{aligned}\mathscr{S}^n(\varphi_0,\psi_0)) &= (\mathcal{L}(\kappa^{(n)}(\psi_0)\eta_0, A_0^n\psi_0), A_0^n\psi_0) = (\mathcal{L}(0, A_0^n\psi_0), A_0^n\psi_0) + O((1-\Gamma)^n|\eta_0|)\\ &= (W(A_0^n\psi_0), A_0^n\psi_0) + O((1-\Gamma)^n|\eta_0|).\end{aligned}$$

Thus, given any $\mathcal{O} \in \mathscr{B}_{\alpha,0}(\Omega,\mathbb{R})$, for any $\varepsilon > 0$ we may fix $n_0 \in \mathbb{N}$ such that for all $n > n_0$ we have $\mathcal{O}(\mathscr{S}^n(\varphi_0,\psi_0)) = \mathcal{O}(W(A_0^n\psi_0), A_0^n\psi_0) + B_n(\eta_0)$, with $|B_n(\eta_0)| < \varepsilon$. Thus, (2.23) follows immediately from Birkhoff's ergodic theorem and the observation that the invariant measure we are considering for the Anosov automorphism is the Lebesgue measure.

Hence, the unique *physical measure* for $\mathscr{S}$ on Ω is given by the image ν_W of m_0 on $\mathcal{W}$. It is thus interesting to study the mixing property of ν_W with respect to ν_0. The following result is proved in Subsection 3.3.

Theorem 3 (**Mixing on the invariant manifold**)**.** *Consider the dynamical system described by $\mathscr{S}$ in (2.11) satisfying Hypotheses 1–3. For any observables $\mathcal{O}_1, \mathcal{O}_2 \in \mathscr{B}_{\alpha_*,1}(\Omega,\mathbb{R})$, with α_* as in Theorem 2, and any $\lambda_*^{-1} > \max\{\lambda^{-\alpha_*}, 1-\Gamma\}$, one has*

$$|\nu_0(\mathcal{O}_1\,\mathcal{O}_2\circ\mathscr{S}^n) - \nu_0(\mathcal{O}_1)\,\nu_W(\mathcal{O}_2)| \le C_1\,|||\mathcal{O}_1|||_{\alpha_*,0}|||\mathcal{O}_2|||_{\alpha_*,1}\lambda_*^{-n},$$

where the constant C_1 depends only on λ_*.

2.4 Averaging

We are interested in the long time evolution generated by the map $\mathscr{S}$ in (2.11), when the component $\mathscr{S}_\varphi$ is close to the identity. To this aim, we consider the family of functions $F(\varphi,\psi) = \rho f(\varphi,\psi)$, with $f \in \mathscr{B}_{\alpha_0,6}(\mathbb{T}\times\mathbb{T}^2,\mathbb{R})$ and $\rho > 0$ a parameter, and study the behaviour of the map

$$\mathscr{S}(\varphi,\psi) = (\varphi + \rho f(\varphi,\psi), A_0\psi) \tag{2.25}$$

when ρ is small. We also define $\gamma := -\sup_{(\varphi,\psi)\in\Lambda}\partial_\varphi f(\varphi,\psi)$, so that $\Gamma = \rho\gamma$ in (2.12).

Remark 2.19. We may and do assume, without loss of generality, that $|||f|||_{\alpha_0,6} = 1$, and hence, since $\gamma < |||f|||_{\alpha_0,6}$, that $\gamma \in (0,1)$. Hypothesis 1 requires ρ not to be arbitrarily large. Therefore, when considering a map $\mathscr{S}$ of the form (2.25), we tacitly assume ρ to be smaller than a suitable value ρ_*, depending on f, such that $\mathscr{S}$ satisfies Hypotheses 1–3 for all $\rho \in (0,\rho_*)$.

We investigate the evolution generated by $\mathscr{S}$, as given in (2.25), when $\rho \to 0^+$. In such a limit the evolution tends to become trivial. For this reason, for given initial conditions (φ_0,ψ_0), one usually studies the dynamics after a linear *rescaling* of time, that is one considers the trajectory $\mathscr{S}^n(\varphi_0,\psi_0)$, with $n = \lfloor t/\rho \rfloor$, where t is fixed and ρ tends to 0^+. In the remaining part of this section we present results for the invariant manifold and the conjugation and discuss properties of $\mathscr{S}^n$ which are uniform in n, where ρ is small (see Subsections 2.4.1 to 2.4.4); finally, we interpret the results in terms of the scaling limit (see Subsection 2.4.5). Throughout the paper, we refer to the regime where ρ is small as the *scaling regime* even when we look for properties which do not involve the rescaled variable t.

Remark 2.20. Since $(\mathscr{S}^n)_\psi(\phi,\psi) = A_0^n\psi$ and the Lebesgue measure m_0 is the physical measure for A_0 (see (2.2)), it is natural, while discussing the scaling limit, to choose the initial conditions (φ,ψ) with φ given and ψ randomly distributed according to m_0 so that also $(\mathscr{S}^n)_\psi(\phi,\psi)$ is distributed according to m_0. For this reason Theorems 4 to 8 below are formulated in term of the measure m_0 on $\mathbb{T}^2$.

2.4.1 Synchronization in the scaling regime

The following two lemmas collect the implications of Theorems 1 and 2 and their proofs for the dynamical system $\mathscr{S}$ in (2.25) with ρ small – for the proofs see the end of Subsections 3.1 and 3.2.3, respectively.

Corollary 2.21. *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. Let $\mathcal{W} = \{(W(\psi), \psi) : \psi \in \mathbb{T}^2\}$ be the invariant manifold for $\mathscr{S}$ as in Theorem 1. One has $\alpha_+ = O(1)$ and $\alpha_- = O(\rho)$, and, furthermore, both $\|W\|_\infty$ and $|W|^-_{\alpha_-}$ are $O(1)$, while $|W|^+_{\alpha_+} = O(\rho)$.*

Corollary 2.22. *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. Let $\mathscr{H}$ be the conjugation in Remark 2.16 for $\mathscr{S}$ as in Theorem 2, and let h and l be defined as in* (2.22)*. One has $\alpha_* = O(\rho)$, while both $|||h|||_{\alpha_*,1}$ and $|||l|||_{\alpha_*,1}$ are $O(1)$ in ρ.*

Corollary 2.21 shows that, when ρ is small, the invariant manifold $\mathcal{W}$ loses most of the smoothness of f in the stable direction while maintaining it and varying slowly in the unstable one.

By using Corollary 2.21, the following corollary follows immediately from Theorem 3.

Corollary 2.23. *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. For any observables $\mathcal{O}_1, \mathcal{O}_2 \in \mathscr{B}_{\alpha_*,1}$, with α_* as in Theorem 2, and for all $n \in \mathbb{N}$, one has*

$$|\nu_0\left(\mathcal{O}_1\, \mathcal{O}_2 \circ \mathscr{S}^n\right) - \nu_0\left(\mathcal{O}_1\right) \nu_W\left(\mathcal{O}_2\right)| \leq C_2\, |||\mathcal{O}_1|||_{\alpha_*,0} |||\mathcal{O}_2|||_{\alpha_*,1} e^{-\xi\rho n},$$

where C_2 and ξ are suitable constants not depending on ρ.

2.4.2 The averaged map

As an intermediate step we consider also the dynamical system described by the *averaged map*

$$\overline{\mathscr{S}}(\varphi,\psi) := (\overline{G}(\varphi), A_0\psi)\,, \qquad \overline{G}(\varphi) := \varphi + \rho\overline{f}(\varphi), \qquad \overline{f}(\varphi) := \langle f\rangle(\varphi), \tag{2.26}$$

and its linearization

$$\overline{\mathscr{S}}_0(\varphi,\psi) := (\overline{\mu}\,\varphi, A_0\psi), \qquad \overline{\mu} := \partial_\varphi \overline{G}(0) = 1 + \rho\, \partial_\varphi \overline{f}(0). \tag{2.27}$$

Noting that $\overline{\Gamma} = \rho\,\overline{\gamma} := -\rho \sup_{\varphi\in\mathcal{U}} \partial_\varphi \overline{f}(\varphi) \geq \rho\,\gamma = \Gamma$, we see that $\overline{\mathscr{S}}$ satisfies Hypotheses 1–3 if $\mathscr{S}$ does.

Remark 2.24. Since $\overline{f}(0) = 0$ (see Remark 2.8), we see that $\overline{W}(\psi) = 0$ solves the equation $\overline{\mathscr{S}}(\overline{W}(\psi), \psi) = (\overline{W}(A_0\psi), A_0\psi)$, and hence $\overline{\mathscr{S}}$ admits the invariant manifold $\overline{\mathcal{W}} = \{(0,\psi), \psi \in \mathbb{T}^2\}$.

Remark 2.25. Although the action of $\overline{\mathscr{S}}$ on each variable is independent of the other one, since one has $(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) = \overline{G}^n(\varphi)$, the notation introduced in (2.26) helps clarify the forthcoming discussion.

In Subsections 2.4.3 and 2.4.4 we compare the evolution generated by $\mathscr{S}$ with the evolution generated by $\overline{\mathscr{S}}$. In the remainder of this subsection, we show that, by adapting the analysis in Subsection 2.3.2 to $\overline{\mathscr{S}}$, the trajectories of the dynamics generated by $\overline{\mathscr{S}}$ can be compared to the solutions $\phi(t)$ of the Cauchy problem

$$\begin{cases} \dot{\phi} = \overline{f}(\phi) \\ \phi(0) = \varphi\,. \end{cases} \tag{2.28}$$

First, we proceed as in Theorem 2, and look for an invertible function $\overline{\mathscr{H}}$ such that

$$\overline{\mathscr{H}} \circ \overline{\mathscr{S}} = \overline{\mathscr{S}}_0 \circ \overline{\mathscr{H}}\,, \tag{2.29}$$

that is a function $\overline{\mathscr{H}}$ that conjugates $\overline{\mathscr{S}}$ to $\overline{\mathscr{S}}_0$. We write

$$\overline{\mathscr{H}}(\varphi,\psi) = (\overline{\mathcal{H}}(\varphi),\psi), \qquad \overline{\mathcal{H}}(\varphi) = \varphi + \varphi^2\overline{h}(\varphi), \qquad \overline{\mathscr{H}}^{-1}(\eta,\psi) = (\eta + \eta^2\overline{l}(\eta),\psi). \tag{2.30}$$

The following result is proved in Subsection 4.5.

Lemma 2.26. *Consider the dynamical system described by $\mathscr{S}$ in* (2.11) *satisfying Hypotheses 1–3, and define $\overline{\mathscr{S}}$ as in* (2.26)*. There exist a closed interval $\mathcal{U}_0 \subset \mathbb{T}$ and a function $\overline{\mathscr{H}} : \mathcal{U} \times \mathbb{T}^2 \to \mathcal{U}_0 \times \mathbb{T}^2$ such that* (2.29) *holds. Moreover, there exists a constant C_3 such that one has $|||\overline{h}|||_{0,3} \leq C_3$ and $|||\overline{l}|||_{0,3} \leq C_3$.*

Remark 2.27. Note that $\Omega_0 \neq \mathcal{U}_0 \times \mathbb{T}^2$: in fact, one has $\Omega_0 = \mathscr{H}(\Omega)$, while $\mathcal{U}_0 = \overline{\mathcal{H}}(\mathcal{U})$.

Let Φ be the flow generated by (2.28), that is the set of the solutions of

$$\begin{cases} \dfrac{d}{dt}\Phi_t(\varphi) = \overline{f}(\Phi_t(\varphi)), \\ \Phi_0(\varphi) = \varphi, \end{cases} \tag{2.31}$$

when varying $\varphi \in \mathcal{U}$. Because of Hypotheses 1–3, all trajectories $\Phi_t(\varphi)$ of the system (2.31), with $\varphi \in \mathcal{U}$, fall inside $[S_m, S_M]$ after a while and hence, since the dynamical system (2.31) is a one-dimensional system and $\varphi\overline{f}(\varphi) > 0$ for $\varphi \in [S_m, S_M] \setminus \{0\}$, they move towards the origin at exponential rate (see for instance [32, exercise 4.40]).

The following result, proved in Subsection 4.6, shows that the trajectories generated by (2.26) and (2.31) remain close and, in fact, merge asymptotically.

Lemma 2.28. *Consider the dynamical system described by $\mathscr{S}$ in (2.25) satisfying Hypotheses 1–3, and define the map $\overline{\mathscr{S}}$ as in (2.26) and the flow Φ as in (2.31). For any $\gamma' \in (0, \gamma)$ there exists a positive constant C_4 such that for all $\varphi \in \mathcal{U}$ and all $n > 0$ one has*

$$|(\overline{\mathscr{S}}^n)_\varphi(\varphi, \psi) - \Phi_{n\rho}(\varphi)| \leq C_4\rho\,(1 - \rho\gamma')^n. \tag{2.32}$$

Remark 2.29. With the notation established after Remark 2.19, when introducing the scaling terminology, we can write (2.32) as $|(\overline{\mathscr{S}}^{\lfloor t/\rho \rfloor})_\varphi(\varphi, \psi) - \Phi_t(\varphi)| \leq C_4'\rho\, e^{-\gamma' t}$, for a suitable constant C_4'.

Remark 2.30. The decay rate in Lemma 2.28 is not in general equal to γ as one could naïvely expect. Indeed, the constant C_4 in Lemma 2.28, as well as the constants C_6, C_8 and C_{10} in the forthcoming Lemma 2.31, Theorem 6 and Theorem 7, respectively, depend on γ' and may diverge as γ' tends to γ.

2.4.3 Oscillations and deviations in the scaling regime

In the next subsection we show that $(\mathscr{S}^n)_\varphi$ remains close to $(\overline{\mathscr{S}}^n)_\varphi$ and hence to $\Phi_{n\rho}$, in the sense that the first and second moments, with respect to ψ, of their difference are $O(\rho)$ uniformly in n. In this subsection we present several preparatory results that are of interest in their own. The proofs of these results form the main technical part of the present work and are reported in Section 4. The main tools used in the proofs are the decay of correlations estimates provided by Propositions 4.3 and 4.26.

We first show that, even though the oscillations of $\mathcal{W}$ around $\overline{\mathcal{W}}$ can be of order 1 in ρ, large oscillations are rare, in the sense of the following result, proved in Subsection 4.3.

Theorem 4 (Oscillations of the invariant manifold). *Consider the dynamical system described by $\mathscr{S}$ in (2.25) satisfying Hypotheses 1–3. Let $\mathcal{W} = \{(W(\psi), \psi) : \psi \in \mathbb{T}^2\}$ be the invariant manifold for the map $\mathscr{S}$ as in Theorem 1, and let $\langle\cdot\rangle$ denote the average with respect to the Lebesgue measure on $\mathbb{T}^2$, according to (2.1). There is a constant C_5 such that*

$$|\langle W\rangle| \leq C_5\rho, \qquad \langle W^2\rangle \leq C_5\rho. \tag{2.33}$$

From Theorem 4 and the standard Chebyshev inequality we obtain that, for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : |W(\psi)| > \delta\Big\}\Big) \leq \frac{C_5\rho}{\delta^2}, \tag{2.34}$$

where m_0 is the Lebesgue measure on $\mathbb{T}^2$. Therefore, the invariant manifold $\mathcal{W}$ for $\mathscr{S}$ converges in probability to the invariant manifold $\overline{\mathcal{W}}$ of $\overline{\mathscr{S}}$.

Next we compare the linearized maps $\mathscr{S}_0$ and $\overline{\mathscr{S}}_0$, defined in (2.20) and (2.27), respectively. Observe that $\overline{\mathscr{S}}_0^n(\varphi, \psi) = (\overline{\mu}^n\varphi, A_0^n\psi)$, while $\mathscr{S}_0^n(\varphi, \psi) = (\kappa^{(n)}(\psi)\varphi, A_0^n\psi)$, with $\kappa^{(n)}\psi)$ given by (2.24).

The following result, proved in Subsection 4.4, shows that the linearized maps $\mathscr{S}_0^n$ and $\overline{\mathscr{S}}_0^n$ stay close to each other uniformly in n; what makes the result not trivial is that the function κ in (2.24) is only in $\mathfrak{B}^*_{\alpha_-,\alpha_+}(\mathbb{T}^2, \mathbb{R})$, with $\alpha_- = O(\rho)$ and $|W|^-_{\alpha_-} = O(1)$ in ρ.

Lemma 2.31. *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. Let $\kappa^{(n)}(\psi)$ and $\overline{\mu}$ be defined in* (2.24) *and in* (2.27)*, respectively. Then for any $\gamma' \in (0,\gamma)$ there is a constant C_6 such that, for all $n \in \mathbb{N}$,*

$$\left|\langle \kappa^{(n)} - \overline{\mu}^n \rangle\right| \leq C_6 \rho\, (1-\rho\,\gamma')^n, \qquad \left|\langle \left(\kappa^{(n)} - \overline{\mu}^n\right)^2 \rangle\right| \leq C_6 \rho\, (1-\rho\gamma')^{2n}. \tag{2.35}$$

Finally, we estimate the deviation of $\mathcal{H}$ from $\overline{\mathcal{H}}$. To this end, it is enough to study the deviations of h from $\overline{h}$. The following result is proved in Subsection 4.7.

Theorem 5 (**Deviations of the conjugation**)**.** *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. Let h and $\overline{h}$ be defined as in* (2.22) *and in* (2.30)*, respectively. There is a constant C_7 such that, for all $\varphi \in \mathcal{U}$,*

$$\left|\left\langle h(\varphi,\cdot) - \overline{h}(\varphi)\right\rangle\right| \leq C_7\rho, \qquad \left\langle (h(\varphi,\cdot) - \overline{h}(\varphi))^2\right\rangle \leq C_7\rho, \tag{2.36a}$$

$$\left|\left\langle \partial_\varphi h(\varphi,\cdot) - \partial_\varphi \overline{h}(\varphi)\right\rangle\right| \leq C_7\rho, \qquad \left\langle (\partial_\varphi h(\varphi,\cdot) - \partial_\varphi \overline{h}(\varphi))^2\right\rangle \leq C_7\rho. \tag{2.36b}$$

Bounds analogous to (2.36) hold for the deviations of l from $\bar{l}$, with l and $\bar{l}$ also defined in (2.22) and (2.30), respectively; we refer to Subsection 4.7 – to Proposition 4.51 in particular – for a precise statement, which requires introducing a suitable extension of the map $\mathscr{S}$ along the lines outlined in Remark 2.10. Note that it is in order to prove the bounds in Theorem 5 and in Proposition 4.51 that we need F to be in $\mathscr{B}_{\alpha_0,6}(\Omega,\mathbb{R})$ (see also Remark 4.46).

2.4.4 Summing up: convergence in square mean and in probability

We can now complete the comparison of the evolution generated by $\mathscr{S}$ with that generated by $\overline{\mathscr{S}}$. From (2.21) and (2.29) we get

$$\left(\mathscr{S}^n\right)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) = \mathcal{H}^{-1}\left(\mathscr{S}_0^n(\mathcal{H}(\varphi,\psi),\psi)\right) - \overline{\mathcal{H}}^{-1}\left(\overline{\mathscr{S}}_0^n(\overline{\mathcal{H}}(\varphi),\psi)\right). \tag{2.37}$$

Combining the estimates in Lemma 2.31 with the bounds in Theorem 5 and the analogous bounds for the inverse conjugation in Proposition 4.51, the following result is proved in Subsection 4.9.

Theorem 6 (**Convergence in square mean**)**.** *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. For any $\gamma' \in (0,\gamma)$ there exists a constant C_8 such that, for all $n \in \mathbb{N}$ and all $\varphi \in \mathcal{U}$,*

$$\left|\left\langle (\mathscr{S}^n)_\varphi(\varphi,\cdot) - \left((\overline{\mathscr{S}}^n)_\varphi(\varphi,\cdot) + W(A_0^n\cdot)\right)\right\rangle\right| \leq C_8\rho\,(1-\rho\,\gamma')^n, \tag{2.38a}$$

$$\left\langle \left((\mathscr{S}^n)_\varphi(\varphi,\cdot) - \left((\overline{\mathscr{S}}^n)_\varphi(\varphi,\cdot) + W(A_0^n\cdot)\right)\right)^2\right\rangle \leq C_8\rho\,(1-\rho\,\gamma')^{2n}, \tag{2.38b}$$

where $\langle\cdot\rangle$ denotes the average with respect to the Lebesgue measure on $\mathbb{T}^2$.

Remark 2.32. The bound (2.38b) for the second moment of the fluctuations and the Cauchy-Schwartz inequality trivially imply a weaker bound than (2.38a). By contrast, proving that also the first moment of the fluctuations is of order ρ requires a substantially greater amount of work. Apart of its intrinsic interest, the stronger bound (2.38a) plays a crucial role in view of some of the extensions proposed in Section 2.6. A similar comment holds for the results in Theorems 4 and 5, and in Lemma 2.31 as well.

The following result is an immediate consequence of Theorems 4 and 6, and of Lemma 2.28.

Corollary 2.33. *Under the hypotheses of Theorem 6, there exists a constant C_9 such that, for all $n \in \mathbb{N}$ and all $\varphi \in \mathcal{U}$, one has*

$$\left|\left\langle (\mathscr{S}^n)_\varphi(\varphi,\cdot) - \Phi_{n\rho}(\varphi)\right\rangle\right| \leq C_9\rho, \qquad \left\langle \left((\mathscr{S}^n)_\varphi(\varphi,\cdot) - \Phi_{n\rho}(\varphi)\right)^2\right\rangle \leq C_9\rho, \tag{2.39}$$

with the flow Φ defined as in (2.31)*.*

From Theorem 6, for fixed φ and n, we obtain that, for any $\gamma' \in (0,\gamma)$ and for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}\big|(\mathscr{S}^n)_\varphi(\varphi,\psi) - \big(\Phi_{n\rho}(\varphi) + W(A_0^n\psi)\big)\big| > \delta\Big\}\Big) \le \frac{C_9\rho}{\delta^2}, \tag{2.40}$$

where m_0 is the Lebesgue measure on $\mathbb{T}^2$ and C_9 is the same constant as in (2.39).

Note that the set of angles ψ considered in (2.40) depends on n, even though its measure is bounded independently of the value of n. In fact, the result may be improved by showing that, for most values of ψ, for all $n \in \mathbb{N}$ the difference between $(\mathscr{S}^n)_\varphi(\varphi,\psi)$ and $\Phi_{n\rho}(\varphi) + W(A_0^n\psi)$ is small, exponentially in n. The idea is essentially the following. All trajectories, for both the map $\mathscr{S}$ and the flow Φ, come closer and closer to the attractor, so that, in fact, it is enough to control the sets of angles appearing in (2.40) only for n of order ρ^{-1}. Moreover, since ρ is very small, when a large deviation appears, it persists for a long time. Thus, only a small number of such deviations may occur for n of order ρ^{-1}. This leads to a bound of the same kind as (2.40), but uniformly in n, only at the cost of an additional factor δ^{-1}. A formal statement of the argument above is provided by the following theorem, proved in Subsection 5.1.

Theorem 7 (**Convergence in probability**). *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. Let the flow Φ be defined as in* (2.31)*. For any $\gamma' \in (0,\gamma)$ there exists a constant C_{10} such that, for all $\varphi \in \mathcal{U}$,*

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : \sup_{n\ge 0}(1-\rho\,\gamma')^{-n}\big|(\mathscr{S}^n)_\varphi(\varphi,\psi) - \big(\Phi_{n\rho}(\varphi) + W(A_0^n\psi)\big)\big| > \delta\Big\}\Big) \le \frac{C_{10}\rho}{\delta^3}, \tag{2.41}$$

where m_0 is the Lebegsue measure on $\mathbb{T}^2$.

In (2.41) we can chose δ as a function of ρ in such a way that both δ and ρ/δ^3 tend to 0 as ρ tends to 0 (a natural choice is $\delta = \rho^{1/4}$ so that also $\rho/\delta^3 = \rho^{1/4}$). Thus, Theorem 7 implies that, up to a set of initial conditions for ψ whose measure vanishes with ρ, the difference between the actual evolution of the neutral variable toward the invariant manifold and its approximate evolution as described by the averaged differential equation vanishes with ρ as well. Moreover, such a difference remains uniformly bounded in time for arbitrarily long times.

2.4.5 Aftermath: the scaling limit

We now give a more probabilistic description of the results in Theorem 7 on the relations between $\mathscr{S}^{\lfloor t/\rho\rfloor}$ and the solutions of (2.28) and express them in terms of the rescaled time t: this provides a rigorous description of what we called the scaling limit in the introduction. To this aim, we fix $\varphi \in \mathbb{T}$ once and for all, and we consider the function $X\colon \mathbb{T}^2 \times \mathbb{R}^+ \to \mathbb{T}$ defined as

$$X_t(\psi) := (\mathscr{S}^{\lfloor t/\rho\rfloor})_\varphi(\varphi,\psi) + (t/\rho - \lfloor t/\rho\rfloor)\big((\mathscr{S}^{\lfloor t/\rho\rfloor+1})_\varphi(\varphi,\psi) - (\mathscr{S}^{\lfloor t/\rho\rfloor})_\varphi(\varphi,\psi)\big). \tag{2.42}$$

For every fixed t, X_t is a random variable on the probability space $(\mathbb{T}^2, B, m_0)$, where B is the standard Borel σ-algebra on $\mathbb{T}^2$ and m_0 is the Lebesgue measure. Moreover, for any ψ, $X_t(\psi)$ is a continuous function of t. Thus, X_t can be seen as a stochastic process on the probability space $(\mathbb{T}^2, B, m_0)$ with trajectories in $C^0(\mathbb{R}^+,\mathbb{T})$. Together with X_t we consider the process $\mathcal{X}_t$, which also has trajectories in $C^0(\mathbb{R}^+,\mathbb{T})$, given by

$$\mathcal{X}_t(\psi) = X_t(\psi) - W(A_0^{\lfloor t/\rho\rfloor}\psi) - (t/\rho - \lfloor t/\rho\rfloor)\big(W(A_0^{\lfloor t/\rho\rfloor+1}\psi) - W(A_0^{\lfloor t/\rho\rfloor}\psi)\big). \tag{2.43}$$

We want to compare the stochastic processes X_t and $\mathcal{X}_t$ with the flow Φ_t defined in (2.31), seen as a stochastic process on $C^0(\mathbb{R}^+,\mathbb{T})$, in the limit $\rho \to 0^+$. Note that, since t is such that $n = \lfloor t/\rho\rfloor$, this means that we are considering what happens in the scaling limit, that is the limit of infinite iterations n of the map and vanishing interaction strength ρ such that the product $n\rho$ is kept finite (we refer to [23, 46] for a more detailed discussion of the scaling limit). To this aim we introduce the norm $\|x\|_\xi^{\exp} := \sup_{t\ge0} e^{\xi t}|x(t)|$, with $\xi \ge 0$. Combining Lemma 2.28 and Theorem 7 gives the following result, proved in Subsection 5.2.

Theorem 8 (**Scaling limit**). *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. Let $\mathcal{X}_t$ and Φ_t the stochastic process* (2.43) *and the flow defined in* (2.31)*, respectively. One has $\lim_{\rho\to 0^+} \mathcal{X}_t = \Phi_t$, where the limit is taken in probability in the topology on $C^0(\mathbb{R}^+, \mathbb{T})$ generated by the norm $\|\cdot\|_\xi^{\exp}$ for any $\xi \in [0, \gamma)$.*

We close our results with an immediate consequence of Theorem 8.

Corollary 2.34. *Assume the hypotheses of Theorem 8 and let X_t be the stochastic process defined in* (2.42)*. One has $\lim_{\rho\to 0^+} X_t = \Phi_t$, where the limit is taken in probability in the topology of the uniform convergence in $C^0(\mathbb{R}^+, \mathbb{T})$.*

2.5 Relation with previous works

Our work is strongly related, as pointed out in Section 1, to the analysis in [22, 24, 23, 46, 16]. In these works a family of one-parameter maps $\mathscr{T}$ on $\mathbb{T} \times \mathbb{T}$ is considered, given by

$$\mathscr{T}(\varphi, \omega) := (\varphi + \rho f(\varphi, \omega), h(\varphi, \omega)), \tag{2.44}$$

where $h(\varphi, \cdot)$ is an expanding circle map for every fixed φ. Systems of the form (2.44), notwithstanding the fact that they are not time-reversible and have no Hamiltonian structure, have been extensively investigated since a full understanding of their behaviour may be considered as an important step toward the study of more realistic models.

Although the hypotheses on h and f in [22, 23] are much weaker then ours, it is interesting to compare their results with ours. For this purpose, we consider the dynamical system $\mathscr{S}_b$ of the form (2.11) but with the baker transformation in place of the automorphism A_0, that is

$$\mathscr{S}_b(\varphi, \psi) := (\varphi + \rho f(\varphi, \psi), B(\psi)), \qquad B(\psi) := \left(2\psi_1 \bmod 2\pi, \frac{\psi_2 + \lfloor 2\psi_1 \rfloor}{2}\right),$$

with $\psi = (\psi_1, \psi_2)$ and $\lfloor\cdot\rfloor$ denoting the lower integer part, and assume that $\mathscr{S}_b$ satisfies Hypotheses 1–3. If $f(\varphi, \psi)$ does not depend on ψ_2, the system can be seen as an extension of the system with $\mathscr{S}_b$ replaced with $\mathscr{T}_b(\varphi, \psi_1) := (\varphi + \rho f(\varphi, \psi), 2\psi_1 \bmod 2\pi)$, that is part of the family of systems studied in [22, 23]. On the other hand, we expect that most of the results in the present paper apply with minor modifications to $\mathscr{S}_b$ and hence that (1) there exists an invariant manifold $\mathcal{W}_b := \{(\psi, W_b(\psi)) : \psi \in \mathbb{T}^2\}$, with $W_b : \mathbb{T}^2 \to \mathbb{T}$ Hölder continuous such that $\mathscr{S}_b(W_b(\psi), \psi) = W_b(B(\psi))$; (2) $\mathcal{W}_b$ is an attractor for $\mathscr{S}_b$, (3) $\mathscr{S}_b$ admits a unique physical measure ν_b such that

$$\nu_b(\mathcal{O}) = \int \mathcal{O}(W_b(\psi), \psi)\, m_0(\mathrm{d}\psi) = \int \mathcal{O}(\varphi, \psi)\, \delta(\varphi - W_b(\psi))\, \nu_0(\mathrm{d}\varphi \mathrm{d}\,\psi),$$

with ν_0 the Lebesgue measure on $\mathbb{T} \times \mathbb{T}^2$. In particular, when $f(\varphi, \psi)$ does not depend on ψ_2, if we take $\mathcal{O}$ independent of ψ_2 as well, we can write

$$\nu_b(\mathcal{O}) = \int \mathcal{O}(\varphi, \psi_1)\, \bar{\nu}_b(\mathrm{d}\varphi\, \mathrm{d}\psi_1),$$

where $\bar{\nu}_b$ is the projection of ν_b along the ψ_2-direction and is the unique physical measure for $\mathscr{T}_b$. In [16] the authors show that, under very general conditions, $\bar{\nu}_b$ is absolutely continuous w.r.t. the Lebesgue measure $\bar{\nu}_0 = \mathrm{d}\psi_1\, \mathrm{d}\varphi$, that is $\nu_b(\mathrm{d}\varphi\, \mathrm{d}\psi_1) = n_b(\varphi, \psi_1)\mathrm{d}\varphi\, \mathrm{d}\psi_1$, where

$$n_b(\varphi, \psi_1) = \int \delta(\varphi - W_b(\psi))\, \mathrm{d}\psi_2$$

is a well-defined integrable function. We think that, in our case, this should follow from the Hölder continuity of $W_b(\psi)$ and the fact that it varies rapidly in the ψ_2 direction; a formal derivation of this property is beyond the scope of the present paper.

Moreover, an adaptation of Corollary 2.23 to the context under consideration should imply that, for every observable $\mathcal{O}_1$ and $\mathcal{O}_2$ on $\mathbb{T} \times \mathbb{T}$ that are C^1 in φ and Hölder continuous in ψ_1, we have

$$\left| \int \mathcal{O}_1(\varphi, \psi_1) \mathcal{O}_2(\mathscr{T}_b^n(\varphi, \psi_1))\, \mathrm{d}\varphi\, \mathrm{d}\psi_1 - \int \mathcal{O}_1(\varphi, \psi_1)\, \mathrm{d}\varphi\, \mathrm{d}\psi_1 \int \mathcal{O}_2(\varphi, \psi_1)\, \bar{\nu}_0(\mathrm{d}\varphi\, \mathrm{d}\psi_1) \right| \le C e^{-c\rho n},$$

for suitable constants C and c, not depending on ρ. Similarly, calling

$$\overline{X}_t(\psi_1) := (\mathscr{T}_b^{\lfloor t/\rho \rfloor})_\varphi(\varphi_0, \psi_1) + (t/\rho - \lfloor t/\rho \rfloor) \Big((\mathscr{T}_b^{\lfloor t/\rho \rfloor + 1})_\varphi(\varphi_0, \psi_1) - (\mathscr{T}_b^{\lfloor t/\rho \rfloor})_\varphi(\varphi_0, \psi_1) \Big),$$

a result analogous to Corollary 2.34 should imply that $\overline{X}_t$ converges in probability to Φ_t in the topology of uniform convergence in $C^0(\mathbb{R}^+, \mathbb{T})$. Thus, the presence of uniform contraction near $\mathcal{W}_b$ should allow to control the dynamics for all positive times and hence to obtain stronger bounds for the decay of correlations with respect to the models considered in [23] (in particular, see [23, Section 3.1]).

2.6 Possible extensions and generalizations

Systems of the form (2.11) are called *skew products* since the fast variable does not depend on the slow variable. Skew product systems are widely studied in the literature, with a view to focus on the major features of the equations and discard the more technical aspects. For averaging in fast-slow systems where the base is an expanding map we may quote [44], beyond the illustrative examples proposed in [22, 23]; a different approach, besides averaging, consists in direct comparison with a dynamical system which is random in the fast variable – see for instance [6, 34]. In this subsection we present a heuristic description of how we think our results might be generalized to the case of a fully coupled system, in the framework of averaging.

As a first step one may consider systems of the form

$$\mathscr{S}(\varphi, \psi) = (\varphi + F_\varphi(\varphi, \psi), \mathcal{A}(\psi)), \qquad \mathcal{A}_0(\psi) := A_0\psi + F_\psi(\psi), \tag{2.45}$$

with $F_\psi \colon \mathbb{T}^2 \to \mathbb{T}^2$ and $F_\varphi(\varphi, \psi) \colon \mathbb{T} \times \mathbb{T}^2 \to \mathbb{T}$ such that $\mathcal{A}_0(\psi)$ describes an Anosov diffeomorphism on $\mathbb{T}^2$ and $\mathscr{S}$ still satisfies Hypotheses 1–3. However, since any Anosov diffeomorphism of the form in (2.45) is conjugated with its linear part A_0 (see also [10, 28, 51] for a more general context), there exists a Hölder continuous map $\mathcal{H}' \colon \mathbb{T}^2 \to \mathbb{T}^2$ such that $\mathcal{H}' \circ \mathcal{A}_0 = A_0 \circ \mathcal{H}'$, so that, setting $\mathscr{H}'(\varphi, \psi) := (\varphi, \mathcal{H}'(\psi))$, we may use $\mathscr{H}'$ to conjugate $\mathscr{S}$ in (2.45) with

$$\mathscr{S}'(\varphi, \psi) = (\varphi + F'(\varphi, \psi), A_0\psi), \tag{2.46}$$

with $F'(\varphi, \mathcal{H}'(\psi)) = F_\varphi(\varphi, \psi)$. Clearly $\mathscr{S}'$ is of the form (2.11) and satisfies Hypotheses 1–3. In this situation it is still natural to choose the initial ψ distributed according to the Lebesgue measure m_0 on $\mathbb{T}^2$. This is essentially equivalent to considering the SRB measure $m_{\mathcal{A}_0}$ associated with $\mathcal{A}_0$, since $(\mathcal{A}_0^*)^n m_0$ converges exponentially fast to $m_{\mathcal{A}_0}$ [29]. Thus, for the results in Section 2.4 – in particular Theorems 7 and 8 on the scaling limit – to extend to the map $\mathscr{S}$ in (2.45), one needs estimates on the decay of correlations like those in Proposition 2.4 but with $(\mathcal{H}')^* m_{\mathcal{A}_0}$ in place of m_0. Observe that $\mathcal{H}'$ is Hölder continuous and $(\mathcal{H}')^* m_{\mathcal{A}_0}$ is invariant under the action of A_0; thus $m_{\mathcal{A}_0}$ can be represented as a Gibbs state on the same subshift of finite type used for A_0 in Appendix A.1. One thus need to extend the proof Proposition 2.4, using, for instance, the properties of the potential that generates such a Gibbs state as discussed in [29] (or, alternatively, the anisotropic space-based transfer operator approach built on Ruelle's original thermodynamic formalism [21]).

The next step is to look at systems with bidirectional perturbations of the form

$$\mathscr{S}(\varphi, \psi) = (\varphi + F_\varphi(\varphi, \psi), \mathcal{A}(\psi; \varphi)), \qquad \mathcal{A}(\psi; \varphi) := A_0\psi + F_\psi(\varphi, \psi)), \tag{2.47}$$

with $F_\varphi(\varphi, \psi)$ satisfying once more Hypotheses 1–3 and $F_\psi(\varphi, \psi)$ such that the dynamical system on $\mathbb{T}^2$ generated by the map $\mathcal{A}(\cdot; \varphi)$ is an Anosov diffeomorphism for any fixed $\varphi \in [\phi_m, \phi_M]$. An invariant manifold is proved to exist for the system (2.47) in [18] in the perturbative regime under stronger conditions. More generally, one looks for a conjugation $\mathscr{H}'(\varphi, \psi) = (\varphi, \mathcal{H}'(\varphi, \psi))$ such that $\mathscr{H}' \circ \mathscr{S} = \mathscr{S}' \circ \mathscr{H}'$, with $\mathscr{S}'$ of the form (2.46). This means that one has

$$\mathcal{H}'(\varphi + F_\varphi(\varphi, \psi), \mathcal{A}(\psi; \varphi)) = A_0 \mathcal{H}'(\varphi, \psi), \qquad F'(\varphi, \mathcal{H}'(\varphi, \psi)) = F(\varphi, \psi). \tag{2.48}$$

It is possible to write a formal solution to (2.48) (see also [12] for a similar argument). Applying the conjugation $\mathscr{H}'$ to the system (2.47) allows us to reduce it to a system of the form of (2.11). To apply the results of the present paper one then needs to prove that the resulting system has the geometric and regularity properties needed to satisfy Hypotheses 1–3. The above construction, assuming it is successful, allows one to extend the results on synchronization to the systems as in (2.47). Then, one has to show that the averaging principle proved for the system (2.11) implies, thanks to the existence of $\mathscr{H}'$, an averaging principle for (2.47), more precisely that, writing $F_\varphi(\varphi, \psi) = \rho\, f_\varphi(\varphi, \psi)$ and starting the evolution at φ_0, with ψ distributed according to m_0, the correct scaling limit for the evolution generated by (2.47) is described by the flow Φ'_t defined by

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\Phi'_t(\varphi_0) = f'(\Phi'_t(\varphi_0)), \qquad\qquad f'(\varphi) := \displaystyle\int_{\mathbb{T}^2} f(\varphi, \psi)\, m_\varphi(\mathrm{d}\psi), \\ \Phi'_0(\varphi) = \varphi_0, \end{cases} \tag{2.49}$$

where the new measure m_φ can be computed from $\mathcal{H}'$ and, as heuristic arguments suggest, it is expected to be the SRB measure of the Anosov diffeomorphism $\mathcal{A}(\cdot; \varphi)$.

A third possible – and harder – generalization to investigate, already in the skew product case, is obtained by weakening the hypotheses on the dissipative nature of the map $\mathscr{S}$, for instance by allowing more general interactions where the rate of contraction of the neutral variable is non-zero only in average, as in [22, 23], where expanding maps are considered instead of Anosov automorphisms. For such a kind of interactions, *a fortiori* also in the case of systems of the form (2.47) new phenomena with respect to those found in the present paper are expected to occur, such as non-uniform hyperbolicity, positive Lyapunov exponents in the central direction, non-absolutely integrable central direction and metastability. Very interesting, from a physical point of view, is the conservative case, where $\langle f(\cdot, \varphi)\rangle$ is assumed to vanish for all φ. In such a situation, on the basis of the results available for related models [19, 27, 39, 49], one expects the scaling to be $k = \lfloor t/\rho^2 \rfloor$ and to lead in the limit to a stochastic differential equation, instead of an ordinary differential equation as in the dissipative case.

Finally, it would be of interest to find a more detailed description of the fluctuations of the process X_t in (2.42) around the flow Φ_t in (2.31). By comparison with the results available in the literature [41, 26, 27, 24, 22, 23], we expect the stochastic process $\Delta_t := (X_t - \Phi_t)/\sqrt{\rho}$ to converge in distribution, as $\rho \to 0^+$, to the solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}\Delta_t = \partial_\varphi \overline{f}(\Phi_t(\varphi_0))\Delta_t \mathrm{d}t + \sigma(\Phi_t(\varphi_0))\, \mathrm{d}B(t), \\ \Delta_0 = 0, \end{cases} \tag{2.50}$$

where $B(t)$ is a standard Brownian motion and $(\sigma(\varphi))^2 := \big\langle \sum_{n=-\infty}^{\infty} \tilde{f}(\varphi, \cdot)\, \tilde{f}(\varphi, A_0^n \cdot)\big\rangle$, with $\tilde{f}$ defined according to (2.1). Such a result is known to hold with $\sigma > 0$ in the case in which the chaotic system is an expanding map, if one requires the function $\varphi \mapsto F(\varphi, \psi,)$ not to be cohomologous to a constant function for any ψ [22, 23]. Since for conservative interactions the scaling limit is expected to lead to a stochastic differential equation, studying how an equation like (2.50) emerges from (2.11) in the scaling regime can be seen as a precursory step before dealing with the more demanding scaling regime for conservative systems. The fact that Δ_t converges to the solution of equation (2.50) implies that $W/\sqrt{\rho}$, with W (see Subsection 2.3.1) seen as a random variable on $\mathbb{T}^2$, converges in distribution to a normal random variable with mean 0 and standard deviation $\sigma(0)$, that is

$$\lim_{\rho \to 0^+} m_0\left(\{\psi \in \mathbb{T}^2 : W(\psi) \le \sqrt{\rho} z\}\right) = \frac{1}{\sqrt{2\pi\sigma(0)^2}} \int_{-\infty}^{z} \mathrm{d}y\, e^{-\frac{y^2}{2\sigma(0)^2}}. \tag{2.51}$$

To start with, as a consistency check we show how to derive (2.51). First of all, using the notation in Remark 2.9, with $F = \rho f$ and $\mu(\psi) := 1 - \nu(\psi)$, set

$$W_0 := \sum_{i=1}^{\infty}\left(\prod_{j=1}^{i-1} \mu \circ A_0^{-j}\right)\beta \circ A_0^{-i}, \qquad W_{00} := \sum_{i=1}^{\infty} \overline{\mu}^{i-1}\beta \circ A_0^{-i}.$$

We prove in Subsection 4.3 that $\langle |W - W_0|\rangle = O(\rho)$, which implies that $(W - W_0)/\sqrt{\rho}$ converges in probability to 0, so that we just need to prove that $W_0/\sqrt{\rho}$, in the limit, has the correct normal distribution. Then, a partial extension of Proposition 4.3 to multi-time correlation functions – whose proof we omit – shows that $\big\langle (W_0 - W_{00})^2 \big\rangle = O(\rho^2)$. On the other hand, the central limit theorem for Anosov system [17] implies that $W_{00}/\sqrt{\rho}$ tends to the correct normal limit as $\rho \to 0^+$. We expect a similar strategy to work for the derivation of (2.50), by proceeding along the lines of Subsection 4.7.

2.7 Content of the paper and strategy of the proof

The rest of the paper, devoted to the proof of the results stated in Subsections 2.3 and 2.4, is organised as follows.

Section 3 contains the proofs of the Theorems 1 to 3, together with the derivation of the properties of both the invariant manifold and the conjugation that are needed for dealing with the scaling regime. The existence of the invariant manifold is formulated as a fixed-point problem in a suitable Banach space (Theorem 1). Thereafter, the conjugation is shown to admit a series representation which is proved to converge to a function which satisfies the properties stated in Theorem 2. This result implies the existence of the inverse conjugation, which satisfies similar properties (Corollary 2.18), and the mixing properties of the system while evolving toward the invariant manifold (Theorem 3). Technically, it turns out to be useful to use the coordinates (θ, ψ), with $\theta := \varphi - W(\psi)$ and $W(\psi)$ as in Subsection 2.3.1, in terms of which the dynamics is described by a map $\mathscr{S}_1$ that we call the *translated map* and the attracting invariant manifold is the flat torus $\mathbb{T}^2$.

In Section 4 we discuss the averaging problem in the scaling regime: setting $F(\varphi, \theta) = \rho f(\varphi, \theta)$, the dynamics of $\mathscr{S}^k(\varphi, \psi)$ is studied for $k = \lfloor t/\rho \rfloor$, with t fixed and $\rho \to 0^+$. After deriving the deviation laws for the invariant manifold (Theorem 4) and the conjugation (Theorem 5), we study the deviations of the dynamics from that of the averaged system (Theorem 6). The analysis is based on delicate correlation inequalities, which exploit the hyperbolicity of A_0 and make use of the map being weakly dissipative and the ensuing fact that any trajectory after a while comes close enough to the attracting invariant manifold. A main issue – and a major source of technical intricacies – is that the correlations inequalities are related to the regularity of the involved functions. In general, we have to deal with averages of the form $\langle g_+ g_- \circ \mathscr{S}_1^n \rangle$, with $g_- \in \mathscr{B}_\alpha^-(\Omega, \mathbb{R})$, for which we can expect decay properties analogous to those of Proposition 2.4. However, the invariant manifold and, hence, the map $\mathscr{S}_1$ are only α-Hölder continuous, with $\alpha = O(\rho)$, so that a naiv̈e generalization of Proposition 2.4 would provide unavailing bounds. If we expand the average $\langle g_+ g_- \circ \mathscr{S}_1^n \rangle$ to second order in W, only the first order contributions require a careful analysis, since the leading terms are regular and the second order terms can be dealt with by using directly the bounds on the variance provided by Theorem 4. Instead, in order to use the bounds on the average in Theorem 4 we need to study in detail the first order contributions. To this aim, we isolate the contributions which do not depend on the dynamics on the torus and show that the remaining contributions admit better dimensional bounds. To implement the scheme outlined above, we introduce a regularized version $\mathscr{S}_2$ of the translated map, that we call the *auxiliary map*, for which we can apply the correlation inequalities, and, in Subsection 4.7.4 and Appendix B, we compare the translated map with the auxiliary map through a series of technical lemmas which aim at extracting and studying the linear dependence on the function W. To do that, we need $\mathscr{S}_2$ to regularize $\mathscr{S}_1$ and, at the same time, still satisfy Hypotheses 1 to 3: achieving both goals leads to contributions which, albeit linear in W so that they be dealt with as outlined before, unfortunately contain an extra factor ρ^{-1}. However such contributions involve sums of terms in which there appear differences of functions W and the sums can be rearranged in such a way that the difference is shifted to more regular functions: this allows us to regain a further factor ρ so as to compensate the factor ρ^{-1}. To implement the idea described above we perform iterated expansions which make the analysis rather intricate: in fact, Subsection 4.7.4 and Appendix B constitute the most technical part of the paper.

In Section 5, we prove Theorems 7 and 8 on the asymptotic behavior of the stochastic process associated to the dynamics. Again, the crucial issue is that in average the deviations are small, and hence the evolution of the system is essentially determined by the averaged map.

Appendix A, besides reviewing some basic facts about Anosov automorphisms, contains the results on the decay of correlations for the evolution generated by A_0 on $\mathbb{T}^2$. Such results form the main toolkit we use in Section 4 to obtain more general correlation inequalities. Appendix B contains the proofs of the more technical results presented in Section 4.

3 Mapping to a simpler model

In this section we prove Theorems 1 and 2 by explicitly solving (2.19) and (2.21). The first one follows by relying on Banach fixed-point theorem, while the second one exploits the dynamics being uniformly contracting around the invariant manifold. Next, the two results together are showed to yield Theorem 3.

3.1 The invariant manifold: proof of Theorem 1

From (2.19) we get

$$W(A_0\psi) = G(W(\psi),\psi) = W(\psi) + F(W(\psi),\psi). \tag{3.1}$$

To show that a solution of (3.1) exists we follow a scheme which is reminiscent of the proof of Hartman-Grobman's theorem [54]: we introduce the map

$$\mathscr{E}[W](\psi) := G(W(A_0^{-1}\psi), A_0^{-1}\psi), \tag{3.2}$$

so that the invariant manifold is the solution of the fixed point equation $\mathscr{E}[W] = W$. If we define the set $B := \{W \in \mathfrak{B}(\mathbb{T}^2,\mathbb{R}) : S_m \le W \le S_M\}$, then $\mathscr{E}(B) \subset B$ by Hypotheses 1 and 2. Moreover

$$\|\mathscr{E}[W_1] - \mathscr{E}[W_2]\|_\infty = \sup_{\psi\in\mathbb{T}^2} \left|G(W_1(A_0^{-1}\psi), A_0^{-1}\psi) - G(W_2(A_0^{-1}\psi), A_0^{-1}\psi)\right| \le (1-\Gamma)\|W_1 - W_2\|_\infty,$$

that is $\mathscr{E}$ is a contraction on B, and thus, by the Banach fixed-point theorem, there is a unique $W \in B$ that satisfies (3.1). We pass to discuss the regularity of W. Since F is α_0-Hölder continuous, we get

$$\begin{aligned}|\mathscr{E}[W]|^+_{\alpha_+} \le \lambda^{-\alpha_+} \sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+}\Big(&|G(W(\psi+xv_+),\psi+xv_+) - G(W(\psi),\psi+xv_+)| \\ &+ |G(W(\psi),\psi+xv_+) - G(W(\psi),\psi)|\Big) \le \lambda^{-\alpha_+}\big((1-\Gamma)|W|^+_{\alpha_+} + |F|_{\alpha_+}\big),\end{aligned} \tag{3.3}$$

for $\alpha_+ \in (0,\alpha_0]$, and similarly we obtain, for $\alpha_- \in (0,\alpha_0]$,

$$|\mathscr{E}[W]|^-_{\alpha_-} \le \lambda^{\alpha_-}\big((1-\Gamma)|W|^-_{\alpha_-} + |F|_{\alpha_-}\big). \tag{3.4}$$

This implies that the set

$$B_{\alpha_-,\alpha_+} := \left\{ W \in \mathfrak{B}_{\alpha_-,\alpha_+}(\mathbb{T}^2,\mathbb{R}) : |W|^+_{\alpha_+} \le \frac{|F|_{\alpha_+}}{\lambda^{\alpha_+} - (1-\Gamma)} \text{ and } |W|^-_{\alpha_-} \le \frac{|F|_{\alpha_-}}{\lambda^{-\alpha_-} - (1-\Gamma)} \right\}$$

is invariant under $\mathscr{E}$. Thus, we need to show that $\mathscr{E}$ is a contraction on B_{α_-,α_+} for suitable α_- and α_+, in order to apply once more the Banach Fixed-Point Theorem. To this end, observe that, for any $W_1, W_2 \in B_{\alpha_-,\alpha_+}$ we bound $|\mathscr{E}[W_1] - \mathscr{E}[W_2]|^+_{\alpha_+}$ by

$$\begin{aligned}\lambda^{-\alpha_+} \sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+}\Big(&\Big|G(W_1(\psi),\psi) - G(W_2(\psi),\psi) - G(W_1(\psi+xv_+),\psi) + G(W_2(\psi+xv_+),\psi)\Big| \\ +&\Big|G(W_1(\psi+xv_+),\psi) - G(W_2(\psi+xv_+),\psi) - G(W_1(\psi+xv_+),\psi+xv_+) + G(W_2(\psi+xv_+),\psi+xv_+)\Big|\Big).\end{aligned}$$

Writing, in the second line of the last equation,

$$G(W_1(\psi_1),\psi_2) - G(W_2(\psi_1),\psi_2) = \int_0^1 \mathrm{d}t\, \partial_\varphi G(tW_1(\psi_1) + (1-t)W_2(\psi_1),\psi_2)(W_1(\psi_1) - W_2(\psi_1)),$$

with either $\psi_1 = \psi$ or $\psi_1 = \psi + xv_+$, we bound

$$\begin{aligned}\sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+}\Big|&G(W_1(\psi+xv_+),\psi) - G(W_2(\psi+xv_+),\psi) \\ &- G(W_1(\psi+xv_+),\psi+xv_+) + G(W_2(\psi+xv_+),\psi+xv_+)\Big| \le |\partial_\varphi G|_{\alpha_+}\|W_1 - W_2\|_\infty.\end{aligned} \tag{3.5}$$

Moreover, to deal with the first line of the equation, we use that, for any C^2 function h,

$$\begin{aligned} & h(x_1) - h(x_2) - h(x_3) + h(x_4) \\ & \quad = \frac{1}{2} \int_0^1 \mathrm{d}t \, \big(h'(tx_1 + (1-t)x_2) + h'(tx_3 + (1-t)x_4) \big) (x_1 - x_2 - x_3 + x_4) \\ & \quad + \frac{1}{2} (x_1 - x_2 + x_3 - x_4) \int_0^1 \mathrm{d}t \, (tx_1 + (1-t)x_2 - tx_3 - (1-t)x_4) \\ & \qquad \times \int_0^1 \mathrm{d}s \, h''\big(s(tx_1 + (1-t)x_2) + (1-s)(tx_3 + (1-t)x_4) \big), \end{aligned}$$

so that we may bound

$$\begin{aligned} & \sup_{\psi \in \mathbb{T}^2} \sup_{x \in \mathbb{R}} |x|^{-\alpha_+} \Big| G(W_1(\psi), \psi) - G(W_2(\psi), \psi) - G(W_1(\psi + xv_+), \psi) \\ & \quad + G(W_2(\psi + xv_+), \psi) \Big| \leq (1 - \Gamma) |W_1 - W_2|^+_{\alpha_+} + \|\partial^2_\varphi G\|_\infty \max_{i=1,2} |W_i|^+_{\alpha_+} \, \|W_1 - W_2\|_\infty. \end{aligned} \tag{3.6}$$

Collecting the bounds (3.5) and (3.6), we obtain

$$|\mathscr{E}[W_1] - \mathscr{E}[W_2]|^+_{\alpha_+} \leq \lambda^{-\alpha_+} \Big((1-\Gamma) |W_1 - W_2|^+_{\alpha_+} + \big(\|\partial^2_\varphi F\|_\infty \max_{i=1,2} |W_i|^+_{\alpha_+} + |\partial_\varphi F|^+_{\alpha_+} \big) \|W_1 - W_2\|_\infty \Big). \tag{3.7}$$

Analogously, we obtain

$$|\mathscr{E}[W_1] - \mathscr{E}[W_2]|^-_{\alpha_-} \leq \lambda^{\alpha_-} \Big((1-\Gamma) |W_1 - W_2|^-_{\alpha_-} + \big(\|\partial^2_\varphi F\|_\infty \max_{i=1,2} |W_i|^-_{\alpha_-} + |\partial_\varphi F|^-_{\alpha_-} \big) \|W_1 - W_2\|_\infty \Big). \tag{3.8}$$

Let now fix $\alpha_-, \alpha_+ \in (0, \alpha_0]$ so as to satisfy $\lambda^{\alpha_-}(1 - \Gamma) < 1$ and

$$\alpha_+ \lambda^{-\alpha_+} \bigg(\frac{\|\partial^2_\varphi F\|_\infty |F|_{\alpha_+}}{\lambda^{\alpha_+} - (1-\Gamma)} + |\partial_\varphi F|^+_{\alpha_+} \bigg) + \alpha_- \lambda^{\alpha_-} \bigg(\frac{\lambda^{\alpha_-} \|\partial^2_\varphi F\|_\infty |F|_{\alpha_-}}{1 - \lambda^{\alpha_-}(1-\Gamma)} + |\partial_\varphi F|^-_{\alpha_-} \bigg) < \Gamma. \tag{3.9}$$

For such α_- and α_+, combining the bounds (3.3), (3.4), (3.7) and (3.8), the map $\mathscr{E}$ turns out to be a contraction on B_{α_-,α_+}. This concludes the proof of Theorem 1.

Recalling that $S_m < 0 < S_M$, by Remark 2.8, the discussion above implies that the sequence $\mathscr{E}^n[0]$ converges to W in $\mathfrak{B}(\mathbb{T}^2, \mathbb{R})$. Moreover the inequalities above are satisfied for α_-, α_+ small enough; in particular, if ρ is small and $|||F|||_{\alpha_0,2} = O(\rho)$, then the condition $\lambda^{\alpha_-}(1 - \Gamma) < 1$ requires $\alpha_- = O(\rho)$, which inserted into (3.9) allows α_+ to be $O(1)$ in ρ.

In the scaling regime, where the map $\mathscr{S}$ is of the form (2.25) and hence F and its derivatives are proportional to ρ, the best we can say about the invariant manifold is that $W \in B \cap B_{\alpha_-,\alpha_+}$ and hence $|W| \leq \max\{|S_m|, S_M\}$, $\alpha_+ = O(1)$ and $\alpha_- = O(\rho)$, so that we have $\|W\|_\infty = O(1)$, $|W|^-_{\alpha_-} = O(1)$ and $|W|^+_{\alpha_+} = O(\rho)$. This proves Corollary 2.21.

Remark 3.1. If ψ_0 is a fixed point of A_0, that is $A_0\psi_0 = \psi_0$, from (3.1) we get $W(\psi_0) = S(\psi_0)$, so that, in the scaling regime, $W(\psi_0)$ does not depend on ρ. In a similar way, if ψ_k is a periodic point of period k, that is $A_0^k \psi_k = \psi_k$, then $W(\psi_k)$ is found between $\min_i S(A_0^i \psi_k)$ and $\max_i S(A_0^i \psi_k)$. This shows that, in the scaling regime, in general $\lim_{\rho \to 0^+} W(\psi) \neq 0$.

3.2 The conjugation: proof of Theorem 2

3.2.1 The translated map

We construct the conjugation $\mathscr{H}$ by linearizing the dynamics around the invariant manifold. We write $\varphi = \theta + W(\psi)$ so that, in terms of the variables (θ, ψ), the dynamics is described by the map

$$\mathscr{S}_1(\theta, \psi) = ((\mathscr{S}_1)_\theta(\theta, \psi), (\mathscr{S}_1)_\psi(\theta, \psi)) := (G_1(\theta, \psi), A_0\psi), \qquad G_1(\theta, \psi) := \theta + F_1(\theta, \psi), \tag{3.10}$$

with

$$F_1(\theta,\psi) := F(\theta+W(\psi),\psi) - F(W(\psi),\psi) = F(\theta+W(\psi),\psi) + W(\psi) - W(A_0\psi). \tag{3.11}$$

We call $\mathscr{S}_1$ the *translated map.* Setting $\Omega_1 := \{(\theta,\psi)\in\mathbb{T}\times\mathbb{T}^2 \,:\, \phi_m - W(\psi)\le\theta\le\phi_M - W(\psi)\}$ and noting that $\varphi\mapsto\theta$ defines a bijection between Ω and Ω_1, we see that $\mathscr{S}_1$ is injective from Ω_1 into itself.

Remark 3.2. One easily checks that $\mathscr{S}_1(0,\psi)=(0,A_0\psi)$, so that, expressed in terms of θ, the invariant manifold reduces to $\overline{\mathcal{W}}=\{(0,\psi):\psi\in\mathbb{T}^2\}$ (see Remark 2.24).

Remark 3.3. The iterations of $\mathscr{S}$ and $\mathscr{S}_1$ are such that $(\mathscr{S}^n)_\varphi(\varphi,\psi)=(\mathscr{S}_1^n)_\theta(\varphi-W(\psi),\psi)+W(A_0^n\psi)$ and $(\mathscr{S}_1^n)_\theta(\theta,\psi)=(\mathscr{S}^n)_\varphi(\theta+W(\psi),\psi)-W(A_0^n\psi)$, while $(\mathscr{S}^n)_\psi(\varphi,\psi)=(\mathscr{S}_1^n)_\psi(\varphi-W(\psi),\psi)=A_0^n\psi$.

In Subsections 3.2.2 and 3.2.3 we study the conjugation relation

$$\mathscr{H}_1\circ\mathscr{S}_1=\mathscr{S}_0\circ\mathscr{H}_1, \tag{3.12}$$

where $\mathscr{H}_1:\Omega_1\to\Omega_0$ is of the form $\mathscr{H}_1(\theta,\psi)=(\mathcal{H}_1(\theta,\psi),\psi)$. We can then write

$$\mathscr{H}(\varphi,\psi)=(\mathcal{H}_1(\varphi-W(\psi),\psi),\psi), \tag{3.13}$$

with $\mathscr{H}$ satisfying the conjugation relation (2.21). Thus, if the conjugation $\mathscr{H}_1$ exists and is invertible, the conjugation $\mathscr{H}$ exists and is invertible as well, and vice versa. Moreover, while $\mathscr{H}$ and $\mathscr{H}_1$ have a different domain (a rectangular strip for $\mathscr{H}$ and a deformed strip for $\mathscr{H}_1$), they have the same image Ω_0 as defined in Theorem 2. In analogy with (3.13), we write $\mathscr{H}_1^{-1}(\eta,\psi)=(\mathcal{L}_1(\eta,\psi),\psi)$.

Considering also Remark 2.16, we look for functions $\mathcal{H}_1:\Omega_1\to\mathbb{R}$ and $\mathcal{L}_1:\Omega_0\to\mathbb{R}$ of the form

$$\mathcal{H}_1(\theta,\psi)=\theta+\theta^2h_1(\theta,\psi), \qquad \mathcal{L}_1(\eta,\psi)=\eta+\eta^2l_1(\eta,\psi). \tag{3.14}$$

If we show that functions of the form (3.14) exist, with $h_1\in\mathscr{B}_{\alpha_*,1}(\Omega_1,\mathbb{R})$ and $l_1\in\mathscr{B}_{\alpha_*,1}(\Omega_0,\mathbb{R})$, for a suitable $\alpha_*\in(0,\alpha_0)$, then we can write the conjugation $\mathscr{H}$ and its inverse as in (2.22), with

$$h(\varphi,\psi)=h_1(\varphi-W(\psi),\psi), \qquad l(\eta,\psi)=l_1(\eta,\psi). \tag{3.15}$$

Remark 3.4. As a consequence of Theorem 1, there exists a closed interval $\Theta:=[\theta_-,\theta_+]$, such that $\Theta\times\mathbb{T}^2\subset\Omega_1$ and hence $\{(\varphi,\psi)\in\mathbb{T}\times\mathbb{T}^2 \,:\, \varphi-W(\psi)\in\Theta\}\subset\Omega$.

3.2.2 Dynamics toward the invariant manifold

From (3.12) we see that $\mathcal{H}_1$ satisfies the equation $\kappa(\psi)\mathcal{H}_1(\theta,\psi)=\mathcal{H}_1(\mathscr{S}_1(\theta,\psi))$. Using (3.14), we find $h_1(\theta,\psi)=(\kappa(\psi)\theta^2)^{-1}\big(G_1(\theta,\psi)+G_1(\theta,\psi)^2h_1(\mathscr{S}_1(\theta,\psi))-\kappa(\psi)\,\theta\big)$, so that, setting

$$p_1(\theta,\psi):=\frac{(G_1(\theta,\psi))^2}{\theta^2\kappa(\psi)}=\frac{1}{\kappa(\psi)}\left(1+\int_0^1 \mathrm{d}t\,\partial_\varphi F(t\theta+W(\psi),\psi)\right)^2, \tag{3.16a}$$

$$q_1(\theta,\psi):=\frac{G_1(\theta,\psi)-\theta\kappa(\psi)}{\theta^2\kappa(\psi)}=\frac{1}{\kappa(\psi)}\int_0^1 \mathrm{d}t\,(1-t)\,\partial_\varphi^2F(t\theta+W(\psi),\psi)\,, \tag{3.16b}$$

we obtain $h_1(\theta,\psi)=q_1(\theta,\psi)+p_1(\theta,\psi)\,h_1(\mathscr{S}_1(\theta,\psi))$, whose solution can be formally written as

$$h_1(\theta,\psi)=\sum_{n=1}^{\infty}p_1^{(n)}(\theta,\psi)\,q_1(\mathscr{S}_1^n(\theta,\psi)), \qquad p_1^{(n)}(\theta,\psi):=\prod_{i=0}^{n-1}p_1(\mathscr{S}_1^i(\theta,\psi)). \tag{3.17}$$

where, according to our conventions (see Remark 2.11),we have $p_1^{(0)}(\theta,\psi)=1$.

Remark 3.5. Both $p_1(\theta,\psi)$ and $q_1(\theta,\psi)$ are well defined at $\theta=0$. Using (3.1) and the definition of $\kappa(\psi)$ after (2.20), we find $\kappa(\psi)\,q_1(0,\psi)=\partial_\varphi^2F(0,\psi)$ and $p_1(0,\psi)=\kappa(\psi)$.

We start by studying the regularity properties of the functions p_1 and q_1. The following result is an easy consequence of the representation in (3.16) together with Lemma 2.7.

Lemma 3.6. *Assume $\mathscr{S}$ in (2.11) to satisfy Hypotheses 1–3. For any $\alpha \in (0, \min\{\alpha_-, \alpha_+\})$, with α_- and α_+ as in Theorem 1, and any $\Gamma' \in (0, \Gamma)$, one has $p_1, q_1 \in \mathscr{B}_{\alpha,5}(\Omega_1, \mathbb{R})$ and $p_1(\theta, \psi) \le 1 - \Gamma'$ as long as $|\theta| \le \theta_1$, with*

$$\theta_1 := \frac{\Gamma - \Gamma'}{2\|\kappa^{-1}\|_\infty (1 + \|\partial_\varphi F\|_\infty)\|\partial_\varphi^2 F\|_\infty}. \tag{3.18}$$

Proof. Observe that $p_1(0, \psi) = 1 + \partial_\varphi F(W(\psi), \psi) \le 1 - \Gamma$. Thus, using that, for any $\theta_1 > 0$ and all $\theta \in [-\theta_1, \theta_1]$, we have $|p_1(\theta, \psi) - p_1(0, \psi)| \le \|\partial_\theta p_1\|_\infty \theta_1$, if we bound $\partial_\theta p_1$ using (3.16a) and fix θ_1 as in (3.18), the bound on $p_1(\theta, \psi)$ for $|\theta| \le \theta_1$ follows immediately. The other bounds too are easily obtained by estimating the derivatives of p_1 and q_1. $\square$

Remark 3.7. In the scaling regime, where $\Gamma = \rho\,\gamma$, one has $|||p_1|||_{\alpha,5} = 1 + O(\rho)$ and $|||q_1|||_{\alpha,5} = O(\rho)$, while $|||\partial_\theta p_1|||_{0,4} = O(\rho)$ and $|p_1|_\alpha = O(\rho)$. Moreover, if in Lemma 3.6 one takes $\Gamma' = \rho\,\gamma'$ with $\gamma' \in (0, \gamma)$, then (3.18) implies that $\theta_1 = O(1)$.

Next, we study the regularity of the iterates of $\mathscr{S}_1$.

Lemma 3.8. *Assume $\mathscr{S}$ in (2.11) to satisfy Hypotheses 1–3. For any $\alpha \in (0, \min\{\alpha_-, \alpha_+\})$, with α_- and α_+ as in Theorem 1, and any $\Gamma' \in (0, \Gamma)$, and all $n \ge 0$, one has $(\mathscr{S}_1^n)_\theta \in \mathscr{B}_{\alpha,2}(\Omega_1, \mathbb{R})$ and*

$$|||\partial_\theta(\mathscr{S}_1^n)_\theta|||_{0,1} \le D_0\,(1 - \Gamma')^n, \qquad |(\mathscr{S}_1^n)_\theta|_\alpha \le D_0 \lambda^{\alpha n}, \qquad |\partial_\theta(\mathscr{S}_1^n)_\theta|_\alpha \le D_0 \lambda^{\alpha n},$$

with the constant D_0 depending on F and Γ' but not on n.

Proof. For a fixed $\Gamma' \in (0, \Gamma)$, let $r = r(\Gamma')$ and N_r be defined as in Lemma 2.7, and observe that $N_r = O((\Gamma')^{-1})$ in such a case. We have, in general, $\|\partial_\theta G_1\|_\infty \le 1 + \|\partial_\varphi F\|_\infty$, while for $n \ge N_r$ we may bound $\|(\partial_\theta G_1)\circ\mathscr{S}_1^n\|_\infty \le \|(1 + \partial_\varphi F)\circ\mathscr{S}^n\|_\infty \le (1 - \Gamma')$. Thus, noting that $(\mathscr{S}_1^n)_\theta = G_1 \circ \mathscr{S}_1^{n-1}$, we get

$$\begin{aligned} \|\partial_\theta(\mathscr{S}_1^n)_\theta\|_\infty &\le \prod_{i=0}^{n-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^i\|_\infty \le \left(\prod_{i=0}^{N_r-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^i\|_\infty\right)\left(\prod_{i=N_r}^{n-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^i\|_\infty\right) \\ &\le \big(1 + \|\partial_\varphi F\|_\infty\big)^{N_r} \big(1 - \Gamma'\big)^{n-N_r} \le c_1\,(1 - \Gamma')^n, \end{aligned} \tag{3.19}$$

with $c_1 := (\,(1+\|\partial_\varphi F\|_\infty)/(1-\Gamma'))^{N_r}$, and

$$|G_1\circ\mathscr{S}_1^{n-1}|_\alpha \le \|(\partial_\theta G_1)\circ\mathscr{S}_1^{n-1}\|_\infty |G_1\circ\mathscr{S}_1^{n-2}|_\alpha + \lambda^{\alpha(n-1)}|G_1|_\alpha,$$

where $|G_1|_\alpha \le \|\partial_\varphi F\|_\infty |W|_\alpha + |F|_\alpha$, so that, iterating, we obtain

$$|(\mathscr{S}_1^n)_\theta|_\alpha \le |G_1|_\alpha \sum_{i=1}^{n} \lambda^{\alpha(n-i)} \prod_{j=1}^{i-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^{n-j}\|_\infty \le C_0 |G_1|_\alpha\, \lambda^{\alpha n} \sum_{i=1}^{n} (1-\Gamma')^{i-1} \lambda^{-\alpha i} \le c_2 \lambda^{\alpha n}, \tag{3.20}$$

with $c_2 := c_1 |G_1|_\alpha\,(\Gamma')^{-1}$. Exploiting (3.19), we find that

$$\|\partial_\theta^2(\mathscr{S}_1^n)_\theta\|_\infty \le \sum_{i=0}^{n-1} \big\|(\partial_\theta^2 G_1)\circ\mathscr{S}_1^i\big\|_\infty \big\|\partial_\theta(\mathscr{S}_1^i)_\theta\big\|_\infty \prod_{\substack{j=0\\ j\ne i}}^{n-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^j\|_\infty \le c_3\,(1 - \Gamma')^n, \tag{3.21}$$

with $c_3 := c_1 \|\partial_\varphi^2 F\|_\infty\,(\Gamma')^{-1}$. Finally, noting that

$$|\partial_\theta(\mathscr{S}_1^n)_\theta|_\alpha \le \sum_{i=0}^{n-1} \Big(\big\|\partial_\theta^2 G_1\big\|_\infty \big|(\mathscr{S}_1^i)_\theta\big|_\alpha + \lambda^{\alpha i} |\partial_\theta G_1|_\alpha \Big) \prod_{\substack{j=1\\ j\ne i}}^{n-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^j\|_\infty,$$

and proceeding as done to get the bound (3.20), we obtain

$$|\partial_\theta(\mathscr{S}_1^n)_\theta|_\alpha \le c_4 \lambda^{\alpha n}, \tag{3.22}$$

with $c_4 := \big(c_2 \|\partial_\varphi^2 F\|_\infty + |\partial_\theta G_1|_\alpha\big)\,(\Gamma')^{-1}$. Collecting together the bounds (3.19)–(3.22), the assertion follows follow with $D_0 = \max\{c_1, c_2, c_3, c_4\}$. $\square$

Remark 3.9. Essentially, the proof of Lemma 3.8 shows that, apart a small number of iterates necessary for the trajectory to reach the set Λ_r (see Lemma 2.7), which are controlled by the constant c_1, depending on r but not on n, all the remaining iterates produce an exponentially decaying factor.

Remark 3.10. In the scaling regime, where $\Gamma = \rho\gamma$, if we choose $\Gamma' = \rho\gamma'$, with $\gamma' \in (0,\gamma)$, we find that the constant D_0 is $O(1)$ in ρ. Indeed both $|||F|||_{0,2}$ and $|F|_\alpha$, and hence $|G_1|_\alpha$ as well, are $O(\rho)$, while one has $N_r = O(\rho^{-1})$ by (2.13).

Remark 3.11. Since $(\mathscr{S}_1^n)_\theta(0,\psi) = 0$ for all $n \in \mathbb{Z}$, we bound also $|(\mathscr{S}_1^n)_\theta(\theta,\psi)| \le \|\partial_\theta(\mathscr{S}_1^n)_\theta\|_\infty|\theta|$ and hence $\|(\mathscr{S}_1^n)_\theta\|_\infty \le c_1(\max\{|\phi_m|,\phi_M\} + \max\{|S_m|, S_M\})(1-\Gamma')^n$.

3.2.3 Existence and regularity of the conjugation

We have all the ingredients to complete the proof of Theorem 2. Fix $\Gamma' \in (0,\Gamma)$ and set $r = r(\Gamma')$, $S_r := \max\{S_M, |S_m|\} + r$ and $M_r := \max\{2\,(\Gamma')^{-1}\log(S_r/\theta_1), 0\}$, with the notation of Lemma 2.7 and θ_1 as in (3.18). If $(\theta + W(\psi),\psi) \in \Lambda_r$ and $n > M_r$, then $|(\mathscr{S}_1^n(\theta,\psi))_\theta| \le \theta_1$ and hence, thanks to the property 3 in Lemma 2.7, $|(\mathscr{S}_1^n(\theta,\psi))_\theta| \le \theta_1$ for all $(\theta,\psi) \in \Omega_1$ and for all $n \ge M_r' := M_r + N_r$. Thus, for $n \ge M_r'$, we obtain $\|p_1^{(n)}\|_\infty \le \|p_1\|_\infty^{M_r'}(1-\Gamma')^{n-M_r'}$ which, inserted in (3.17), implies

$$\|h_1\|_\infty \le D_1 \sum_{n=0}^{\infty} (1-\Gamma')^n \|q_1\|_\infty \le D_1(\Gamma')^{-1}\|q_1\|_\infty, \qquad D_1 := \left(\frac{\|p_1\|_\infty}{1-\Gamma'}\right)^{M_r'}. \tag{3.23}$$

Furthermore, using Lemma 3.8, we see that

$$\left\|\partial_\theta p_1^{(n)}\right\|_\infty \le D_0 D_1 \|\partial_\theta p_1\|_\infty (1-\Gamma')^{n-1} \sum_{i=0}^{n-1} (1-\Gamma')^i \le D_0 D_1 (\Gamma')^{-1} \|\partial_\theta p_1\|_\infty (1-\Gamma')^n \,, \tag{3.24}$$

with D_0 as in Lemma 3.8 and D_1 as in (3.23), so that, summing over n, we get

$$\|\partial_\theta h_1\|_\infty \le \sum_{n=1}^{\infty} \left(\left\|\partial_\theta p_1^{(n)}\right\|_\infty \|q_1\|_\infty + \left\|p_1^{(n)}\right\|_\infty \|\partial_\theta(q_1 \circ \mathscr{S}_1^n)\|_\infty \right) \le D_2\,(\Gamma')^{-1}|||q|||_{0,1}\,, \tag{3.25}$$

for some constant D_2. Proceeding along the same lines, we obtain

$$\|\partial_\theta^2 h_1\|_\infty \le \sum_{n=1}^{\infty} \left(\left\|\partial_\theta^2 p_1^{(n)}\right\|_\infty \|q_1\|_\infty + \left\|\partial_\theta p_1^{(n)}\right\|_\infty \|\partial_\theta(q_1 \circ \mathscr{S}_1^n)\|_\infty + \left\|p_1^{(n)}\right\|_\infty \|\partial_\theta^2(q_1 \circ \mathscr{S}_1^n)\|_\infty \right),$$

where $\left\|\partial_\theta p_1^{(n)}\right\|_\infty$ can be bounded as in (3.24). Thus, noting that $\partial_\theta(q_1 \circ \mathscr{S}_1^n) = (\partial_\theta q_1 \circ \mathscr{S}_1^n)\,\partial_\theta \mathscr{S}_1^n$ and $\partial_\theta^2(q_1 \circ \mathscr{S}_1^n) = (\partial_\theta^2 q_1 \circ \mathscr{S}_1^n)\,(\partial_\theta \mathscr{S}_1^n)^2 + (\partial_\theta q_1 \circ \mathscr{S}_1^n)\,\partial_\theta^2 \mathscr{S}_1^n$, and writing $\partial_\theta^2 p_1^{(n)}$ according to (2.18b), we can use Lemma 3.8 to obtain

$$\|\partial_\theta^2 h_1\|_\infty \le D_3\,(\Gamma')^{-1}|||q|||_{0,2}\,, \tag{3.26}$$

for a suitable constant D_3. Finally, for any $\alpha \in (0, \min\{\alpha_-,\alpha_+\}]$, we get, again relying on Lemma 3.8,

$$\left|p_1^{(n)}\right|_\alpha \le D_0 D_1 (|p_1|_\alpha + \|\partial_\theta p_1\|_\infty)\,(1-\Gamma')^{n-1}\,\frac{\lambda^{\alpha n} - 1}{\lambda^\alpha - 1} \le D_4\,\lambda^{\alpha n}\,(1-\Gamma')^n \,,$$

for a some constant D_4 proportional to D_1. To sum over n we take $\alpha = \alpha_*$ in Lemma 3.8, with α_* such that $\lambda^{\alpha_*}(1-\Gamma') < 1-\Gamma''$, with $\Gamma'' < \Gamma'$, so as to obtain

$$|h_1|_{\alpha_*} \le \sum_{n=1}^{\infty} \left(\left|p_1^{(n)}\right|_{\alpha_*} \|q\|_\infty + \left\|p_1^{(n)}\right\|_\infty |q_1 \circ \mathscr{S}_1^n|_{\alpha_*} \right) \le D_5\,(\Gamma'')^{-1}|||q_1|||_{\alpha_*,1}, \tag{3.27}$$

for some constant D_5 proportional to D_1. We bound $|\partial_\theta h_1|_{\alpha_*}$ by reasoning in a similar way: using that

$$\partial_\theta h_1 = \sum_{n=1}^{\infty} \left((\partial_\theta p_1^{(n)})\,q_1 + p_1^{(n)}\,\partial_\theta(q_1 \circ \mathscr{S}_1^n) \right),$$

we find

$$|\partial_\theta h_1|_{\alpha_*} \le \sum_{n=1}^{\infty}\Big(\big|\partial_\theta p_1^{(n)}\big|_{\alpha_*}\|q_1\|_\infty + \big\|\partial_\theta p_1^{(n)}\big\|_\infty |q_1\circ\mathscr{S}_1^n|_{\alpha_*} + \big|p_1^{(n)}\big|_{\alpha_*}\|\partial_\theta q\|_\infty + \big\|p_1^{(n)}\big\|_\infty |\partial_\theta q_1\circ\mathscr{S}_1^n|_{\alpha_*}\Big),$$

where the derivatives with respect to θ are bounded once more by relying on (2.18a) and the bounds in Lemma 3.8, while the non-differentiated functions are bounded as to obtain (3.27). This leads to

$$|\partial_\theta h_1|_{\alpha_*} \le D_6\,(\Gamma'')^{-1}|||q_1|||_{\alpha_*,1}, \tag{3.28}$$

with D_6 proportional to D_2. The bounds (3.23) and (3.25)–(3.28) ensure that $\|h_1\|_{\alpha_*,2}$ is bounded. Therefore, recalling the first relation in (3.15), Theorem 2 is proved, with $\Omega_0 = \mathscr{H}_1(\Omega_1) = \mathscr{H}(\Omega)$.

Remark 3.12. The argument above shows that it is enough to prove the existence of the conjugation inside Λ. Indeed, once the conjugation has been defined in Λ, it can be extended to the whole Ω by using that all trajectories fall inside a neighborhood of the attracting invariant manifold in a finite time.

With the notation of Lemma 2.7, for $r = r(\Gamma')$ we have $N_r = O((\Gamma')^{-1})$. Thus, from Remark 3.10 it is easy to see that in the scaling regime the constant D_0 in Lemma 3.8 is $O(1)$ in ρ. We can take $\Gamma' = \rho\gamma'$ and $\Gamma'' = \rho\gamma''$, with $0 < \gamma'' < \gamma' < \gamma$. Using that $\theta_1 = O(1)$ in ρ (see Remark 3.7), so that $M_r = O(\rho^{-1})$ and hence $M'_r = O((\Gamma')^{-1})$ as well, and requiring α_* in Theorem 2 to be such that $\lambda^{\alpha_*}(1-\rho\,\gamma') < (1-\rho\gamma'')$, so that $\alpha_* = O(\rho)$, we easily check that D_1 in (3.23) and, as a consequence, the constants D_2, D_3 and D_4 as well are all $O(1)$ in ρ. This implies that both $|||h|||_{0,1}$ and $|h|_{\alpha_*}$ are $O(1)$ in ρ. Collecting together the bounds above immediately implies Corollary 2.22.

3.2.4 The inverse conjugation

Since $\mathcal{H}_1(\theta,\psi) = \theta + \theta^2 h_1(\theta,\psi)$ and $h_1 \in \mathscr{B}_{\alpha_*,1}(\Omega_1,\mathbb{R})$ for a suitable $\alpha_* > 0$, there exists $\theta_* \le \theta_1$ such that $\partial_\theta\mathcal{H}_1(\theta,\psi) > 1/2$ for $|\theta| < \theta_*$. We have $\mathcal{H}_1(\theta,\psi) = (\kappa^{(n)}(\psi))^{-1}\mathcal{H}_1(\mathscr{S}_1^n(\theta,\psi))$ for all $(\theta,\psi)\in\Omega_1$ and all $n\in\mathbb{N}$. If we reason like deriving the bound on $\|p_1^{(n)}\|_\infty$ in Subsection 3.2.3 we see that there exists M_2 such that $|(\mathscr{S}_1^{M_2}(\theta,\psi))_\theta| \le \theta_*$. Thus, by Lemma 2.7 and Hypothesis 1, we get, for all $(\theta,\psi)\in\Omega_1$,

$$\partial_\theta\mathcal{H}_1(\theta,\psi) \ge \|\kappa\|_\infty^{-M_2} \inf_{|\theta|\le\theta^\star}\partial_\theta\mathcal{H}_1(\theta,\psi)\left(\inf_{(\theta,\psi)\in\Omega_1}\partial_\theta G_1(\theta,\psi)\right)^{M_2} =: \tau_1 > 0. \tag{3.29}$$

By the inverse function theorem, there is $l_1\in\mathscr{B}_{0,1}(\mathscr{H}_1(\Omega_1,\mathbb{R}))$ such that $(\mathscr{H}_1^{-1})_\eta(\eta,\psi) = \eta+\eta^2 l_1(\eta,\psi)$. Furthermore, for any $\psi,\psi'\in\mathbb{T}^2$, we can write $\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi'),\psi') - \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) = 0$, so that $|\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi'),\psi') - \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi')| = |\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi') - \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)|$, which, divided by $|\psi'-\psi|^{\alpha_*}$, gives $\partial_\theta\mathcal{H}_1(\theta,\psi)\,|\mathcal{H}_1^{-1}|_{\alpha_*} \le |\mathcal{H}_1|_{\alpha_*}$ for all $(\theta,\psi)\in\Omega_1$, so that we obtain $|\mathcal{H}_1^{-1}|_{\alpha_*} \le \tau_1^{-1}|\mathcal{H}_1|_{\alpha_*}$. Therefore, also $l_1 \in \mathscr{B}_{\alpha_*,1}(\mathscr{H}_1(\Omega_1),\mathbb{R})$ and, by the second relation of (3.15), Corollorary 2.18 follows.

Remark 3.13. By reasoning as in the derivation of (3.29), we find that there exists $\bar\tau > 0$ such that $\partial_\theta\overline{\mathcal{H}}(\theta) \ge \bar\tau$ for all $\theta\in\mathcal{U}$. This will be used later on (see Subsection 4.8).

3.3 Physical measure: proof of Theorem 3

Let $\mathcal{O}_1$ and $\mathcal{O}_2$ be two observables in $\mathscr{B}_{\alpha_*,1}(\Omega,\mathbb{R})$. By Remark 3.11, we have

$$|\mathcal{O}_2(\mathscr{S}^n(\varphi,\psi)) - \mathcal{O}_2(W(A_0^n\psi),A_0^n\psi)| \le C(1-\Gamma')^n|||\mathcal{O}_2|||_{0,1},$$

for a suitable constant C. Thus, if ν_W and ν_0 are defined as in Subsection 2.3.3, we find

$$\begin{aligned}&|\nu_0\,(\mathcal{O}_1\,\mathcal{O}_2\circ\mathscr{S}^n) - \nu_0\,(\mathcal{O}_1)\,\nu_W\,(\mathcal{O}_2)|\\ &\quad\le \int d\varphi\,|\langle\mathcal{O}_1(\varphi,\cdot)\mathcal{O}_2(W(A_0^n\cdot),A_0^n\cdot)\rangle - \langle\mathcal{O}_1(\varphi,\cdot)\rangle\langle\mathcal{O}_2(W(\cdot),\cdot)\rangle| + C(1-\Gamma')^n|||\mathcal{O}_2|||_{0,1}\|\mathcal{O}_1\|_\infty.\end{aligned}$$

Since $W\in\mathfrak{B}^-_{\alpha_-}(\mathbb{T}^2,\mathbb{R})$, relying on Remark 2.5, we obtain the bound in Theorem 3.

4 Averaging and deviations

The study of the convergence of the dynamics of $\mathscr{S}$ to the deterministic dynamics $\Phi_{n\rho}$ is structured in several steps: we start with the first and second moments of W (Subsections 4.2 and 4.3); then we consider the moments of the functions h and l and their derivatives (Subsections 4.4 to 4.8); eventually we draw the conclusions about the deviations of the dynamics from the averaged system (Subsection 4.9).

Remark 4.1. Throughout this section, we assume ρ to be as discussed in Remark 2.19. In the previous section, to help follow the bounds we discussed therein, we have made explicit the dependence of all the involved constants on the various parameters. This was also needed in the proofs of Corollaries 2.21, 2.22 and 2.23 to extract the dependence on ρ. In the forthcoming discussion we are mainly interested in what happens to the long time dynamics when ρ is small. Thus, from now on, in order not to overwhelm the notation, we call C any constant independent of ρ whose numerical value is not relevant.

4.1 The extended map

Even though the set Ω is positively invariant for the map $\mathscr{S}$ in (2.25), it is useful to extend the map $\mathscr{S}|_\Omega$ outside Ω. To do this, we proceed as follows:

- let $\chi_{\rm ext} : \mathbb{R} \to \mathbb{R}$ be a C^∞ function such that $\chi_{\rm ext}(x) = 1$ for $x \le 0$ and $\chi_{\rm ext}(x) = 0$ for $x \ge 1$ while $\partial_x \chi_{\rm ext}(x) \le 0$ for every $x \in \mathbb{R}$;
- set $s_M := \min\{1, \nu_M/2\}$, where $\nu_M := \inf_{\psi\in\mathbb{T}^2} |f(\varphi_M, \psi)|$;
- define $f_{\rm ext} \in \mathscr{B}_{\alpha_0,6}(\mathbb{R} \times \mathbb{T}^2, \mathbb{R})$, by setting $f_{\rm ext}(\varphi, \psi) = f(\varphi, \psi)$ for $\varphi_m \le \varphi \le \varphi_M$ while
$$f_{\rm ext}(\varphi,\psi) = \left(\sum_{i=0}^{6} \partial^i_\varphi f(\varphi_M,\psi)(\varphi-\varphi_M)^i\right)\chi_{\rm ext}\left(\frac{\varphi-\varphi_M}{s_M}\right) - \frac{\nu_M}{2}\left(1-\chi_{\rm ext}\left(\frac{\varphi-\varphi_M}{s_M}\right)\right),$$
for $\varphi > \varphi_M$, and an analogous expression for $\varphi < \varphi_m$.

One has $|||f_{\rm ext}|||_{\alpha_0,6} \le s_M^{-6} |||\chi_{\rm ext}|||_{0,6} |||F|||_{\alpha_0,6}$ and $\inf_{(\varphi,\psi)\in\mathbb{R}\times\mathbb{T}^2\setminus\Lambda_r} |f_{\rm ext}(\varphi,\psi)| \ge \frac{1}{2} \inf_{(\varphi,\psi)\in\Omega\setminus\Lambda_r} |f(\varphi,\psi)|$ for any $r > 0$. Moreover the map

$$\mathscr{S}_{\rm ext}(\varphi,\psi) := (\varphi + \rho f_{\rm ext}(\varphi,\psi), A_0\psi)$$

is defined on $\mathbb{R}\times\mathbb{T}^2$, coincides with $\mathscr{S}(\varphi,\psi)$ for $(\varphi,\psi) \in \Omega$ and, restricted to any $\Omega_{\rm ext} = \mathcal{U}_{\rm ext}\times\mathbb{T}^2$, with $\mathcal{U}_{\rm ext} \supseteq \mathcal{U}$ a closed interval, satisfies Hypotheses 1–3 with $\Omega_{\rm ext}$ instead of Ω. Extending $\mathscr{S}(\varphi,\psi)$ on $\Omega_{\rm ext}$ naturally defines a map $\mathscr{S}_{1,\rm ext}(\theta,\psi) := \mathscr{S}_{\rm ext}(\theta + W(\psi), \psi) - W(A_0\psi)$ on $\Omega_{1,\rm ext} := \{(\theta,\psi) \in \mathbb{R}\times\mathbb{T}^2 : (\theta + W(\psi), \psi) \in \Omega_{\rm ext}\}$. In the following discussion we need to choose $\Omega_{\rm ext}$ in such a way that $\Omega \subset \Omega_{1,\rm ext}$ (see Subsection 4.7). The conjugation $\mathcal{H}$ as well can be extended to a function $\mathcal{H}_{\rm ext} \in \mathscr{B}_{\alpha_0,1}(\Omega_{\rm ext}, \mathbb{R})$, by reasoning as in Subsection 3.2 with $\mathscr{S}_{\rm ext}$ instead of $\mathscr{S}$ everywhere.

Remark 4.2. The reason why we need to extend $\mathscr{S}|_\Omega$ outside Ω to a map $\mathscr{S}_{\rm ext}$, potentially different from the original $\mathscr{S}$ on $(\mathbb{T}\times\mathbb{T}^2)\backslash\Omega$, is that we aim to compare $\mathscr{S}$ with other maps which, albeit being constructed starting from $\mathscr{S}$, not only may fail to admit Ω as an invariant set (see Remark 4.6), but also are not necessarily defined in the whole Ω. In order not to discuss separately the dynamics near the boundary of Ω, it turns out to be easier to extend the maps to a larger domain $\Omega_{\rm ext}$ in such a way that they satisfy automatically the same properties as $\mathscr{S}$. We stress again that all the functions appearing in Theorems 3 to 8 depend only on $\mathscr{S}|_\Omega$. When proving (2.36) in Theorem 5 we write $h(\varphi,\psi) - \bar{h}(\varphi)$ as a sum of different contributions potentially depending on the extension chosen (see (4.26), (4.33) and also Remarks 4.6 and 4.13). Nevertheless, these contributions are uniquely determined up to corrections which, besides fading away when approaching $\mathcal{W}$, are bounded proportionally to ρ uniformly everywhere (see Propositions 4.28 and 4.44 below).

4.2 A correlation inequality

This subsection is dedicated to a generalization of Proposition 2.4 that plays a central role in the proof of Theorem 4 and Lemma 2.31. Together with the linearization around $\mathcal{W}$, the splitting (4.2) will be the core of the forthcoming discussion because it allows replace the study of the dynamics to any power n with the dynamics along the unperturbed part plus the effect of the dynamics on the part controlled by the parameter ρ, so as to estimate the decay of correlations of the coupled part.

For any two given sets of functions $\mathfrak{v}_0,\dots,\mathfrak{v}_{n-1}$ and $\mathfrak{w}_0,\dots,\mathfrak{w}_{n-1}$ on $\mathbb{T}^2$, set $\mathfrak{u}_i = \mathfrak{v}_i + \rho\,\mathfrak{w}_i$ for $i = 0,\dots,n-1$. Define, for $k = 0,\dots,n$ and $i = 0,\dots,n-k$,

$$\mathfrak{u}_i^{(k)}(\psi) := \prod_{j=0}^{k-1} \mathfrak{u}_{i+j}(A_0^j\psi), \qquad \mathfrak{u}^{(k)}(\psi) := \mathfrak{u}_0^{(k)}(\psi), \qquad \mathfrak{u}^{(-k)}(\psi) := \prod_{j=1}^{k} \mathfrak{u}_{k-j}(A_0^{-j}\psi) = \mathfrak{u}^{(k)}\circ A_0^{-k}, \tag{4.1}$$

with ${\mathfrak{u}_i}^{(0)}(\psi) = 1$ according to Remark 2.11, and define similarly $\mathfrak{v}_i^{(k)}(\psi)$, $\mathfrak{v}^{(k)}(\psi)$ and $\mathfrak{v}^{(-k)}(\psi)$.

We write $\mathfrak{u}_i^{(k)} - \mathfrak{v}_i^{(k)} = \rho\,\mathfrak{u}_i^{(k-1)}\,\mathfrak{w}_{i+k-1}\circ A_0^{k-1} + (\mathfrak{u}_i^{(k-1)} - \mathfrak{v}_i^{(k-1)})\mathfrak{v}_{i+k-1}\circ A_0^{k-1}$, so that, iterating, we get

$$\begin{aligned} \mathfrak{u}_i^{(k)} &= \mathfrak{v}_i^{(k)} + \rho\sum_{j=0}^{k-1} \mathfrak{u}_i^{(j)}\,\mathfrak{w}_{i+j}\circ A_0^j\,\mathfrak{v}_{i+j+1}^{(k-j-1)}\circ A_0^{j+1} = \mathfrak{v}_i^{(k)} + \rho\sum_{j=0}^{k-1} \mathfrak{v}_i^{(j)}\,\mathfrak{w}_{i+j}\circ A_0^j\,\mathfrak{v}_{i+j+1}^{(k-j-1)}\circ A_0^{j+1} \\ &\quad + \rho^2 \sum_{j=1}^{k-1}\sum_{j'=0}^{j-1} \mathfrak{u}_i^{(j')}\,\mathfrak{w}_{i+j'}\circ A_0^{j'}\,\mathfrak{v}_{i+j'+1}^{(j-j'-1)}\circ A_0^{j'+1}\,\mathfrak{w}_{i+j}\circ A_0^j\,\mathfrak{v}_{i+j+1}^{(k-j-1)}\circ A_0^{j+1}. \end{aligned} \tag{4.2}$$

where the second equality follows applying the first equality to the factor $\mathfrak{u}_i^{(j)}$ in the first line; according to the convention in Remark 2.11, the sum in the second line of (4.2) vanishes for $k = 0$, while the sum in the last line vanishes for $k = 0, 1$. We can also proceed "in the opposite direction" to get

$$\mathfrak{u}_i^{(k)} = \mathfrak{v}_i^{(k)} + \rho\sum_{j=0}^{k-1} \mathfrak{v}_i^{(j)}\,\mathfrak{w}_{i+j}\circ A_0^j\,\mathfrak{u}_{i+j+1}^{(k-j-1)}\circ A_0^{j+1}, \tag{4.3}$$

where we avoid writing the equivalent of the second expansion in (4.2) since we will not need it.

The following result plays a key role in the forthcoming analysis. The proof is based on (4.2).

Proposition 4.3. *Let $\mathfrak{u}_0,\dots,\mathfrak{u}_{n-1}$ be any functions in $\mathfrak{B}_\alpha(\mathbb{T}^2,\mathbb{R})$, with $\alpha\in(0,1]$, such that*

1. *$0 < \langle\mathfrak{u}_i\rangle \le \|\mathfrak{u}_i\|_\infty \le 1-\rho\,\gamma$ for all $i = 0,\dots,n-1$,*
2. *$\tilde{\mathfrak{u}}_i = O(\rho)$ for all $i = 0,\dots,n-1$,*

and set $\langle\mathfrak{u}\rangle^{(n)} := \langle\mathfrak{u}_0\rangle\langle\mathfrak{u}_1\rangle\dots\langle\mathfrak{u}_{n-1}\rangle$. Given $g_+\in\mathfrak{B}_\alpha^+(\mathbb{T}^2,\mathbb{R})$ and $g_-\in\mathfrak{B}_\alpha^-(\mathbb{T}^2,\mathbb{R})$, one has

$$\begin{aligned} &\Big|\langle g_+\,\mathfrak{u}^{(n)}\circ A_0\,g_-\circ A_0^{n+1}\rangle - \langle\mathfrak{u}\rangle^{(n)}\langle g_+\rangle\langle g_-\rangle\Big| \\ &\quad \le C(1-\rho\,\gamma)^n\Big((1+\alpha n)\lambda^{-\alpha n}\|\tilde g_+\|_\alpha^+\|\tilde g_-\|_\alpha^- + \rho\big(\|\tilde g_+\|_\alpha^+\|g_-\|_\alpha^- + \|g_+\|_\alpha^+\|\tilde g_-\|_\alpha^-\big) + \rho^2 n|\langle g_+\rangle||\langle g_-\rangle|\Big). \end{aligned}$$

Proof. For any function $\mathfrak{u}_i$, with $i = 0,\dots,n-1$, let $\mathfrak{v}_i := \langle\mathfrak{u}_i\rangle$ and $\rho\,\mathfrak{w}_i := \tilde{\mathfrak{u}}_i$ and introduce the notation $\langle\mathfrak{u}\rangle_i^{(k)} := \langle\mathfrak{u}_i\rangle\dots\langle\mathfrak{u}_{k-1}\rangle$, so that $\langle\mathfrak{u}\rangle^{(k)} = \langle\mathfrak{u}\rangle_0^{(k)}$. Then we write

$$\langle g_+\,\mathfrak{u}^{(n)}\circ A_0\,g_-\circ A_0^{n+1}\rangle = \langle g_+\,\mathfrak{u}^{(n)}\circ A_0\,\tilde g_-\circ A_0^{n+1}\rangle + \langle\tilde g_+\,\mathfrak{u}^{(n)}\circ A_0\rangle\langle g_-\rangle + \langle g_+\rangle\langle\mathfrak{u}^{(n)}\rangle\langle g_-\rangle, \tag{4.4}$$

so that, using the first line of (4.2), with $\mathfrak{u}^{(n)}\circ A_0$ instead of $\mathfrak{u}^{(n)}$, we rewrite the first term in (4.4) as

$$\langle g_+\,\mathfrak{u}^{(n)}\circ A_0\,\tilde g_-\circ A_0^{n+1}\rangle = \langle\mathfrak{u}\rangle^n\langle g_+\,\tilde g_-\circ A_0^{n+1}\rangle + \rho\sum_{k=0}^{n-1}\langle\mathfrak{u}\rangle_{k+1}^{n-k-1}\langle g_+\,\mathfrak{u}^{(k)}\circ A_0\,\tilde{\mathfrak{u}}_k\circ A_0^{k+1}\,\tilde g_-\circ A_0^{n+1}\rangle.$$

Using (2.6b) and (2.30), we get

$$\begin{aligned} \big|g_+\circ A_0^{-(k+1)}\mathfrak{u}^{(k)}\circ A_0^{-k}\,\tilde{\mathfrak{u}}_k\big|_\alpha^+ &\le \|g_+\circ A_0^{-(k+1)}\,\tilde{\mathfrak{u}}_k\|_\infty\big|\mathfrak{u}^{(k)}\circ A_0^{-k}\big|_\alpha^+ + \big|g_+\circ A_0^{-(k+1)}\,\tilde{\mathfrak{u}}_k\big|_\alpha^+\|\mathfrak{u}^{(k)}\|_\infty \\ &\le C(1-\rho\,\gamma)^k\|g_+\circ A_0^{-(k+1)}\,\tilde{\mathfrak{u}}_k\|_\alpha^+, \end{aligned}$$

so that, from (2.9) and Proposition 2.4, we obtain in (4.4)

$$\big|\langle g_+\,\mathfrak{u}^{(n)}\circ A_0\,\tilde g_-\circ A_0^{n+1}\rangle\big| \le C(1-\rho\,\gamma)^n\lambda^{-\alpha n}(1+\alpha n)\|\tilde g_+\|_\alpha^+\|\tilde g_-\|_\alpha^- + C\rho(1-\rho\,\gamma)^n\|g_+\|_\alpha^+\|\tilde g_-\|_\alpha^- . \quad (4.5)$$

Analogously to (4.5), using (4.3), with $\mathfrak{u}^{(n)}\circ A_0$ instead of $\mathfrak{u}^{(n)}$, and Proposition 2.4, we get

$$\big|\langle \tilde g_+\,\mathfrak{u}^{(n)}\circ A_0\rangle\big| \le \rho\sum_{k=0}^{n-1}\langle\mathfrak{u}\rangle_0^k|\langle\tilde g_+\,\tilde{\mathfrak{u}}_k\circ A_0^k\,\mathfrak{u}^{(n-k-1)}\circ A_0^{k+1}\rangle\,| \le C\rho(1-\rho\,\gamma)^n\|\tilde g_+\|_\alpha^+ . \quad (4.6)$$

Finally, for $n\ge 2$, since $\langle\tilde{\mathfrak{u}}_k\rangle = 0$ for all $k=0,\dots,n-1$, using the third line of (4.2), once more with $\mathfrak{u}^{(n)}\circ A_0$ instead of $\mathfrak{u}^{(n)}$, we have

$$\begin{aligned}\big|\langle\,\mathfrak{u}^{(n)}\rangle - \langle\mathfrak{u}\rangle^n\big| &= \Big|\rho^2\sum_{k=1}^{n-1}\sum_{j=0}^{k-1}\langle\mathfrak{u}\rangle_{j+1}^{k-j-1}\langle\mathfrak{u}\rangle_{k+1}^{n-k-1}\langle\mathfrak{u}^{(j)}\circ A_0^{-j}\,\tilde{\mathfrak{u}}_j\,\tilde{\mathfrak{u}}_k\circ A_0^{k-j}\rangle\Big| \\ &\le C\rho^2\sum_{k=1}^{n-1}\sum_{j=0}^{n-1} j^2\lambda^{-\alpha j}(1-\rho\,\gamma)^n\|\tilde{\mathfrak{u}}_j\|_\alpha^+\|\tilde{\mathfrak{u}}_k\|_\alpha^- \le C\rho^2 n(1-\rho\,\gamma)^{n-2}\|\tilde{\mathfrak{u}}_j\|_\alpha^+\|\tilde{\mathfrak{u}}_k\|_\alpha^-,\end{aligned}$$

where we have used Proposition 2.4 and (4.1), so that we obtain

$$\big|\langle\,\mathfrak{u}^{(n)}\rangle - \langle\mathfrak{u}\rangle_0^n\big| \le C\rho^2 n(1-\rho\,\gamma)^n . \quad (4.7)$$

Collecting all contributions (4.5)–(4.7) gives the thesis. □

Remark 4.4. The factor $\rho\big(\|\tilde g_+\|_\alpha^+\|g_-\|_\alpha^- + \|g_+\|_\alpha^+\|\tilde g_-\|_\alpha^-\big)$ appearing in the estimate of Proposition 4.3, in principle, might be replaced more pragmatically with $2\rho\,\|g_+\|_\alpha^+\|g_-\|_\alpha^-$. However, it may be useful in some cases. In particular, for $g_+ = g_- = 1$, we obtain $|\langle\mathfrak{u}^{(n)}\rangle - \langle\mathfrak{u}\rangle^{(n)}| \le C\rho^2 n(1-\rho\,\gamma)^n$ and hence $|\langle\mathfrak{u}^{(n)}\rangle - \langle\mathfrak{u}\rangle^{(n)}| \le C\rho(1-\rho\,\gamma')^n$ for any $\gamma'\in(0,\gamma)$, with the constant C depending on γ'.

4.3 Oscillations of the invariant manifold: proof of Theorem 4

Following Remark 2.9, we set $b(\psi) := f(0,\psi)$ and $v(\psi) := \partial_\varphi f(0,\psi)$, with $\|v\|_\infty \le C$ by Remark 2.19, and write

$$f(\varphi,\psi) = b(\psi) + v(\psi)\,\varphi + d(\varphi,\psi)\varphi^2, \quad (4.8)$$

which implicitly defines the function $d(\varphi,\psi)$. We also set $\mu(\psi) := 1+\rho v(\psi) = 1+\rho\partial_\varphi f(0,\psi)$ and, for $i\ge 0$, according to (4.1),

$$\mu^{(i)}(\psi) = \prod_{j=0}^{i-1}\mu(A_0^j\psi), \qquad \mu^{(-i)}(\psi) = \prod_{j=1}^{i}\mu(A_0^{-j}\psi) = \mu^{(i)}(A^{-i}\psi). \quad (4.9)$$

Decomposing $\mu(\psi) = \langle\mu\rangle + \tilde\mu(\psi)$ according to (2.1), we have $\langle\mu\rangle = \overline{\mu}$, with $\overline{\mu}$ as in (2.27), and $\tilde\mu(\psi) = \rho\tilde v(\psi)$. We get $\langle\mu\rangle \le \|\mu\|_\infty \le 1-\rho\,\gamma$, since $\partial_\varphi f(\varphi,\psi) \le -\gamma$ for all $\varphi\in[S_m,S_M]$ and $\bar\varphi = 0\in[S_m,S_M]$ (see the definition of γ after (2.25) and Remark 2.8) and $|\mu|_{\alpha_0}^+ = O(\rho)$.

In Section 3.1 we proved that $W = \lim_{n\to\infty}\mathscr{E}^n[0]$ uniformly on $\mathbb{T}^2$, with $\mathscr{E}$ defined in (3.2). Thus, to prove (2.33), we show that, for all $n\in\mathbb{N}$, one has $\big|\langle\mathscr{E}^n[0]\rangle\big| \le C\rho$ and $\langle(\mathscr{E}^n[0])^2\rangle \le C\rho$, for some constant C independent of n. To this end, analogously to (3.2), we set

$$\mathscr{E}_0[W](\psi) := G_0(W(A_0^{-1}\psi), A_0^{-1}\psi), \qquad G_0(\varphi,\psi) := \varphi + \rho f_0(\varphi,\psi), \qquad f_0(\varphi,\psi) := b(\psi) + v(\psi)\varphi$$

and observe that, for any $h:\mathbb{T}^2\to\mathbb{R}$,

$$\mathscr{E}_0^n[h] = \rho\sum_{i=1}^{n}\mu^{(-i+1)}b\circ A_0^{-i} + \mu^{(-n)}h\circ A_0^{-n}. \quad (4.10)$$

Remark 4.5. Since $\mathscr{E}_0$ is a contraction on the space of bounded continuous functions from $\mathbb{T}^2$ to $\mathbb{R}$, the fixed point equation $\mathscr{E}_0[W] = W$ admits a unique solution W_0 and $\mathscr{E}_0^n[0]$ converges uniformly to W_0. We can write

$$W_0 = \rho \sum_{i=1}^{n} \mu^{(-i+1)} b \circ A_0^{-i} + \mu^{(-n)} W_0 \circ A_0^{-n} = \rho \sum_{i=1}^{\infty} \mu^{(-i+1)} b \circ A_0^{-i}\,, \tag{4.11}$$

from which we get $\|W_0\|_\infty \leq \gamma^{-1}\|b\|_\infty$, while $|W_0|^+_{\alpha_0} = O(\rho)$ by (2.10).

To prove Theorem 4 we show, first, by explicit computation, that $|\langle W_0\rangle| = O(\rho)$ and $\langle W_0^2\rangle = O(\rho)$, then that $\langle |W - W_0|\rangle = O(\rho)$, by comparing the iterates of $\mathscr{E}_0$ and of $\mathscr{E}$ and using their contractivity.

From the last definition in (4.1) we get $\langle \mu^{(-i+1)} b \circ A_0^{-i}\rangle = \langle b\, \mu^{(i-1)} \circ A_0\rangle$. Since, by definition, $\langle b\rangle = 0$, using Proposition 4.3, with $n = i-1$, $\mathfrak{u}_k = \mu$ for $k = 0, \ldots, n-1$, $g_+ = b$ and $g_- = 1$, we find

$$\big|\langle \mathscr{E}_0^n[0]\rangle\big| \leq C\rho^2 \sum_{i=0}^{n-1} (1-\rho\,\gamma)^i \|b\|_{\alpha_0} \leq C\rho\,.$$

Moreover we have

$$(\mathscr{E}_0^n[0])^2 = \rho^2 \sum_{i=1}^{n} (\mu^{(-i+1)})^2 (b\circ A_0^{-i})^2 + 2\rho^2 \sum_{1\leq i<j\leq n} (\mu^{(-i+1)})^2\, b\circ A_0^{-i}\, \mu \circ A_0^{-i}\, \mu^{-(j-i-1)} \circ A_0^{-i}\, b \circ A_0^{-j},$$

where

$$\big\langle (\mu^{(-i+1)})^2\, b\circ A_0^{-i}\, \mu\circ A_0^{-i}\, \mu^{(-j+i+1)}\circ A_0^{-i}\, b\circ A_0^{-j}\big\rangle = \big\langle b\,\mu^{(j-i-1)}\circ A_0 (\mu^{(i-1)})^2 \circ A_0^{j-i+1}\, b\circ A_0^{j-i}\, \mu \circ A_0^{j-i}\big\rangle,$$

so that, first using (2.9) and (2.10) to estimate

$$\big|(\mu^{(i-1)})^2 \circ A_0\, b\,\mu\big|^-_{\alpha_0} \leq \big|(\mu^{(i-1)})^2\circ A_0\big|^+_{\alpha_0} \big\|b\,\mu\big\|_\infty + \big\|(\mu^{(i-1)})^2\big\|_\infty \big|b\,\mu\big|^+_{\alpha_0} \leq C(1-\rho\,\gamma)^{2(i-1)},$$

then applying once more Proposition 4.3, with $n = j-i-1$, $\mathfrak{u}_k = \mu$ for $k = 0, \ldots, n-1$, $g_+ = b$ and $g_- = (\mu^{(i-1)})^2 \circ A_0\, b\,\mu$, we get

$$\langle \mathscr{E}_0^n[0]^2\rangle \leq C\rho\|b\|^2_\infty + C\rho^3 \sum_{1\leq i<j\leq n} (1-\rho\gamma)^{j-i} \|b\|_{\alpha_0} \|(\mu^{(i-1)})^2\circ A_0\, b\,\mu\|^-_{\alpha_0} \leq C\rho + C\rho^3 \sum_{1\leq i<j\leq n}^{n} (1-\rho\gamma)^{j+i} \leq C\rho.$$

Thus, if W_0 is defined as in Remark 4.5, we obtain

$$\langle W_0\rangle = \lim_{n\to\infty} \langle \mathscr{E}_0^n[0]\rangle = O(\rho), \qquad \langle W_0^2\rangle = \lim_{n\to\infty} \langle \mathscr{E}_0^n[0]^2\rangle = O(\rho). \tag{4.12}$$

For any $n \geq 0$, we can write

$$\mathscr{E}^n[0] - \mathscr{E}_0^n[0] = \sum_{i=0}^{n-1} \mathscr{E}^{n-i}[\mathscr{E}_0^i[0]] - \mathscr{E}^{n-i-1}[\mathscr{E}_0^{i+1}[0]] = \sum_{i=0}^{n-1} \mathscr{E}^{n-i-1}[\mathscr{E}[\mathscr{E}_0^i[0]]] - \mathscr{E}^{n-i-1}[\mathscr{E}_0[\mathscr{E}_0^i[0]]]\,. \tag{4.13}$$

Remark 4.6. It may happen that $\mathscr{E}_0^i[0] \not\subset \Omega$ for some i. In such a case, we observe that there exists an interval $\mathcal{U}' := [\phi'_m, \phi'_M] \supset \mathcal{U}$ such that $\Omega' := \mathcal{U}' \times \mathbb{T}^2$ and $\mathscr{E}_0^n[0](\psi) \in \Omega'$ for all $n \geq 0$ and all $\psi \in \mathbb{T}^2$. Thus for the r.h.s. of (4.13) to be well defined, we compute $\mathscr{E}$ by replacing the map $\mathscr{S}$ with the extension $\mathscr{S}_{\mathrm{ext}}$ introduced in Subsection 4.1 such that $\Omega_{\mathrm{ext}} \supset \Omega'$ (see also Remark 4.2). The forthcoming analysis still applies if a different extensions is taken, as far as it satisfies the Hypotheses considered in Section 2.2 (see also Remark 4.13).

By taking $\mathscr{S}_{\mathrm{ext}}$ according to Remark 4.6 and reasoning as in the proof of Lemma 2.7, we find that for any $r > 0$ there exists $N'_r \leq \max\{\phi'_M - S_M, S_m - \phi'_m\}/(\rho \inf_{\Omega'\setminus\Lambda_r} |f_{\mathrm{ext}}(\varphi,\psi)|)$ such that $\mathscr{S}_{\mathrm{ext}}^{N'_r}(\Omega') \subset \Lambda_r$, and, given $W_1, W_2 : \mathbb{T}^2 \to [\phi'_m, \phi'_M]$, we have $\mathscr{E}^{N'_r}[W_i] \in [S_m - r, S_M + r]$ for $i = 1, 2$, while $|\mathscr{E}^{N'_r}[W_1] - \mathscr{E}^{N'_r}[W_2]| \leq (1 + \rho\|\partial_\varphi f\|_\infty)^{N'_r} |W_1 - W_2|$. Thus, for any $\gamma' \in (0, \gamma)$, with the notation in property 4 of Lemma 2.7, we can choose $r = r(\rho\,\gamma')$ and obtain $N'_r = O(1/\rho\,\gamma')$ and, for any $W_1, W_2 \in [S_m - r, S_M + r]$ and any $k \geq 0$, we may bound $|\mathscr{E}^k[W_1] - \mathscr{E}^k[W_2]| \leq (1-\rho\,\gamma')^k |W_1 - W_2|$. Summing up, (4.13) gives

$$|\mathscr{E}^n[0](\psi) - \mathscr{E}_0^n[0](\psi)| \leq \left(\frac{1+\rho\|\partial_\varphi f\|_\infty}{1-\rho\,\gamma'}\right)^{N'_r} \rho \sum_{i=0}^{n-1} (1-\rho\,\gamma')^{n-i-1} \big|d(\mathscr{E}_0^i[0](A_0^{-1}\psi), A_0^{-1}\psi)\big|\, (\mathscr{E}_0^i[0](A_0^{-1}\psi))^2,$$

with $d(\varphi,\psi)$ implicitly defined in (4.8). Integrating over ψ and using the bound on $\langle \mathscr{E}_0^n[0]^2\rangle$, we obtain

$$\langle |\mathscr{E}^n[0]-\mathscr{E}_0^n[0]|\rangle \le C\rho \sum_{i=0}^{n-1} (1-\rho\,\gamma')^i\, \|d\|_\infty \langle \mathscr{E}_0^i[0]^2\rangle \le C\rho,$$

which yields

$$\langle |W-W_0|\rangle \le C\rho, \tag{4.14}$$

which, together with (4.12), gives $|\langle W\rangle| \le |\langle W-W_0\rangle| + |\langle W_0\rangle| \le \langle |W-W_0|\rangle + |\langle W_0\rangle| \le C\rho$, that is the first bound in (2.33). Finally, we have $|\langle W^2-W_0^2\rangle| \le \langle |W-W_0|\,|W+W_0|\rangle \le C\rho$, because of (4.14), and hence, thanks to (4.12), $\langle W^2\rangle \le |\langle W^2-W_0^2\rangle| + \langle W_0^2\rangle \le C\rho$, that is the second bound in (2.33).

4.4 Fluctuations of the linearized dynamics: proof of Lemma 2.31

Having studied the oscillations of the invariant manifold we now focus our attention on the linearized dynamics. Using that $\kappa(\psi) = 1+\rho\,\partial_\varphi f(W(\psi),\psi)$ and $\mu(\psi) = 1+\rho\,\partial_\varphi f(0,\psi) = 1-\rho\, v(\psi)$, we write

$$\kappa(\psi) = \mu(\psi) + \rho\,\xi(\psi), \qquad \xi(\psi) := d_1(\psi)\, W(\psi) + d_2(\psi)\, W^2(\psi), \qquad d_1(\psi) := \partial_\varphi^2 f(0,\psi), \tag{4.15}$$

which implicitly defines the function $d_2 \in \mathfrak{B}^*_{\alpha_-,\alpha_+}(\mathbb{T}^2,\mathbb{R})$.

From Proposition 4.3, with $g_+ = g_- = 1$ and either $\mathfrak{u}_i = \mu$ or $\mathfrak{u}_i = \mu^2$ for all i, and Remark 4.4, it follows that $\left|\langle \mu^{(n)} - \overline{\mu}^n\rangle\right| \le C\rho^2 n(1-\rho\gamma)^n$ and $\left\langle (\mu^{(n)}-\overline{\mu}^n)^2\right\rangle \le C\rho^2 n (1-\rho\gamma)^{2n}$, and hence, in order to prove (2.35), it is enough to show that

$$\left|\langle \kappa^{(n)} - \mu^{(n)}\rangle\right| \le C\rho(1-\rho\gamma)^n \sum_{k=0}^{2} \rho^k n^k, \qquad \left\langle (\kappa^{(n)}-\mu^{(n)})^2\right\rangle \le C\rho (1-\rho\gamma)^{2n} \sum_{k=0}^{2} \rho^k n^k. \tag{4.16}$$

Remark 4.7. The proof of (4.16) represents a first instance of the strategy outlined in Section 2.7. In order to bypass the low regularity of κ due to its dependence on the invariant manifold, we expand it up to the second order in W so as to use Theorem 4 to estimate the average of the quadratic contributions, the bound (4.14) to reduce the analysis of the average of the linear contributions to the more manageable function W_0 and the explicit expression for W_0 in (4.11) to obtain the desired estimates.

Thus, we proceed with the proof of (4.16). By Remark (4.2) we can write

$$\begin{aligned}
\kappa^{(n)} - \mu^{(n)} &= \rho \sum_{j=0}^{n-1} \kappa^{(j)}\, \xi\circ A_0^j\, \mu^{(n-j-1)}\circ A^{j+1} \\
&= \rho \sum_{j=0}^{n-1} \mu^{(j)}\, \xi\circ A_0^j\, \mu^{(n-j-1)}\circ A^{j+1} + \rho^2 \sum_{j=1}^{n-1}\sum_{k=0}^{j-1} \kappa^{(k)}\, \xi\circ A_0^k\, \mu^{(j-k-1)}\circ A_0^{k+1}\, \xi\circ A_0^j\, \mu^{(n-j-1)}\circ A_0^{j+1}.
\end{aligned} \tag{4.17}$$

Since $|\langle \kappa^{(k)}\xi\circ A_0^k \mu^{(j-k-1)}\circ A_0^{k+1} \xi\circ A_0^j \mu^{(n-j-1)}\circ A_0^{j+1}\rangle| \le (1-\rho\,\gamma)^{n-2}\langle |\xi\circ A_0^k\, \xi\circ A_0^j|\rangle \le (1-\rho\,\gamma)^{n-2}\langle \xi^2\rangle$ and $\langle \xi^2\rangle \le C\rho$, by (4.15) and (2.33), eventually we get, for $n\ge 2$,

$$\left| \sum_{j=1}^{n-1}\sum_{k=0}^{j-1} \left\langle \kappa^{(k)}\, \xi\circ A_0^k\, \mu^{(j-k-1)}\circ A_0^{k+1}\, \xi\circ A_0^j\, \mu^{(n-j-1)} A_0^{j+1} \right\rangle \right| \le C\rho\, n^2 (1-\rho\,\gamma)^n.$$

On the other hand, since we have, again by (2.33),

$$\left|\left\langle \mu^{(j)}\, (d_2 W^2)\circ A_0^j\, \mu^{(n-j-1)}\circ A^{j+1}\right\rangle\right| \le (1-\rho\,\gamma)^{n-1} \|d_2\|_\infty \langle W^2\rangle \le C\rho (1-\rho\,\gamma)^n.$$

we need to estimate $\langle \mu^{(j)}\, (d_1 W)\circ A_0^j\, \mu^{(n-j-1)}\circ A_0^{j+1}\rangle = \langle \mu^{(j)}\circ A_0^{-j}\, d_1 W\, \mu^{(n-j-1)}\circ A_0\rangle$. By Proposition 4.3, with $\alpha=\alpha_+$, $n-j-1$ instead of n, $\mathfrak{u}_i = \mu$ for $i=0,\dots,n-j-2$, $g_+ = \mu^{(j)}\circ A_0^{-j}\, d_1 W$ and $g_- = 1$, we find

$$\left|\left\langle \mu^{(j)}\, (d_1 W)\circ A_0^j\, \mu^{(n-j-1)}\circ A_0^{j+1}\right\rangle\right| \le C(1-\rho\,\gamma)^{n-j} \left|\left\langle \mu^{(j)}\right\rangle\circ A_0^{-j}\, d_1 W\right| + C\rho\,(1-\rho\,\gamma)^n\,(1+\rho\,(n-j)),$$

where we have used also that $\|W\|^+_{\alpha_+} \le C$ and $\|\mu^{(j)}\circ A_0^{-j}\|^+_{\alpha_0} \le C$, because of Theorem 2 and (2.10), respectively. Then, by (4.14) and Remark 4.5, we find

$$\begin{aligned}\left|\left\langle \mu^{(j)}\circ A_0^{-j} d_1 W\right\rangle\right| &= \left|\left\langle \mu^{(j)}\,(d_1 W)\circ A_0^{j}\right\rangle\right| \le \left|\left\langle \mu^{(j)}\,(d_1 W_0)\circ A_0^{j}\right\rangle\right| + C\rho(1-\rho\gamma)^j \\ &\le \left|\left\langle W_0\,(\mu^{(j)})^2\, d_1\circ A_0^j\right\rangle\right| + \rho\sum_{k=1}^{j}\left|\left\langle \mu^{(j-k)} b\circ A_0^{j-k}(\mu^{(k-1)})^2\circ A_0^{j-k}\mu\circ A_0^{j-1} d_1\circ A_0^j\right\rangle\right| + C\rho(1-\rho\gamma)^j,\end{aligned}$$

where, since $\langle W_0\rangle = O(\rho)$, we have $|\langle W_0\,(\mu^{(j)})^2\, d_1\circ A_0^j\rangle| \le C\,(1-\rho\gamma)^{2j}\,(1+\alpha_0 j)\lambda^{-\alpha_0 j} + \rho + \rho^3 j)$, by Proposition 4.3, with $n = j$, $\mathfrak{u}_i = \mu^2$ for $i = 0,\dots,j-1$, $g_+ = W_0\circ A_0^{-1}$ and $g_- = d_1$, while

$$\begin{aligned}&\left|\left\langle \mu^{(j-k)} b\circ A_0^{j-k}(\mu^{(k-1)})^2\circ A_0^{j-k}\mu\circ A_0^{j-1}\, d_1\circ A_0^j\right\rangle\right| \\ &\qquad\le (1-\rho\gamma)^{j-k}\left|\left\langle b(\mu^{(k-1)})^2\,\mu\circ A_0^{k-1}\, d_1\circ A_0^k\right\rangle\right| + C\rho(1-\rho\gamma)^{j+k}\,(1+\rho\,(j-k)) \\ &\qquad\le C(1-\rho\gamma)^{j-k}\big((1+\alpha_0 k)\lambda^{-\alpha_0 k} + \rho\big)\|b\|_{\alpha_0}\|d_1\|_{\alpha_0} + C\rho(1-\rho\gamma)^{j+k},\end{aligned}$$

where we have used twice Proposition 4.3, first with $n = j-k$, $\mathfrak{u}_i = \mu$ for $i = 0,\dots,j-k-1$, $g_+ = 1$ and $g_- = b\,(\mu^{(k-1)})^2\mu\circ A_0^{k-1} d_1\circ A_0^k$, then with $n = k$, $\mathfrak{u}_i = \mu^2$ for $i = 0,\dots,k-1$ and $\mathfrak{u}_k = \mu$, $g_+ = b$ and $g_- = d_1$. Inserting all the bounds into (4.17), we obtain the first of (4.16).

For the second of (2.35), we use the first line of (4.17) to write

$$\left\langle\left(\kappa^{(n)} - \mu^{(n)}\right)^2\right\rangle \le \rho^2\sum_{i,j=1}^{n-1}\left|\left\langle \kappa^{(i)}\,\xi\circ A_0^i\,\mu^{(n-i-1)}\circ A^{i+1}\kappa^{(j)}\,\xi\circ A_0^j\,\mu^{(n-j-1)}\circ A^{j+1}\right\rangle\right| \le C\,n^2\rho^2\,(1-\rho\gamma)^{2(n-1)}\langle\xi^2\rangle.$$

Thus, the second bound in (4.16) follows using again the second bound in (2.33).

4.5 Conjugation of the averaged dynamics: proof of Lemma 2.26

Proceeding as in Subsection 3.2.2 we see that the function $\overline{h}\colon \mathcal{U}\to\mathcal{U}_0$ introduced in (2.30) satisfies the equation $\overline{h}(\varphi) = \overline{q}(\varphi) + \overline{p}(\varphi)\,\overline{h}(\overline{G}(\varphi))$, with $\overline{G}(\varphi) = \overline{\mathscr{S}}_\varphi(\varphi,\psi)$ defined in (2.26) and

$$\overline{p}(\varphi) := \frac{(\overline{G}(\varphi))^2}{\varphi^2\,\overline{\mu}} = \frac{1}{1+\rho\,\partial_\varphi\overline{f}(0)}\left(\frac{\varphi+\rho\overline{f}(\varphi)}{\varphi}\right)^2, \tag{4.18a}$$

$$\overline{q}(\varphi) := \frac{\overline{G}(\varphi) - \varphi\,\overline{\mu}}{\varphi^2\,\overline{\mu}} = \frac{\rho}{1+\rho\,\partial_\varphi\overline{f}(0)}\,\frac{\overline{f}(\varphi) - \partial_\varphi\overline{f}(0)\varphi}{\varphi^2}. \tag{4.18b}$$

Thus we can write (see (2.17) for the notation)

$$\overline{h}(\varphi) = \sum_{n=1}^{\infty}\overline{p}^{(n)}(\varphi)\,\overline{q}(\overline{G}^n(\varphi)), \qquad \overline{p}^{(n)}(\varphi) = \overline{p}^{(n)}(\overline{\mathscr{S}};\varphi) = \prod_{i=0}^{n-1}\overline{p}(\overline{G}^i(\varphi)). \tag{4.19}$$

Analogously to Lemma 3.6 the following result holds.

Lemma 4.8. *In $\mathcal{U}$ one has $|||\overline{p}|||_{0,5} = 1 + O(\rho)$, $|||\overline{q}|||_{0,5} = O(\rho)$ and $|||\partial_\theta\overline{p}|||_{0,4} = O(\rho)$. Moreover, for any $\gamma' \in (0,\gamma)$, there exists $\overline{\theta} = O(1)$ such that $\overline{p}(\varphi) \le 1-\rho\gamma'$ for $|\varphi| \le \overline{\theta}$.*

Remark 4.9. By reasoning as in the proof of Lemma 3.8 (and hence, actually, Lemma 2.7), we find that, for any $\gamma' \in (0,\gamma)$, there exist $r > 0$ and $\overline{N} = O(\rho^{-1})$ such that $\overline{\mathscr{S}}^n(\mathcal{U}) \subset [S_m - r, S_M + r]$ and $(1+\rho\partial_\varphi\overline{f})\circ\overline{\mathscr{S}}^n \le (1-\rho\gamma')$ for all $n \ge \overline{N}$.

Lemma 4.10. *Given $\gamma' \in (0,\gamma)$ and $\overline{\theta}$ as in Lemma 4.8, consider any functions $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ in $\mathscr{B}_{0,3}(\Omega,\mathbb{R})$, independent of ψ, such that*

1. *$\|\mathfrak{p}_i - 1\|_{0,3} = O(\rho)$ for all $i = 0,\dots,n-1$,*
2. *$|\mathfrak{p}_i(\varphi)| \le 1-\rho\gamma'$ for any $|\varphi| \le \overline{\theta}$ and for all $i = 0,\dots,n-1$,*

and define $\mathfrak{p}^{(n)}(\varphi) = \mathfrak{p}^{(n)}(\overline{\mathscr{S}};\varphi)$ *as in* (2.16) *with* $\mathscr{S} = \overline{\mathscr{S}}$. *One has*

$$|||\mathfrak{p}^{(n)}|||_{0,3} \leq C(1-\rho\,\gamma')^n, \qquad \|\partial_\varphi^k \overline{G}^n\|_\infty \leq C(1-\rho\,\gamma')^n\,, \qquad k=1,2,3. \tag{4.20}$$

where the constant C *does not depend on* n.

Proof. Let $\overline{N} = O(\rho^{-1})$ be defined as in Remark 4.9. Then, using that $|\overline{G}^n(\varphi)| \leq \overline{\theta}$ for $n \geq \overline{N}$ and that

$$\partial_\varphi \overline{G}^n(\varphi) = \prod_{i=0}^{n-1} \partial_\varphi \overline{G}(\overline{G}^i(\varphi)),$$

we get

$$\|\partial_\varphi \overline{G}^n\|_\infty \leq C(1-\rho\,\gamma')^n\,. \tag{4.21}$$

Then, combining (2.18a) and (2.16) together with the bound (4.21), we obtain

$$\left|\prod_{i=0}^{n-1} \mathfrak{p}_i(\overline{G}^i(\varphi)\right| \leq C(1-\rho\gamma')^n, \qquad \left|\partial_\varphi\bigg(\prod_{i=0}^{n-1} \mathfrak{p}_i(\overline{G}^i(\varphi))\bigg)\right| \leq C\rho(1-\rho\gamma')^{n-1}\sum_{i=0}^{n-1}(1-\rho\gamma')^i \leq C(1-\rho\gamma')^n,$$

so that $|||\mathfrak{p}^{(n)}|||_{0,0} = \|\mathfrak{p}^{(n)}\|_\infty \leq C(1-\rho\,\gamma')^n$ and

$$|||\mathfrak{p}^{(n)}|||_{0,1} \leq C(1-\rho\,\gamma')^n, \qquad \|\partial_\varphi^2 \overline{G}^n\|_\infty \leq C(1-\rho\,\gamma')^n, \tag{4.22}$$

where the second one is a special case of the first one, with $\mathfrak{p}_i = \partial_\varphi \overline{G}$ $\forall i = 0,\dots,n-1$. Inserting the bounds (4.22) in (2.18b) (again see (2.16) for notations) and reasoning in a similar way, we get

$$|||\mathfrak{p}^{(n)}|||_{0,2} \leq C(1-\rho\,\gamma')^n, \qquad \|\partial_\varphi^3 \overline{G}^n\|_\infty \leq C(1-\rho\,\gamma')^n. \tag{4.23}$$

Using (4.23) in (2.18c) gives (4.20), while from (4.21)–(4.23) we get the second bound in (4.20). ☐

Remark 4.11. An immediate consequence of the functions $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ satisfying condition 1 in Lemma 4.10 is that $|||\partial_\varphi \mathfrak{p}_i|||_{0,2} = O(\rho)$ for all $i = 0,\dots,n-1$.

Remark 4.12. In the following we have to consider cases in which condition 2 of Lemma 4.10 holds for all $i = 0,\dots,n-1$ except at most n_* values, for some $n_* < n$ independent of n. However, such cases are easily reduced to Lemma 4.10. Indeed, if $\mathfrak{p}_{i_1},\dots,\mathfrak{p}_{i_{n_*}}$ are the functions which do not satisfy condition 2, setting $I_* = \{i_1,\dots i_{n_*}\}$ and $K_* = \max\{\|\mathfrak{p}_i\|_\infty : i \in I_*\}$, we may define

$$\tilde{\mathfrak{p}}_i(\varphi) = \begin{cases} (1-\rho\,\gamma')^{-1}K_*^{-1}\mathfrak{p}_i(\varphi), & i \in I_*, \\ \mathfrak{p}_i(\varphi), & i \notin I_*, \end{cases}$$

so that $|\tilde{\mathfrak{p}}_i(\varphi)| \leq (1-\rho\,\gamma')^{-1}$ for $i = 1,\dots,n-1$ and hence, since $\|\tilde{\mathfrak{p}}_i - 1\|_{0,3} = O(\rho)$ if $\|\mathfrak{p}_i - 1\|_{0,3} = O(\rho)$, the functions $\tilde{\mathfrak{p}}_0,\dots,\tilde{\mathfrak{p}}_{n-1}$ satisfy the hypotheses of Lemma 4.10. Therefore, we can write $\mathfrak{p}_i(\varphi) = (1-\rho\,\gamma')\,K_*\tilde{\mathfrak{p}}_i(\varphi)$ for $i \in I_*$ and incorporate the factor $((1-\rho\,\gamma')\,K_*)^{n_*}$ into the constant C in (4.20).

The bounds for $\overline{h}$ now follow easily considering (4.19), reasoning like in Subsection 3.2.3 and using the bound (4.20) with $\mathfrak{p}_i = \overline{p}$ for $i = 0,\dots,n-1$. Invertibility of $\overline{\mathscr{H}}$ follows by the same argument used in the proof of Corollary 2.18 (see Subsection 3.2.4), and the bounds for the function $\overline{l}$ are easily obtained using the inverse function theorem.

4.6 The averaged and the continuous time system: proof of Lemma 2.28

Let $\Phi_t(\varphi)$ be the solution of (2.31). Observe first that

$$\Phi_\rho(\varphi) = \varphi + \rho\overline{f}(\varphi) + \rho^2\int_0^1 (1-t)\,\partial_\varphi\overline{f}(\Phi_{t\rho}(\varphi))\,\overline{f}(\Phi_{t\rho}(\varphi))\,\mathrm{d}t\,. \tag{4.24}$$

Moreover, reasoning as in Lemma 3.8 and in particular Remark 3.11, we get that for any $\gamma'' \in (0,\gamma)$ there exists a constant C independent of ρ such that, for all $n \geq 0$,

$$|\Phi_{n\rho}(\varphi)| \leq C(1-\rho\,\gamma'')^n, \qquad |(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi)| \leq C(1-\rho\,\gamma'')^n. \tag{4.25}$$

Then, using (2.26) and (4.24), we obtain, for some φ_n between $(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi)$ and $\Phi_{n\rho}(\varphi)$,

$$\left|(\overline{\mathscr{S}}^{n+1})_\varphi(\varphi,\psi) - \Phi_{(n+1)\rho}(\varphi)\right| \leq (1+\rho\partial_\varphi \overline{f}(\varphi_n))|(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) - \Phi_{n\rho}(\varphi)| + C\rho^2(1-\rho\,\gamma'')^n$$

where we have used (4.25) and the fact that $\overline{f}(0) = 0$ in order to bound $|\overline{f}(\Phi_{t\rho}(\varphi))| \leq C|\Phi_{t\rho}(\varphi))|$. Iterating we get, for suitable $\varphi_1,\dots,\varphi_{n-1}$,

$$|(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) - \Phi_{n\rho}(\varphi)| \leq C\rho^2 \sum_{i=0}^{n-1}(1-\rho\,\gamma'')^{n-1-i}\prod_{j=n-i}^{n}\left(1+\rho\,\partial_\varphi\overline{f}(\varphi_j)\right).$$

Let $\overline{N}$ be defined as in Remark 4.9. By the first of (4.25), there exists $\overline{M} \geq \overline{N}$, with $\overline{M} = O(\rho^{-1})$, such that both $(\overline{\mathscr{S}}^k)_\varphi(\varphi,\psi)$ and $\Phi_{k\rho}(\varphi)$ – and hence φ_k as well – are in $[S_m - r, S_M + r]$ for $k \geq M$, with r such that $|1+\rho\,\partial_\varphi\overline{f}(\varphi)| \leq C(1-\rho\gamma'')$ for all $\varphi \in [S_m - r, S_M + r]$. Thus, we get

$$\rho^2\sum_{i=0}^{n-1}(1-\rho\,\gamma'')^{n-1-i}\prod_{j=n-i}^{n}(1+\rho\,\partial_\varphi\overline{f}(\varphi_j)) \leq C\rho\left(\frac{1+\rho\|\partial_\varphi f\|_\infty}{1-\rho\,\gamma''}\right)^{\overline{M}} n\rho(1-\rho\,\gamma'')^n,$$

so that (2.32) follows by taking $\gamma'' \in (0,\gamma)$ and $\gamma' \in (0,\gamma'')$ such that $n\rho(1-\rho\,\gamma'')^n \leq (1-\rho\,\gamma')^n$. $\square$

4.7 Deviations of the conjugation: proof of Theorem 5

To compare $\mathscr{S}^n$ with $\overline{\mathscr{S}}^n$, according to (2.37) we need to control the deviations of h from $\overline{h}$ and of l with respect to $\overline{l}$. Since $\overline{G}(0) = 0$, the map $\overline{\mathscr{S}}$ admits the same invariant manifold $\overline{\mathcal{W}} := \{(0,\psi)\,|\,\psi \in \mathbb{T}^2\}$ as $\mathscr{S}_1$ (see Remarks 2.24 and 3.2). Thus, in order to complete the program outlined at the beginning of Subsection 2.4.2, it is more convenient to compare first h_1 with $\overline{h}$ and show that, in average, they are close to each other, and then show that the same happens when comparing h with h_1.

The domain Ω_1 of $\mathscr{S}_1(\varphi,\psi)$ and $h_1(\varphi,\psi)$ has the form considered in Remark 2.3, with $a_+(\psi) = \phi_M - W(\psi)$ and $a_-(\psi) = \phi_m - W(\psi)$. For $\langle h_1(\varphi,\cdot)\rangle$ to make sense we need $h_1(\varphi,\psi)$ to be defined for every ψ, but this happens only for $\varphi \in \Theta$ with Θ strictly contained in $\mathcal{U}$ (see Remark 3.4). Since we want to compare $\overline{h}(\varphi)$ with $\langle h(\varphi,\cdot)\rangle$ for all $\varphi \in \mathcal{U}$, we need to extend $h_1(\varphi,\psi)$ to the whole set Ω. One way to accomplish this is to extend both F and $\mathcal{H}$, as described in Subsection 4.1, to functions F_{ext} and $\mathcal{H}_{\text{ext}}$ defined on a larger domain Ω_{ext} such that the extended map $\mathscr{S}_{1,\text{ext}}(\theta,\psi)$ is defined for all $(\theta,\psi) \in \Omega_{\text{ext}}$. To this end we set $\Omega_{\text{ext}} := \mathcal{U}_{\text{ext}} \times \mathbb{T}^2$, with $\mathcal{U}_{\text{ext}} = [\phi_m + \min_{\psi\in\mathbb{T}^2} W(\psi), \phi_M + \max_{\psi\in\mathbb{T}^2} W(\psi)]$, so that we have $\Omega_{\text{ext}} \supset \{(\varphi,\psi) \in \mathbb{R}\times\mathbb{T}^2 : \phi_m + \min\{0,W(\psi)\} \leq \varphi \leq \phi_M + \max\{0,W(\psi)\}\}$. Then, when considering $h(\varphi,\psi) - \overline{h}(\varphi)$, we may write, for $\varphi \in \mathcal{U}$,

$$\begin{aligned} h(\varphi,\psi) &= h_{\text{ext}}(\varphi,\psi) = h_{\text{ext}}(\varphi+W(\psi),\psi) - \big(h_{\text{ext}}(\varphi+W(\psi),\psi) - h_{\text{ext}}(\varphi,\psi)\big) \\ &= h_{1,\text{ext}}(\varphi,\psi) - W(\psi)\,\partial_\varphi h_{\text{ext}}(\varphi,\psi) - (W(\psi))^2\int_0^1 \mathrm{d}t\,(1-t)\,\partial_\varphi^2 h_{\text{ext}}(\varphi - tW(\psi),\psi) \end{aligned} \tag{4.26}$$

and start by studying the average of $h_{1,\text{ext}}(\varphi,\psi) - \overline{h}(\varphi)$. The next step is to show, thanks to Theorem 4, that the average of the other terms appearing in (4.26) produce corrections of order ρ.

Remark 4.13. Although we use the extended maps along the proof of Theorem 6, the final result does not depend on the extension we take. This is due to the fact that the difference between Ω_1 and Ω has measure of order ρ and hence any extended map we may consider produces corrections which are at most of order ρ. A different choice could be $h_1(\varphi,\psi) = 0$ for $(\varphi,\psi) \in \Omega\setminus\Omega_1$; the advantage of taking $h_{1,\text{ext}}$ as in Subsection 4.1 is that it has the same regularity of the original h_1.

Remark 4.14. Throughout the rest of the section, we work with the extended functions, but we drop the subscript 'ext' not to overwhelm the notation. Since we first compare $\overline{h}$ with h_1, to avoid confusion, we call θ the first variable not only of $\mathscr{S}_1$ and h_1, but also of $\overline{h}$ and $\overline{\mathscr{S}}$. From the above discussion, it follows that, for all $\psi \in \mathbb{T}^2$, the range of the variable θ contains the whole interval $\mathcal{U}$. Only at the end, when comparing h with $\bar{h}$, we compute $\overline{h}$ and h_1 at $\theta = \varphi$.

The rest of the subsection is mostly devoted to the proof of the following proposition and some of its implications.

Proposition 4.15. *Let h_1 and $\overline{h}$ be as in* (3.14) *and in* (2.30)*, respectively. For all $\theta \in \mathcal{U}$, one has*

$$\left|\left\langle h_1(\theta,\cdot) - \overline{h}(\theta)\right\rangle\right| \le C\rho, \qquad \left|\left\langle (h_1(\theta,\cdot) - \overline{h}(\theta))^2\right\rangle\right| \le C\rho, \tag{4.27a}$$

$$\left|\left\langle \partial_\theta h_1(\theta,\cdot) - \partial_\theta\overline{h}(\theta)\right\rangle\right| \le C\rho, \qquad \left|\left\langle (\partial_\theta h_1(\theta,\cdot) - \partial_\theta\overline{h}(\theta))^2\right\rangle\right| \le C\rho. \tag{4.27b}$$

After proving Proposition 4.15, to complete the proof of Theorem 5 we need to reexpress h_1 in terms of h. This is done in the last Subsection 4.7.5.

4.7.1 The auxiliary map

Let r_2 be such that $\theta(f(\theta,\psi) - f(0,\psi)) < 0$ for $(\theta,\psi) \in \Lambda_{2r_2}$ (see Lemma 2.7 for notation) and let $\chi : \mathcal{U} \to \mathbb{R}$ be a C^∞ function such that $\chi(\theta) = 1$ for $\theta \in [S_m - r_2, S_M + r_2]$ and $\chi(\theta) = 0$ for $\theta \notin [S_m - 2r_2, S_M + 2r_2]$. We introduce the *auxiliary map*

$$\mathscr{S}_2(\theta,\psi) := (G_2(\theta,\psi), A_0\psi), \qquad G_2(\theta,\psi) := \theta + \rho f_2(\theta,\psi), \qquad f_2(\theta,\psi) := f(\theta,\psi) - \chi(\theta) f(0,\psi). \tag{4.28}$$

Remark 4.16. Instead of $\mathscr{S}_2$, one might like to consider the simpler map $\mathscr{S}_3(\theta,\psi) := (G_3(\theta,\psi), A_0\psi)$, with $G_3(\theta,\psi) := \theta + \rho f_3(\theta,\psi)$ and $f_3(\theta,\psi) := f(\theta,\psi) - f(0,\psi)$. However, it could happen that $f_3(\theta,\psi) = 0$ also for some $\theta \neq 0$, so that $\mathscr{S}_3$ would not satisfy Hypothesis 2 and $\overline{\mathcal{W}}$ might fail to be a global attractor for $(\Omega, \mathscr{S}_3)$; hence, in general, $\mathscr{S}_3$ could not be conjugated with $\mathscr{S}_0$. On the other hand, if one is willing to restrict the map $\mathscr{S}$ to a smaller set inside Ω, say the set Λ_{2r_2} defined above, then one can define $f_2(\theta,\psi) = f(\theta,\psi) - f(0,\psi)$, without introducing the function χ, and the corresponding map $\mathscr{S}_2$ satisfies all Hypotheses 1–3. The same goal might be achieved by assuming stronger hypotheses on the map $\mathscr{S}$, say by requiring it to be uniformly contracting along the direction of the slow variable on the whole Ω, but the conditions so introduced would be too restrictive and unnecessary for the result to hold.

If we set $F_1(\theta,\psi) = \rho f_1(\theta,\psi)$ in (3.11), with $f_1(\theta,\psi) := f(\theta + W(\psi),\psi) - f(W(\psi),\psi)$, we can write $f_1(\theta) = f_2(\theta,\psi) + \zeta(\theta,\psi)$, with

$$\zeta(\theta,\psi) := c_0(\theta)\,\mathfrak{f}(\psi) + c_1(\theta,\psi)\,W(\psi) + c_2(\theta,\psi)\,W(\psi)^2, \tag{4.29a}$$

$$c_0(\theta) := \chi(\theta) - 1, \tag{4.29b}$$

$$c_1(\theta,\psi) := \left(\partial_\theta f(\theta,\psi) - \chi(\theta)\,\partial_\theta f(0,\psi)\right), \tag{4.29c}$$

$$c_2(\theta,\psi) := \int_0^1 \mathrm{d}t\,(1-t)\left(\partial_\theta^2 f(\theta + tW(\psi),\psi) - \chi(\theta)\,\partial_\theta^2 f(tW(\psi),\psi)\right), \tag{4.29d}$$

$$\mathfrak{f}(\psi) := f(W(\psi),\psi) = \rho^{-1}(W(A_0\psi) - W(\psi)). \tag{4.29e}$$

where the identity (4.29e) derives from (3.1), with $F = \rho f$, and is used at length in the following.

Remark 4.17. The definition of the function χ ensures that $\mathscr{S}_2$ satisfies Hypotheses 1–3, provided ρ is such that $\rho\,\partial_\theta f_2(\theta,\psi) > -1$ for all $(\theta,\psi) \in \Omega$. Thus, if we wish to use the results in Subsections 2.3 and 3.2 with $\mathscr{S}_2$ in place of $\mathscr{S}$ we need to restrict ρ to an interval $(0,\rho_0)$ with $\rho_0 > 0$ possibly smaller than ρ_*, as defined in Remark 2.19. This is not a problem since we are mainly interested in the regime in which ρ tends to zero. Moreover, for any fixed $\rho_0 \in (0,\rho_*)$, the bounds in Theorem 5 become trivial for $\rho \in [\rho_0,\rho_*)$ by taking, if needed, larger values for the involved constants C (see also Remark 4.19).

Remark 4.18. By construction $\langle f_2(\theta,\cdot)\rangle = \langle f(\theta,\cdot)\rangle$ for $\theta\in\mathcal{U}$ and $\partial_\theta f_2(0,\psi)=\partial_\theta f(0,\psi)$. On the other hand, the fact that $f_1(\theta)=f_2(\theta,\psi)+\zeta(\theta,\psi)$ shows that $\mathscr{S}_2$ can be seen as a regularization of $\mathscr{S}_1$. In particular, one has $\mathscr{S}_2(0,\psi)=(0,A_0\psi)$, so that also the invariant manifold of $\mathscr{S}_2$ is given by $\overline{\mathcal{W}}=\{(0,\psi):\psi\in\mathbb{T}^2\}$, and it is thus the same as that of $\overline{\mathscr{S}}$ and $\mathscr{S}_1$ (see Remarks 2.24 and 3.2).

From Theorem 2, with $\mathscr{S}_2$ instead of $\mathscr{S}$, and Remark 4.18 it follows that there exists a set $\Omega_2\subset\mathbb{R}\times\mathbb{T}^2$ and a map $\mathscr{H}_2\colon\Omega\to\Omega_2$ of the form

$$\mathscr{H}_2(\theta,\psi):=(\mathcal{H}_2(\theta,\psi),\psi),\qquad \mathcal{H}_2(\theta,\psi):=\theta+\theta^2 h_2(\theta,\psi), \tag{4.30}$$

which conjugates $\mathscr{S}_2$ to its linearization $(\mu(\psi)\theta,A_0\psi)$, i.e. such that $\mu(\psi)\mathcal{H}_2(\theta,\psi)=\mathcal{H}_2(\mathscr{S}_2(\theta,\psi))$, with $\mu(\psi)$ defined after (4.8). Introducing the functions

$$p_2(\theta,\psi):=\frac{1}{1+\rho\partial_\theta f_2(0,\psi)}\left(\frac{\theta+\rho f_2(\theta,\psi)}{\theta}\right)^2, \tag{4.31a}$$

$$q_2(\theta,\psi):=\frac{\rho}{1+\rho\,\partial_\theta f_2(0,\psi)}\frac{f_2(\theta,\psi)-\partial_\theta f_2(0,\psi)\theta}{\theta^2}, \tag{4.31b}$$

we get

$$h_2(\theta,\psi):=\sum_{n=0}^{\infty}p_2^{(n)}(\theta,\psi)\,q_2(\mathscr{S}_2^n(\theta,\psi)),\qquad p_2^{(n)}(\theta,\psi):=\prod_{i=0}^{n-1}p_2(\mathscr{S}_2^i(\theta,\psi)). \tag{4.32}$$

In order to study the average of $h_1(\theta,\psi)-\overline{h}(\theta)$ and of its derivative, we split

$$h_1(\theta,\psi)-\overline{h}(\theta)=(h_1(\theta,\psi)-h_2(\theta,\psi))+\left(h_2(\theta,\psi)-\overline{h}(\theta)\right), \tag{4.33a}$$

$$\partial_\theta h_1(\theta,\psi)-\partial_\theta\overline{h}(\theta)=(\partial_\theta h_1(\theta,\psi)-\partial_\theta h_2(\theta,\psi))+\left(\partial_\theta h_2(\theta,\psi)-\partial_\theta\overline{h}(\theta)\right), \tag{4.33b}$$

and study separately the two contributions in each of (4.33a) and (4.33b). This is the content of Propositions 4.28, 4.29, 4.44, and 4.45 below, which combined immediately imply Proposition 4.15. For both maps $\mathscr{S}_1$ and $\mathscr{S}_2$, the θ-component vanishes at $\theta=0$. However, while $\mathscr{S}_1$ depends on W and hence inherits the low regularity of the invariant manifold, $\mathscr{S}_2$ has the same regularity as $\mathscr{S}$. Therefore, through the splitting (4.33), we control first the deviations of h_2 from $\overline{h}$ (see Subsections 4.7.2 and 4.7.3) using the regularity of $\mathscr{S}_2$ and the fact that $\mathscr{S}$ and $\mathscr{S}_2$ share the same averaged map $\overline{\mathscr{S}}$; next we show that the deviations of h_1 from h_2 are small thanks to (4.29a) and Theorem 4 (see Subsection 4.7.4).

Remark 4.19. For $\mathscr{S}_2$ to satisfy Hypotheses 1–3 we need to restrict the maximum value allowed for ρ to ρ_0 since the condition $\rho\,\partial_\theta f_2(\theta,\psi)>-1$ is more stringent than the condition $\rho\,\partial_\varphi f(\varphi,\psi)>-1$ (see Remark 4.17). However, since the bounds in Theorem 5 are trivially satisfied for any fixed ρ, by possibly taking a larger constant C_7, we may and do take for granted that the bounds hold for $\rho\geq\rho_0$. Thus, we confine ourselves to consider $\rho\in(0,\rho_0)$ and assume that both $\mathscr{S}$ and $\mathscr{S}_2$ satisfy Hypotheses 1–3.

4.7.2 A new correlation inequality

To fulfill the program outlined at the end of the previous section, we start by comparing h_2 with $\overline{h}$. To this aim, by using the expansion (4.32), we find useful to compare first $\langle {p_2}^{(n)}q_2\circ\mathscr{S}_2^n\rangle$ with $\langle p_2\rangle^{(n)}\langle q_2\rangle\circ\overline{G}^n$ (see (4.52) below). This comparison is similar to the comparison in Proposition 4.3, the main difference being that the analogues of g_+, $\mathfrak{u}$ and g_- now depend also on θ. Thus we need a correlation inequality that generalizes the previous one to this case.

The following preliminary result that bounds analogous to those in Lemma 4.10, which hold for functions depending only on the slow variable, extend to functions depending also on the fast variables, as far as the latter dependence is regular enough.

Lemma 4.20. *Let $\mathfrak{p}_0,\ldots,\mathfrak{p}_{n-1}$ be any functions in $\mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$ such that, for some $\gamma'\in(0,\gamma)$,*

1. *$|||\mathfrak{p}_i-1|||_{\alpha_0,3}=O(\rho)$ for all $i=0,\ldots,n-1$,*
2. *$|\mathfrak{p}_i(\theta)|\leq 1-\rho\,\gamma'$ for $|\theta|\leq\theta'$ for some θ' independent of ρ and for all $i=0,\ldots,n-1$,*

and set $\mathfrak{p}^{(n)}(\theta,\psi) := \mathfrak{p}^{(n)}(\mathscr{S}_2;\theta,\psi)$*, with* $\mathfrak{p}^{(n)}(\mathscr{S}_2;\theta,\psi)$ *defined according to* (2.15)*. One has*

$$|||\mathfrak{p}^{(n)}|||^-_{\alpha_0,3} \le C\,(1-\rho\,\gamma')^n, \qquad |||\partial_\theta(\mathscr{S}_2^n)_\theta|||^-_{\alpha_0,4} \le C\,(1-\rho\,\gamma')^n, \tag{4.34}$$

where the constant C *does not depend on* n*.*

Proof. Reasoning as in the proof of Lemma 4.10, we find $|||\mathfrak{p}^{(n)}|||_{0,3} \le C(1-\rho\,\gamma')^n$. Since $\partial_\theta(\mathscr{S}_2^n)_\theta = (\partial_\theta G_2)^{(n)}$, where $(\partial_\theta G_2)^{(n)}$ is given by (2.15) with $\mathfrak{p}_i = \partial_\theta G_2$ for all $i=0,\dots,n-1$, and $(\mathscr{S}_2)_\theta(0,\psi)=0$, we also have

$$\|\partial_\theta^k(\mathscr{S}_2^n)_\theta\|_\infty \le C(1-\rho\,\gamma')^n \qquad k=0,1,2,3. \tag{4.35}$$

On the other hand, writing

$$\begin{aligned} G_2(\mathscr{S}_2^n(\theta,\psi)) - G_2(\mathscr{S}_2^n(\theta,\psi')) &= G_2((\mathscr{S}_2^n)_\theta(\theta,\psi), A_0^n\psi) - G_2((\mathscr{S}_2^n)_\theta(\theta,\psi'), A_0^n\psi) \\ &\quad + (\mathscr{S}_2^n)_\theta(\theta,\psi') \int_0^1 (\partial_\theta G_2(t(\mathscr{S}_2^n)_\theta(\theta,\psi'), A_0^n\psi) - \partial_\theta G_2(t(\mathscr{S}_2^n)_\theta(\theta,\psi'), A_0^n\psi'))\,\mathrm{d}t, \end{aligned}$$

we have, for $n\ge 2$,

$$|(\mathscr{S}_2^n)_\theta|^-_{\alpha_0} = |G_2\circ\mathscr{S}_2^{n-1}|^-_{\alpha_0} \le \|(\partial_\theta G_2)\circ\mathscr{S}_2^{n-1}\|_\infty |G_2\circ\mathscr{S}_2^{n-2}|^-_{\alpha_0} + (1-\rho\,\gamma')^{n-1}\lambda^{-\alpha_0(n-1)}|\partial_\theta G_2|^-_{\alpha_0},$$

so that, iterating, thanks to the condition assumed on ρ, we get

$$|(\mathscr{S}_2^n)_\theta|^-_{\alpha_0} \le C|||G_2|||_{\alpha_0,1}(1-\rho\,\gamma')^{n-1}\sum_{i=0}^{n-1}\lambda^{-\alpha_0(n-1-i)} \le C(1-\rho\,\gamma')^n\,, \tag{4.36}$$

which holds true also for $n=1$. We can now write, by (2.6b),

$$|\mathfrak{p}^{(n)}|^-_{\alpha_0} \le \sum_{i=0}^{n-1}|\mathfrak{p}_i\circ\mathscr{S}_2^i|^-_{\alpha_0}\prod_{\substack{j=0\\ j\neq i}}^{n-1}\|\mathfrak{p}_i\circ\mathscr{S}_2^j\|_\infty,$$

where $|\mathfrak{p}_i\circ\mathscr{S}_2^i|^-_{\alpha_0} \le \|\partial_\theta\mathfrak{p}_i\|_\infty|(\mathscr{S}_2^i)_\theta|^-_{\alpha_0} + \lambda^{-\alpha_0 i}|\mathfrak{p}_i|_{\alpha_0}$, with the first term missing if $i=0$, so that we get

$$|\mathfrak{p}^{(n)}|^-_{\alpha_0} \le C(1-\rho\,\gamma')^{n-1}\left(\rho\sum_{i=1}^{n-1}(1-\rho\,\gamma')^i + \rho\sum_{i=0}^{n-1}\lambda^{-\alpha_0 i}\right) \le C(1-\rho\,\gamma')^n, \tag{4.37}$$

which implies, taking $\mathfrak{p}_i = \partial_\theta G_2\ \forall i=0,\dots,n-1$, also the bound

$$|\partial_\theta(\mathscr{S}_2^n)_\theta|^-_{\alpha_0} \le C(1-\rho\,\gamma')^n. \tag{4.38}$$

Analogously, using the expression (2.18b) for $\partial_\theta^2\mathfrak{p}^{(n)}(\theta)$ and the bound (4.35), we get

$$\begin{aligned} |\partial_\theta\mathfrak{p}^{(n)}|^-_{\alpha_0} &\le \sum_{i=0}^{n-1}\Big(|\partial_\theta\mathfrak{p}_i\circ\mathscr{S}_2^i|^-_{\alpha_0}\|\partial_\theta(\mathscr{S}_2^i)_\theta\|_\infty + \|\partial_\theta\mathfrak{p}_i\circ\mathscr{S}_2^i\|_\infty|\partial_\theta(\mathscr{S}_2^i)_\theta|^-_{\alpha_0}\Big)\prod_{\substack{j=0\\ j\neq i}}^{n-1}\|\mathfrak{p}_j\circ\mathscr{S}_2^j\|_\infty \\ &\quad + \sum_{\substack{i,j=0\\ i\neq j}}^{n-1}\|\partial_\theta\mathfrak{p}_i\circ\mathscr{S}_2^i\|_\infty\|\partial_\theta(\mathscr{S}_2^i)_\theta\|_\infty|\mathfrak{p}_j\circ\mathscr{S}_2^j|^-_{\alpha_0}\prod_{\substack{k=0\\ k\neq i,j}}^{n-1}\|\mathfrak{p}_k\circ\mathscr{S}_2^k\|_\infty, \end{aligned}$$

that, by the same argument used in (4.37), delivers

$$|\partial_\theta\mathfrak{p}^{(n)}|^-_{\alpha_0} \le C\,(1-\rho\,\gamma')^n, \qquad |\partial_\theta^2(\mathscr{S}_2^n)_\theta|^-_{\alpha_0} \le C\,(1-\rho\,\gamma')^n. \tag{4.39}$$

It is now easy, by using the expression (2.18c) for $\partial_\theta^3 p^{(n)}(\theta)$ and the bounds (4.35)–(4.39), and reasoning once more as in in (4.37), to obtain

$$|\partial_\theta^2\mathfrak{p}^{(n)}|^-_{\alpha_0} \le C\,(1-\rho\,\gamma')^n, \tag{4.40}$$

so as to complete the proof of the first bound in (4.34). Finally the second bound in (4.34) follow collecting together the bounds (4.35)–(4.40). □

Remark 4.21. Condition 1 in Lemma 4.20 implies that $|\mathfrak{p}_i|_{\alpha_0}$, $|||\partial_\theta \mathfrak{p}_i|||_{\alpha_0,2}$ and $\|\mathfrak{p}_i - \langle \mathfrak{p}_i\rangle\|_\infty$ are $O(\rho)$.

Remark 4.22. A comment analogous to Remark 4.12 applies also to Lemma 4.20 and the forthcoming Proposition 4.40: if we assume that property 2 in Lemma 4.20 holds for all $i = 0, \ldots, n-1$ except at most n_* values, with $n_* < n$ independent of n, then both results still hold.

We are now ready to discuss the new correlation inequality involving the average

$$\langle g_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \rangle - \langle g_+ \rangle \langle \mathfrak{p} \rangle^{(n)} \langle g_- \rangle \circ \overline{G}^n. \tag{4.41}$$

Before stating and proving the result in the general case, we start by introducing some notation and discussing some preliminary results akin to the Doeblin-Fortet (or Lasota-Yorke) inequalities.

Define $\mathfrak{p}^{(k)}_{[i]} := \mathfrak{p}^{(k)}_{[i]}(\mathscr{S}_2; \theta, \psi)$ and $\langle \mathfrak{p}\rangle^{(k)}_{[i]} := \langle \mathfrak{p}\rangle^{(k)}_{[i]}(\overline{\mathscr{S}}; \theta)$, for $k = 0, \ldots, n-1$ and $i = 0, \ldots, n-k$. according to (2.15) and (2.17), respectively, and set $\mathfrak{p}^{(n)} = \mathfrak{p}^{(n)}_{[0]}$ and $\langle \mathfrak{p}\rangle^{(n)} = \langle \mathfrak{p}\rangle^{(n)}_{[0]}$. Observe that

$$\begin{aligned}
&\mathfrak{p}^{(n)}(\theta,\psi)\, g_-(\mathscr{S}_2^n(\theta,\psi)) - \langle \mathfrak{p}\rangle^{(n)}(\theta)\, g_-(\overline{\mathscr{S}}^n(\theta,\psi)) \\
&\qquad = \sum_{i=0}^{n-1} \Big(\langle \mathfrak{p}\rangle^{(i)}(\theta)\, \mathfrak{p}^{(n-i)}_{[i]}(\overline{\mathscr{S}}^i(\theta,\psi))\, g_-(\mathscr{S}_2^{n-i}(\overline{\mathscr{S}}^i(\theta,\psi))) \\
&\qquad\quad - \langle \mathfrak{p}\rangle^{(i+1)}(\theta)\, \mathfrak{p}^{(n-i-1)}_{[i+1]}(\overline{\mathscr{S}}^{i+1}(\theta,\psi))\, g_-(\mathscr{S}_2^{n-i-1}(\overline{\mathscr{S}}^{i+1}(\theta,\psi))) \Big).
\end{aligned} \tag{4.42}$$

Moreover we have $\mathfrak{p}^{(n-i)}_{[i]}(\overline{\mathscr{S}}^i(\theta,\psi))\, g_-(\mathscr{S}_2^{n-i}(\overline{\mathscr{S}}^i(\theta,\psi))) = \mathfrak{p}^{(n-i)}_{[i]}(\overline{G}^i(\theta), A_0^i\psi)\, g_-(\mathscr{S}_2^{n-i}(\overline{G}^i(\theta), A_0^i\psi))$, that is $\langle \mathfrak{p}\rangle^{(i)}(\theta)\, \mathfrak{p}^{(n-i)}_{[i]}(\overline{\mathscr{S}}^i(\theta,\psi))\, g_-(\mathscr{S}_2^{n-i}(\overline{\mathscr{S}}^i(\theta,\psi)))$ depends on ψ only through $\psi' := A_0^i\psi$; a similar consideration holds for the second term inside the summation in (4.42). Defining

$$\Delta_i^{(k)}(\theta,\psi) := \mathfrak{p}^{(k)}_{[i]}(\theta,\psi)\, g_-(\mathscr{S}_2^k(\theta,\psi)) - \langle \mathfrak{p}_i\rangle(\theta)\, \mathfrak{p}^{(k-1)}_{[i+1]}(\overline{G}(\theta), A_0\psi)\, g_-(\mathscr{S}_2^{k-1}(\overline{G}(\theta), A_0\psi)), \tag{4.43}$$

we get

$$\mathfrak{p}^{(n)}(\theta,\psi)\, g_-(\mathscr{S}_2^n(\theta,\psi)) - \langle \mathfrak{p}\rangle^{(n)}(\theta)\, g_-(\overline{\mathscr{S}}^n(\theta,\psi)) = \sum_{i=0}^{n-1} \langle \mathfrak{p}\rangle^{(i)}(\theta)\, \Delta_i^{(n-i)}(\overline{\mathscr{S}}^i(\theta,\psi)). \tag{4.44}$$

We start with a particular case, by assuming the function g_+ in (4.41) to have zero average. Eventually we extend the result to any $g_+ \in \mathscr{B}^+_{\alpha_0}(\Omega,\mathbb{R})$.

Lemma 4.23. *Assume ρ to be such that the map $\mathscr{S}_2$ satisfies Hypotheses 1–3. Let $\mathfrak{p}_0, \ldots, \mathfrak{p}_{n-1}$ be any functions in $\mathscr{B}_{0,3}(\Omega,\mathbb{R})$ such that, for some $\rho' \in (0,1)$,*

1. *$\|\mathfrak{p}_i - 1\|_\infty = O(\rho)$ for all $i = 0, \ldots, n-1$,*
2. *$|\mathfrak{p}_i(\theta)| \le 1 - \rho\gamma'$ for $|\theta| \le \theta'$ for some $\theta' = O(1)$ in ρ and for all $i = 0, \ldots, n-1$,*

and set $\mathfrak{p}^{(n)}(\theta,\psi) = \mathfrak{p}^{(n)}(\mathscr{S}_2; \theta,\psi)$, with $\mathfrak{p}^{(n)}(\mathscr{S}_2;\theta,\psi)$ as in (2.15). For any $g_+ \in \mathscr{B}^+_{\alpha_0}(\Omega,\mathbb{R})$ with $\langle g_+\rangle(\theta) = 0$ and any $g_- \in \mathscr{B}^-_{\alpha_0,2}(\Omega,\mathbb{R})$ one has

$$\left| \left\langle g_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle \right| \le C(1-\rho\,\gamma')^n \left(\lambda^{-\alpha_0 n} \|g_+\|^+_{\alpha_0} \|\tilde{g}_-\|^-_{\alpha_0} + \rho\, \|g_+\|^+_{\alpha_0} |||g_-|||^-_{\alpha_0,2} \right),$$

where the constant C does not depends on n.

Proof. In (4.44), $\Delta_i^{(n-i)}(\overline{\mathscr{S}}^i(\theta,\psi)) = \Delta_i^{(n-i)}(\overline{G}^i(\theta), A_0\psi))$ depends on ψ only through $\psi' := A_0^i\psi$. Thus, using Proposition 2.4, we bound

$$\begin{aligned}
&\left| \left\langle g_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle \right| \le \left| \left\langle g_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle - \langle \mathfrak{p}\rangle^{(n)} \left\langle g_+ g_- \circ \overline{\mathscr{S}}^n \right\rangle \right| + \left| \langle \mathfrak{p}\rangle^{(n)} \right| \left| \left\langle g_+\, g_- \circ \overline{\mathscr{S}}^n \right\rangle \right| \\
&\quad \le C \sum_{i=0}^{n-1} \|\mathfrak{p}^{(i)}\|_\infty (1+\alpha_0 i)\, \lambda^{-\alpha_0 i} \|g_+\|^+_{\alpha_0} \|\Delta_i^{(n-i)}\|^-_{\alpha_0} + C\|\langle \mathfrak{p}\rangle^{(n)}\|_\infty (1+\alpha_0 n)\, \lambda^{-\alpha_0 n} \|g_+\|^+_{\alpha_0} \|\tilde{g}_-\|^-_{\alpha_0}.
\end{aligned} \tag{4.45}$$

Moreover, $\mathscr{S}_2(\theta,\psi) = (\overline{G}(\theta) + \rho\Delta f(\theta,\psi), A_0\psi)$, with $\Delta f(\theta,\psi) := \tilde{f}(\theta,\psi) - \chi(\theta)\, f(0,\psi)$, so that we have

$$\begin{aligned}\Delta_i^{(k)}(\theta,\psi) &= \big(\mathfrak{p}_i(\theta,\psi) - \langle \mathfrak{p}_i\rangle(\theta)\big)\mathfrak{p}_{[i+1]}^{(k-1)}(\mathscr{S}_2(\theta,\psi))\, g_-(\mathscr{S}_2^k(\theta,\psi)) \\ &\quad + \langle \mathfrak{p}_i\rangle(\theta)\Big(\mathfrak{p}_{[i+1]}^{(k-1)}(\overline{G}(\theta) + \rho\Delta f(\theta,\psi), A_0\psi) - \mathfrak{p}_{[i+1]}^{(k-1)}(\overline{G}(\theta), A_0\psi)\Big) g_-(\mathscr{S}_2^k(\theta,\psi)) \\ &\quad + \langle \mathfrak{p}_i\rangle(\theta)\, \mathfrak{p}_{[i+1]}^{(k-1)}(\overline{G}(\theta), A_0\psi)\big(g_-(\mathscr{S}_2^{k-1}(\overline{G}(\theta,\psi) + \rho\Delta f(\theta,\psi)) - g_-(\mathscr{S}_2^{k-1}(\overline{G}(\theta), A_0\psi))\big).\end{aligned}$$

Thus, writing

$$\begin{aligned}&\mathfrak{p}_{[i+1]}^{(k-1)}(\overline{G}(\theta) + \rho\Delta f(\theta,\psi), A_0\psi) - \mathfrak{p}_{[i+1]}^{(k-1)}(\overline{G}(\theta), A_0\psi) \\ &\qquad = \rho\Delta f(\theta,\psi)\int_0^1 \mathrm{d}t\, \partial_\theta \mathfrak{p}_{[i+1]}^{(k-1)}(\overline{G}(\theta) + t\rho\Delta f(\theta,\psi), A_0\psi), \\ &g_-(\mathscr{S}_2^{k-1}(\overline{G}(\theta) + \rho\Delta f(\theta,\psi)) - g_-(\mathscr{S}_2^{k-1}(\overline{G}(\theta), A_0\psi)) \\ &\qquad = \rho\Delta f(\theta,\psi)\int_0^1 \mathrm{d}t\, \big((\partial_\theta g_- \circ \mathscr{S}_2^{k-1})(\partial_\theta(\mathscr{S}_2^{k-1})_\theta)\big)\, (\overline{G}(\theta) + t\rho\Delta f(\theta,\psi), A_0\psi),\end{aligned} \tag{4.46}$$

and using that $|\partial_\theta g_- \circ \mathscr{S}_2^k|_{\alpha_0}^- \leq \|\partial_\theta^2 g_- \circ \mathscr{S}_2^k\|_\infty |(\mathscr{S}_2^k)_\theta|_{\alpha_0}^- + \lambda^{-\alpha_0 k}|\partial_\theta g_-|_{\alpha_0}^-$, the assumptions on the functions $\mathfrak{p}_0, \ldots, \mathfrak{p}_{n-1}$ and the bounds provided in Lemma 4.20 and Remark 4.21 provide the bound $\|\Delta_i^{(k)}\|_{\alpha_0}^- \leq C\rho(1-\rho\,\gamma')^{k-1}|||g_-|||_{\alpha_0,2}^-$, which, inserted into (4.45), completes the proof. □

Remark 4.24. If g_- is in $\mathscr{B}_{\alpha_0,3}^-(\Omega,\mathbb{R})$, then, using (4.46) and still relying on Lemma 4.20 in the same way as we obtained the bound on $\|\Delta_i^{(k)}\|_{\alpha_0}^-$, we find that $|||\Delta_i^{(k)}|||_{\alpha_0,2}^- \leq C\rho(1-\rho\,\gamma')^{k-1}|||g_-|||_{\alpha_0,3}^-$.

Lemma 4.25. *Assume ρ to be such that the map $\mathscr{S}_2$ satisfies Hypotheses 1–3. Let $\mathfrak{p}_0, \ldots, \mathfrak{p}_{n-1}$ be any functions in $\mathscr{B}_{0,3}(\Omega,\mathbb{R})$ verifying the hypotheses in Lemma 4.23, and define $\langle\mathfrak{p}\rangle^{(n)}$ as in* (2.15) *and $\Xi^{(k)}(\theta,\psi) := \langle\mathfrak{p}\rangle^{(k)}(\theta)\, g_-(\overline{\mathscr{S}}^k(\theta,\psi))$. For any $g_- \in \mathscr{B}_{\alpha_0,2}^-(\Omega,\mathbb{R})$ one has*

$$\left|\left\langle \Xi^{(k)}\circ\mathscr{S}_2 - \Xi^{(k)}\circ\overline{\mathscr{S}}\right\rangle\right| \leq C\rho(1-\rho\,\gamma')^k\big((1+\alpha_0 k)\,\lambda^{-\alpha_0 k}|||g_-|||_{\alpha_0,2}^- + \rho|||g_-|||_{0,2}\big)\,.$$

Proof. Writing

$$\begin{aligned}&\Xi^{(k)}(\mathscr{S}_2(\theta,\psi)) - \Xi^{(k)}(\overline{\mathscr{S}}(\theta,\psi)) \\ &\quad = \rho\Delta f(\theta,\psi)\,\partial_\theta\Xi^{(k)}(\overline{\mathscr{S}}(\theta,\psi)) + (\rho\Delta f(\theta,\psi))^2\int_0^1 \mathrm{d}t\,(1-t)\,\partial_\theta^2\Xi^{(k)}(\overline{G}(\theta,\psi) + t\rho\Delta f(\theta,\psi), A_0\psi)\end{aligned}$$

and observing that $\Xi^{(k)}(\overline{\mathscr{S}}(\theta,\psi)) = \langle\mathfrak{p}\rangle^{(k)}(\overline{G}(\theta))\, g_-(\overline{G}^{k+1}(\theta), A_0^{k+1}\psi))$, we get, by using Lemmas 4.10 and 4.23, $|\langle\Delta f\,\partial_\theta\Xi^{(k)}\circ\overline{\mathscr{S}}\rangle| \leq C(1+\alpha_0 k)\,\lambda^{-\alpha_0 k}\|\Delta f\|_{\alpha_0}(1-\rho\,\gamma')^k|||g_-|||_{\alpha_0,2}^-$, while, using (2.8), we find $\|\partial_\theta^2\Xi^{(k)}\|_\infty \leq C|||\langle\mathfrak{p}\rangle^{(k)}|||_{0,2}\,|||g_-\circ\overline{\mathscr{S}}^{(k)}|||_{0,2} \leq C(1-\rho\,\gamma')^k|||g_-|||_{0,2}$, again by Lemma 4.10. Combining the above estimates we obtain the desired bound. □

Finally, we can discuss how to bound (4.41) in full generality. This leads to the following new correlation inequality.

Proposition 4.26. *Let $\mathfrak{p}_0, \ldots, \mathfrak{p}_{n-1}$ be any functions in $\mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$ satisfying, for some $\gamma' \in (0,\gamma)$, properties 1 and 2 in Lemma 4.20. For any $g_+ \in \mathscr{B}_{\alpha_0}^+(\Omega,\mathbb{R})$ and $g_- \in \mathscr{B}_{\alpha_0,3}^-(\Omega,\mathbb{R})$ one has*

$$\begin{aligned}&\left|\left\langle g_+\mathfrak{p}^{(n)} g_-\circ\mathscr{S}_2^n\right\rangle - \langle g_+\rangle\langle\mathfrak{p}\rangle^{(n)}\langle g_-\rangle\circ\overline{G}^n\right| \\ &\quad \leq C\,(1-\rho\,\gamma')^n\,\Big((1+\alpha_0 n)\lambda^{-\alpha_0 n}\|\tilde{g}_+\|_{\alpha_0}^+\|\tilde{g}_-\|_{\alpha_0}^- + \rho\|g_+\|_{\alpha_0}^+|||g_-|||_{\alpha_0,2}^- + \rho^2 n\|\langle g_+\rangle\|_\infty|||g_-|||_{\alpha_0,3}^-\Big),\end{aligned}$$

where $\mathfrak{p}^{(n)}(\theta,\psi) := \mathfrak{p}^{(n)}(\mathscr{S}_2;\theta,\psi)$ and $\langle\mathfrak{p}\rangle^{(n)}(\theta) := \langle\mathfrak{p}\rangle^{(n)}(\overline{\mathscr{S}};\theta)$, with $\mathfrak{p}^{(n)}(\mathscr{S}_2;\theta,\psi)$ and $\langle\mathfrak{p}\rangle^{(n)}(\overline{\mathscr{S}};\theta)$ defined according to (4.34) *and* (2.17), *respectively, and C is a constant independent of n.*

Proof. Observe that $\langle g_+\mathfrak{p}^{(n)}g_-\circ\mathscr{S}_2^n\rangle=\langle g_+\rangle\langle\mathfrak{p}^{(n)}g_-\circ\mathscr{S}_2^n\rangle+\langle\tilde g_+\mathfrak{p}^{(n)}g_-\circ\mathscr{S}_2^n\rangle$. The second term is bounded using Lemma 4.23. If we write the first term using (4.44), we need to estimate $\langle{\Delta_i}^{(n-i)}\rangle$. We write, for any $k\geq 1$,

$$\begin{aligned}\Delta_i^{(k)} &= (\mathfrak{p}_i-\langle\mathfrak{p}_i\rangle)\Big(\mathfrak{p}_{[i+1]}^{(k-1)}\circ\mathscr{S}_2\,g_-\circ\mathscr{S}_2^k-\mathfrak{p}_{[i+1]}^{(k-1)}\circ\overline{\mathscr{S}}\,g_-\circ\mathscr{S}_2^{k-1}\circ\overline{\mathscr{S}}\Big)\\ &\quad+(\mathfrak{p}_i-\langle\mathfrak{p}_i\rangle)\,\mathfrak{p}_{[i+1]}^{(k-1)}\circ\overline{\mathscr{S}}\,g_-\circ\mathscr{S}_2^{k-1}\circ\overline{\mathscr{S}}+\langle\mathfrak{p}_i\rangle\Big(\mathfrak{p}_{[i+1]}^{(k-1)}\circ\mathscr{S}_2\,g_-\circ\mathscr{S}_2^k-\mathfrak{p}_{[i+1]}^{(k-1)}\circ\overline{\mathscr{S}}\,g_-\circ\mathscr{S}_2^{k-1}\circ\overline{\mathscr{S}}\Big).\end{aligned}\tag{4.47}$$

We have $\|(\mathfrak{p}_i(\theta,\psi)-\langle\mathfrak{p}_i\rangle(\theta))\|_\infty\leq C\rho$, while, reasoning like when studying (4.46), we get

$$\|\mathfrak{p}_{[i+1]}^{(k-1)}(\mathscr{S}_2(\theta,\psi))\,g_-(\mathscr{S}_2^k(\theta,\psi))-\mathfrak{p}_{[i+1]}^{(k-1)}(\overline{\mathscr{S}}(\theta,\psi))\,g_-(\mathscr{S}_2^{k-1}(\overline{\mathscr{S}}(\theta,\psi)))\|_\infty\leq C\rho(1-\rho\,\gamma')^k|||g_-|||_{0,1},$$

so that the average of the first contribution in the r.h.s. of (4.47) is bounded by $C\rho^2(1-\rho\,\gamma')^k|||g_-|||_{0,1}$. A similar bound for the second contribution is obtained by using (4.46), as in the proof of Lemma 4.23, with the gain of a further factor ρ because of the factor $\mathfrak{p}_i-\langle\mathfrak{p}_i\rangle$. As to the last contribution, we use (4.44) twice and write

$$\begin{aligned}\mathfrak{p}_{[i+1]}^{(k-1)}\circ\overline{\mathscr{S}}\,g_-\circ\mathscr{S}_2^{k-1}\circ\overline{\mathscr{S}} &= \sum_{j=0}^{k-2}\Big(\langle\mathfrak{p}\rangle_{[i+1]}^{(j)}\circ G_2\Delta_j^{(k-1-j)}\circ\overline{\mathscr{S}}^j\circ\mathscr{S}_2-\langle\mathfrak{p}\rangle_{[i+1]}^{(j)}\circ\overline{G}\,\Delta_j^{(k-1-j)}\circ\overline{\mathscr{S}}^{j+1}\Big)\\ &\quad+\langle\mathfrak{p}\rangle_{[i+1]}^{(k-1)}\circ G_2\,g_-\circ\overline{\mathscr{S}}^{k-1}\circ\mathscr{S}_2-\langle\mathfrak{p}\rangle_{[i+1]}^{(k-1)}\circ\overline{G}\,g_-\circ\overline{\mathscr{S}}^k.\end{aligned}$$

Thus, Lemma 4.25, applied to the last line, gives

$$\big|\langle\langle\mathfrak{p}\rangle_{[i+1]}^{(k-1)}\circ G_2\,g_-\circ\overline{\mathscr{S}}^{k-1}\circ\mathscr{S}_2-\langle\mathfrak{p}\rangle_{[i+1]}^{(k-1)}\circ\overline{G}\,g_-\circ\overline{\mathscr{S}}^k\rangle\big|\leq C\rho(1-\rho\,\gamma')^k\big((1+\alpha_0k)\lambda^{-\alpha_0k}|||g_-|||^-_{\alpha_0,2}+\rho|||g_-|||_{0,2}\big),$$

and, applied to the second line, with $g_-=\Delta_j^{(k-1-j)}$, gives

$$\begin{aligned}&\big|\langle\langle\mathfrak{p}\rangle_{[i+1]}^{(j)}\circ G_2\,\Delta_j^{(k-1-j)}\circ\overline{\mathscr{S}}^j\circ\mathscr{S}_2-\langle\mathfrak{p}\rangle_{[i+1]}^{(j)}\circ\overline{G}\,\Delta_j^{(k-1-j)}\circ\overline{\mathscr{S}}^{j+1}\rangle\big|\\ &\qquad\leq C\rho(1-\rho\,\gamma')^j\big((1+\alpha_0j)\lambda^{-\alpha_0j}|||\Delta_j^{(k-1-j)}|||^-_{\alpha_0,2}+\rho|||\Delta_j^{(k-1-j)}|||^-_{0,2}\big)\,,\end{aligned}$$

where $|||\Delta_j^{(k-1-j)}|||^-_{\alpha_0,2}$ and hence also $|||\Delta_j^{(k-1-j)}|||^-_{0,2}$ are bounded as discussed in Remark 4.24. This concludes the proof. □

Remark 4.27. Proposition 4.26 can be seen as a generalization of Proposition 4.3 to functions which also depend smoothly on the slow variable. This will be exploited in Subsection 4.7.3 to compare the averaged map with the auxiliary map, by using that all the involved functions are regular – i.e. at least α_0-Hölder continuous for some α_0 independent of ρ – in the fast variable. The next step, to be achieved in Subsection 4.7.4, is to compare the auxiliary map with the translated map, where the dependence on the fast variable is only α_*-Hölder continuous, with $\alpha_*=O(\rho)$.

4.7.3 Deviations of the conjugation of the auxiliary map

As discussed in Remark 4.19, we assume, without loss of generality, that $\rho\leq\rho_0$ and hence that $\mathscr{S}_2$ satisfies Hypotheses 1–3. Recall also that we are working with the extension of $\mathscr{S}$, although not explicitly indicated (see Remark 4.14); thus, all functions appearing in what follows refer to such an extension.

We start by comparing h_2 with $\overline{h}$. By relying on the expansions (4.32) and (4.19), we may write

$$\langle p_2\rangle^{(n)}\langle q_2\rangle\circ\overline{G}^n-\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n=\langle q_2\rangle\circ\overline{G}^n\sum_{i=0}^{n-1}\langle p_2\rangle^{(i)}\big(\langle p_2\rangle-\overline{p}\big)\circ\overline{G}^i\,\overline{p}^{(n-i-2)}\circ\overline{G}^{i+1}+\overline{p}^{(n)}\big(\langle q_2\rangle\circ\overline{G}^n-\overline{q}\circ\overline{G}^n\big).\tag{4.48}$$

We can now prove the following result.

Proposition 4.28. *Let h_2 and $\overline{h}$ be defined according to* (4.30) *and* (2.30)*, respectively. For all $\theta \in \mathcal{U}$, one has*

$$\bigl|\bigl\langle h_2(\theta,\cdot)-\overline{h}(\theta)\bigr\rangle\bigr| \le C\rho, \qquad \bigl\langle (h_2(\theta,\cdot)-\overline{h}(\theta))^2\bigr\rangle \le C\rho\,. \tag{4.49}$$

Proof. Observe that $p_2(\theta,\psi)=1+O(\rho)$, so that (compare with Remark 4.21)

$$\|p_2-\langle p_2\rangle\|_\infty \le C\rho, \qquad |p_2|_{\alpha_0}\le C\rho, \qquad |||\partial_\theta p_2|||_{\alpha_0,2}\le C\rho\,. \tag{4.50}$$

Since $p_2(0,\psi)=1+\rho\partial_\theta f_2(0,\psi)\le 1-\rho$, for any $\rho'\in(1,\rho)$ there exists θ_2 such that $p_2(\theta,\psi)\le 1-\rho'$ for $|\theta|\le\theta_2$. Moreover, we easily check that

$$|||q_2|||_{\alpha_0,3}\le C\rho, \tag{4.51}$$

so that we can apply Proposition 4.26 and obtain

$$\bigl|\langle p_2^{(n)} q_2\circ\mathscr{S}_2^n\rangle - \langle p_2\rangle^{(n)}\langle q_2\rangle\circ\overline{G}^n\bigr| \le C(1-\rho\,\gamma')^n\rho^2\,(1+\rho n), \tag{4.52}$$

where $\langle p_2\rangle^{(n)}(\theta)$ is given by (2.17), with $\mathfrak{p}_i=p_2$ $\forall i=0,\dots,n-1$. Since

$$|||\langle p_2\rangle-\overline{p}|||_{0,1}\le C\rho^2, \qquad |||\langle q_2\rangle-\overline{q}|||_{0,1}\le C\rho^2, \tag{4.53}$$

we have also, by (4.48),

$$\bigl|\langle p_2\rangle^{(n)}\langle q_2\rangle\circ\overline{G}^n-\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n\bigr| \le C(1-\rho\,\gamma')^n\left(n\rho^2\|q_2\|_\infty+\|\langle q_2\rangle-\overline{q}\|_\infty\right)\le C(1-\rho\,\gamma')^n\rho^2(1+\rho n), \tag{4.54}$$

so that, combining (4.52) and (4.54), we obtain $|\langle p_2^{(n)}q_2\circ\mathscr{S}_2^n\rangle-\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n|\le C(1-\rho\,\gamma')^n\rho^2(1+\rho n)$. Summing over n gives the first bound in (4.49).

For the second bound, we start considering

$$\begin{aligned}
&\bigl(p_2^{(n_1)}q_2\circ\mathscr{S}_2^{n_1}-\langle p_2\rangle^{(n_1)}\langle q_2\rangle\circ\overline{G}^{n_1}\bigr)\bigl(p_2^{(n_2)}q_2\circ\mathscr{S}_2^{n_2}-\langle p_2\rangle^{(n_2)}\langle q_2\rangle\circ\overline{G}^{n_2}\bigr)\\
&\quad= p_2^{(n_1)}q_2\circ\mathscr{S}_2^{n_1}\,p_2^{(n_2)}q_2\circ\mathscr{S}_2^{n_2}-\langle p_2\rangle^{(n_1)}\langle p_2\rangle^{(n_2)}\langle q_2\rangle\circ\overline{G}^{n_1}\langle q_2\rangle\circ\overline{G}^{n_2}\\
&\quad-\bigl(p_2^{(n_1)}q_2\mathscr{S}_2^{n_1}-\langle p_2\rangle^{(n_1)}\langle q_2\rangle\circ\overline{G}^{n_1}\bigr)\langle p_2\rangle^{(n_2)}\langle q_2\rangle\circ\overline{G}^{n_2}-\langle p_2\rangle^{(n_1)}\langle q_2\rangle\circ\overline{G}^{n_1}\bigl(p_2^{(n_2)}q_2\mathscr{S}_2^{n_1}-\langle p_2\rangle^{(n_2)}\langle q_2\rangle\circ\overline{G}^{n_2}\bigr),
\end{aligned}$$

and observe that, thanks to (4.52), the averages of both contributions in the last line are bounded by $C\rho^3(1+\rho(n_1+n_2))(1-\rho\,\gamma')^{n_1+n_2}$. As to the difference in the second line, assuming $n_1\le n_2$, we can write

$$p_2^{(n_1)}(\theta,\psi)\,q_2(\mathscr{S}_2^{n_1}(\theta,\psi))\,p_2^{(n_2)}(\theta,\psi)\,q_2(\mathscr{S}_2^{n_2}(\theta,\psi))=\left(\prod_{i=0}^{n_1-1}p_3(\mathscr{S}_2^i(\theta,\psi))\right)q_{3,n_2-n_1}(\mathscr{S}_2^{n_1}(\theta,\psi)),$$

with $p_3(\theta,\psi):=(p_2(\theta,\psi))^2$ and $q_{3,n}(\theta,\psi):=q_2(\theta,\psi)p_2^{(n)}(\theta,\psi)\,q_2(\mathscr{S}_2^n(\theta,\psi))$. Proposition 4.26 gives

$$\bigl|p_2^{(n_1)}q_2\circ\mathscr{S}_2^{n_1}\,p_2^{(n_2)}q_2\circ\mathscr{S}_2^{n_2}-\langle p_3\rangle^{(n_1)}\langle q_{3,n_2-n_1}\rangle\circ\overline{G}^{n_1}\bigr|\le C\,(1-\rho\,\gamma')^{n_1}\rho\,(1+n_1\rho)|||q_{3,n_2-n_1}|||^-_{\alpha_0,3}\,,$$

where we have $|||q_{3,n_2-n_1}|||^-_{\alpha_0,3}\le C\rho^2(1-\rho\,\gamma')^{n_2-n_1}$, as a consequence of Lemma 4.20, of the bound (4.51) and of the inequality (2.8). Thus we are left with studying

$$\begin{aligned}
&\langle p_3\rangle^{(n_1)}\langle q_{3,n_2-n_1}\rangle\circ\overline{G}^{n_1}-\langle p_2\rangle^{(n_1)}\langle p_2\rangle^{(n_2)}\langle q_2\rangle\circ\overline{G}^{n_1}\langle q_2\rangle\circ\overline{G}^{n_2}\\
&\qquad=\langle p_3\rangle^{(n_1)}\bigl(\langle q_{3,n_2-n_1}\rangle\circ\overline{G}^{n_1}-\langle q_2\rangle\circ\overline{G}^{n_1}\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1}\langle q_2\rangle\circ\overline{G}^{n_2}\bigr)\\
&\qquad+\bigl(\langle p_3\rangle^{(n_1)}\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1}-\langle p_2\rangle^{(n_1)}\langle p_2\rangle^{(n_2)}\bigr)\langle q_2\rangle\circ\overline{G}^{n_1}\langle q_2\rangle\circ\overline{G}^{n_2}.
\end{aligned}$$

Using again Proposition 4.26 and (4.51), we obtain

$$\begin{aligned}
&\bigl|\langle q_{3,n_2-n_1}\rangle-\langle q_2\rangle\langle p_2\rangle^{(n_2-n_1)}\langle q_2\rangle\circ\overline{G}^{n_2-n_1}\bigr|\\
&\qquad\le C\rho^2(1-\rho\,\gamma')^{n_2-n_1}\bigl((1+\alpha_0(n_2-n_1))\lambda^{-\alpha_0(n_2-n_1)}+\rho+\rho^2(n_2-n_1)\bigr),
\end{aligned}$$

while the first bound in (4.50) yields $|\langle p_3\rangle - \langle p_2\rangle^2| \le C\rho^2$, which in turn gives

$$\big|\langle p_3\rangle^{(n_1)}\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1} - \langle p_2\rangle^{(n_1)}\langle p_2\rangle^{(n_2)}\big| = \big|\big(\langle p_3\rangle^{(n_1)} - \langle p_2\rangle^{(n_1)}\langle p_2\rangle^{(n_1)}\big)\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1}\big|$$
$$= \big|\sum_{i=0}^{n_1}\langle p_3\rangle^{(i)}\big(\langle p_3\rangle - \langle p_2\rangle^2\big)\circ\overline{G}^i\,\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1}\big| \le C\,(1-\rho\,\gamma')^{n_1+n_2}n_1\rho^2,$$

Combining all the estimates together and proceeding like in (4.48), we get, respectively,

$$\Big\langle\Big(\sum_{n=0}^{\infty}\Big(p_2^{(n)}\,q_2\circ\mathscr{S}_2^n - \langle p_2\rangle^{(n)}\langle q_2\rangle\circ\overline{G}^n\Big)\Big)^2\Big\rangle \le C\rho, \qquad \Big(\sum_{n=0}^{\infty}\Big(\langle p_2\rangle^{(n)}\langle q_2\rangle\circ\overline{G}^n - \overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n\Big)\Big)^2 \le C\rho.$$

which, together, provides the second bound in (4.49). □

The following result extends the analysis above to the first derivatives of h_2 and $\overline{h}$; the proof, based on the same ideas used for Proposition 4.28 up to technical intricacies, is deferred to Appendix B.1.

Proposition 4.29. *Let h_2 and $\overline{h}$ be defined as in* (4.32) *and* (4.19)*, respectively. For all $\theta\in\mathcal{U}$, one has*

$$\big|\big\langle\partial_\theta h_2(\theta,\cdot) - \partial_\theta\overline{h}(\theta)\big\rangle\big| \le C\rho, \qquad \big\langle(\partial_\theta h_2(\theta,\cdot) - \partial_\theta\overline{h}(\theta))^2\big\rangle \le C\rho\,. \tag{4.55}$$

Remark 4.30. To prove Proposition 4.29 we need $f \in \mathscr{B}_{\alpha_0,5}$ while to prove Theorems 1 to 4 it would be enough to assume $f \in \mathscr{B}_{\alpha_0,2}$. The full regularity assumed in our hypotheses is required to prove the forthcoming Proposition 4.45 (see Remark 4.46).

4.7.4 Comparison between the translated map and the auxiliary map

We start by introducing some notation. If, for $\mathfrak{p}_1,\mathfrak{p}_2 \in \mathscr{B}_{0,3}(\Omega,\mathbb{R})$, we define

$$\mathfrak{p}_1^{(n)} := \prod_{i=0}^{n-1}\mathfrak{p}_1\circ\mathscr{S}_1^i, \qquad \mathfrak{p}_2^{(n)} := \prod_{i=0}^{n-1}\mathfrak{p}_2\circ\mathscr{S}_2^i, \tag{4.56}$$

reasoning as in Subsection 4.4 we may write

$$\mathfrak{p}_1^{(n)} - \mathfrak{p}_2^{(n)} = \sum_{k=0}^{n-1}\mathfrak{p}_1^{(k)}\big(\mathfrak{p}_1\circ\mathscr{S}_1^k - \mathfrak{p}_2\circ\mathscr{S}_2^k\big)\,\mathfrak{p}_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}, \tag{4.57a}$$

$$\mathfrak{p}_1^{(n)}\mathfrak{q}_1\circ\mathscr{S}_1^n - \mathfrak{p}_2^{(n)}\,\mathfrak{q}_2\circ\mathscr{S}_2^n = \big(\mathfrak{p}_1^{(n)} - \mathfrak{p}_2^{(n)}\big)\,\mathfrak{q}_2\circ\mathscr{S}_2^n + \mathfrak{p}_1^{(n)}\,\big(\mathfrak{q}_1\circ\mathscr{S}_1^n - \mathfrak{q}_2\circ\mathscr{S}_2^n\big)\,, \tag{4.57b}$$

and, in a similar way, we find

$$\mathfrak{p}_2^{(n)}\circ\mathscr{S}_1^i - \mathfrak{p}_2^{(n)}\circ\mathscr{S}_2^i = \sum_{k=0}^{n-1}\mathfrak{p}_2^{(k)}\circ\mathscr{S}_1^i\big(\mathfrak{p}_2\circ\mathscr{S}_2^k\circ\mathscr{S}_1^i - \mathfrak{p}_2\circ\mathscr{S}_2^k\circ\mathscr{S}_2^i\big)\,\mathfrak{p}_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1+i} \tag{4.58a}$$

$$\mathfrak{p}_1^{(n)}\circ\mathscr{S}_1^i - \mathfrak{p}_2^{(n)}\circ\mathscr{S}_2^i = \sum_{k=0}^{n-1}\mathfrak{p}_1^{(k)}\circ\mathscr{S}_1^i\big(\mathfrak{p}_1\circ\mathscr{S}_1^{k+i} - \mathfrak{p}_2\circ\mathscr{S}_2^{k+i}\big)\,\mathfrak{p}_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1+i}, \tag{4.58b}$$

with the latter reducing to (4.57a) for $i = 0$.

Now, we come back to Proposition 4.15. To complete the proof we are left to study the contributions $h_1 - h_2$ and $\partial_\theta h_1 - \partial_\theta h_2$ in (4.33). In the light of (3.17) and of the the expansion (4.57a), we may write

$$\begin{aligned} h_1 - h_2 &= \sum_{n=0}^{\infty}\big(p_1^{(n)}\,q_1\circ\mathscr{S}_1^n - p_2^{(n)}q_2\circ\mathscr{S}_2^n\big) = \sum_{n=0}^{\infty}p_2^{(n)}\big(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n\big) \\ &+ \sum_{n=0}^{\infty}\big(p_1^{(n)} - p_2^{(n)}\big)\,q_2\circ\mathscr{S}_2^n + \sum_{n=0}^{\infty}\sum_{k=0}^{n-1}p_1^{(k)}\,\big(p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k\big)\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\big(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n\big). \end{aligned} \tag{4.59}$$

We estimate $h_1 - h_2$ by studying the differences that appear in (4.59). First, we prove a series a technical lemmas based on the structure of the difference $f_1 - f_2$, and come back to (4.59) in Lemma 4.41 below. To start with, we write ζ as in (4.29a) and note that $|||\zeta|||_{0,2} \le C$, while, setting $\zeta_0 = c_1(\theta,\psi)\,W(\psi) + c_2(\theta,\psi)\,W(\psi)^2$, we have $\langle\zeta_0^2\rangle \le C\rho$ and $\langle(\partial_\theta\zeta_0)^2\rangle \le C\rho$ by (2.33) in Theorem 4.

Remark 4.31. A key observation in the argument used below and in the related appendices is the following. According to (4.29a), the function ζ can be written as sum of three terms. While the last two terms, which depend linearly or quadratically on W, are controlled by Theorem 4, the first one depends on W through the function $\mathfrak{f}$. One can write $\mathfrak{f}$ in terms of the difference $W\circ A_0 - W$ (and hence linearly in W) as in (4.29e), but, in doing so, a factor ρ is lost. However, in order to compare $\mathscr{S}_1$ with $\mathscr{S}_2$, one deals with sums over i of contributions of the form $\Xi_i\circ\mathscr{S}_2^i(W\circ A_0^{i+1} - W\circ A_0^i)$, with Ξ_i more regular than W (see for instance the proof of Lemma 4.41 below). Thus, one can rearrange the sums and obtain summands of the form $(\Xi_{i+1}\circ\mathscr{S}_2^{i-1} - \Xi_i\circ\mathscr{S}_2^i)W\circ A_0^i$ (see Lemmas 4.37 and 4.39), where the differences $\Xi_{i+1}\circ\mathscr{S}_2^{i-1} - \Xi_i\circ\mathscr{S}_2^i$ allow to gain a compensating factor ρ.

The following result plays a crucial role in the forthcoming discussion. The proof, based on the idea illustrated in Remark 4.31, is given in Appendix B.2.1 (which we refer to for more details).

Lemma 4.32. *For any $\mathfrak{p}$ in $\mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$ one has*

$$\mathfrak{p}\circ\mathscr{S}_1^n - \mathfrak{p}\circ\mathscr{S}_2^n = \Delta_{\mathfrak{p},W,n} := \rho\sum_{i=0}^n \mathfrak{C}_{\mathfrak{p},n,n-i}\circ\mathscr{S}_2^i\, W\circ A_0^i + \rho\,\mathfrak{R}_{\mathfrak{p},n}, \tag{4.60}$$

for suitable functions $\mathfrak{C}_{\mathfrak{p},n,0},\ldots,\mathfrak{C}_{\mathfrak{p},n,n} \in \mathscr{B}_{\alpha_0,2}(\Omega,\mathbb{R})$ and $\mathfrak{R}_{\mathfrak{p},n} \in \mathscr{B}_{0,2}(\Omega,\mathbb{R})$ such that

$$|||\mathfrak{C}_{\mathfrak{p},n,k}|||^-_{\alpha_0,2} \le C\,(1-\rho\,\gamma')^k\, |||\partial_\theta\mathfrak{p}|||_{\alpha_0,2}, \qquad k=1,\ldots,n-1, \tag{4.61a}$$

$$\sum_{k=0}^n |||\mathfrak{C}_{\mathfrak{p},n,k}|||^-_{\alpha_0,2} \le C\,\rho^{-1}|||\partial_\theta\mathfrak{p}|||_{\alpha_0,2}, \tag{4.61b}$$

$$\rho\|\mathfrak{R}_{\mathfrak{p},n}\|_\infty + \langle|\mathfrak{R}_{\mathfrak{p},n}|\rangle \le C\,|||\partial_\theta\mathfrak{p}|||_{\alpha_0,2}. \tag{4.61c}$$

The two next results are immediate consequences of Lemma 4.32.

Lemma 4.33. *For any $\mathfrak{p}_2 \in \mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$ and any $\mathfrak{p}_1 \in \mathscr{B}_{0,3}(\Omega,\mathbb{R})$ such that*

$$(\mathfrak{p}_1-\mathfrak{p}_2)(\theta,\psi) = \rho\,\mathfrak{c}_1(\theta,\psi)W(\psi) + \rho\,\mathfrak{c}_2(\theta,\psi) + \rho\,\mathfrak{c}_3(\theta)\,\mathfrak{f}(\psi), \tag{4.62}$$

with $\mathfrak{c}_1 \in \mathscr{B}_{\alpha_0,2}(\Omega,\mathbb{R})$, $\mathfrak{c}_2 \in \mathscr{B}_{0,2}(\Omega,\mathbb{R})$, $\mathfrak{c}_3 \in C^1(\mathcal{U},\mathbb{R})$ and $\mathfrak{f}$ as in (4.29e), one has

$$\mathfrak{p}_1\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n = \Delta_{\mathfrak{p}_1,\mathfrak{p}_2,W,n} + \rho\,\mathfrak{c}_3\circ\mathscr{S}_1^n\,\mathfrak{f}\circ A_0^n, \tag{4.63}$$

with

$$\Delta_{\mathfrak{p}_1,\mathfrak{p}_2,W,n} := \rho\sum_{i=0}^n \mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,n,n-i}\circ\mathscr{S}_2^i\, W\circ A_0^i + \rho\,\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n} \tag{4.64a}$$

$$\mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,n,k} := \mathfrak{C}_{\mathfrak{p}_2,n,k} + \delta_{k,0}\mathfrak{c}_1, \qquad k=0,\ldots,n, \tag{4.64b}$$

$$\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n} := \mathfrak{R}_{\mathfrak{p}_2,n} + \mathfrak{c}_2\circ\mathscr{S}_1^n + (\mathfrak{c}_1\circ\mathscr{S}_1^n - \mathfrak{c}_1\circ\mathscr{S}_2^n)\,W\circ A_0^n. \tag{4.64c}$$

Proof. Write $\mathfrak{p}_1\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n = \mathfrak{p}_2\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n + (\mathfrak{p}_1-\mathfrak{p}_2)\circ\mathscr{S}_1^n$ and $\mathfrak{c}_1\circ\mathscr{S}_1^n = \mathfrak{c}_1\circ\mathscr{S}_1^n - \mathfrak{c}_1\circ\mathscr{S}_2^n + \mathfrak{c}_1\circ\mathscr{S}_2^n$, and use Lemma 4.32 to deal with the contribution $\mathfrak{p}_2\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n$. □

Corollary 4.34. *Let $\mathfrak{p}_1,\mathfrak{p}_2 \in \mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$ be such that $|||\partial_\theta\mathfrak{p}_2|||_{\alpha_0,2} \le C\rho$ and (4.62) holds, with $|||\mathfrak{c}_1|||^-_{\alpha_0,2} \le C$, and $\rho\|\mathfrak{c}_2\|_\infty + \langle|\mathfrak{c}_2|\rangle \le C\rho$. The functions (4.64) satisfy the bounds*

$$\sum_{k=0}^n |||\mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,n,k}|||_{\alpha_0,2} \le C, \qquad\qquad \rho\|\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n}\|_\infty + \langle|\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n}|\rangle \le C\rho.$$

Proof. The assertion follows immediately from the bounds (4.61) and from Theorem 4. □

Remark 4.35. Taking either $\mathfrak{p}_1 = p_1$ and $\mathfrak{p}_2 = p_2$ or $\mathfrak{p}_1 = q_1$ and $\mathfrak{p}_2 = q_2$, with p_1 and q_1 as in (3.16) and p_2 and q_2 as in (4.31), the hypotheses of Lemma 4.33 are verified, with the functions $\mathfrak{c}_1$ and $\mathfrak{c}_2$ satisfying the estimates in Corollary 4.34 in both cases. A straightforward computation gives $\mathfrak{c}_3(\theta) = \mathfrak{a}_1(\theta) := -2\theta^{-1}c_0(\theta)$ and $\mathfrak{c}_3(\theta) = \mathfrak{a}_2(\theta) := -\theta^{-2}c_0(\theta)$, respectively, so that

$$p_1\circ\mathscr{S}_1^n - p_2\circ\mathscr{S}_2^n = \Delta_{p_1,p_2,W,n} + \rho\,\mathfrak{a}_1\circ\mathscr{S}_1^n\,\mathfrak{f}\circ A_0^n, \qquad q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n = \Delta_{q_1,q_2,W,n} + \rho\,\mathfrak{a}_2\circ\mathscr{S}_1^n\,\mathfrak{f}\circ A_0^n, \tag{4.65}$$

with $\mathfrak{a}_1, \mathfrak{a}_2 \in C^\infty(\mathcal{U}, \mathbb{R})$, $\|\mathfrak{a}_1\|_\infty \le C$ and $\|\mathfrak{a}_2\|_\infty \le C$. Moreover, Theorem 4, for any $n \ge 0$ one has

$$|||\Delta_{p_1,p_2,W,n}|||^-_{\alpha_0,2} \le C\rho, \qquad |||\Delta_{q_1,q_2,W,n}|||^-_{\alpha_0,2} \le C\rho. \tag{4.66}$$

If one restricts $\mathscr{S}$ to Λ_{2r_2} (see Remark 4.16), c_0 is replaced with 0 and both functions $\mathfrak{a}_1$ and $\mathfrak{a}_2$ vanish, so that the terms with $\mathfrak{c}_3$ in (4.62) and (4.63) disappear. In particular, the coming Lemmas 4.37 and 4.39 are not needed, and, in the discussion of the remaining results, all terms involving the functions $\mathfrak{a}_1$ and $\mathfrak{a}_2$ vanish, with a substantial simplification of the proofs.

Remark 4.36. Bounds analogous to (4.66) hold for $\Delta_{p_2,W,n}$ and $\Delta_{q_2,W,n}$. Indeed, for any $n \ge 0$, one has $|||\Delta_{p_2,W,n}|||^-_{\alpha_0,2} \le C\rho$ and $|||\Delta_{q_2,W,n}|||^-_{\alpha_0,2} \le C\rho$. Moreover, by Theorem 4, for any $n, k \ge 0$ one has

$$\big\langle \big|\Delta_{p_2,W,n}\, W\circ A_0^k\big|\big\rangle \le C\rho^2, \qquad \big\langle \big|\Delta_{q_2,W,n}\, W\circ A_0^k\big|\big\rangle \le C\rho^2, \tag{4.67a}$$

$$\big\langle \big|\Delta_{p_1,p_2,W,n}\, W\circ A_0^k\big|\big\rangle \le C\rho^2, \qquad \big\langle \big|\Delta_{p_1,p_2,W,n}\, \Delta_{q_1,q_2,W,k}\big|\big\rangle \le C\rho^3. \tag{4.67b}$$

Corollary 4.34 allows us to deal with the first two contributions in the r.h.s. of (4.63). In order to deal with the last contribution we need also the following two results.

Lemma 4.37. *For $\mathfrak{p}_1, \mathfrak{p}_2 \in \mathscr{B}_{0,3}(\Omega, \mathbb{R})$, such that (4.62) is satisfied and $\|\mathfrak{p}_1 - 1\|_\infty \le C\rho$, let $\mathfrak{p}_1^{(n)}$ and $\mathfrak{p}_2^{(n)}$ be defined as in (4.56). For any function $\mathfrak{a} \colon \mathcal{U} \to \mathbb{R}$ of class C^1 and any $n \ge 1$ and $j \ge 0$, one has*

$$\begin{aligned}\sum_{k=0}^{n-1} &\mathfrak{p}_1^{(k)}\circ\mathscr{S}_1^j\, \mathfrak{a}\circ\mathscr{S}_1^{k+j} \mathfrak{p}_2^{(n-1-k)}\circ\mathscr{S}_2^{k+1+j}\big(W\circ A_0^{k+1+j} - W\circ A_0^{k+j}\big) \\ &= \sum_{k=0}^{n} \mathfrak{p}_1^{(k-1)}\circ\mathscr{S}_1^j\, \mathfrak{D}_{\mathfrak{a},n,k,j}\circ\mathscr{S}_1^{k+j}\, \mathfrak{p}_2^{(n-1-k)}\circ\mathscr{S}_2^{k+1+j}\, W\circ A_0^{k+j},\end{aligned} \tag{4.68}$$

with

$$\mathfrak{D}_{\mathfrak{a},n,k,j} := \begin{cases} -\mathfrak{a}, & k = 0, \\ \mathfrak{a}\circ\mathscr{S}_1^{-1}\, \mathfrak{p}_2\circ\mathscr{S}_2^{k+j}\circ\mathscr{S}_1^{-(k+j)} - \mathfrak{p}_1\circ\mathscr{S}_1^{-1}\, \mathfrak{a}, & k = 1, \dots, n-1, \\ \mathfrak{a}\circ\mathscr{S}_1^{-1}, & k = n, \end{cases} \tag{4.69}$$

such that

$$\|\mathfrak{D}_{\mathfrak{a},n,k,j}\|_\infty \le C\rho\, |||\mathfrak{a}|||_{0,1}, \qquad k = 1, \dots, n-1. \tag{4.70}$$

Proof. The identity (4.68) is easily checked. To bound $\mathfrak{D}_{\mathfrak{a},n,k,j}$ for $k = 1, \dots, n-1$, write

$$\begin{aligned}&\mathfrak{a}\circ\mathscr{S}_1^{-1}\, \mathfrak{p}_2\circ\mathscr{S}_2^{k+j}\circ\mathscr{S}_1^{-(k+j)} - \mathfrak{a}\, \mathfrak{p}_1\circ\mathscr{S}_1^{-1} \\ &\quad = \big(\mathfrak{a}\circ\mathscr{S}_1^{-1} - \mathfrak{a}\big)\mathfrak{p}_2\circ\mathscr{S}_2^{k+j}\circ\mathscr{S}_1^{-(k+j)} + \mathfrak{a}\big(\mathfrak{p}_2\circ\mathscr{S}_2^{k+j} - \mathfrak{p}_1\circ\mathscr{S}_1^{k+j}\big)\circ\mathscr{S}_1^{-(k+j)} + \mathfrak{a}\big(\mathfrak{p}_1 - \mathfrak{p}_1\circ\mathscr{S}_1^{-1}\big),\end{aligned}$$

and use that

$$\begin{aligned}\big\|\big(\mathfrak{a}\circ\mathscr{S}_1^{-1} - \mathfrak{a}\big)\mathfrak{p}_2\big\|_\infty &\le \|\partial_\theta\mathfrak{a}\|_\infty \|(\mathscr{S}_1)_\theta - \mathbb{1}\|_\infty \le C\rho\, \|\partial_\theta\mathfrak{a}\|_\infty, \\ \big\|\mathfrak{p}_2\circ\mathscr{S}_2^{k+j} - \mathfrak{p}_1\circ\mathscr{S}_1^{k+j}\big\|_\infty &\le \rho \sum_{i=0}^{k+j} \|\mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,k+j,k+j-i}\|_\infty + \rho\|\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,k+j}\|_\infty + \rho\|\mathfrak{c}_3\|_\infty\|\mathfrak{f}\|_\infty \le C\rho, \\ \big\|\mathfrak{p}_1 - \mathfrak{p}_1\circ\mathscr{S}_2^{-1}\big\|_\infty &\le \|(1 + O(\rho)) - (1 + O(\rho))\|_\infty \le C\rho,\end{aligned}$$

with the second inequality following from (4.63) in Lemma 4.33 and the bounds in Corollary 4.34. $\square$

Remark 4.38. The coefficients $\mathfrak{D}_{\mathfrak{a},n,k,j}$ with $k = 0, \dots, n-1$ in (4.69) do not depend on n, in the sense that $\mathfrak{D}_{\mathfrak{a},n,k,j} := \mathfrak{D}_{\mathfrak{a},n',k,j}$ for all $k < \min\{n, n'\}$. Thus, we may define, for future convenience,

$$\mathfrak{D}_{\mathfrak{a},k,j} := \mathfrak{a}\circ{\mathscr{S}_1}^{-1}\, \mathfrak{p}_2\circ{\mathscr{S}_2}^{k+j}\circ{\mathscr{S}_1}^{-(k+j)} - \mathfrak{p}_1\circ{\mathscr{S}_1}^{-1}\mathfrak{a}.$$

Lemma 4.39. *Let $\mathfrak{p}_1, \mathfrak{p}_2 \in \mathscr{B}_{0,3}(\Omega, \mathbb{R})$ be as in Lemma 4.37, and let $\mathfrak{p}_1^{(n)}$ and $\mathfrak{p}_2^{(n)}$ be defined as in* (4.56)*. For any function $\mathfrak{a}\colon \mathcal{U} \to \mathbb{R}$ of class C^3 one has*

$$\sum_{n=0}^{\infty} \mathfrak{p}_1^{(n)}\, \mathfrak{a}\circ\mathscr{S}_1^n \big(W\circ A_0^{n+1} - W\circ A_0^n\big) = \sum_{n=0}^{\infty} \mathfrak{p}_1^{(n-1)}\, \mathfrak{D}_{0,\mathfrak{a},n}\circ\mathscr{S}_1^n\, W\circ A_0^n, \tag{4.71}$$

with

$$\mathfrak{D}_{0,\mathfrak{a},n} := \begin{cases} -\mathfrak{a}, & n=0, \\ \mathfrak{a}\circ\mathscr{S}_1^{-1} - \mathfrak{p}_1\circ\mathscr{S}_1^{-1}\,\mathfrak{a}, & n\geq 1, \end{cases} \tag{4.72}$$

and, similarly, for $r=1,2$,

$$\sum_{n=0}^{\infty} \mathfrak{p}_2^{(n)}\, \mathfrak{a}\circ\mathscr{S}_r^n \big(W\circ A_0^{n+1} - W\circ A_0^n\big) = \sum_{n=0}^{\infty} \mathfrak{p}_2^{(n-1)}\, \mathfrak{D}_{r,\mathfrak{a},n}\circ\mathscr{S}_r^n\, W\circ A_0^n, \tag{4.73}$$

with

$$\mathfrak{D}_{r,\mathfrak{a},n} := \begin{cases} -\mathfrak{a}, & n=0, \\ \mathfrak{a}\circ\mathscr{S}_r^{-1} - \mathfrak{p}_2\circ\mathscr{S}_2^{n-1}\circ\mathscr{S}_r^{-n}\,\mathfrak{a}, & n\geq 1. \end{cases} \tag{4.74}$$

Furthermore, if ρ is such that $\mathscr{S}_2$ satisfies Hypotheses 1–3, one has

$$\|\mathfrak{D}_{r,\mathfrak{a},n}\|_\infty \leq C\rho\, |||\mathfrak{a}|||_{0,1}, \qquad n\geq 1, \quad r=0,1,2, \tag{4.75a}$$

$$|||\mathfrak{D}_{2,\mathfrak{a},n}|||_{0,2} \leq C\rho\, |||\mathfrak{a}|||_{0,3}, \qquad n\geq 1. \tag{4.75b}$$

Proof. After checking (4.71) and (4.73) by direct computation, the bounds (4.75a) are easily obtained by writing $\mathfrak{a}\circ\mathscr{S}_1^{-1} - \mathfrak{p}_2\circ\mathscr{S}_2^{n-1}\circ\mathscr{S}_1^{-n}\,\mathfrak{a} = \mathfrak{a}\circ\mathscr{S}_1^{-1} - \mathfrak{p}_1\circ\mathscr{S}_1^{-1}\,\mathfrak{a} + \big(\mathfrak{p}_1\circ\mathscr{S}_1^{n-1} - \mathfrak{p}_2\circ\mathscr{S}_2^{n-1}\big)\circ\mathscr{S}_1^{-n}$ in (4.74) and using that $\|\mathfrak{a}\circ\mathscr{S}_r^{-1} - \mathfrak{p}_r\circ\mathscr{S}_r^{-1}\,\mathfrak{a}\|_\infty \|\mathfrak{a}\circ\mathscr{S}_r^{-1} - \mathfrak{a}\|_\infty + \|(1-\mathfrak{p}_r\circ\mathscr{S}_r^{-1})\|_\infty \|\mathfrak{a}\|_\infty \leq C\rho\, |||\mathfrak{a}|||_{0,1}$ for $r=1,2$, and $\|\mathfrak{p}_1\circ\mathscr{S}_1^{n-1} - \mathfrak{p}_2\circ\mathscr{S}_2^{n-1}\|_\infty \leq C\rho$, with the second bound holding by Lemma 4.33.

The bound (4.75b) is obtained by writing $\mathfrak{a}\circ\mathscr{S}_2^{-1} - \mathfrak{p}_2\circ\mathscr{S}_2^{-1}\,\mathfrak{a} = \big(\mathfrak{a}\circ\mathscr{S}_2^{-1} - \mathfrak{a}\big) + \big(1 - \mathfrak{p}_2\circ\mathscr{S}_2^{-1}\big)\mathfrak{a}$ and using the bounds of Lemma 4.20. □

Also the next result, essentially based on Proposition 4.26, is used at length in what follows.

Proposition 4.40. *Let $\mathfrak{p}_0, \ldots, \mathfrak{p}_k \in \mathscr{B}_{\alpha_0,3}(\Omega, \mathbb{R})$ satisfy the properties 1 and 2 in Lemma 4.20, and let $\mathfrak{p}^{(k)}(\theta,\psi) := \mathfrak{p}_0^{(k)}(\mathscr{S}_2;\theta,\psi)$ be defined according to* (2.15)*. For any $\mathfrak{C} \in \mathscr{B}_{\alpha_0,2}(\Omega,\mathbb{R})$ and $\mathfrak{u} \in \mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$, and for any $\gamma' \in (0,\gamma)$ and $i=0,\ldots,k$, one has*

$$\left|\left\langle \mathfrak{p}^{(k)}\mathfrak{C}\circ\mathscr{S}_2^i\, W\circ A_0^i\, \mathfrak{u}\circ\mathscr{S}_2^{k+1}\right\rangle\right| \leq C(1-\rho\,\gamma')^k \left((1+\alpha_0 i)\lambda^{-\alpha_0 i} + \rho + i\rho^2 + i^2\rho^3\right) |||\mathfrak{C}|||^-_{\alpha_0,2} |||\mathfrak{u}|||^-_{\alpha_0,3}.$$

Proof. For $i=0,\ldots,k$, let $\mathfrak{p}_{[i]}^{(k)}(\theta,\psi) := \mathfrak{p}_{[i]}^{(k)}(\mathscr{S}_2;\theta,\psi)$ be as in (2.15) with $\mathscr{S} = \mathscr{S}_2$. We have

$$\begin{aligned} \mathfrak{p}^{(k)}\mathfrak{C}\circ\mathscr{S}_2^i\, W\circ A_0^i\, \mathfrak{u}\circ\mathscr{S}_2^{k+1} &= \mathfrak{p}^{(k)}\, \mathfrak{C}\circ\mathscr{S}_2^i\, (W-W_0)\circ A_0^i\, \mathfrak{u}\circ\mathscr{S}_2^{k+1} + \mathfrak{p}^{(i)}\, \mathfrak{p}_{[i]}^{(k-i)}\circ\mathscr{S}_2^i\, \mathfrak{C}\circ\mathscr{S}_2^i\, \mu^{(i)}\, W_0\, \mathfrak{u}\circ\mathscr{S}_2^{k+1} \\ &\quad + \rho\sum_{j=1}^{i} \mathfrak{p}^{(i-j)}\, \mathfrak{p}_{[i-j]}^{(k-i+j)}\circ\mathscr{S}_2^{i-j}\, \mathfrak{C}\circ\mathscr{S}_2^i\, \mu^{(j-1)}\circ A_0^{i-(j-1)}\, b\circ A_0^{i-j}\, \mathfrak{u}\circ\mathscr{S}_2^{k+1}, \end{aligned}$$

by (4.1) and Remark 4.5, so that its average can be bounded as

$$\begin{aligned} \left|\left\langle \mathfrak{p}^{(k)}\mathfrak{C}\circ\mathscr{S}_2^i\, W\circ A_0^i\, \mathfrak{u}\circ\mathscr{S}_2^{k+1}\right\rangle\right| &\leq C(1-\rho\,\gamma')^k\rho\, \|\mathfrak{C}\|_\infty \|\mathfrak{u}\|_\infty \\ &\quad + C(1-\rho\,\gamma')^k \Big(\big((1+\alpha_0 i)\lambda^{-\alpha_0 i} + \rho + i\rho^2\big)\big) + \rho\sum_{j=1}^{i}\big(\rho + (i-j)\rho^2\big)\Big) |||\mathfrak{C}|||^-_{\alpha_0,3} |||\mathfrak{u}|||^-_{\alpha_0,3} \\ &\quad + C\rho\sum_{j=1}^{i}\left|\langle\mathfrak{p}\rangle^{(i-j)}\left\langle b\,\mathfrak{p}_{i-j}(\mathfrak{p}\,\mu)^{(j-1)}_{[i-j+1]}\circ\mathscr{S}_2\,(\mathfrak{p}_{[i]}^{(k-i)}\,\mathfrak{C})\circ\mathscr{S}_2^j\, \mathfrak{u}\circ\mathscr{S}_2^{k+1-i+j}\right\rangle\right|, \end{aligned}$$

where we have used (4.14) to obtain the first line and applied Proposition 4.26 twice: first with $g_+ = W_0$, $g_- = \mathfrak{p}^{(k-i)}_{[i]} \mathfrak{C}\, \mathfrak{u}\circ\mathscr{S}_2^{k+1-i}$ and $\mathfrak{p}^{(n)}$ replaced with $(\mathfrak{p}\mu)^{(i)}$ to obtain the second line; then, after writing $\mathfrak{p}^{(k-i+j)}_{[i-j]} = \mathfrak{p}_{i-j}\mathfrak{p}^{(j-1)}_{[i-j+1]}\circ\mathscr{S}_2\mathfrak{p}^{(k-i)}_{[i]}\circ\mathscr{S}_2^j$, with $g_+ = 1$, $g_- = b\,\mathfrak{p}_{i-j}(\mathfrak{p}\,\mu)^{(j-1)}_{[i-j+1]}\circ\mathscr{S}_2\,(\mathfrak{p}^{(k-i)}_{[i]}\,\mathfrak{C})\circ\mathscr{S}_2^j\,\mathfrak{u}\circ\mathscr{S}_2^{k+1-i+j}$ and $\mathfrak{p}^{(n)}$ replaced with $\mathfrak{p}^{(i-j)}$, to obtain the last two lines.

Using once more Proposition 4.26, with $g_+ = b$, $g_- = (\mathfrak{p}^{(k-i)}_{[i]}\,\mathfrak{C})\,\mathfrak{u}\circ\mathscr{S}_2^{k+1-i}$ and $\mathfrak{p}^{(n)}$ replaced with $\mathfrak{p}_{i-j}(\mathfrak{p}\,\mu)^{(j-1)}_{[i-j+1]}\circ\mathscr{S}_2$, in the last line we bound

$$\begin{aligned}&\Big|\Big\langle b\,\mathfrak{p}_{i-j}(\mathfrak{p}\,\mu)^{(j-1)}_{[i-j+1]}\circ\mathscr{S}_2\,(\mathfrak{p}^{(k-i)}_{[i]}\,\mathfrak{C})\circ\mathscr{S}_2^j\,\mathfrak{u}\circ\mathscr{S}_2^{k+1-i+j}\Big\rangle\Big|\\&\qquad\le C(1-\rho\gamma')^{k-(i-j)}\left((1+\alpha_0 j)\lambda^{-\alpha_0 j}+\rho+j\rho^2\right)|||\mathfrak{C}|||^-_{\alpha_0,3}|||\mathfrak{u}|||^-_{\alpha_0,3}.\end{aligned}$$

Collecting all the bounds together, we obtain the assertion. □

We can now come back to the study of $h_1 - h_2$ as represented in (4.59). We study separately the averages of the three contributions in (4.59), starting from the last one.

Lemma 4.41. *For p_1 and q_1 as in* (3.16)*, and p_2 and q_2 as in* (4.31)*, let $p_1^{(n)}$ and $p_2^{(n)}$ be defined as in* (3.17) *and in* (4.32)*, respectively. One has*

$$\left|\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\Big\langle p_1^{(k)}(p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k)\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n)\Big\rangle\right| \le C\rho.$$

Proof. Using (4.29e) and the notation in (4.60) and in (4.63), we get

$$\begin{aligned}&p_1^{(k)}(p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k)\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n)\\&\quad = p_1^{(k)}\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\\&\quad + p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1} - W\circ A_0^k)\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\\&\quad + p_1^{(k)}\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^n\,(W\circ A_0^{n+1} - W\circ A_0^n)\\&\quad + p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1} - W\circ A_0^k)\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^n\,(W\circ A_0^{n+1} - W\circ A_0^n).\end{aligned}\tag{4.76}$$

The bounds (4.67) give immediately

$$\Big|\Big\langle p_1^{(k)}\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\Big\rangle\Big| \le C\rho^3\,n\,(1-\rho\gamma')^n.$$

Furthermore, by applying Lemma 4.37 using the bounds (4.67) and (4.70), we obtain

$$\left|\sum_{k=0}^{n-1}\Big\langle p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1} - W\circ A_0^k)\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\Big\rangle\right| \le C\rho^2\,n\,(1-\rho\gamma')^n.$$

Analogously, using, firstly, Lemma 4.39 and the expansion (4.58a) together with the notation in (4.63), then (4.67) and (4.75a), we bound

$$\left|\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\Big\langle p_1^{(k)}\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^n\,(W\circ A_0^{n+1} - W\circ A_0^n)\Big\rangle\right| \le C\rho,$$

Finally, to deal with the contribution arising from the last line of (4.76), after using first Lemma 4.37 and Remark 4.38, then Lemma 4.39 and the expansions (4.58), we write

$$\begin{aligned}&\sum_{n=1}^{\infty} p_1^{(n-1)}\,\mathfrak{a}_1\circ\mathscr{S}_1^{n-1}\,W\circ A_0^n\,\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^n\mathfrak{f}\circ A_0^n,\\&\qquad = \frac{1}{2}\sum_{n=0}^{\infty} p_1^{(n)}\,\big(\mathfrak{a}_1\,\mathfrak{a}_2\circ\mathscr{S}_1\big)\circ\mathscr{S}_1^n\Big((W\circ A_0^{n+2})^2 - (W\circ A_0^{n+1})^2 - \big(\rho\,\mathfrak{f}\circ A_0^{n+1}\big)^2\Big),\end{aligned}\tag{4.77}$$

so that, eventually, we rewrite

$$\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\Big\langle p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1}-W\circ A_0^k)\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^n\,(W\circ A_0^{n+1}-W\circ A_0^n)\Big\rangle$$

$$=\sum_{k=0}^{\infty}\Big\langle p_1^{(k-1)}\,\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^k\,W\circ A_0^k\,\sum_{n=0}^{\infty}\big(p_2^{(n-1)}\,\mathfrak{D}_{1,\mathfrak{a}_2,n}\circ\mathscr{S}_1^n\,W\circ A_0^n\big)\circ\mathscr{S}_1^{k+1}\Big\rangle$$

$$-\sum_{k=0}^{\infty}\Big\langle p_1^{(k-1)}\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^kW\circ A_0^k\sum_{n=0}^{\infty}\sum_{i=0}^{n-1}p_2^{(i)}\circ\mathscr{S}_1^{k+1}\Delta_{p_2\circ\mathscr{S}_2^i,W,k+1}p_2^{(n-1-i)}\circ\mathscr{S}_2^{k+2+i}\rho\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\mathfrak{f}\circ A_0^{n+k+1}\Big\rangle$$

$$+\frac{1}{2}\sum_{n=0}^{\infty}\Big\langle p_1^{(n-1)}\,\Delta_{1,\mathfrak{a}_1\mathfrak{a}_2\circ\mathscr{S}_1,n}\circ\mathscr{S}_1^n\big(W\circ A_0^{n+1}\big)^2\Big\rangle-\frac{1}{2}\sum_{n=0}^{\infty}\Big\langle p_1^{(n)}\,\big(\mathfrak{a}_1\,\mathfrak{a}_2\circ\mathscr{S}_1\big)\circ\mathscr{S}_1^n\big(\rho\,\mathfrak{f}\circ A_0^{n+1}\big)^2\Big\rangle,$$

and proceed as in the proof of Lemma 4.37. Thus, relying once more on (4.67), (4.70) and (4.75a), we get the assertion. □

The first application of Proposition 4.40 is to estimate the other two contributions in (4.59). This leads to the two following lemmas, whose proof makes also use of the argument given in Remark 4.31 (in particular of Lemmas 4.32, 4.37 and 4.39, which are based on the latter).

Lemma 4.42. *Let $p_2^{(n)}$ be defined as in* (4.32)*, with q_1 as in* (3.16) *and p_2,q_2 as in* (4.31)*. One has*

$$\bigg|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\big(q_1\circ\mathscr{S}_1^n-q_2\circ\mathscr{S}_2^n\big)\Big\rangle\bigg|\le C\rho.$$

Proof. Using Lemma 4.33 and Remark 4.35, we write

$$\begin{aligned}\bigg|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\,(q_1\circ\mathscr{S}_1^n-q_2\circ\mathscr{S}_2^n)\Big\rangle\bigg|&\le\bigg|\rho\sum_{n=0}^{\infty}\sum_{k=0}^{n}\Big\langle p_2^{(n)}\mathfrak{C}_{q_1,q_2,n,n-k}\circ\mathscr{S}_2^k\,W\circ A_0^k\Big\rangle\bigg|+\bigg|\rho\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}R_{q_1,q_2,n}\Big\rangle\bigg|\\&+\bigg|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\mathfrak{a}_2\circ\mathscr{S}_2^n\big(W\circ A_0^{n+1}-W\circ A_0^n\big)\Big\rangle\bigg|+\bigg|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\,\rho\big(\mathfrak{a}_2\circ\mathscr{S}_1^n-\mathfrak{a}_2\circ\mathscr{S}_2^n\big)\mathfrak{f}\circ A_0^n\Big\rangle\bigg|.\end{aligned}\tag{4.78}$$

We use Proposition 4.40, with n and k instead of k and i, respectively, $\mathfrak{p}_j=p_2\ \forall j=1,\dots,k-1$, $\mathfrak{C}=\mathfrak{C}_{q_1,q_2,n,n-k}$ and $\mathfrak{u}=1$, and the estimates (4.61) in Lemma 4.32 to bound the two contributions in first line in (4.78). In the first contribution of the second line, using Lemma 4.39, we write, for $n\ge1$,

$$\bigg|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\mathfrak{a}_2\circ\mathscr{S}_2^n\big(W\circ A_0^{n+1}-W\circ A_0^n\big)\Big\rangle\bigg|=\bigg|\sum_{n=0}^{\infty}\Big\langle p_2^{(n-1)}\mathfrak{D}_{2,\mathfrak{a}_2,n}\circ\mathscr{S}_2^n\,W\circ A_0^n\Big\rangle\bigg|\le C\rho,\tag{4.79}$$

where the last bound follows from Proposition 4.40, with $k=n-1$, $i=n$, $\mathfrak{p}_j=p_2$ for $j=0,\dots,n-2$ (see Remark 4.22), $\mathfrak{C}=\mathfrak{D}_{2,\mathfrak{a}_2,n}$, so that $|||\mathfrak{C}|||^-_{\alpha_0,2}=|||\mathfrak{D}_{2,\mathfrak{a}_2,n}|||^-_{\alpha_0,2}\le C\rho$ for $n\ge1$, and $\mathfrak{u}=1$. To deal with the second contribution in the second line of (4.78), setting

$$\mathfrak{f}_0(\psi):=f(0,\psi),\qquad \mathfrak{f}\circ A_0^n-\mathfrak{f}_0\circ A_0^n=\mathfrak{g}_n\,W\circ A_0^n,\qquad \mathfrak{g}_n(\psi):=\int_0^1\partial_\theta\mathfrak{f}(tW(A_0^n\psi),A_0\psi),\tag{4.80}$$

and using Lemma 4.32, we rewrite

$$\big(\mathfrak{a}_2\circ\mathscr{S}_1^n-\mathfrak{a}_2\circ\mathscr{S}_2^n\big)\mathfrak{f}\circ A_0^n=\bigg(\rho\sum_{k=0}^{n}\mathfrak{C}_{\mathfrak{a}_2,n,n-k}\circ\mathscr{S}_2^k\,W\circ A_0^k+\rho R_{\mathfrak{a}_2,n}\bigg)\big(\mathfrak{f}_0\circ A_0^n+\mathfrak{g}_n\,W\circ A_0^n\big),$$

so that, by exploiting Proposition 4.40, with k and i replaced with n and k, respectively, and with $\mathfrak{u}=\mathfrak{f}_0\circ A_0^{-1}$, Theorem 4 and the bounds of Lemma 4.32, we obtain, for any $\gamma'\in(0,\gamma)$

$$\bigg|\sum_{n=0}^{\infty}\sum_{k=0}^{n}\rho^2\Big\langle p_2^{(n)}\mathfrak{C}_{\mathfrak{a}_2,n,n-k}\circ\mathscr{S}_2^k\,W\circ A_0^k\,\mathfrak{f}_0\circ A_0^n\Big\rangle\bigg|=\bigg|\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\rho^2\Big\langle p_2^{(n+k)}\mathfrak{C}_{\mathfrak{a}_2,n+k,n}\circ\mathscr{S}_2^k\,W\circ A_0^k\,\mathfrak{f}_0\circ A_0^{n+k}\Big\rangle\bigg|$$

$$\le\sum_{k=0}^{\infty}\rho^2(1-\rho\,\gamma')^k\big((1+\alpha_0k)\,\lambda^{-\alpha_0k}+\rho+k\rho^2+k\rho^3\big)\sum_{n=0}^{\infty}|||\mathfrak{C}_{\mathfrak{a}_2,n+k,n}|||^-_{\alpha_0,2}\le C\rho.$$

Similarly, we find

$$\Big|\sum_{n=0}^{\infty}\rho^2\Big\langle p_2^{(n)}\,R_{\mathfrak{a}_2,n}\,\mathfrak{f}_0\circ A_0^n\Big\rangle\Big|\le C\sum_{n=0}^{\infty}\rho^2(1-\rho\,\gamma')^n\big\langle|\mathfrak{R}_{\mathfrak{a}_2,n}|\big\rangle\le C\rho,$$
$$\Big|\sum_{n=0}^{\infty}\rho^2\Big\langle p_2^{(n)}\mathfrak{g}_n\Big(\sum_{k=0}^{n}\mathfrak{C}_{\mathfrak{a}_2,n,n-k}\circ\mathscr{S}_2^k\,W\circ A_0^k+R_{\mathfrak{a}_2,n}\Big)W\circ A_0^n\Big\rangle\Big|\le C\rho,$$

which, together with the previous bound, imply the desired bound. □

Lemma 4.43. *Let $p_1^{(n)}$ and $p_2^{(n)}$ be defined as in* (3.17) *and in* (4.32)*, respectively, with p_1 and q_1 as in* (3.16)*, and p_2 and q_2 as in* (4.31)*. One has*

$$\Big|\sum_{n=0}^{\infty}\Big\langle\big(p_1^{(n)}-p_2^{(n)}\big)\,q_2\circ\mathscr{S}_2^n\Big\rangle\Big|\le C\rho.$$

Proof. We expand $p_1^{(n)}-p_2^{(n)}$ as in (4.57a), hence split $p_1^{(k)}=p_2^{(k)}+(p_1^{(k)}-p_2^{(k)})$ and thence expand again $p_1^{(k)}-p_2^{(k)}$, so as to obtain

$$\begin{aligned}\sum_{n=0}^{\infty}\Big\langle\big(p_1^{(n)}-p_2^{(n)}\big)\,q_2\circ\mathscr{S}_2^n\Big\rangle&=\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\Big\langle p_2^{(k)}\big(p_1\circ\mathscr{S}_1^k-p_2\circ\mathscr{S}_2^k\big)p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle\\&+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\sum_{j=0}^{k-1}\Big\langle p_1^{(j)}\big(p_1\circ\mathscr{S}_1^j-p_2\circ\mathscr{S}_2^j\big)p_2^{(k-1-j)}\circ\mathscr{S}_2^{j+1}\,\big(p_1\circ\mathscr{S}_1^k-p_2\circ\mathscr{S}_2^k\big)p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle,\end{aligned}\tag{4.81}$$

To bound the second line of (4.81), we write $p_1^{(j)}\big(p_1\circ\mathscr{S}_1^j-p_2\circ\mathscr{S}_2^j\big)p_2^{(k-1-j)}\circ\mathscr{S}_2^{j+1}\,\big(p_1\circ\mathscr{S}_1^k-p_2\circ\mathscr{S}_2^k\big)$ as in (4.76), with k, j, p_1 and p_2 instead of n, k, q_1 and q_2, respectively, and we reason as in the proof of Lemma 4.41, by relying on Lemmas 4.37 and 4.39, and using (4.77). Thus, the second line becomes

$$\begin{aligned}&\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\sum_{j=0}^{k-1}\Big\langle p_1^{(j)}\Delta_{p_1,p_2,W,j}\,p_2^{(k-1-j)}\circ\mathscr{S}_2^{j+1}\,\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle\\&+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\sum_{j=0}^{k-1}\Big\langle p_1^{(j-1)}\,\mathfrak{D}_{\mathfrak{a}_1,j,0}\circ\mathscr{S}_1^j\,W\circ A_0^j\,p_2^{(k-j-1)}\circ\mathscr{S}_2^{j+1}\,\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle\\&+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\Big\langle p_1^{(k-1)}\,\mathfrak{a}_1\circ\mathscr{S}_1^{k-1}\,W\circ A_0^k\,\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle\\&+\sum_{j=0}^{\infty}\Big\langle p_1^{(j)}\,\Delta_{p_1,p_2,W,j}\sum_{k=0}^{\infty}\big(p_2^{(k-1)}\,\mathfrak{D}_{1,\mathfrak{a}_2,k}\circ\mathscr{S}_1^k\,W\circ A_0^k\big)\circ\mathscr{S}_1^{j+1}\mathcal{Q}_{k+1}\Big\rangle\\&-\sum_{j=0}^{\infty}\Big\langle p_1^{(j)}\,\Delta_{p_1,p_2,W,j}\sum_{k=0}^{\infty}\sum_{i=0}^{k}p_2^{(i)}\circ\mathscr{S}_1^{j+1}\Delta_{p_2\circ\mathscr{S}_2^i,W,j+1}\mathcal{P}_{j,k,i}(\mathfrak{a}_2\rho\,\mathfrak{f})\circ\mathscr{S}_1^{k+j+1}\mathcal{Q}_{k+1}\Big\rangle\\&+\sum_{j=0}^{\infty}\Big\langle p_1^{(j-1)}\,\mathfrak{D}_{\mathfrak{a}_1,j,0}\circ\mathscr{S}_1^j\,W\circ A_0^j\sum_{k=0}^{\infty}\big(p_2^{(k-1)}\mathfrak{D}_{1,\mathfrak{a}_2,k}\circ\mathscr{S}_1^k\,W\circ A_0^k\big)\circ\mathscr{S}_1^{j+1}\mathcal{Q}_{k+1}\circ\mathscr{S}_2^{j+1}\Big\rangle\\&-\sum_{j=0}^{\infty}\Big\langle p_1^{(j-1)}\,\mathfrak{D}_{\mathfrak{a}_1,j,0}\circ\mathscr{S}_1^j\,W\circ A_0^j\sum_{k=0}^{\infty}\sum_{i=0}^{k}p_2^{(i)}\circ\mathscr{S}_1^{j+1}\Delta_{p_2\circ\mathscr{S}_2^i,W,j+1}\mathcal{P}_{j,k,i}(\mathfrak{a}_2\rho\,\mathfrak{f})\circ\mathscr{S}_1^{k+j+1}\mathcal{Q}_{k+1}\circ\mathscr{S}_2^{j+1}\Big\rangle\\&+\frac12\sum_{k=0}^{\infty}\Big\langle p_1^{(k-1)}\mathfrak{D}_{\mathfrak{a}_1\mathfrak{a}_2\circ\mathscr{S}_1^{-1},k,0}\circ\mathscr{S}_1^k\big(W\circ A_0^{k+1}\big)^2\mathcal{Q}_{k+1}\Big\rangle-\frac12\sum_{k=0}^{\infty}\Big\langle p_1^{(k)}\big(\mathfrak{a}_1\mathfrak{a}_2\circ\mathscr{S}_1^{-1}\big)\circ\mathscr{S}_1^k\big(\rho\,\mathfrak{f}\circ A_0^{k+1}\big)^2\mathcal{Q}_{k+1}\Big\rangle,\end{aligned}$$

with $\mathcal{Q}_k:=\sum_{n=0}^{\infty}\big(p_2^{(n)}\,q_2\circ\mathscr{S}_2^n\big)\circ\mathscr{S}_2^k$ and $\mathcal{P}_{j,k,i}:=p_2^{(k-1-i)}\circ\mathscr{S}_2^{j+i+2}$, so that all contributions are found to be bounded as $C\rho$.

The contribution in the first line of the r.h.s. of (4.81), after writing $p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k$ once more according to Remark 4.35, becomes

$$\begin{aligned}&\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\Big\langle p_2^{(k)}\Big(\rho\sum_{i=0}^{k}\mathfrak{C}_{p_1,p_2,k,k-i}\circ\mathscr{S}_2^i\,W\circ A_0^i+\rho\,\mathfrak{R}_{p_1,p_2,k}\Big)p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle\\&\qquad+\sum_{n=0}^{\infty}\sum_{k=0}^{n}\Big\langle p_2^{(k)}\,\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^k\,\mathfrak{f}\circ A_0^k\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle.\end{aligned}\tag{4.82}$$

In (4.82) the first line is bounded by Proposition 4.40, with $\mathfrak{p}_i=p_2\ \forall i=1,\dots,k-1$, $\mathfrak{C}=\mathfrak{C}_{p_1,p_2,k,k-i}$ and $\mathfrak{u}=p_2^{(n-k-1)}\,q_2\circ\mathscr{S}^{n-k-1}$, while in order to bound the second line we proceed as follows. After splitting $\mathfrak{a}_2\circ\mathscr{S}_1^k=\mathfrak{a}_2\circ\mathscr{S}_2^k+(\mathfrak{a}_2\circ\mathscr{S}_1^k-\mathfrak{a}_2\circ\mathscr{S}_2^k)$, we reason as done for the last line of (4.78): the contribution with $\mathfrak{a}_2\circ\mathscr{S}_2^k$ is dealt with as (4.79), the only difference being that, when applying Proposition 4.40, one sets $\mathfrak{u}=p_2^{(n-k-1)}\circ\mathscr{S}_2\,q_2\circ\mathscr{S}_2^{n-k}$; in the second contribution we expand $\mathfrak{a}_2\circ\mathscr{S}_1^k-\mathfrak{a}_2\circ\mathscr{S}_2^k$ as in (4.60) and write $\mathfrak{f}\circ A_0^k$ as in (4.80), so as to obtain the three contributions

$$\begin{aligned}&\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\Big\langle\rho\,p_2^{(k+i)}\,\rho\,\mathfrak{C}_{\mathfrak{a}_2,k+i,k}\circ\mathscr{S}_2^i\,W\circ A_0^i\,\mathfrak{f}_0\circ A_0^{k+1}\,p_2^{(n-1)}\circ\mathscr{S}_2^{k+i+1}\,q_2\circ\mathscr{S}_2^{n+k+i}\Big\rangle,\\&\sum_{n=0}^{\infty}\sum_{k=0}^{n}\Big\langle\rho\,p_2^{(k)}\,\rho\,\mathfrak{R}_{\mathfrak{a}_2,k}\,\mathfrak{f}_0\circ A_0^{k+1}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^n\Big\rangle,\\&\sum_{n=0}^{\infty}\sum_{k=0}^{n}\Big\langle\rho\,p_2^{(k)}\Big(\rho\sum_{i=0}^{k}\mathfrak{C}_{\mathfrak{a}_2,k,k-i}\circ\mathscr{S}_2^i\,W\circ A_0^i+\rho\,\mathfrak{R}_{\mathfrak{a}_2,k}\Big)\mathfrak{g}_k\,W\circ A_0^k\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^{ni}\Big\rangle,\end{aligned}$$

which are all bounded proportionally to ρ. □

The following result puts together the bounds obtained above and extends the analysis to the square of h_1-h_2. Together with the forthcoming Proposition 4.45, it completes the first step in order to prove Proposition 4.29, as outlined at the beginning of the present subsection.

Proposition 4.44. *Let h_2 and h_1 be as in* (4.32) *and in* (3.13)*, respectively. For all $\theta\in\mathcal{U}$ one has*

$$|\langle h_1(\theta,\cdot)-h_2(\theta,\cdot)\rangle|\le C\rho,\qquad\big\langle(h_1(\theta,\cdot)-h_2(\theta,\cdot))^2\big\rangle\le C\rho\,.\tag{4.83}$$

Proof. The first bound in (4.83) follows immediately from (4.59) and from the estimates in Lemma 4.41, 4.42 and 4.43. To obtain the second bound in (4.83), we expand $(h_1-h_2)^2$ as

$$\begin{aligned}&\sum_{n_1,n_2=0}^{\infty}\Big(p_1^{(n_1)}\big(q_1\circ\mathscr{S}_1^{n_1}-q_2\circ\mathscr{S}_2^{n_1}\big)p_1^{(n_2)}\big(q_1\circ\mathscr{S}_1^{n_2}-q_2\circ\mathscr{S}_2^{n_2}\big)+p_1^{(n_1)}\big(q_1\circ\mathscr{S}_1^{n_1}-q_2\circ\mathscr{S}_2^{n_1}\big)\big(p_1^{(n_2)}-p_2^{(n_2)}\big)q_2\circ\mathscr{S}_2^{n_2}\\&+\big(p_1^{(n_1)}-p_2^{(n_1)}\big)\,q_2\circ\mathscr{S}_2^{n_1}\,p_1^{(n_2)}\,\big(q_1\circ\mathscr{S}_1^{n_2}-q_2\circ\mathscr{S}_2^{n_2}\big)+\big(p_1^{(n_1)}-p_2^{(n_1)}\big)\,q_2\circ\mathscr{S}_2^{n_1}\,\big(p_1^{(n_2)}-p_2^{(n_2)}\big)q_2\circ\mathscr{S}_2^{n_2}\Big),\end{aligned}$$

where, writing, for $i=1,2$, $q_1\circ\mathscr{S}_1^{n_i}-q_2\circ\mathscr{S}_2^{n_i}$ according to (4.65) and $p_1^{(n_i)}-p_2^{(n_i)}$ using first (4.57a) and then once more (4.65), we obtain a sum of contributions which can be dealt with as the contributions in (4.76). Then, using also that $|||q_2|||_{\alpha_0,3}\le C\rho$, the second bound follows. □

Finally, the following results shows that bounds analogous to those of Proposition 4.44 extend to the derivatives of the two functions h_1 and h_2. The proof, which follows the same lines of the proof of Proposition 4.44, is given in Appendix B.2.2.

Proposition 4.45. *Let h_1 and h_2 be as in* (3.13) *and in* (4.32)*, respectively. For all $\theta\in\mathcal{U}$, one has*

$$|\langle\partial_\theta h_1(\theta,\cdot)-\partial_\theta h_2(\theta,\cdot)\rangle|\le C\rho,\qquad\big\langle(\partial_\theta h_1(\theta,\cdot)-\partial_\theta h_2(\theta,\cdot))^2\big\rangle\le C\rho\,.\tag{4.84}$$

Remark 4.46. By looking at the proof of Proposition 4.45, we see that $F=\rho f$ has to be required to belong to $\mathscr{B}_{\alpha_0,5}$, because we need to apply Proposition 4.40. The further condition that F be in $\mathscr{B}_{\alpha_0,6}$ is needed in order to control the deviations of the second derivative of the conjugation, which in turn is used to estimate the deviations of the first derivative of the inverse conjugation (see Remark 4.48).

4.7.5 Deviations of the conjugation: proof of the bounds (2.36a) and (2.36b)

We now come back to $h(\varphi,\psi)-\overline{h}(\varphi)$, with $h(\varphi,\psi)$ written as in (4.26) (taking in mind Remark 4.14 for the notation), and estimate the whole average. After rewriting

$$\langle h(\varphi,\cdot)-\overline{h}(\varphi)\rangle=\big\langle h_1(\varphi,\cdot)-\overline{h}(\varphi)\big\rangle+\big\langle W(\cdot)\,\partial_\varphi h_1(\varphi,\cdot)\big\rangle+\int_0^1 \mathrm{d}t\,(1-t)\,\big\langle (W(\cdot))^2\,\partial_\varphi^2 h_1(\varphi+tW(\cdot),\cdot)\big\rangle,$$

we bound $|\langle h_1(\varphi,\cdot)-\overline{h}(\varphi)\rangle|\le C\rho$ by Proposition 4.15, and $\langle (W(\cdot))^2\rangle\le C\rho$ by Theorem 4. Writing $\big\langle W(\cdot)\,\partial_\varphi h_1(\varphi,\cdot)\big\rangle=\big\langle W(\cdot)\,\partial_\varphi (h_1(\varphi,\cdot)-\overline{h}(\varphi))\big\rangle+\langle W\rangle\partial_\varphi\overline{h}(\varphi)$, Theorem 4 gives $|\langle W\rangle\partial_\varphi\overline{h}(\varphi)|\le C\rho$ while we use first the Cauchy-Schwarz inequality, then Proposition 4.15, together again with Theorem 4, to obtain $|\big\langle W(\cdot)\,\partial_\varphi(h_1(\varphi,\cdot)-\overline{h}(\varphi))\big\rangle|\le (\big\langle (W(\cdot))^2\big\rangle\langle(\partial_\varphi h_1(\varphi,\cdot)-\partial_\varphi\overline{h}(\varphi))^2\rangle)^{\frac12}\le C\rho$. This concludes the proof of the first bound in (2.36a).

The second bound in (2.36a) is easily obtained by writing

$$\begin{aligned}\big(h(\varphi,\psi)-\overline{h}(\varphi)\big)^2&=\left(h_1(\varphi,\psi)-\overline{h}(\varphi)-W(\psi)\int_0^1\mathrm{d}t\,\partial_\varphi h(\varphi-tW(\psi),\psi)\right)^2\\&\le 2\left(\big(h_1(\varphi,\psi)-\overline{h}(\varphi)\big)^2+(W(\psi))^2\,\|\partial_\varphi h\|^2\right)\\&\le 4\,(h_1(\varphi,\psi)-h_2(\varphi,\psi))^2+4\,\big(h_2(\varphi,\psi)-\overline{h}(\varphi)\big)^2+2(W(\psi))^2\,\|\partial_\varphi h\|^2,\end{aligned}$$

which in turn is bounded using once more Proposition 4.15 and Theorem 4.

Now we pass to the bounds (2.36). In order to study the average of $\partial_\varphi h(\varphi,\psi)-\partial_\varphi\overline{h}(\varphi)$, we write

$$\begin{aligned}\partial_\varphi h(\varphi,\psi)&=\partial_\varphi h_1(\varphi,\psi)-W(\psi)\int_0^1\mathrm{d}t\,\partial_\varphi^2 h(\varphi-tW(\psi),\psi)=\partial_\varphi h_1(\varphi,\psi)\\&\quad-W(\psi)\left(\partial_\varphi^2 h_1(\varphi,\psi)-\partial_\varphi^2\overline{h}(\varphi)\right)-W(\psi)\,\partial_\varphi^2\overline{h}(\varphi)+(W(\psi))^2\int_0^1\mathrm{d}t\,\partial_\varphi^3 h(\varphi-tW(\psi),\psi).\end{aligned}\tag{4.85}$$

Therefore, using the second expansion in (4.85), we can bound

$$\begin{aligned}\big|\big\langle\partial_\varphi h(\varphi,\cdot)-\partial_\varphi\overline{h}(\varphi)\big\rangle\big|&\le\big|\big\langle\partial_\varphi h_1(\varphi,\cdot)-\partial_\varphi\overline{h}(\varphi)\big\rangle\big|\\&\quad+\big|\big\langle W(\cdot)\left(\partial_\varphi^2 h_1(\varphi,\cdot)-\partial_\varphi^2\overline{h}(\varphi)\right)\big\rangle\big|+\langle W(\psi)\rangle\partial_\varphi^2\overline{h}(\varphi)+\big\langle(W(\psi))^2\big\rangle\|\partial_\varphi^3 h\|_\infty.\end{aligned}\tag{4.86}$$

The term in the first line and the last two terms in the second line are bounded by (4.27b) and by Theorem 4, respectively. To bound the first term in the second line of (4.86) we need the following result, which is proved in Appendix B.3.

Lemma 4.47. *Let h_1 and $\overline{h}$ be defined as in* (3.13) *and in* (4.19)*, respectively. For all $\theta\in\mathcal{U}$ one has*

$$\big|\big\langle(\partial_\theta^2 h_1(\theta,\cdot)-\partial_\theta^2\overline{h}(\theta))^2\big\rangle\big|\le C\rho.$$

Remark 4.48. As mentioned in Remark 4.46, the condition $F\in\mathscr{B}_{\alpha_0,6}$ is required to obtain the bounds in Lemma 4.47. A bound like $|\big\langle\partial_\theta^2 h_1(\theta,\cdot)-\partial_\theta^2\overline{h}(\theta)\big\rangle|\le C\rho$ could be obtained as well, but actually we need only the bound on the squared deviations because the latter will be needed in order to estimate the deviations of first derivative of the inverse conjugation.

By the Cauchy-Schwarz inequality and Lemma 4.47, we get $\big|\big\langle W(\cdot)\left(\partial_\varphi^2 h_1(\varphi,\cdot)-\partial_\varphi^2\overline{h}(\varphi)\right)\big\rangle\big|\le C\rho$, which inserted into (4.86) yields the first bound in (2.36b).

Finally, we pass to the second bound in (2.36b). By using the first expansion in (4.85), we may bound $|\partial_\varphi h(\varphi,\psi)-\partial_\varphi\overline{h}(\varphi)\le|\partial_\varphi h_1(\varphi,\psi)-\partial_\varphi\overline{h}(\varphi)|+|W(\psi)|\,\|\partial_\varphi^2 h\|_\infty$, which allows us to estimate $\langle(\partial_\varphi h(\varphi,\cdot)-\partial_\varphi\overline{h}(\varphi))^2\rangle\le 2\langle(\partial_\varphi h_1(\varphi,\cdot)-\partial_\varphi\overline{h}(\varphi))^2\rangle+2\|\partial_\varphi^2 h\|_\infty^2\langle(W(\cdot))^2\rangle$. This immediately implies the second bound in (2.36b) by (4.55) and by Theorem 4.

4.8 Deviations of the inverse conjugation

To deal with the inverse conjugation, we exploit the following trivial identity.

$$\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) = \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta)) = \eta, \tag{4.87}$$

which holds for all $(\eta,\psi)\in\Omega_0 = \mathcal{U}_0\times\mathbb{T}^2 := \overline{\mathscr{H}}(\Omega)$, provided $\Omega_{\rm ext}$ is such that $\mathscr{H}(\Omega_{\rm ext})\supset\overline{\mathscr{H}}(\Omega)$ and $\Omega_{1,\rm ext}\supset\Omega$ (recall that we are working with the extended maps).

Lemma 4.49. *Let l_1 and $\bar l$ be defined as in* (3.14) *and in* (2.30)*, respectively. For all $\eta\in\mathcal{U}_0$, one has*

$$|\langle l_1(\eta,\cdot)-\bar l(\eta)\rangle| \le C\rho, \qquad \langle (l_1(\eta,\cdot)-\bar l(\eta))^2\rangle \le C\rho\,. \tag{4.88}$$

Proof. By using (4.87) we get

$$\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) - \mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi) = \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta)) - \mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi). \tag{4.89}$$

If we write the l.h.s. in (4.89) as

$$\big(\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta)\big)\int_0^1 {\rm d}t\,\partial_\theta\mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta)+t(\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta)),\psi)$$

and use the lower bound (3.29), we obtain from (4.89)

$$\gamma_1^2\big\langle\big(\mathcal{H}_1^{-1}-\overline{\mathcal{H}}^{-1}\big)^2\big\rangle(\eta) \le \big\langle\big(\mathcal{H}_1-\overline{\mathcal{H}}\big)^2\big\rangle\big(\overline{\mathcal{H}}^{-1}(\eta)\big). \tag{4.90}$$

Analogously, if we write l.h.s. of (4.89) as

$$\begin{aligned}&\big(\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta)\big)\partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta)) + \big(\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta)\big)\Big(\partial_\theta\mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi)-\partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\Big)\\ &\quad + \big(\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta)\big)^2\int_0^1(1-t)\,{\rm d}t\,\partial_\theta^2\mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta)+t(\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta)),\psi),\end{aligned}$$

then we find

$$\begin{aligned}\bar\gamma\big|\big\langle\mathcal{H}_1^{-1}-\overline{\mathcal{H}}^{-1}\big\rangle(\eta)\big| &\le \partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\big|\big\langle\mathcal{H}_1^{-1}-\overline{\mathcal{H}}^{-1}\big\rangle(\eta)\big| \le \big|\big\langle\mathcal{H}_1-\overline{\mathcal{H}}\big\rangle\big(\overline{\mathcal{H}}^{-1}(\eta)\big)\big|\\ &+\|\partial_\theta^2\mathcal{H}_1\|_\infty\big\langle\big(\mathcal{H}_1^{-1}-\overline{\mathcal{H}}^{-1}\big)^2\big\rangle(\eta) + \Big(\big\langle\big(\mathcal{H}_1^{-1}-\overline{\mathcal{H}}^{-1}\big)^2\big\rangle(\eta)\big\langle\big(\partial_\theta\mathcal{H}_1-\partial_\theta\overline{\mathcal{H}}\big)^2\big\rangle\big(\overline{\mathcal{H}}^{-1}(\eta)\big)\Big)^{1/2},\end{aligned} \tag{4.91}$$

with $\bar\gamma = O(1)$ as in Remark 3.13. Thus, using that $\mathcal{H}_1(\theta,\psi)-\overline{\mathcal{H}}(\theta) = \theta^2\big(h_1(\theta,\psi)-\bar h(\theta)\big)$ and $\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta) = \eta^2\big(l_1(\eta,\psi)-\bar l(\eta)\big)$ and that, for $\theta=\overline{\mathcal{H}}^{-1}(\eta)$, there exists a constant d_0 such that $d_0^{-1}\le\theta/\eta\le d_0$ (see Section 3.2.4), the bound (4.90), together with the second bound in (4.27a), gives the second bound in (4.88), which, in turn, inserted into (4.91), together with the first bound in (4.27a) and the second bound in (4.27b), yields the first bound in (4.88). □

Lemma 4.50. *Let l_1 and $\bar l$ be defined as in* (3.14) *and in* (2.30)*, respectively. For all $\eta\in\mathcal{U}_0$, one has*

$$|\langle \partial_\eta l_1(\eta,\cdot)-\partial_\eta\bar l(\eta)\rangle| \le C\rho, \qquad \langle (\partial_\eta l_1(\eta,\cdot)-\partial_\eta\bar l(\eta))^2\rangle \le C\rho\,. \tag{4.92}$$

Proof. Differentiating (4.87) with respect to η, we obtain

$$\partial_\theta\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)\,\partial_\eta\mathcal{H}_1^{-1}(\eta,\psi) = \partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\,\partial_\eta\overline{\mathcal{H}}^{-1}(\eta) = 1, \tag{4.93}$$

that implies

$$\partial_\eta\mathcal{H}_1^{-1}(\eta,\psi)-\partial_\eta\overline{\mathcal{H}}^{-1}(\eta) = -\frac{\partial_\theta\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)-\partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))}{\partial_\theta\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)\,\partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))}. \tag{4.94}$$

Thus, we find that square $\big(\partial_\eta\mathcal{H}_1^{-1}(\eta,\psi)-\partial_\eta\overline{\mathcal{H}}^{-1}(\eta)\big)^2$ is bounded by

$$\frac{2}{\tau_1^2\bar\tau^2}\big(\|\partial_\theta^2\mathcal{H}_1\|_\infty^2\big(\mathcal{H}_1^{-1}(\eta,\psi)-\overline{\mathcal{H}}^{-1}(\eta)\big)^2 + \big(\partial_\theta\mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi)-\partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\big)^2\big),$$

where τ_1 and $\overline{\tau}$ are on lower bounds, respectively, on $\partial_\theta \mathcal{H}_1$ and $\partial_\theta \overline{\mathcal{H}}$ (see Subsection 3.2.4), so that

$$\left\langle \left(\partial_\eta l_1 - \partial_\eta \bar{l}\right)^2 \right\rangle \le \frac{2}{\tau_1^2 \overline{\tau}^2} \left(\|\partial_\theta^2 \mathcal{H}_1\|_\infty^2 \left\langle \left(l_1 - \bar{l}\right)^2 \right\rangle + \left\langle \left(\partial_\theta h_1 - \partial_\theta \overline{h}\right)^2 \right\rangle \right),$$

which, together with (4.27b) and (4.88), implies the second of (4.92). Rewriting (4.94) as

$$-\frac{\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) - \partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))}{\left(\partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\right)^2} + \frac{\left(\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) - \partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\right)^2}{\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) \left(\partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\right)^2},$$

and expanding

$$\begin{aligned}
\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) &= \partial_\theta \mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi) + \partial_\theta^2 \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta)) \left((\mathcal{H}_1^{-1}(\eta,\psi) - \overline{\mathcal{H}}^{-1}(\eta) \right) \\
&\quad + \left(\partial_\theta^2 \mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi) - \partial_\theta^2 \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta)) \right) \left((\mathcal{H}_1^{-1}(\eta,\psi) - \overline{\mathcal{H}}^{-1}(\eta) \right) \\
&\quad + \left((\mathcal{H}_1^{-1}(\eta,\psi) - \overline{\mathcal{H}}^{-1}(\eta) \right)^2 \int_0^1 \mathrm{d}t\, (1-t) \partial_\theta^3 \mathcal{H}_1\big(\overline{\mathcal{H}}^{-1}(\eta),\psi) + t\big((\mathcal{H}_1^{-1}(\eta,\psi) - \overline{\mathcal{H}}^{-1}(\eta)\big)\big),
\end{aligned}$$

we obtain eventually

$$\begin{aligned}
\left|\left\langle \partial_\eta l_1 - \partial_\eta \bar{l}(\eta)\right\rangle\right| &\le \frac{1}{\overline{\tau}^2} \Big(\left|\left\langle \partial_\theta h_1 - \partial_\theta \overline{h}\right\rangle\right| + \|\partial_\theta^2 \overline{\mathcal{H}}\|_\infty \left|\left\langle l_1 - \bar{l}\right\rangle\right| + \left(\left\langle (\partial_\theta^2 h_1 - \partial_\theta^2 \overline{h})^2 \right\rangle \left\langle (l_1 - \bar{l})^2 \right\rangle \right)^{1/2} \\
&\quad + \|\partial_\theta^2 \mathcal{H}_1\|_\infty \left|\left\langle (l_1 - \bar{l})^2 \right\rangle\right| \Big) + \frac{2}{\tau_1^2 \overline{\tau}} \Big(\|\partial_\theta^3 \mathcal{H}_1\|_\infty \left|\left\langle (l_1 - \bar{l})^2 \right\rangle\right| + \left|\left\langle (\partial_\theta h_1 - \partial_\theta \overline{h})^2 \right\rangle\right| \Big),
\end{aligned}$$

so that the first of (4.92) follows from the estimates of Propositions 4.15 and Lemmas 4.47 and 4.50. $\square$

Recalling that $l = l_1$, by (3.15), Lemmas 4.49 and 4.50 immediately imply the following result.

Proposition 4.51. *Let l be defined as in* (2.22). *There is a constant C such that, for all $\eta \in \mathcal{U}_0$,*

$$\left|\left\langle l(\eta,\cdot) - \bar{l}(\eta)\right\rangle\right| \le C\rho, \qquad \left\langle (l(\eta,\cdot) - \overline{h}(\eta))^2 \right\rangle \le C\rho \tag{4.95a}$$

$$\left|\left\langle \partial_\eta l(\eta,\cdot) - \partial_\eta \bar{l}(\eta)\right\rangle\right| \le C\rho, \qquad \left\langle (\partial_\eta l(\eta,\cdot) - \partial_\eta \bar{l}(\eta))^2 \right\rangle \le C\rho\,. \tag{4.95b}$$

4.9 Fluctuations of the dynamics: proof of Theorem 6

We start by proving the following result for the extension of the translated map $\mathscr{S}_1$.

Lemma 4.52. *For any $\gamma' \in (0,\gamma)$ there exists a constant C such that for all $\theta \in \mathcal{U}$ and all $n \in \mathbb{N}$*

$$\left|\left\langle (\mathscr{S}_1^n)_\theta(\theta,\cdot) - (\overline{\mathscr{S}}^n)_\theta(\theta)\right\rangle\right| \le C\rho(1-\rho\,\gamma')^n, \qquad \left\langle \left((\mathscr{S}_1^n)_\theta(\theta,\cdot) - (\overline{\mathscr{S}}^n)_\theta(\theta)\right)^2 \right\rangle \le C\rho\,(1-\rho\,\gamma')^{2n}.$$

Proof. From (2.29), (3.12) and (3.14), we find

$$\begin{aligned}
(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta) &= \mathcal{H}_1^{-1}(\mathscr{S}_0^n(\mathcal{H}_1(\theta,\psi),\psi)) - \overline{\mathcal{H}}^{-1}(\overline{\mathscr{S}}_0^n(\overline{\mathcal{H}}(\theta),\psi)) \\
&= \kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n \overline{\mathcal{H}}(\theta) + \Big(\big(\kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi)\big)^2 - \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2 \Big)\, \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big) \\
&\quad + \Big(\big(\kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi)\big)^2 - \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2 \Big) \Big(l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big) \Big) \\
&\quad + \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2 \Big(l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big) \Big).
\end{aligned} \tag{4.96}$$

Shorten for notational simplicity

$$\begin{aligned}
\mathcal{A}_n(\theta,\psi) &:= \kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}, & \lambda_1(\theta,\psi) &:= l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big), \\
\bar{\lambda}_1(\theta,\psi) &:= l_1\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta), A_0^n\psi)\big), & \lambda_2(\theta,\psi) &:= \partial_\eta l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big), \\
\bar{\lambda}(\theta) &:= \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta))\big), & \bar{\lambda}_2(\theta) &:= \partial_\eta \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta))\big),
\end{aligned}$$

$$I_n(\theta,\psi) := \int_0^1 \mathrm{d}t\,(1-t)\,\partial_\eta^2 l_1\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta) + t\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big), A_0^n\psi\big),$$

and observe that $(\kappa^{(n)}\mathcal{H}_1)^2 - (\overline{\mu}^n\,\overline{\mathcal{H}})^2 = \mathcal{A}_n^2 + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}\,\mathcal{A}_n$ and

$$\begin{aligned} \mathcal{A}_n &= \big(\kappa^{(n)} - \overline{\mu}^n\big)\big(\mathcal{H}_1 - \overline{\mathcal{H}}\big) + \big(\kappa^{(n)} - \overline{\mu}^n\big)\overline{\mathcal{H}} + \overline{\mu}^n\big(\mathcal{H}_1 - \overline{\mathcal{H}}\big) \\ &= \theta^2\big(\kappa^{(n)} - \overline{\mu}^n\big)\big(h_1 - \overline{h}\big) + \big(\kappa^{(n)} - \overline{\mu}^n\big)\overline{\mathcal{H}} + \theta^2\overline{\mu}^n\big(h_1 - \overline{h}\big), \\ \lambda_1 - \bar{\lambda} &= \lambda_2\mathcal{A}_n + I_n\,\mathcal{A}_n^2 + \big(\bar{\lambda}_1 - \bar{\lambda}\big) = \overline{\lambda}_2\,\mathcal{A}_n + \big(\lambda_2 - \overline{\lambda}_2\big)\,\mathcal{A}_n + I_n\mathcal{A}_n^2 + \big(\overline{\lambda}_1 - \overline{\lambda}\big)\,. \end{aligned} \tag{4.97}$$

Then, we rewrite (4.96) as

$$(\mathscr{S}_1^n)_\theta - (\overline{\mathscr{S}}^n)_\theta = \mathcal{A}_n + \big(\mathcal{A}_n^2 + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}\,\mathcal{A}_n\big)\overline{\lambda} + \big(\mathcal{A}_n^2 + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}\,\mathcal{A}_n\big)\big(\lambda_1 - \overline{\lambda}\big) + \big(\overline{\mu}^n\,\overline{\mathcal{H}}\big)^2\big(\lambda_1 - \bar{\lambda}\big),$$

with $\mathcal{A}_n$ and $\lambda_1 - \overline{\lambda}$ as in the second and fourth line of (4.97), respectively. Thus, using Lemma 2.31, Propositions 4.15 and 4.50, and the Cauchy-Schwarz inequality, the bound (2.38a) follows.

To obtain the second bound in (2.38) we write (4.96) as

$$(\mathscr{S}_1^n)_\theta - (\overline{\mathscr{S}}^n)_\theta = \mathcal{A}_n + \big(\mathcal{A}_n^2 + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}\,\mathcal{A}_n\big)\lambda_1 + \big(\overline{\mu}^n\,\overline{\mathcal{H}}\big)^2\big(\lambda_1 - \bar{\lambda}\big),$$

with $\mathcal{A}_n$ and $\lambda_1 - \overline{\lambda}$ as in the second and third line of (4.97), respectively, so that, bounding

$$\begin{aligned} \big((\mathscr{S}_1^n)_\theta - (\overline{\mathscr{S}}^n)_\theta\big)^2 &\le 3\,\Big(\mathcal{A}_n^2 + 2\,\mathcal{A}_n^2\big(\mathcal{A}_n^2 + 4\,(\overline{\mu}^n\,\overline{\mathcal{H}})^2\big)\lambda_1^2 + \big(\overline{\mu}^n\,\overline{\mathcal{H}}\big)^4\big(\lambda_1 - \bar{\lambda}\big)^2\Big)\,, \\ \big(\lambda_1 - \bar{\lambda}\big)^2 &\le 3\,\Big(\lambda_2^2\mathcal{A}_n^2 + I_n^2\,\mathcal{A}_n^4 + \big(\bar{\lambda}_1 - \bar{\lambda}\big)^2\Big)\,, \\ \mathcal{A}_n^2 &\le 2\,\Big(\big(\kappa^{(n)} - \overline{\mu}^n\big)^2\overline{\mathcal{H}} + \theta^4\overline{\mu}^{2n}\big(h_1 - \overline{h}\big)^2\Big)\,, \end{aligned}$$

we obtain that $\langle(\,(\mathscr{S}_1^n)_\theta(\theta,\cdot) - (\overline{\mathscr{S}}^n)_\theta(\theta)\,)^2\rangle$ is bounded by

$$C\Big(\big\langle\big(\kappa^{(n)}(\cdot) - \overline{\mu}^n\big)^2\big\rangle + \overline{\mu}^{2n}\big\langle\big(h_1(\theta,\cdot) - \bar{h}(\theta)\big)^2\big\rangle + \overline{\mu}^{4n}\big\langle\big(l_1(\overline{\mu}^n\overline{\mathcal{H}}(\theta),\cdot) - \bar{l}(\overline{\mu}^n\overline{\mathcal{H}}(\theta)\big)^2\big\rangle\Big).$$

Then, using Lemma 2.31 and Propositions 4.15 and 4.50 gives immediately the bound (2.38b). □

Finally, we come back to the map $\mathscr{S}$. In order to compare the full dynamics generated by $\mathscr{S}$ with that generated by $\overline{\mathscr{S}}$, we proceed along the same lines as the proof of Lemma 4.52. Thus, we decompose $\big(\mathscr{S}^n\big)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi)$ as done for $\big(\mathscr{S}_1^n\big)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta)$, relying on (2.21) and (2.22) instead of (3.12) and (3.14), and obtain

$$\begin{aligned} &(\mathscr{S}^n)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi) = \mathcal{H}^{-1}(\mathscr{S}_0^n(\mathcal{H}(\theta,\psi),\psi)) - \overline{\mathcal{H}}^{-1}(\overline{\mathscr{S}}_0^n(\overline{\mathcal{H}}(\theta),\psi)) \\ &\quad = W(A_0^n\psi) + \kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta) + \Big(\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\Big)^2\,\bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big) \\ &\quad + \Big(\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\Big)^2\,\Big(l\big(\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi),\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big) \\ &\quad + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\,\Big(\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\Big)\,\bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big) \\ &\quad + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\,\Big(\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\Big)\,\Big(l\big(\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi),\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big) \\ &\quad + \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2\Big(\partial_\eta\bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\,\big(\kappa^{(n)}(\psi)\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big) \\ &\quad + \big(\partial_\eta l\big(\overline{\mu}^n\overline{\mathcal{H}}(\theta),A_0^n\psi\big)\big) - \partial_\eta\bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\,\big(\kappa^{(n)}(\psi)\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big) \\ &\quad + \big(\kappa^{(n)}(\psi)\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2\int_0^1 \mathrm{d}t\,(1-t)\,\partial_\eta^2 l\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta) + t\big(\kappa^{(n)}(\psi)\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big), A_0^n\psi\big) \\ &\quad + l\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta),A_0^n\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big), \end{aligned}$$

where we expand

$$\begin{aligned} &\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta) \\ &\quad = \theta^2\big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\big(h(\theta,\psi) - \overline{h}(\theta)\big) + \big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\overline{\mathcal{H}}(\theta) + \theta^2\overline{\mu}^n\big(h(\theta,\psi),A_0^n\psi) - \overline{h}(\theta)\big). \end{aligned}$$

Then, using the bounds of Lemma 2.31, Theorem 4, Proposition 4.51 and the Cauchy-Schwarz inequality, we obtain the bound (2.38a) of Theorem 6. Again, the bound (2.38b) is obtained in a similar way.

5 Convergence in probability

In this section we collect the previous results to prove Theorems 7 and 8, so as to provide the probabilistic description of the results discussed in Subsection 2.4.4. Below, as in Subsections 4.7 to 4.9, for notational simplicity, $\mathscr{S}_1$ denotes the extended map $\mathscr{S}_{1,\mathrm{ext}}$.

5.1 The probability of deviations: proof of Theorem 7

Recall that m_0 denotes the Lebesgue measure on $\mathbb{T}^2$. From Theorem 1 and a direct application of Chebyshev inequality we obtain that, for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : |W(\psi)| > \delta\Big\}\Big) \le \frac{C\rho}{\delta^2},$$

for a suitable positive constant C independent of n. Therefore, the invariant manifold $\mathcal{W}$ for $\mathscr{S}$ converges in probability to $\overline{\mathcal{W}} = \{(0,\psi) : \psi \in \mathbb{T}^2\}$.

Similarly, from Lemma 4.52, we obtain that, if $\gamma' \in (0,\gamma)$, then, for fixed θ and n, and for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \delta\Big\}\Big) \le \frac{c_1(\gamma')\,\rho}{\delta^2}, \tag{5.1}$$

for a suitable positive constant $c_1(\gamma')$ depending on γ' but neither on θ nor n.

The following result provides a uniform version of the estimate (5.1) and shows that, for most values of $\psi \in \mathbb{T}^2$, the quantity $(1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)|$ remains small for all $n \ge 0$.

Lemma 5.1. *Consider the dynamical system described by $\mathscr{S}$ in* (2.25) *satisfying Hypotheses 1–3. Let m_0 denote the Lebesgue measure on $\mathbb{T}^2$, and let $\mathscr{S}_1$ be defined as in* (3.10)*. For ρ small enough and any $\gamma' \in (0,\gamma)$ there exists a constant C such that, for all $\theta \in \mathcal{U}$,*

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : \sup_{n\ge 0}(1-\rho\,\gamma')^{-n}|\,(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \delta\Big\}\Big) \le \frac{C\rho}{\delta^3}.$$

Proof. For all $\theta, \theta' \in \mathcal{U}$ we have

$$|(\mathscr{S}_1)_\theta(\theta,\psi) - (\overline{\mathscr{S}})_\theta(\theta',\psi)| \ge (1-\rho)|\theta-\theta'| - \rho|f_1(\theta,\psi) - \bar f(\theta)| \ge (1-\rho)|\theta-\theta'| - C\rho|\theta|$$

where we have used that $|||f|||_{\alpha_0,6} = 1$ and $f_1(0,\psi) = \overline{f}(0) = 0$, as it follows from the definition of f_1 in Subsection 4.7.1 and from Remark 2.8. Iterating we obtain, for any $\theta,\theta' \in \mathcal{U}$ and for any $k \ge 0$,

$$|(\mathscr{S}_1^k)_\theta(\theta,\psi) - (\overline{\mathscr{S}}_\theta^k(\theta',\psi)| \ge (1-\rho)^k|\theta-\theta'| - Ck(1-\rho\,\gamma')^k\rho|\theta|,$$

where we have used that $(1-\rho) < (1-\rho\,\gamma')$, since $\gamma' < \gamma < 1$ (see Remark 2.19), and that, by Remark 3.10, $|(\mathscr{S}_1^k)_\theta(\theta,\psi)| \le C(1-\rho\,\gamma')^k|\theta|$. Assuming that $(1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \delta$ for some $n > 0$, we find that

$$(1-\rho\,\gamma')^{-(n+k)}|(\mathscr{S}_1^{n+k})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n+k})_\theta(\theta,\psi)| \ge \left(\frac{1-\rho}{1-\rho\,\gamma'}\right)^k \delta - Ck\rho\,.$$

For ρ small enough, we can now find $M \in \mathbb{N}$ of the form $M = M_0\delta/\rho$, with M_0 independent of ρ and δ, such that $Ck\rho \le \delta/4$ and $((1-\rho)/(1-\rho\,\gamma'))^k \ge 1/2$ for all $k \le M$. This means that for $k \le M$ we have $(1-\rho\,\gamma')^{-(n+k)}|(\mathscr{S}_1^{n+k})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n+k})_\theta(\theta,\psi)| \ge \delta/4$. We thus obtain that

$$\begin{aligned}&\sum_{n=0}^{\infty} m_0\left(\left\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right)\\ &\qquad \ge \frac{M_0\delta}{\rho}\, m_0\Big(\Big\{\psi \in \mathbb{T}^2 : \sup_{n\ge 0}(1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \delta\Big\}\Big).\end{aligned} \tag{5.2}$$

On the other hand, for any $\gamma', \gamma'' \in (0,\gamma)$ we find, by using (5.1),

$$m_0\left(\left\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma'')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right) \le \left(\frac{1-\rho\,\gamma''}{1-\rho\,\gamma'}\right)^{2n} \frac{16\,c_1(\gamma')\,\rho}{\delta^2},$$

and in the same way, by inverting the roles of γ' and γ'',

$$m_0\left(\left\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right) \le \left(\frac{1-\rho\,\gamma'}{1-\rho\,\gamma''}\right)^{2n} \frac{16\,c_1(\gamma'')\,\rho}{\delta^2},$$

so that, fixing $\gamma' \in (0,\gamma)$ and taking $\gamma'' = \gamma'/2$, we obtain

$$\sum_{n=0}^{\infty} m_0\left(\left\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right) \le \frac{c_2(\gamma')}{\delta^2}, \tag{5.3}$$

with

$$c_2(\gamma') = 16\,c_1(\gamma'/2)\,\rho \sum_{n=0}^{\infty}\left(\frac{1-\rho\,\gamma'}{1-\rho\,\gamma'/2}\right)^{2n}.$$

Inserting (5.3) into (5.2) delivers the thesis. □

To deal with $\mathscr{S}$, we follow the same argument used for $\mathscr{S}_1$. First of all, from Theorem 6 and Chebyshev inequality, if $\gamma' \in (0,\gamma)$, then, for fixed θ and n, and for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}^n)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\theta(\varphi,\psi) - W(A_0^n\psi)| > \delta\Big\}\Big) \le \frac{c_3(\gamma')\,\rho}{\delta^2},$$

with the constant $c_3(\gamma')$ independent of both θ and n. Next, observe that we can write

$$\begin{aligned}
\big|\mathscr{S}_\varphi(\varphi,\psi) - \overline{\mathscr{S}}_\varphi(\varphi',\psi) - W(A_0\psi)\big| &= \big|\mathscr{S}_\varphi(\varphi,\psi) - \overline{\mathscr{S}}_\varphi(\varphi',\psi) - W(\psi) - \rho\, f(W(\psi),\psi)\big| \\
&= \big|\varphi - W(\psi) - \varphi' + \rho\big(f(\varphi,\psi) - \rho\, f(W(\psi),\psi) - \rho\,\overline{f}(\varphi - W(\psi)\big) + \rho\big(\overline{f}(\varphi - W(\psi)) - \overline{f}(\varphi')\big)\big| \\
&\ge (1-\rho)\big|\varphi - W(\psi) - \varphi'\big| - C\rho\big|\varphi - W(\psi)\big|,
\end{aligned}$$

so that, iterating and using that $|(\mathscr{S}^i)_\varphi(\varphi,\psi) - W(A_0^i\psi)| = |(\mathscr{S}_1^i)_\theta(\theta,\varphi)| \le C(1-\rho\,\gamma')^i|\theta|$, with $\theta = \varphi - W(\psi)$, we find, once again thanks to Lemma 3.8,

$$\big|(\mathscr{S}^k)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^k)_\varphi(\varphi',\psi) - W(A_0^n\psi)\big| \ge (1-\rho)^k\big|\varphi - W(\psi) - \varphi'\big| - Ck(1-\rho\,\gamma')^k\rho\big|\varphi - W(\psi)\big|.$$

Therefore, assuming that $(1-\rho\,\gamma')^{-n}|(\mathscr{S}^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi) - W(A_0^n\psi)| > \delta$ for some $n > 0$ and proceeding as done for $\mathscr{S}_1$ leads to

$$\begin{aligned}
\frac{c_4(\gamma')}{\delta^2} &\ge \sum_{n=0}^{\infty} m_0\left(\left\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}^n)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) - W(A_0^n\psi)| > \frac{\delta}{4}\right\}\right) \\
&\ge \frac{M_0\delta}{\rho}\, m_0\Big(\Big\{\psi \in \mathbb{T}^2 : \sup_{n\ge 0}\,(1-\rho\,\gamma')^{-n}|(\mathscr{S}^n)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) - W(A_0^n\psi)| > \delta\Big\}\Big).
\end{aligned}$$

for a suitable constant $c_4(\gamma')$. This completes the proof of Theorem 7.

5.2 Continuous time: proof of Theorem 8

Relying on Remark 3.3, we get

$$\begin{aligned}
\mathcal{X}_t(\psi) = {} & (\mathscr{S}_1^{\lfloor t/\rho\rfloor})_\theta(\varphi_0 - W(\psi),\psi) \\
& + (t/\rho - \lfloor t/\rho\rfloor)\left((\mathscr{S}_1^{\lfloor t/\rho\rfloor+1})_\theta(\varphi_0 - W(\psi),\psi) - (\mathscr{S}_1^{\lfloor t/\rho\rfloor})_\theta(\varphi_0 - W(\psi),\psi)\right).
\end{aligned}$$

By Lemma 3.8, we obtain the bound $\big|(\mathscr{S}_1^k)_\theta(\varphi_0 - W(\psi),\psi) - (\mathscr{S}_1^k)_\theta(\varphi_0,\psi)\big| \le C(1-\rho\,\gamma')^k|W(\psi)|$, which, combined with Lemmas 4.52 and 2.28, allows us to estimate, for $\xi \in (0,\gamma)$,

$$\sup_{t\ge 0} e^{\xi t}|\mathcal{X}_t(\psi) - \Phi_t(\varphi_0)| \le \sup_{n\ge 0}\,(1-\rho\,\xi)^{-n}\left|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)\right| + C\left(|W(\psi)| + \rho\right),$$

so that, for every $\delta > 0$, we have

$$\lim_{\rho\to 0^+} m_0\Big(\Big\{\psi \in \mathbb{T}^2 : \sup_{t\geq 0} e^{\xi t}|\mathcal{X}_t(\psi) - \Phi_t(\varphi_0)| \geq \delta\Big\}\Big) = 0,$$

because of (2.34) and Lemma 5.1.

Delarations

Conflict of Interest

The authors have no conflicts of interest to disclose.

Author Contributions

Federico Bonetto: Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal). Guido Gentile: Formal analysis (equal); Writing – original draft (equal); Writing – review & editing (equal).

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Acknowledgements

F.B. has been partially supported by the NSF grant DMS-1907643. G.G. has been partially supported by the research project PRIN 20223J85K3_002 "Mathematical Interacting Quantum Fields" of the Italian Ministry of Education and Research (MIUR).

A Decay of correlations for Hölder continuous functions

A.1 Symbolic dynamics

Consider a Markov partition $\mathcal{Q} = \{Q_1, Q_2, \dots, Q_s\}$ for the Anosov automorphism A_0 in (2.11) and the corresponding compatibility matrix T, that is the $s \times s$ matrix T such that $T_{ij} = 1$ if $A_0Q_i \cap Q_j \neq \emptyset$ and $T_{ij} = 0$ otherwise. Let a denote the mixing time, i.e. the minimum value $a \in \mathbb{N}$ such that all the entries of T^{a+1} are different from 0 (since every Anosov automorphism is transitive, a is finite), and let λ_i be the eigenvalues of T ordered according to their modulus. Then λ_1, which is real and simple by the Perron-Frobenius Theorem, equals $\lambda = |\lambda_+|$, with λ_+ being the eigenvalue of A_0 with the largest modulus [29, Ch. 6]. Set $\mathcal{N} = \{1, 2, \dots, s\}$. We say that a sequence $\underline{\sigma} \in \mathcal{N}^{\mathbb{Z}}$ is T-compatible if $T_{\sigma_k,\sigma_{k+1}} = 1$ for all $k \in \mathbb{Z}$; let $\mathcal{N}_T^{\mathbb{Z}}$ denote the set of T-compatible sequences. We use the symbol $\underline{\sigma}$ also to denote subsequences, i.e. elements of $\mathcal{N}^{[n,m]}$, with finite n and m. For $\underline{\sigma}, \underline{\sigma}' \in \mathcal{N}_T^{\mathbb{Z}}$, set $\nu(\underline{\sigma}, \underline{\sigma}') := \max\{k \in \mathbb{N} : \sigma_i = \sigma_i', |i| < k\}$. On $\mathcal{N}_T^{\mathbb{Z}}$ we consider the topology generated by the distance $d(\underline{\sigma}, \underline{\sigma}') = \lambda^{-\nu(\underline{\sigma},\underline{\sigma}')}$. Given $\underline{\sigma} \in \mathcal{N}_T^{\mathbb{Z}}$, let $\psi = \psi(\underline{\sigma}) \in \mathbb{T}^2$ be the unique point whose symbolic representation is $\underline{\sigma}$; then we have $A_0\psi(\underline{\sigma}) = \psi(\tau\underline{\sigma})$, where τ is the left translation, that is $(\tau\underline{\sigma})_i = \sigma_{i+1}$, and $|\psi(\underline{\sigma}) - \psi(\underline{\sigma}')| \leq D_s\, d(\underline{\sigma}, \underline{\sigma}')$, where D_s is the maximum diameter of the sets $Q_1, \dots, Q_s$. We call $C^J_{\underline{\sigma}}$ the T-compatible cylinder with base $J = (j_1, \dots, j_q) \subset \mathbb{Z}$ and specification $\underline{\sigma} = (\sigma_{j_1}, \dots, \sigma_{j_q}) \in \mathcal{N}_T^q$, i.e. the set of sequences $\underline{\sigma}' \in \mathcal{N}_T^{\mathbb{Z}}$ such that $\sigma'_{j_k} = \sigma_{j_k}$ for $k = 1, \dots, q$. Let $\mathcal{F}^{\leq k}$ be the σ-algebra generated by the cylinder sets $C^J_{\underline{\sigma}}$ with $J \subset (-\infty, k]$ and $\mathcal{F}^{\geq k}$ be the σ-algebra generated by the cylinders $C^J_{\underline{\sigma}}$ with $J \subset [k, \infty)$. The following result is implied by the mixing properties of T.

Lemma A.1. *Given $k, k' \in \mathbb{Z}$, with $k' > k$, let g_- be a bounded $\mathcal{F}^{\leq k}$-measurable function and g_- be a bounded $\mathcal{F}^{\geq k'}$-measurable function. One has*

$$|\langle g_- g_+ \rangle - \langle g_- \rangle \langle g_+ \rangle| \leq K_1 \lambda_0^{-(k'-k)} \|g_-\|_\infty \|g_+\|_\infty \, .$$

with $\lambda_0 := \lambda / |\lambda_2|$ and K_1 a suitable constant.

Remark A.2. The eigenvalues of the compatibility matrix of any Anosov automorphism on $\mathbb{T}^2$ are λ, λ^{-1}, together with 0's and roots of unity [58]. Therefore we have $\lambda_0 \geq \lambda$ in Lemma A.1.

Remark A.3. If $g_-(\underline{\sigma})$ is $\mathcal{F}^{\leq k}$-measurable, then $g_-(\underline{\sigma})$ depends only on σ_i with $i \leq k$, that is $g_-(\underline{\sigma}) = g_-(\underline{\sigma}')$ if $\sigma_i = \sigma_i'$ for every $i \leq k$. Similarly, if $g_+(\underline{\sigma})$ is $\mathcal{F}^{\geq k}$-measurable, then $g_+(\underline{\sigma})$ depends only on σ_i with $i \geq k'$, that is $g_+(\underline{\sigma}) = g_+(\underline{\sigma}')$ if $\sigma_i = \sigma_i'$ for every $i \geq k'$.

Let the sequences $\underline{\sigma}, \underline{\sigma}' \in \mathcal{N}_T^{\mathbb{Z}}$ be such that $\sigma_i = \sigma_i'$ for $i \leq n$. If $n > 0$ then $d(\underline{\sigma}, \underline{\sigma}') \leq \lambda^{-n}$ and, for $N \geq 0$, $d(\tau^{-N}\underline{\sigma}, \tau^{-N}\underline{\sigma}') = \lambda^{-N} d(\underline{\sigma}, \underline{\sigma}')$. Thus $\psi(\underline{\sigma}')$ is on the unstable manifold of $\psi(\underline{\sigma})$ and, since the unstable manifold is a straight line, we have $\psi(\underline{\sigma}') = \psi(\underline{\sigma}) + x\, v_+$, with $|x| \leq D_s \lambda^{-n}$. If $n \leq 0$ we still have that $d(\tau^{-N}(\underline{\sigma}), \tau^{-N}(\underline{\sigma}')) \to 0$ as $N \to \infty$, so that even in this case we have $\psi(\underline{\sigma}') = \psi(\underline{\sigma}) + x v_+$ for some $x \in \mathbb{R}$. Observe that $\psi(\tau^{2n}\underline{\sigma}') = \psi(\tau^{2n}\underline{\sigma}) + \lambda^{2n} x\, v_+$ while, since $(\tau^{2n}\underline{\sigma}')_i = (\tau^{2n}\underline{\sigma})_i$ for $i \leq -n$, we have $\|\psi(\tau^{2n}\underline{\sigma}') - \psi(\tau^{2n}\underline{\sigma})\| \leq D_s \lambda^{-n}$, so that, by the previous argument, we find again that $|x| \leq D_s \lambda^{-n}$. In the same way one shows that, if $\underline{\sigma}$ and $\underline{\sigma}'$ are such that $\sigma_i = \sigma_i'$ for $i \geq n$, then $\psi(\underline{\sigma}') = \psi(\underline{\sigma}) + x v_-$, with $|x| \leq D_s \lambda^n$. For $\underline{\sigma}, \underline{\sigma}' \in \mathcal{N}_T^{\mathbb{Z}}$ let $\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)}$ be the sequence that agrees with $\underline{\sigma}$ on $(-\infty, n]$ and with $\underline{\sigma}'$ outside such an interval and call $m(\mathrm{d}\underline{\sigma}'_{(n,\infty)} | \underline{\sigma}_{(-\infty,n]})$ the conditional probability measure on $\underline{\sigma}'_{(n,\infty)}$ given $\underline{\sigma}_{(-\infty,n]}$ [29, prop. 5.3.2]. Calling $\mathcal{N}(\underline{\sigma}) = \{\underline{\sigma}' \,|\, \underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)} \in \mathcal{N}_T^{\mathbb{Z}}\}$ we have $m(\mathcal{N}(\underline{\sigma}) | \underline{\sigma}_{(-\infty,n]}) = 1$.

Given $g \in \mathfrak{B}_{\alpha_-,\alpha_+}(\mathbb{T}^2, \mathbb{R})$ write $\hat{g}(\underline{\sigma}) := g(\psi(\underline{\sigma}))$ and observe that $\hat{g}(\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)})$ is almost everywhere well defined with respect to $m(\mathrm{d}\underline{\sigma}'_{(n,\infty)} | \underline{\sigma}_{(-\infty,n]})$. We can thus define

$$\hat{g}^{(\leq n)}(\underline{\sigma}) := \int \hat{g}(\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)}) m(\mathrm{d}\underline{\sigma}'_{(n,\infty)} | \underline{\sigma}_{(-\infty,n]}), \tag{A.1a}$$

$$\hat{g}^{(\geq n)}(\underline{\sigma}) := \int \hat{g}(\underline{\sigma}'_{(-\infty,n)} \vee \underline{\sigma}_{[n,\infty)}) m(\mathrm{d}\underline{\sigma}'_{(-\infty,n)} | \underline{\sigma}_{[n,\infty)}) \tag{A.1b}$$

By construction $\hat{g}^{(\leq n)}(\underline{\sigma})$ is $\mathcal{F}^{(\leq n)}$-measurable and $\langle \hat{g}^{(\geq n)} \rangle = \langle \hat{g} \rangle$, where, with a slight abuse of notation, we use $\langle \cdot \rangle$ also for the average with respect to m. Analogous considerations hold for $\hat{g}^{(\leq n)}(\underline{\sigma})$ as well. Finally, we define

$$\hat{g}^{(n,+)}(\underline{\sigma}) := \hat{g}^{(\leq \lfloor n/\alpha_+ \rfloor)}(\underline{\sigma}) - g^{(\leq \lfloor (n-1)/\alpha_+ \rfloor)}(\underline{\sigma}), \qquad \hat{g}^{(n,-)}(\underline{\sigma}) := \hat{g}^{(\geq \lfloor n/\alpha_- \rfloor)}(\underline{\sigma}) - g^{(\geq \lfloor (n+1)/\alpha_- \rfloor))}(\underline{\sigma}),$$

which allows us to decompose

$$\hat{g}(\underline{\sigma}) = \hat{g}^{(\leq 0)}(\underline{\sigma}) + \sum_{k=1}^{\infty} \hat{g}^{(k,+)}(\underline{\sigma}),$$

with $\hat{g}^{(k,+)}(\underline{\sigma})$ depending only on $\underline{\sigma}_{(-\infty, \lfloor k/\alpha_+ \rfloor)}$.

Lemma A.4. *Let g be (α_-, α_+)-Hölder continuous on $\mathbb{T}^2$. One has*

$$|\hat{g}^{(\leq n)}(\underline{\sigma}) - g(\underline{\sigma})| \leq K_2 \lambda^{-\alpha_+ n} |g|^+_{\alpha_+} \, , \qquad |\hat{g}^{(\geq n)}(\underline{\sigma}) - g(\underline{\sigma})| \leq K_2 \lambda^{\alpha_- n} |g|^-_{\alpha_-} \, , \tag{A.2a}$$

$$|\hat{g}^{(n,+)}(\underline{\sigma})| \leq K_2 \lambda^{-n} |g|^+_{\alpha_+}, \qquad |\hat{g}^{(n,-)}(\underline{\sigma})| \leq K_2 \lambda^{n} |g|^-_{\alpha_-}, \tag{A.2b}$$

for some positive constant K_2.

Proof. If $\underline{\sigma}$ and $\underline{\sigma}'$ are T-compatible sequences and $\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)}$ is T-compatible as well, we have $(\underline{\sigma})_i = (\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)})_i$ for $i \leq n$, so that we get $\psi(\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)}) = \psi(\sigma) + x v_+$ with $|x| \leq D_s \lambda^{-n}$ and hence $|\hat{g}(\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)}) - \hat{g}(\underline{\sigma})| \leq D_s \lambda^{-n\alpha_+} |g|^+_{\alpha_+}$. Integrating over $m(\mathrm{d}\underline{\sigma}'_{(n,\infty)} | \underline{\sigma}_{(-\infty,n]})$ gives the first bound in (A.2a). The second bound is derived in a similar way.

Let now $\underline{\omega}$ be a T-compatible sequence and, for any pair of symbols $\sigma, \sigma' \in \{1, \dots, s\}$, let $\pi = \pi(\sigma, \sigma')$ be a T-compatible sequence of length a such that $T_{\sigma\pi_1} = T_{\pi_a\sigma'} = 1$. For every T-compatible sequence $\underline{\sigma}$, define the T-compatible sequence $\underline{\omega}_n(\underline{\sigma}) := \underline{\sigma}_{(-\infty,n]} \vee \pi(\sigma_n, \omega_{n+a}) \vee \underline{\omega}_{[n+a,\infty)}$. Reasoning as above and using that $\lfloor (n-1)/\alpha_+ \rfloor \alpha_+ \geq n - 1 - \alpha_+$, we get

$$|\hat{g}(\underline{\sigma}_{(-\infty,\lfloor (n-1)/\alpha_+ \rfloor]} \vee \underline{\sigma}'_{(\lfloor (n-1)/\alpha_+ \rfloor,\infty)}) - \hat{g}(\underline{\omega}_{\lfloor (n-1)/\alpha_+ \rfloor}(\underline{\sigma}))| \leq D_s \lambda^{\alpha_+} \lambda^{-(n-1)} |g|^+_{\alpha_+} \tag{A.3}$$

and, analogously,

$$|\hat{g}(\underline{\sigma}_{(-\infty,\lfloor n/\alpha_+ \rfloor]} \vee \underline{\sigma}'_{(\lfloor n/\alpha_+ \rfloor],\infty)}) - \hat{g}(\underline{\omega}_{\lfloor (n-1)/\alpha_+ \rfloor}(\underline{\sigma}))| \leq D_s \lambda^{\alpha_+} \lambda^{-(n-1)} |g|^+_{\alpha_+}. \tag{A.4}$$

The first bound in (A.2b) follows by integrating (A.3) over $m(\mathrm{d}\underline{\sigma}'_{(\lfloor (n-1)/\alpha_+ \rfloor,\infty)} | \underline{\sigma}_{(-\infty,\lfloor (n-1)/\alpha_+ \rfloor]})$ and (A.4) over $m(\mathrm{d}\underline{\sigma}'_{(\lfloor n/\alpha_+ \rfloor,\infty)} | \underline{\sigma}_{(-\infty,\lfloor n/\alpha_+ \rfloor]})$, so as to obtain

$$\begin{aligned}
|\hat{g}^{(\leq \lfloor (n-1)/\alpha_+ \rfloor)}(\underline{\sigma}) - \hat{g}(\underline{\omega}_{\lfloor (n-1)/\alpha_+ \rfloor}(\underline{\sigma}))| &\leq D_s \lambda^{\alpha_+} \lambda^{-(n-1)} |g|^+_{\alpha_+}, \\
|\hat{g}^{(\leq \lfloor n/\alpha_+ \rfloor)}(\underline{\sigma}) - \hat{g}(\underline{\omega}_{\lfloor (n-1)/\alpha_+ \rfloor}(\underline{\sigma}))| &\leq D_s \lambda^{\alpha_+} \lambda^{-(n-1)} |g|^+_{\alpha_+},
\end{aligned}$$

which gives with $K_2 = 2D_s \lambda^{1+\alpha_+}$. In a similar way we obtain the second bound. □

A.2 A correlation inequality: proof of Proposition 2.4

Since $\langle g_+ \, g_- \circ A_0^n \rangle - \langle g \rangle_+ \langle g \rangle_- = \langle \tilde{g}_+ \, \tilde{g}_- \circ A_0^n \rangle$, with the notation in (2.1), we can assume that $\langle g_+ \rangle = \langle g_- \rangle = 0$. Then, calling $\tilde{n} := \lfloor \alpha n \rfloor$, we write, with obvious meaning of the symbols,

$$\begin{aligned}
\hat{g}_+(\underline{\sigma}) &= \hat{g}_+^{(\leq 0)}(\underline{\sigma}) + \sum_{k=1}^{\tilde{n}} \hat{g}_+^{(k,+)}(\underline{\sigma}) + (\hat{g}_+(\underline{\sigma}) - \hat{g}_+^{(\leq \lfloor \tilde{n}/\alpha \rfloor)}(\underline{\sigma})) =: \sum_{k=0}^{\tilde{n}+1} \check{g}_+^{(k)}(\underline{\sigma}), \\
\hat{g}_-(\underline{\sigma}) &= \hat{g}_-^{(\geq 0)}(\underline{\sigma}) + \sum_{k=1}^{\tilde{n}} \hat{g}_-^{(-k,-)}(\underline{\sigma}) + (\hat{g}_-(\underline{\sigma}) - \hat{g}_-^{(\geq - \lfloor \tilde{n}/\alpha \rfloor}(\underline{\sigma})) =: \sum_{k=0}^{\tilde{n}+1} \check{g}_-^{(k)}(\underline{\sigma}),
\end{aligned}$$

with $\|\check{g}_+^{(k)}\|_\infty \leq K_2 \lambda^{-k} \|g_+\|^+_\alpha$ and $\|\check{g}_-^{(k)}\|_\infty \leq K_2 \lambda^{-k} \|g_-\|^-_\alpha$ by Lemma A.4. Since $(\tau^n \underline{\sigma})_{(-k,\infty)} = \underline{\sigma}_{(n-k,\infty)}$, so that $\check{g}_-^{(k)}(\tau^n \underline{\sigma})$ depends only on $\underline{\sigma}_{(n-k,\infty)}$, we have

$$\begin{aligned}
|\langle \check{g}_+^{(k)} \, \check{g}_-^{(k')} \circ \tau^n \rangle| &\leq K_1 K_2 \|g_+\|^+_\alpha \|g_-\|^-_\alpha \lambda^{-(k+k')} \lambda_0^{-(n - \lfloor k/\alpha \rfloor - \lfloor k'/\alpha \rfloor)}, && \text{if } k + k' < \tilde{n}, \\
|\langle \check{g}_+^{(k)} \, \check{g}_-^{(k')} \circ \tau^n \rangle| &\leq K_2 \|g_+\|^+_\alpha \|g_-\|^-_\alpha \lambda^{-(k+k)}, && \text{if } k + k' \geq \tilde{n},
\end{aligned}$$

by Lemmas A.1 and A.4. Summing over k and k', using that $(n - \lfloor k/\alpha \rfloor - \lfloor k'/\alpha \rfloor) \geq (\tilde{n} - k - k')/\alpha - 3$, so that $(k + k') + (n - \lfloor k/\alpha \rfloor - \lfloor k'/\alpha \rfloor) \geq \tilde{n}/\alpha - (k + k')(1 - 1/\alpha) - 3 \geq \tilde{n} + (1 - \alpha)\alpha^{-1}(\tilde{n} - k - k') - 3$, and relying on Remark A.2, we obtain

$$\begin{aligned}
|\langle g_+ \, g_- \circ A_0^n \rangle| &\leq K_2 \lambda^{-\tilde{n}} \|g_+\|^+_\alpha \|g_-\|^-_\alpha \left(\sum_{\substack{k,k'=0,\dots,\tilde{n}+1 \\ k+k' < \tilde{n}}} K_1 \lambda^3 \lambda^{-(\tilde{n}-k-k')\frac{1-\alpha}{\alpha}} + \sum_{\substack{k,k'=0,\dots,\tilde{n}+1 \\ k+k' \geq \tilde{n}}} \lambda^{\tilde{n}-k-k'} \right) \\
&\leq \max\{1, K_1\} K_2 \lambda^3 \|g_+\|^+_\alpha \|g_-\|^-_\alpha \lambda^{-\tilde{n}} \left(\tilde{n} \sum_{q=0}^{\tilde{n}} \lambda^{-q\frac{1-\alpha}{\alpha}} + \sum_{q=0}^{\tilde{n}+2} (\tilde{n} + q + 1) \lambda^{-q} \right),
\end{aligned}$$

from which the estimate in Proposition 2.4 follows observing that $\lambda^{-\tilde{n}} \leq \lambda^\alpha \lambda^{-\alpha n} \leq \lambda \, \lambda^{-\alpha n}$.

B Proof of some technical results

We use $\mathscr{S}$, $\mathscr{S}_1$ and $\mathscr{S}_2$ as a shortened notation for the corresponding extended maps. In the light of Remark 4.19, we also assume $\rho < \rho_0$.

B.1 Derivative of the auxiliary map: proof of Proposition 4.29

In order to bound $\langle \partial_\theta\,(p_2^{(n)}\,q_2\circ\mathscr{S}_2^n) - \partial_\theta\,(\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n)\rangle$, we split

$$\partial_\theta\,(p_2^{(n)}\,q_2\circ\mathscr{S}_2^n) = \partial_\theta p_2^{(n)}\,q_2\circ\mathscr{S}_2^n + p_2^{(n)}\,\partial_\theta(q_2\circ\mathscr{S}_2^n), \tag{B.1a}$$

$$\partial_\theta\,(\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n) = \partial_\theta \overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n + \overline{p}^{(n)}\,\partial_\theta(\overline{q}\circ\overline{G}^n), \tag{B.1b}$$

with both $\partial_\theta p_2^{(n)}$ and $\partial_\theta\overline{p}^{(n)}$ written according to (2.18a), and we bound separately the two contributions $\langle\partial_\theta p_2^{(n)}\,q_2\circ\mathscr{S}_2^n\rangle - \partial_\theta\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n$ and $\langle p_2^{(n)}\,\partial_\theta(q_2\circ\mathscr{S}_2^n)\rangle - \overline{p}^{(n)}\,\partial_\theta(\overline{q}\circ\overline{G}^n)$. Since $\partial_\theta(\mathscr{S}_2)^n_\theta = (\partial_\theta G_2)^{(n)}$ and $\partial_\theta\overline{G}^n = (\partial_\theta\overline{G})^{(n)}$, we bound the first contribution as

$$\begin{aligned}
&\big|\langle\partial_\theta p_2^{(n)}\,q_2\circ\mathscr{S}_2^n\rangle - \partial_\theta\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n\big| \\
&\leq \sum_{k=0}^{n-1}\Big|\langle(p_2\,\partial_\theta G_2)^{(k)}\big(\partial_\theta p_2\,P_{n-1-k}Q_{n-k}\big)\circ\mathscr{S}_2^k\rangle - \langle p_2\,\partial_\theta G_2\rangle^{(k)}\langle\partial_\theta p_2\,P_{n-1-k}Q_{n-k}\rangle\circ\overline{G}^k\Big| \\
&\quad+\sum_{k=0}^{n-1}\big|\langle p_2\,\partial_\theta G_2\rangle^{(k)}\big|\big|\big(\langle\partial_\theta p_2\,P_{n-1-k}Q_{n-k}\rangle\circ\overline{G}^k - \langle\partial_\theta p_2\rangle\langle P_{n-1-k}\rangle\langle q_2\rangle\circ\overline{G}^{n-k}\big)\circ\overline{G}^k\big| \\
&\quad+\sum_{k=0}^{n-1}\big|\langle p_2\partial_\theta G_2\rangle^{(k)}\big(\langle\partial_\theta p_2\rangle\langle P_{n-1-k}\rangle\langle q_2\rangle\circ\overline{G}^{n-k}\big)\circ\overline{G}^k - (\overline{p}\,\partial_\theta\overline{G})^{(k)}\big(\partial_\theta\overline{p}\,\overline{P}_{n-1-k}\,\overline{Q}_{n-k}\big)\circ\overline{G}^k\big|,
\end{aligned} \tag{B.2}$$

where we have set $P_n := p_2^{(n)}\circ\mathscr{S}_2$, $Q_n := q_2\circ\mathscr{S}_2^n$, $\overline{P}_n := \overline{p}^{(n)}\circ\overline{G}$ and $\overline{Q}_n := \overline{q}\circ\overline{G}^n$. We use Proposition 4.26, with k instead of n, $g_+ = 1$, $\mathfrak{p}_i = p_2\partial_\theta G_2$ for $i = 1,\dots,k$ and $g_- = \partial_\theta p_2\,p_2^{(n-1-k)}\circ\mathscr{S}_2\,q_2\circ\mathscr{S}^{n-k}$, to bound the second line of (B.2) by

$$\sum_{k=1}^{n} C(1-\rho\,\gamma')^k\big(\rho + k\rho^2\big)\|\!|\partial_\theta p_2\,p_2^{(n-1-k)}\circ\mathscr{S}_2\,q_2\circ\mathscr{S}_2^{n-k}\|\!|^-_{\alpha_0,3}, \tag{B.3}$$

and with $n-1-k$ instead of n, $g_+ = \partial_\theta p_2$, $\mathfrak{p}_i = p_2\circ\mathscr{S}_2$ for $i = 1,\dots,n-k-1$ and $g_- = q_2$, to bound the third line of (B.2) by

$$\sum_{k=0}^{n-1} C(1-\rho\,\gamma')^k\Big((n-1-k)\lambda^{-\alpha_0(n-1-k)} + \big(\rho + (n-1-k)\rho^2\big)\|\!|\partial_\theta p_2\|\!|^+_{\alpha_0}\|\!|q_2\|\!|^-_{\alpha_0,3}\Big). \tag{B.4}$$

After estimating

$$\begin{aligned}
&\big|\langle(p_2\,\partial_\theta G_2)^{(k)}\rangle\big(\langle\partial_\theta p_2\rangle\langle P_{n-1-k}\rangle\langle q_2\rangle\circ\overline{G}^{n-k}\big)\circ\overline{G}^k - (\overline{p}\,\partial_\theta\overline{G})^{(k)}\big(\partial_\theta\overline{p}\;\overline{P}_{n-1-k}\,\overline{Q}_{n-k}\big)\circ\overline{G}^k\big| \\
&\quad\leq \big|\langle p_2\,\partial_\theta G_2\rangle^{(k)} - (\overline{p}\,\partial_\theta\overline{G})^{(k)}\big|\big|\big(\langle\partial_\theta p_2\rangle\langle P_{n-1-k}\rangle\langle q_2\rangle\circ\overline{G}^{n-k}\big)\circ\overline{G}^k\big| \\
&\qquad+ \big|(\overline{p}\,\partial_\theta\overline{G})^{(k)}\big|\big|\big(\langle\partial_\theta p_2\rangle - \partial_\theta\overline{p}\big)\circ\overline{G}^k\big|\big|\big(\langle P_{n-1-k}\rangle\langle q_2\rangle\circ\overline{G}^{n-k}\big)\circ\overline{G}^k\big| \\
&\qquad+ \big|(\overline{p}\,\partial_\theta\overline{G})^{(k)}\,\partial_\theta\overline{p}\circ\overline{G}^k\big|\big|\big(\langle P_{n-1-k}\rangle - \overline{P}_{n-1-k}\big)\circ\overline{G}^k\big|\big|\langle q_2\rangle\circ\overline{G}^n\big| \\
&\qquad+ \big|(\overline{p}\,\partial_\theta\overline{G})^{(k)}\,\big(\partial_\theta\overline{p}\;\overline{P}_{n-1-k}\big)\circ\overline{G}^k\big|\big|\langle q_2\rangle\circ\overline{G}^n - \overline{Q}_n\big|,
\end{aligned}$$

we rely once more Proposition 4.26 to bound the fourth line of (B.2) with

$$\begin{aligned}
\sum_{k=0}^{n-1} C(1-\rho\gamma')^k\Big(&(\rho + k\rho^2)\|\partial_\theta p_2\|_\infty\|p_2^{(n-1-k)}\|_\infty\|q_2\|_\infty + \rho\,\|\langle\partial_\theta p_2\rangle - \partial_\theta\overline{p}\|_\infty\|p_2^{(n-1-k)}\|_\infty\|q_2\|_\infty \\
&+ (\rho + (n-k-1)\,\rho^2)\,(1-\rho\,\gamma')^{n-k-1}\|q_2\|_\infty + \rho\,(1-\rho\,\gamma')^{n-1-k}\|\langle q_2\rangle - \overline{q}\|_\infty\Big).
\end{aligned} \tag{B.5}$$

Therefore, thanks to (4.50) and (4.53), which yield

$$\|\!|\partial_\theta p_2\|\!|_{\alpha_0,2} \leq C\rho, \qquad \|\!|\langle p_2\rangle - \overline{p}\|\!|_{0,1} \leq C\rho^2, \qquad \|q_2\|_\infty \leq C\rho, \qquad \|\langle q_2\rangle - \overline{q}\|_\infty \leq C\rho^2,$$

by collecting together the bounds (B.3)–(B.5) and using (2.8), we obtain

$$\big|\langle \partial_\theta p_2^{(n)}\, q_2\circ\mathscr{S}_2^n\rangle - \partial_\theta \overline{p}^{(n)}\, \overline{q}\circ\overline{G}^n\big| \le \sum_{k=0}^{n-1} C(1-\rho\,\gamma')^n\rho^3\left(1+k\rho+k^2\rho^2\right). \tag{B.6}$$

Next, we consider the second contribution, that we rewrite as

$$\big(\langle p_2^{(n)}\,\partial_\theta(q_2\circ\mathscr{S}_2^n)\rangle - \langle p_2\rangle^{(n)}\langle\partial_\theta(q_2\circ\overline{G}^n)\rangle\big) + \big(\langle p_2\rangle^{(n)}\langle\partial_\theta(q_2\circ\overline{G}^n\rangle - \overline{p}^{(n)}\,\partial_\theta(\overline{q}\circ\overline{G}^n)\big),$$

and, using that $\partial_\theta(q_2\circ\mathscr{S}_2^n) = (\partial_\theta(\mathscr{S}_2)_\theta)^{(n)}(\partial_\theta q_2)\circ\mathscr{S}_2^n$, $\partial_\theta(q_2\circ\overline{G}^n) = (\partial_\theta\overline{G})^{(n)}\partial_\theta q_2\circ\overline{G}^n$ and $\partial_\theta(\overline{q}\circ\overline{G}^n) = (\partial_\theta\overline{G})^{(n)}(\partial_\theta\overline{q})\circ\overline{G}^n$, we apply first Proposition 4.26, with $\mathfrak{p}_i = p_2\,\partial_\theta\mathscr{S}_2$, $g_+ = 1$ and $g_- = q_2$ to bound

$$\big|\langle(p_2\,\partial_\theta(\mathscr{S}_2)_\theta)^{(n)}(\partial_\theta q_2)\circ\mathscr{S}_2^n\rangle - \langle p_2\,\partial_\theta(\mathscr{S}_2)_\theta\rangle^{(n)}\langle\partial_\theta q_2\rangle\circ\overline{G}^n\big| \le C(1-\rho\,\gamma')^n\big(\rho+\rho^2 n\big)|\!|\!|\partial_\theta q_2|\!|\!|_{\alpha_0,3},$$

then we reason as when dealing with (4.48) to get

$$\begin{aligned}
&\Big|\langle p_2\partial_\theta(\mathscr{S}_2)_\theta\rangle^{(n)}\langle\partial_\theta q_2\rangle\circ\overline{G}^n - (\overline{p}\,\partial_\theta\overline{G})^{(n)}\,(\partial_\theta\overline{q})\circ\overline{G}^n\Big| \\
&\quad\le \langle\partial_\theta q_2\rangle\circ\overline{G}^n\sum_{i=0}^{n-1}\langle p_2\,\partial_\theta(\mathscr{S}_2)_\theta\rangle^{(i)}\left(\langle p_2\partial_\theta(\mathscr{S}_2)_\theta\rangle - \overline{p}\,\partial_\theta\overline{G}\right)\circ\overline{G}^i\,(\overline{p}\,\partial_\theta\overline{G})^{(n-i-2)}\circ\overline{G}^{i+1} \\
&\qquad + (\overline{p}\,\partial_\theta\overline{G})^{(n)}\left(\langle\partial_\theta q_2\rangle\circ\overline{G}^n - \partial_\theta\overline{q}\circ\overline{G}^n\right) \le C(1-\rho\,\gamma')^n n\,\|p_2\partial_\theta(\mathscr{S}_2)_\theta - \overline{p}\,\partial_\theta\overline{G}\|_\infty\,\|\partial q_2\|_\infty \\
&\qquad + C(1-\rho\,\gamma')^n\,\|\langle\partial_\theta q_2\rangle - \partial_\theta\overline{q}\|_\infty \le C(1-\rho\,\gamma')^n n\rho^2,
\end{aligned}$$

so that eventually we obtain

$$\big|\langle p_2^{(n)}\,\partial_\theta(q_2\circ\mathscr{S}_2^n) - \overline{p}^{(n)}\,\partial_\theta(\overline{q}\circ\overline{G}^n)\rangle\big| \le C\rho. \tag{B.7}$$

Thus, summing together (B.6) and (B.7), the first bound in (4.55) follows.

The second bound in (4.55) is obtained in a similar way by reasoning as in the second part of the proof of Proposition 4.28 and using (B.6) and (B.7) instead of (4.52).

B.2 Comparing the translated and auxiliary maps

B.2.1 Proof of Lemma 4.32

Following the strategy outlined in Remark 4.31, we rearrange the sums containing terms $W\circ A_0^{i+1} - W\circ A_0^i$ in such a way that the new summands contain differences of more regular functions (see Lemmas B.1 and B.3 below). We define, for $k \ge 0$,

$$D_{\mathfrak{p},1,k}(\theta,\psi) := \int_0^1 \mathrm{d}t\,\partial_\theta(\mathfrak{p}\circ\mathscr{S}_1^k)\,(G_2(\theta,\psi)+t\,\rho\,\zeta(\theta,\psi),A_0\psi), \tag{B.8a}$$

$$D_{\mathfrak{p},2,k}(\theta,\psi) := \int_0^1 \mathrm{d}t\,(1-t)\,\partial_\theta^2(\mathfrak{p}\circ\mathscr{S}_1^k)\,(G_2(\theta,\psi)+t\,\rho\,\zeta(\theta,\psi),A_0\psi). \tag{B.8b}$$

$$R_{\mathfrak{p},1,k}(\theta,\psi) := \sum_{i=0}^{k}\left(\partial_\theta\big((D_{\mathfrak{p},1,k-i}\,\zeta)\circ\mathscr{S}_2^i\big)\right)(\theta,\psi), \tag{B.8c}$$

$$R_{\mathfrak{p},2,k}(\theta,\psi) := \sum_{i=0}^{k}\left(\left(D_{\mathfrak{p},2,k-i}\,\zeta^2\right)\circ\mathscr{S}_2^i\right)(\theta,\psi), \tag{B.8d}$$

and expand, for $n \ge 1$,

$$\mathfrak{p}\circ\mathscr{S}_1^n - \mathfrak{p}\circ\mathscr{S}_2^n = \rho\sum_{i=0}^{n-1}\big(D_{\mathfrak{p},1,n-1-i}\,\zeta\big)\circ\mathscr{S}_2^i = \rho\sum_{i=0}^{n-1}\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_1^{n-1-i})\circ\mathscr{S}_2\,\zeta\big)\circ\mathscr{S}_2^i + \rho^2 R_{\mathfrak{p},2,n-1}. \tag{B.9}$$

Using the first expansion in (B.9), with $n-i-1$ instead of n, to write

$$\partial_\theta(\mathfrak{p}\circ\mathscr{S}_1^{n-1-i}) = \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1-i}) + \rho\sum_{i=0}^{n-1-i}\partial_\theta\big(\big(D_{\mathfrak{p},1,n-1-i-j}\,\zeta\big)\circ\mathscr{S}_2^j\big)$$

in the second expansion, we obtain

$$\mathfrak{p}\circ\mathscr{S}_1^n - \mathfrak{p}\circ\mathscr{S}_2^n = \rho\sum_{i=0}^{n-1}\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1-i})\circ\mathscr{S}_2\,\zeta\big)\circ\mathscr{S}_2^i + \rho^2\sum_{i=0}^{n-1}\left(R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta\right)\circ\mathscr{S}_2^i + \rho^2 R_{\mathfrak{p},2,n-1}, \tag{B.10}$$

where, by Lemma 3.6 and Remark 3.11, for all $n\ge 1$ we bound

$$\|R_{\mathfrak{p},2,n-1}\|_\infty \le C, \qquad \langle|R_{\mathfrak{p},2,n-1}|\rangle \le C. \tag{B.11}$$

We start with two abstract results, easy to check, in the spirit of some lemmas in Subsection 4.7.4.

Lemma B.1. *Let $\{E_k\}_{k=0}^{n-1}$ be a set of functions $E_k\colon\Omega\to\Omega$, with $n\ge 1$. Define, for $m=1,\dots,n$,*

$$\Delta E_{m,0} = E_0\circ\mathscr{S}_2^{-1}, \quad \Delta E_{m,k} = E_k\circ\mathscr{S}_2^{-1} - E_{k-1},\ k=1,\dots,m-1, \quad \Delta E_{m,m} = -E_{m-1}, \tag{B.12}$$

where $\Delta E_{m,k}$ with $0<k<m$ are meaningful only if $m\ge 2$. For any m as above one has

$$\sum_{i=0}^{m-1} E_{m-1-i}\circ\mathscr{S}_2^i\,(W\circ A_0^{i+1} - W\circ A_0^i) = \sum_{i=0}^{m}\Delta E_{m,m-i}\circ\mathscr{S}_2^i\,W\circ A_0^i. \tag{B.13}$$

Remark B.2. If the functions E_k in Lemma B.1 are such that

$$|||E_k|||^-_{\alpha_0,3} \le (1-\rho\,\gamma')^k\,\Gamma_0, \quad |||E_k\circ\mathscr{S}_2^{-1} - E_{k-1}|||^-_{\alpha_0,3} \le \rho\,(1-\rho\,\gamma')^k\,\Gamma_0, \qquad k=1,\dots,n-1, \tag{B.14}$$

for some constant Γ_0, then one has $\sum_{i=0}^n |||\Delta E_{n,n-i}|||^-_{\alpha_0,3} \le C\,\Gamma_0$. Analogously, if

$$\|E_k\|_\infty \le (1-\rho\,\gamma')^k\,\Gamma_0, \quad \|E_k\circ\mathscr{S}_2^{-1} - E_{k-1}\|_\infty \le \rho\,(1-\rho\,\gamma')^k\,\Gamma_0, \qquad k=1,\dots,n-1, \tag{B.15}$$

one has $\sum_{i=0}^n \|\Delta E_{n,n-i}\|_\infty \le C\,\Gamma_0$.

Lemma B.3. *Let $\{E_k\}_{k=0}^{n-1}$ and, for $1\le m\le n$, $\{\Delta E_{m,k}\}_{k=1}^m$ be as in Lemma B.1, and let $\{F_k\}_{k=0}^n$ be a set of functions $F_k\colon\Omega\to\Omega$. For any m as above one has*

$$\sum_{i=0}^{m-1} E_{m-1-i}\circ\mathscr{S}_2^i\,F_i(W\circ A_0^{i+1} - W\circ A_0^i) = \sum_{i=0}^{m}\Delta E_{m,m-i}\circ\mathscr{S}_2^i\,F_i W\circ A_0^i + \sum_{i=1}^{m} E_{m-i}\circ\mathscr{S}_2^{i-1}\big(F_{i-1}-F_i\big)W\circ A_0^i.$$

Next, for $r=0,1,2$, we introduce the functions

$$E_{\mathfrak{p},r,k} := \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^k)\circ\mathscr{S}_2\,c_r, \qquad Q_{\mathfrak{p},r,k} := \partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big), \qquad Q_{r,k} := \partial_\theta(\mathscr{S}_2^k)_\theta\circ\mathscr{S}_2\,c_r, \tag{B.16a}$$

and define $\Delta E_{\mathfrak{p},0,m,k}$, $\Delta Q_{\mathfrak{p},r,m,k}$ and $\Delta Q_{r,m,k}$ as in (B.12) with $E_k = E_{\mathfrak{p},0,k}$, $E_k = Q_{\mathfrak{p},r,k}$ and $E_k = Q_{r,k}$, respectively. Set also $\Delta_2 Q_{0,m} := \Delta Q_{0,m} - \Delta Q_{0,m+1}$ and $\Delta_2 Q_{0,m,k} := \Delta Q_{0,m,k} - \Delta Q_{0,m+1,k+1}$.

Using that $\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^k) = \partial_\theta\mathfrak{p}\circ\mathscr{S}_2^k\partial_\theta(\mathscr{S}_2^k)_\theta$ and the bounds (4.34) in Lemma 4.20, we may bound

$$\begin{aligned}
&|||E_{\mathfrak{p},0,k}\circ\mathscr{S}_2^{-1} - E_{\mathfrak{p},0,k-1}|||^-_{\alpha_0,3} \le |||\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^k)|||^-_{\alpha_0,3}|||c_0\circ\mathscr{S}_2^{-1} - c_0|||_{\alpha_0,3}\\
&\quad + |||\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^k)\circ\mathscr{S}_2^{-1} - \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{k-1})|||^-_{\alpha_0,3}|||c_0|||_{\alpha_0,3} \le |||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}|||\partial_\theta(\mathscr{S}_2^k)_\theta|||^-_{\alpha_0,3}|||c_0\circ\mathscr{S}_2^{-1} - c_0|||_{\alpha_0,3}\\
&\quad + |||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}|||\partial_\theta(\mathscr{S}_2^{k-1})_\theta|||^-_{\alpha_0,3}|||\partial_\theta(\mathscr{S}_2)_\theta - \mathbb{1}|||^-_{\alpha_0,3}|||c_0|||_{\alpha_0,3} \le C\rho\,(1-\rho\,\gamma')^k\,|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3},\\
&\|\partial_\theta(\mathscr{S}_2^k)_\theta\,c_r\circ\mathscr{S}_2^{-1} - \partial_\theta(\mathscr{S}_2^k)_\theta\circ\mathscr{S}_2\,c_r\|_\infty\\
&\quad \le \|\partial_\theta(\mathscr{S}_2^k)_\theta\|_\infty\|c_r\circ(\mathscr{S}_2^{-1})_\theta - c_r\|_\infty + \|\partial_\theta^2(\mathscr{S}_2^k)_\theta\|_\infty\|(\mathscr{S}_2)_\theta - \mathbb{1}\|_\infty\|\,c_r\|_\infty \le C\rho(1-\rho\,\gamma')^k,\\
&\|\partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big)\circ\mathscr{S}_2^{-1} - \partial_\theta\big(D_{\mathfrak{p},1,k-1}\,c_r\big)\|_\infty\\
&\quad \le \|\partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big)\|_\infty\|\circ\mathscr{S}_2^{-1} - \mathbb{1}\|_\infty + \|\partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big) - \partial_\theta\big(D_{\mathfrak{p},1,k-1}\,c_r\big)\|_\infty \le C\rho\,(1-\rho\,\gamma')^{k+j}|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3},
\end{aligned}$$

so that in (B.16) the functions $E_{\mathfrak{p},0,k}$ satisfy the bounds (B.14), with $\Gamma_0 = C|||\partial_\theta \mathfrak{p}|||_{\alpha_0,3}$, while the functions $Q_{r,k}$ and $Q_{\mathfrak{p},r,k}$, for $r = 0,1,2$, satisfy the bounds (B.15), with $\Gamma_0 = C$ and $\Gamma_0 = C|||\partial_\theta \mathfrak{p}|||_{\alpha_0,3}$, respectively. Thus, for $r = 0,1$, $m \geq 2$ and $k = 1,\dots,m-1$, we obtain

$$\sum_{i=0}^{n} |||\Delta E_{\mathfrak{p},0,n,n-i}|||^{-}_{\alpha_0,3} \leq C\, |||\partial_\theta \mathfrak{p}|||_{\alpha_0,3}, \tag{B.17a}$$

$$\|\Delta Q_{r,m,k}\|_\infty \leq C\rho(1-\rho\,\gamma')^k \qquad \|\Delta Q_{\mathfrak{p},r,m,k}\|_\infty \leq C\rho(1-\rho\,\gamma')^k\, |||\partial_\theta \mathfrak{p}|||_{\alpha_0,3}. \tag{B.17b}$$

Let us come back to (B.10). Recalling (4.29a) and (4.29e), we write $\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1-i})\circ\mathscr{S}_2\,\zeta\big)\circ\mathscr{S}_2^i$ as

$$\rho^{-1}E_{\mathfrak{p},0,n-1-i}\circ\mathscr{S}_2^i\,(W\circ A_0^{i+1} - W\circ A_0^i) + \sum_{r=1}^{2} E_{\mathfrak{p},r,n-1-i}\circ\mathscr{S}_2^i\,\left(W\circ A_0^i\right)^r, \tag{B.18}$$

which, thanks to Lemma B.1, allows to write the first sum in (B.10) as

$$\sum_{i=0}^{n} \Delta E_{\mathfrak{p},0,n,n-i}\circ\mathscr{S}_2^i\, W\circ A_0^i + \rho\sum_{i=1}^{n-1}\sum_{r=1}^{2} E_{\mathfrak{p},r,n-1-i}\circ\mathscr{S}_2^i\,\left(W\circ A_0^i\right)^r. \tag{B.19}$$

The second sum on the r.h.s. of (B.10), in terms of the functions in (B.16a), becomes

$$\rho^2\sum_{i=0}^{n-1}(R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta)\circ\mathscr{S}_2^i = \rho^2\sum_{i=0}^{n-1}\sum_{j=0}^{i}\big(\partial_\theta\big((D_{\mathfrak{p},1,n-1-i}\,\zeta)\circ\mathscr{S}_2^{i-j}\big)\circ\mathscr{S}_2\,\zeta\big)\circ\mathscr{S}_2^j = \rho^2\mathscr{A}_{n,2} + \rho\mathscr{A}_{n,1} + \mathscr{A}_{n,0},$$

with

$$\begin{aligned}
\mathscr{A}_2 &:= \sum_{i=0}^{n-1}\sum_{j=0}^{i}\sum_{r,s=1}^{2} Q_{\mathfrak{p},r,n-1-i}\circ\mathscr{S}_2^{i+1}Q_{s,i-j}\circ\mathscr{S}_2^j\,\left(W\circ A_0^{i+1}\right)^r\left(W\circ A_0^j\right)^s \\
\mathscr{A}_1 &:= \sum_{i=0}^{n-1}\sum_{j=0}^{i+1}\sum_{r=1}^{2} Q_{\mathfrak{p},r,n-1-i}\circ\mathscr{S}_2^{i+1}\Delta Q_{0,i+1,i+1-j}\circ\mathscr{S}_2^j\,\left(W\circ A_0^{i+1}\right)^r W\circ A_0^j \\
&\quad + \sum_{j=0}^{n-1}\sum_{i=0}^{n-j}\sum_{s=1}^{2}\big(\big(\Delta Q_{\mathfrak{p},0,n-j,n-j-i}\circ\mathscr{S}_2^{i+1}Q_{s,i} + Q_{\mathfrak{p},0,n-j-i}\circ\mathscr{S}_2^i\Delta_2 Q_{s,i-1}\big)W\circ A_0^{i+1}W^s\big)\circ\mathscr{S}_2^j \\
\mathscr{A}_0 &:= \sum_{j=0}^{n-1}\sum_{i=0}^{n-j}\big(\big(\Delta Q_{\mathfrak{p},0,n-j,n-j-i}\circ\mathscr{S}_2^{i+1}\Delta Q_{0,i+1,i+1} + Q_{\mathfrak{p},0,n-j-i}\circ\mathscr{S}_2^i\Delta_2 Q_{0,i,i}\big)W\circ A_0^{i+1}W\big)\circ\mathscr{S}_2^j \\
&\quad + \sum_{i=0}^{n-1} Q_{\mathfrak{p},0,n-1-i}\circ\mathscr{S}_2^{i+1}\Delta Q_{0,i+1,0}\circ\mathscr{S}_2^{i+1}\big(W\circ A_0^{i+2} - W\circ A_0^{i+1}\big)W\circ A_0^{i+1},
\end{aligned}$$

where the terms with negative integer labels are meant as 0 and we have used Lemma B.1 and Lemma B.3 to obtain the first and second line of $\mathscr{A}_1$, respectively, and both Lemmas B.1 and B.3 to obtain $\mathscr{A}_0$.

Remark B.4. To bound the contributions to $\mathscr{A}_0$ with $\Delta_2 Q_{0,i,i}$ we use that, for $m > k \geq 0$, we have $\|\Delta_2 Q_{0,m,k}\|_\infty = \|\partial_\theta(\mathscr{S}_2^{k-1})_\theta\,(\mathbb{1}-\partial_\theta(\mathscr{S}_2)_\theta)\,\big((\partial_\theta(\mathscr{S}_2)_\theta-\mathbb{1})\,c_0\circ\mathscr{S}_2^{-1} + \big(c_0\circ\mathscr{S}_2^{-1}-c_0\big)\big)\,\|_\infty \leq C\rho^2(1-\rho\,\gamma')^k$.

The bounds (4.34), the second bound in (2.33), the bounds (B.17b) and Remark B.4 yield that

$$\|\rho^2\mathscr{A}_2 + \rho\mathscr{A}_1 + \mathscr{A}_0\| \leq C\,\rho^{-2}|||\partial_\theta \mathfrak{p}|||_{\alpha_0,3}, \qquad \langle|\rho^2\mathscr{A}_2 + \rho\mathscr{A}_1 + \mathscr{A}_0|\rangle \leq C\rho^{-1}|||\partial_\theta \mathfrak{p}|||_{\alpha_0,3}. \tag{B.20}$$

In conclusion, if we set $C_{\mathfrak{p},n,0} = \Delta E_{\mathfrak{p},1,n,0}$ and $C_{\mathfrak{p},n,k} = E_{\mathfrak{p},1,k-1} + \rho^{-1}\Delta E_{\mathfrak{p},0,n,k}$ for $k = 1,\dots,n$, i.e.

$$\begin{aligned}
C_{\mathfrak{p},n,0} &:= \rho^{-1}\partial_\theta\mathfrak{p}\,c_0\circ\mathscr{S}_2^{-1}, \\
C_{\mathfrak{p},n,k} &:= \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{k-1})\circ\mathscr{S}_2\,c_1 + \rho^{-1}\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{k})\,c_0\circ\mathscr{S}_2^{-1} - \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{k-1})\circ\mathscr{S}_2\,c_0\big), \quad k = 1,\dots,n-1, \\
C_{\mathfrak{p},n,n} &:= \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1})\circ\mathscr{S}_2\,\mathfrak{c}_1 - \rho^{-1}\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1})\,c_0\circ\mathscr{S}_2^{-1},
\end{aligned}$$

and define

$$R_{\mathfrak{p},n} := \sum_{i=0}^{n-1} \Big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1-i})\circ\mathscr{S}_2\, c_2\Big)\circ\mathscr{S}_2^i\, W^2\circ A_0^i + \rho \sum_{i=0}^{n-1} \left(R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta\right)\circ\mathscr{S}_2^i + \rho\, R_{\mathfrak{p},2,n-1}, \tag{B.21}$$

we obtain (4.60), with the bounds (4.61a) and (4.61b) following from the estimates (4.34) and (B.17a), while the bound (4.61c) follows from the bounds (B.11) and (B.20).

B.2.2 Proof of Proposition 4.45

We have which, together with (B.1a), allows us to write the first contribution in (4.33b) as

$$\begin{aligned}\partial_\theta h_1 - \partial_\theta h_2 = \sum_{n=0}^{\infty} \Big(&(\partial_\theta p_1^{(n)} - \partial_\theta p_2^{(n)})\,(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n) + (\partial_\theta p_1^{(n)} - \partial_\theta p_2^{(n)})\, q_2\circ\mathscr{S}_2^n \\ &+ \partial_\theta p_2^{(n)}\,(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n) + (p_1^{(n)} - p_2^{(n)})\,(\partial_\theta q_1\circ\mathscr{S}_1^n - \partial_\theta q_2\circ\mathscr{S}_2^n) \\ &+ (p_1^{(n)} - p_2^{(n)})\,\partial_\theta q_2\circ\mathscr{S}_2^n + p_2^{(n)}\,(\partial_\theta q_1\circ\mathscr{S}_1^n - \partial_\theta q_2\circ\mathscr{S}_2^n)\Big),\end{aligned} \tag{B.22}$$

where we expand

$$p_1^{(n)} - p_2^{(n)} = \sum_{k=0}^{n-1} p_1^{(k)}\big(p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k\big)\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}, \tag{B.23a}$$

$$\begin{aligned}\partial_\theta p_1^{(n)} - \partial_\theta p_2^{(n)} = &\sum_{k=0}^{n-1}\sum_{i=0}^{k-1} \pi_1^{(i)}\Delta_{\pi_1,\pi_2,W,i}\pi_2^{(k-i-1)}\circ\mathscr{S}_2^{i+1}\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\ &+ \sum_{k=0}^{n-1}\sum_{i=0}^{k-1} \pi_1^{(i)}\rho\mathfrak{a}_3\circ\mathscr{S}_1^i\mathfrak{f}\circ A_0^i\pi_2^{(k-i-1)}\circ\mathscr{S}_2^{i+1}\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\ &+ \sum_{k=0}^{n-1}\pi_1^{(k)}\Delta_{\partial_\theta p_1,\partial_\theta p_2,W,k}p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} + \sum_{k=0}^{n-1}\pi_1^{(k)}\rho\mathfrak{a}_4\circ\mathscr{S}_1^k\,\mathfrak{f}\circ A_0^k p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\ &+ \sum_{k=0}^{n-1}\pi_1^{(k)}\partial_\theta p_1\circ\mathscr{S}_1^k \sum_{i=0}^{n-k-2} p_1^{(i)}\circ\mathscr{S}_1^{k+1}\Delta_{p_1,p_2,W,i+k+1}p_2^{(n-k-2-i)}\circ\mathscr{S}_2^{i+k+2}, \\ &+ \sum_{k=0}^{n-1}\pi_1^{(k)}\partial_\theta p_1\circ\mathscr{S}_1^k \sum_{i=0}^{n-k-2} p_1^{(i)}\circ\mathscr{S}_1^{k+1}\rho\mathfrak{a}_1\circ\mathscr{S}_1^{i+k+1}\mathfrak{f}\circ A_0^{i+k+1}p_2^{(n-k-2-i)}\circ\mathscr{S}_2^{i+k+2},\end{aligned} \tag{B.23b}$$

$$\partial_\theta q_1\circ\mathscr{S}_1^n - \partial_\theta q_2\circ\mathscr{S}_2^n = \Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n} + \rho\,\mathfrak{a}_5\circ\mathscr{S}_1^n\,\mathfrak{f}\circ A_0^n, \tag{B.23c}$$

with $\mathfrak{a}_3(\theta) := \mathfrak{a}_1(\theta) + \partial_\theta c_0(\theta)$, $\mathfrak{a}_4(\theta) := \partial_\theta\mathfrak{a}_1(\theta)$ and $\mathfrak{a}_5(\theta) = \partial_\theta\mathfrak{a}_2(\theta)$. Finally, in (B.22) and (B.23a) we express $q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n$ and $p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k$, respectively, according to (4.65), while in (B.23b) we apply Lemma 4.37 to the summands containing the functions $\mathfrak{a}_1$, $\mathfrak{a}_3$ and $\mathfrak{a}_4$.

Remark B.5. Bounds analogous to (4.67) hold also for $\Delta_{\pi_1,\pi_2,W,n}$, $\Delta_{\partial_\theta p_1,\partial_\theta p_2,W,n}$ and $\Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n}$. Thus, for any $\Delta,\Delta'\in\{\Delta_{p_2,W,n},\Delta_{q_2,W,n},\Delta_{p_1,p_2,W,n},\Delta_{q_1,q_2,W,n},\Delta_{\pi_1,\pi_2,W,n},\Delta_{\partial_\theta p_1,\partial_\theta p_2,W,n},\Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n}\}$ and for any $k\geq 0$, one has $|||\Delta|||^-_{\alpha_0,3}\leq C\rho$, $\langle|\Delta\, W\circ A_0^k|\rangle\leq C\rho^2$ and $\langle|\Delta\Delta'|\rangle\leq C\rho^3$.

By collecting together the expansions arising from (B.22) and (B.23), we write $\partial_\theta h_1 - \partial_\theta h_2$ as a sum of 28 contributions, which, despite being long and tedious to estimate, can all be treated along similar lines. For example, the sum of the contributions $(p_1^{(n)} - p_2^{(n)})\,(\partial_\theta q_1\circ\mathscr{S}_1^n - \partial_\theta q_2\circ\mathscr{S}_2^n)$ in the second line of (B.22), by taking into account (B.23a), (4.65) and (B.23c), leads to

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{k=0}^{n-1} p_1^{(k)}\, \Delta_{p_1,p_2,W,k}\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\, \Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n} \\
&\quad+\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} p_1^{(k-1)}\big(\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^k\, p_2^{(n)}\circ\mathscr{S}_2^{k+1}\Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n+k+1}+\delta_{k,n}\mathfrak{a}_1\circ\mathscr{S}_1^{n-1}\Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n}\big)W\circ A_0^k \\
&\quad+\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} p_1^{(k)}\, \Delta_{p_1,p_2,W,k}\, p_2^{(n)}\circ\mathscr{S}_2^{k+1}\, \rho\, \mathfrak{a}_5\circ\mathscr{S}_1^{n+k+1}\, \mathfrak{f}\circ A_0^{n+k+1} \\
&\quad+\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} p_1^{(k-1)}\rho\big(\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^k p_2^{(n)}\circ\mathscr{S}_2^{k+1}\mathfrak{a}_5\circ\mathscr{S}_1^{n+k+1}\mathfrak{f}\circ A_0^{n+k+1}+\delta_{k,n}\mathfrak{a}_1\circ\mathscr{S}_1^{n-1}\mathfrak{a}_5\circ\mathscr{S}_1^{n}\mathfrak{f}\circ A_0^{n}\big)W\circ A_0^n,
\end{aligned}$$

where $\delta_{k,n}$is the Kronecker delta. The average of all contributions is bounded by $C\rho$: in the first two lines, directly because of Remark B.5 and (4.50); in the third line, because of Remark B.5, (4.70) and (4.50), after writing $p_2^{(n)}\circ\mathscr{S}_2^{k+1}=p_2^{(n)}\circ\mathscr{S}_1^{k+1}+(p_2^{(n)}\circ\mathscr{S}_2^{k+1}-p_2^{(n)}\circ\mathscr{S}_1^{k+1})$, expanding $p_2^{(n)}\circ\mathscr{S}_2^{k+1}-p_2^{(n)}\circ\mathscr{S}_1^{k+1}$ as in (4.58a) to apply Lemma 4.32 and using Lemma 4.39 to obtain a factor $\mathfrak{D}_{1,\mathfrak{a}_5,n}\circ\mathscr{S}_1^n$ from the terms with $p_2^{(n)}\circ\mathscr{S}_1^{k+1}$; in the fourth line, because of (4.77) for the terms where only the sum over n appears and once more because of Remark B.5 for all the other terms, after decomposing $\mathfrak{a}_5\circ\mathscr{S}_1^{n+k+1}=\mathfrak{a}_5\circ\mathscr{S}_2^{n+k+1}+(\mathfrak{a}_5\circ\mathscr{S}_1^{n+k+1}-\mathfrak{a}_5\circ\mathscr{S}_2^{n+k+1})$ and applying Lemma 4.39 to obtain a factor $\mathfrak{D}_{5,\mathfrak{a}_2,n+k+1}$ from the terms containing $\mathfrak{a}_5\circ\mathscr{S}_2^{n+k+1}$ and Lemma 4.32 to extract a further function W from the terms with the difference. All the other contributions can be dealt with in a similar way and eventually a bound $C\rho$ is found for all of them, so that the first bound in (4.84) follows.

Next, consider $(\partial_\theta h_1-\partial_\theta h_2)\,(\partial_\theta h_1-\partial_\theta h_2)$ and write both differences as in (B.22). Then, we observe that all the contributions obtained from (B.22) either are $O(1)$ in ρ or contain at least a function W. This yields that $|\langle\mathcal{A}_i\mathcal{A}_j\rangle|\le C\rho$ for all $i,j=1,\dots,8$, so that the second bound in (4.84) follows as well.

B.3 Second derivative of the conjugation: proof of Lemma 4.47

The proof of Lemma 4.47 follows the same scheme as Propositions 4.44 and 4.45. In particular, we prove that $|\langle(\partial_\theta^2 h_2(\theta,\cdot)-\partial_\theta^2\overline{h}(\theta))^2\rangle|\le C\rho$ and $\langle(\partial_\theta^2 h_1(\theta,\cdot)-\partial_\theta^2 h_2(\theta,\cdot))^2\rangle\le C\rho$. To this end, we write

$$\begin{aligned}
\partial_\theta^2(p_2^{(n)}\, q_2\circ\mathscr{S}_2^n)-\partial_\theta^2(\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n) &= \partial_\theta^2 p_2^{(n)}\, q_2\circ\mathscr{S}_2^n-\partial_\theta^2\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n \\
&\quad+2\partial_\theta p_2^{(n)}\,\partial_\theta q_2\circ\mathscr{S}_2^n-2\partial_\theta\overline{p}^{(n)}\,\partial_\theta\overline{q}\circ\overline{G}^n+p_2^{(n)}\,\partial_\theta^2 q_2\circ\mathscr{S}_2^n-\overline{p}^{(n)}\,\partial_\theta^2\overline{q}\circ\overline{G}^n, \\
\partial_\theta^2(p_1^{(n)}\, q_1\circ\mathscr{S}_1^n)-\partial_\theta^2(p_2^{(n)}\, q_2\circ\mathscr{S}_2^n) &= \partial_\theta^2 p_1^{(n)}\, q_1\circ\mathscr{S}_1^n-\partial_\theta^2 p_2^{(n)}\, q_2\circ\mathscr{S}_2^n \\
&\quad+2\partial_\theta p_1^{(n)}\,\partial_\theta q_1\circ\mathscr{S}_1^n-2\partial_\theta p_2^{(n)}\,\partial_\theta q_2\circ\mathscr{S}_2^n+p_1^{(n)}\,\partial_\theta^2 q_1\circ\mathscr{S}_1^n-p_2^{(n)}\,\partial_\theta^2 q_2\circ\mathscr{S}_2^n,
\end{aligned}$$

so that the terms in the second and fourth lines can be studied as in Appendices B.1 and B.2.2, with q_1, q_2 and $\overline{q}$ replaced, respectively, either with $\partial_\theta q_1$, $\partial_\theta q_2$ and $\partial_\theta\overline{q}$ or $\partial_\theta^2 q_1$, $\partial_\theta^2 q_2$ and $\partial_\theta^2\overline{q}$. The terms in the first and third lines can be dealt with by first writing, according to (2.18b),

$$\begin{aligned}
\partial_\theta^2 p_2^{(n)} &= \sum_{k=0}^{n-1}(p_2\,(\partial_\theta G_2)^2)^{(k)}\,\partial_\theta^2 p_2\circ\mathscr{S}_2^k\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\
&\quad+\sum_{k=0}^{n-1}\sum_{i=0}^{k-1}(p_2\,(\partial_\theta G_2)^2)^{(i)}\,(\partial_\theta p_2\partial_\theta G_2)\circ\mathscr{S}_2^i\,(p_2\partial_\theta G_2)^{(k-i-1)}\circ\mathscr{S}_2^{i+1}\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\
&\quad+\sum_{k=0}^{n-1}\sum_{i=0}^{k-1}(p_2\,(\partial_\theta G_2)^2)^{(i)}\,(p_2\,\partial_\theta^2 G_2)\circ\mathscr{S}_2^i\,(p_2\partial_\theta G_2)^{(k-i-1)}\circ\mathscr{S}_2^{i+1}\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\
&\quad+\sum_{k=0}^{n-1}\sum_{i=0}^{n-k-2}(p_2\,\partial_\theta G_2)^{(k)}\,\partial_\theta p_2\circ\mathscr{S}_2^k\,(p_2\partial_\theta G_2)^{(i)}\circ\mathscr{S}_2^{k+1}\partial_\theta p_2\circ\mathscr{S}_2^{k+1+i}\, p_2^{(n-k-i-2)}\circ\mathscr{S}_2^{k+2+i},
\end{aligned} \tag{B.24}$$

and analogous expressions for $\partial_\theta^2 \overline{p}^{(n)}$ and $\partial_\theta^2 p_1^{(n)}$, with $\overline{\mathscr{S}}$ and $\mathscr{S}_1$ instead of $\mathscr{S}_2$, respectively, and then expanding $\partial_\theta^2 p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n - \partial_\theta^2 \overline{p}^{(n)}\, \overline{q} \circ \overline{G}^n$ and $\partial_\theta^2 p_1^{(n)}\, q_1 \circ \mathscr{S}_1^n - \partial_\theta^2 p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n$ by proceeding as in Appendices B.1 and B.2.2, with respect to which besides the functions $\pi_1 = p_1 \partial_\theta G_1$, $\pi_2 = p_2 \partial_\theta G_2$ and $\overline{p}\, \partial_\theta \overline{G}$, the functions $p_1 \partial_\theta^2 G_1$, $\partial_\theta p_1 \partial_\theta G_1$, $p_1 (\partial_\theta G_1)^2$, $p_2 \partial_\theta^2 G_2$, $\partial_\theta p_2 \partial_\theta G_2$, $p_2 (\partial_\theta G_2)^2$ $\overline{p}\, \partial_\theta^2 \overline{G}$, $\partial_\theta \overline{p}\, \partial_\theta \overline{G}$ and $\overline{p}\, (\partial_\theta \overline{G})^2$ also appear. So, when considering $\partial_\theta^2 p_1^{(n)}\, q_1 \circ \mathscr{S}_1^n - \partial_\theta^2 p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n$, together with the differences in (B.23), we have to expand also the differences $(p_1 \partial_\theta^2 G_1)^{(n)} - (p_2 \partial_\theta^2 G_2)^{(n)}$, $(\partial_\theta p_1 \partial_\theta G_1)^{(n)} - (\partial_\theta p_2 \partial_\theta G_2)^{(n)}$ and $(p_1 (\partial_\theta G_1)^2)^{(n)} - (p_2 (\partial_\theta G_2)^2)^{(n)}$, while, inserting (B.24) and the analogous expression for $\partial_\theta^2 \overline{p}^{(n)}$ into $\partial_\theta^2 p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n - \partial_\theta^2 \overline{p}^{(n)}\, \overline{q} \circ \overline{G}^n$, we obtain a sum of contributions with the same structure as (B.2).

Comparing (2.18b) with (2.18a) we note that, because of the presence of an extra derivative with respect to θ, a further sum may appear. However, when this happens, the two derivatives act on two distinct factors; this produces a factor ρ, which compensates the factor $1/\rho$ arising from the sum.

References

[1] R.L. Adler, B. Weiss, *Similarity of automorphisms of the torus*, Mem. Amer. Math. Soc. Vol. 98, American Mathematical Society, Providence, 1970.

[2] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, Ch. Zhou, *Synchronization in complex networks*, Phys. Rep. **469** (2008), no. 3, 93–153.

[3] V.I. Arnold, A. Avez, *Ergodic Problems of Classical Mechanics*, W.A. Benjamin, New York, 1968.

[4] V.I. Arnold, V.N. Kozlov, A.I. Neishtadt, *Dynamical Systems III. Mathematical Aspects of Classical and Celestial Mechanics*, Springer, Berlin, 1988.

[5] M.A. Aziz-Alaoui, *Synchronization of Chaos*, Encyclopedia of Mathematical Physics Vol. 5, p. 213–226, Elsevier, 2006.

[6] K. Bjerklöv, R. Krikorian, *Monotone families of circle diffeomorphisms driven by expanding circle maps*, Comm. Math. Phys **405** (2024), no. 9, Paper 205, 21 pp.

[7] M. Blank, G. Keller, C. Liverani, *Ruelle-Perron-Frobenius spectrum for Anosov maps*, Nonlinearity **15** (2002), no. 6, 1905–1973.

[8] S. Boccaletti, J. Kurths, G. Osipov, D.L. Valladares, C.S. Zhou *The synchronization of chaotic systems*, Phys. Rep. **366** (2002), no. 1-2, 1–101.

[9] N.N. Bogoliubov, Y.A. Mitropolsky, *Asymptotic Methods in the Theory of Non-linear Oscillations*, Gordon and Breach Science Publishers, New York, 1961.

[10] F. Bonetto, P. Falco, A. Giuliani, *Analyticity of the SRB measures of a lattice of coupled Anosov diffeomorphisms of the torus*, J. Math. Phys. **45** (2004), no. 8, 3282–3309.

[11] F. Bonetto, J.L. Lebowitz, L. Rey-Bellet, *Fourier's law: a challenge to theorists*, Mathematical Physics 2000, 128–150, Imperial College Press, London, 2000.

[12] F. Bonetto, A. Kuipianen, J.L. Lebowitz, *Absolute continuity of projected SRB measures of coupled Arnold cat map lattices*, Ergodic Theory Dynam. Systems **24** (2004), no. 1, 59–88.

[13] J. Bricmont, A. Kuipianen, *Diffusion in energy conserving coupled maps*, Comm. Math. Phys. **321** (2013), no. 2, 311–369.

[14] M. Brin, G. Stuck, *Introduction to Dynamical Systems*, Cambridge University Press, Cambridge, 2002.

[15] A. Carati, L. Galgani, F. Gangemi, R. Gangemi, *Electronic trajectories in atomic physics: the chemical bond in the H_2^+ ion*, Chaos **30** (2020), no. 6, 063109, 8pp.

[16] R. Castorrini, C. Liverani, *Quantitative statistical properties of two-dimensional partially hyperbolic systems*, Adv. Math. **409** (2022), Paper 108625, 122 pp.

[17] N.I. Chernov, *Limit theorems and Markov approximations for chaotic dynamical systems*, Probab. Theory Related Fields **101** (1995), no. 3, 321–362.

[18] L. De Carlo, G. Gentile, A. Giuliani, *Construction of the Lyapunov spectrum in a chaotic system displaying phase synchronization*, Math. Phys. Anal. Geom. **19** (2016), no. 2, Art. 10, 23 pp.

[19] G.F. Dell'Antonio, *The van Hove limit in classical and quantum mechanics*, Stochastic processes in quantum theory and statistical physics (Marseille, 1981), 75–110, Lecture Notes in Phys. Vol. 173, Springer, Berlin,

1982.

[20] M.F. Demers, *A gentle introduction to anisotropic Banach spaces*, Chaos Solitons Fractals **116** (2018), 29–42.

[21] M.F. Demers, N. Kiamari, C. Liverani, *Transfer operators in hyperbolic dynamics – an introduction*, 33o Colóq. Bras. Mat., Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2021.

[22] J. De Simoi, C. Liverani, *Statistical properties of mostly contracting fast-slow partially hyperbolic systems*, Invent. Math. **206** (2016), no. 1, 147–227.

[23] J. De Simoi, C. Liverani, *Limit theorems for fast-slow partially hyperbolic systems*, Invent. Math. **213** (2018), no. 3, 811–1016.

[24] J. De Simoi, C. Liverani, Ch. Poquet, D. Volk, *Fast-slow partially hyperbolic systems versus Freidlin-Wentzell random systems*, J. Stat. Phys. **166** (2017), no. 3-4, 650–679.

[25] H.A. Dijkstra, *Nonlinear Climate Dynamics*, Cambridge University Press, Cambridge, 2013.

[26] D. Dolgopyat, *Averaging and invariant measures*, Mosc. Math. J. **5** (2005), no. 3, 537–576.

[27] D. Dolgopyat, C. Liverani, *Energy transfer in a fast-slow Hamiltonian system*, Comm. Math. Phys. **308** (2011), no. 1, 201–225.

[28] J. Franks, *Anosov diffeomorphisms. Global Analysis*, Proc. Sympos. Pure Math. Vol. 14, American Mathematical Society, Providence, R.I., 1970.

[29] G. Gallavotti, F. Bonetto, G. Gentile, *Aspects of the Ergodic, Qualitative and Statistical Theory of Motion*, Springer, Berlin, 2004.

[30] G. Gallavotti, G. Gentile, A. Giuliani, *Resonances within chaos*, Chaos **22** (2012), no. 52, 026108, 6 pp.

[31] G. Gallavotti, E.G.D. Cohen, *Dynamical ensembles in stationary states*, J. Statist. Phys. **80** (1995), no. 5-6, 931–970.

[32] G. Gentile, *Introduzione ai sistemi dinamici – Vol. 1. Equazioni differenziali, analisi qualitativa e alcune applicazioni*, Springer, Milano, 2021.

[33] M. Ghil, V. Lucarini, *The physics of climate variability and climate change*, Rev. Modern Phys. **92** (2020), no. 3, 035002, 77 pp.

[34] P. Giulietti, S. Marmi, M. Tanzi, *Random-like properties of chaotic forcing*, J. Lond. Math. Soc. **106** (2022), no. 3, 2804–2845.

[35] J.M. González-Miranda, *Synchronization and Control of Chaos. An Introduction for Scientists and Engineers*, Imperial College Press, London, 2004.

[36] A. Grigo, *A rigorous derivation of Haff's law for a periodic two-disk fluid*, J. Statist. Phys. **176** (2019), 806–835.

[37] J. Hale, *Ordinary Differential Equations*, John Wiley & Sons, New York, 1980.

[38] Ph.J. Holmes, *Averaging and chaotic motions in forced oscillations*, SIAM J. Appl. Math. **38** (1980), no. 1, 65–80.

[39] D. Kelly, I. Melbourne, *Smooth approximation of stochastic differential equations*, Ann. Probab. **44** (2016), no. 1, 479–520.

[40] Yu. Kifer, *Averaging and climate models*, Progr. Probab. **49** (2001), 171–188.

[41] Yu. Kifer, *Large deviations and adiabatic transitions for dynamical systems and Markov processes in fully coupled averaging*, Mem. Amer. Math. Soc. Vol. 201, American Mathematical Society, Providence, 2009.

[42] N.M. Krylov, N.N. Bogolyubov, *Introduction to Non-Linear Mechanics*, Princeton University Press, Princeton, NJ, 1943.

[43] C.A. Kitio Kwuimy, P. Woafo, *Dynamics, chaos and synchronization of self-sustained electromechanical systems with clamped-free flexible arms*, Nonlinear Dynam. **53** (2008), no. 3, 201–213.

[44] A. Korepanov, Z. Kosloff, I. Melbourne, *Averaging and rates of averaging for uniform families of deterministic fast-slow skew product systems*, Studia Math. **238** (2017), no. 1, 59–89.

[45] J.L. Lagrange, *Mécanique Analitique*, La Veuve Desaint, Paris, 1788.

[46] C. Liverani, *Transport in partially hyperbolic fast-slow systems*, Proceedings of the International Congress of Mathematicians, Rio de Janeiro 2018, Vol. III, 2643-2667, World Sci. Publ., Hackensack, NJ, 2018.

[47] V. Lucarini, R. Blender, C. Herbert, F. Ragone, S. Pascale, J. Wouters, *Mathematical and physical ideas for climate science*, Rev. Geophys. **52** (2014), 809–859.

[48] V. Lucarini, F. Ragone, F. Lunkeit, *Predicting climate change using response theory: global averages and*

spatial patterns, J. Stat. Phys. **166** (2017), no. 3-4, 1036–1064.

[49] T.J. Lyons, *Differential equations driven by rough signals*, Rev. Mat. Iberoam. **14** (1998), no. 2, 215–310.

[50] C.D. Murray, S.F. Dermott, *Solar System Dynamics*, Cambridge University Press, Cambridge, 1999.

[51] S.E. Newhouse, *On codimension one Anosov diffeomorphisms*, Amer. J. Math. **92** (1970), 761–770.

[52] L.M. Pecora, Th.L. Carroll, *Synchronization in chaotic systems*, Phys. Rev. Lett. **64** (1990), no. 8, 821–825.

[53] L.M. Pecora, Th.L. Carroll, G.A. Johnson, D.J. Mar, J.F. Heagy, *Fundamentals of synchronization in chaotic systems, concepts, and applications*, Chaos **7** (1997), no. 4, 520–543.

[54] L. Perko, *Differential equations and dynamical systems*, Texts Appl. Math. 7, Springer, New York, 2001.

[55] W. Parry, M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque 187-188, 1990.

[56] A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization. A Universal Concept in Nonlinear Sciences*, Cambridge University Press, Cambridge, 2001.

[57] J.H. Sanders, F. Verhulst, J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*, Springer, New York, 2007.

[58] M.R. Snavely, *Markov partitions for the two-dimensional torus*, Proc. Amer. Math. Soc. **113** (1991), no. 2, 517–527.

[59] S. Soldatenko, A. Bogomolov, A. Ronzhin, *Mathematical modelling of climate change and variability in the context of outdoor ergonomics*, Mathematics **9** (2021), 2920, 26pp.

[60] H. Spohn, *Large Scale Dynamics of Interacting Particles*, Springer, Berlin, 1991.