Synchronization and averaging in dynamical systems with fast/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, guido.gentile@uniroma3.it

February 19, 2024

Abstract

We study a family of dynamical systems obtained by coupling a chaotic (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. We show that the two systems synchronize: the trajectories evolve toward an attracting invariant manifold, and the full dynamics is conjugated to its linearization around the invariant manifold. When the interaction is small, the evolution of the neutral variable, that is the variable which describes the neutral system, is very close to the identity; hence the neutral variable appears as a *slow* variable with respect to 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.

Contents

1 Introduction

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 [43]. On contrast, in chaotic systems, where trajectories starting at close initial conditions tend to diverge from each other, synchronization may appear as an unlikely phenomenon. Nonetheless, the presence of negative Lypaunov exponents due to a dissipative coupling can still produce synchronization [45].

One of the simplest models one can think of is 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 through a dissipative perturbation which is unidirectional, that is which affects only the motion of the neutral system – by 'neutral system' we mean a system which just remains at rest in the absence of the interaction. Such a model has been explicitly considered in ref. [29], under suitable assumptions which simplify the analysis to a great extent (see Subsection 3.1 for further details). More complicated and realistic models can be easily envisaged [47, 48, 7, 32, 2, 46], however the main advantage of the simple model studied in ref. [29] 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 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 ref. [7] and references therein for an introduction to the topics). The invariant manifold is no more than Hölder continuous, but, with the hypotheses considered in ref. [29], 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 ref. [29], 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. Moreover, we extend the analysis beyond the perturbative regime, by requiring the coupling only to be dissipative in a finite region. 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 both the average and the variance of the function whose graph describes the invariant manifold is of the order of the perturbation.

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. In particular, the invariant manifold is proved to be an attractor. The conjugation too, in general, is no more than Hölder continuous. In the perturbative regime, also the average deviations of the conjugating function are found to be of the order of the perturbation, despite the fact that deviations of order 1 are possible. The results discussed above yield, as a byproduct, that any system in the class we consider admits a unique physical measure given by the lift to the invariant manifold of the normalized Lebesgue measure of $\mathbb{T}^2$. This measure is exponentially mixing, with mixing rate limited by the low regularity of the invariant manifold.

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 [39]. In quasi-integrable 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 method of averaging [33, 4, 52, 49]: once the oscillations of the fast variables have been integrated out, the approximate solution one finds provides, in general, a reliable description of the dynamics up to a time which is inversely proportional to the slow time scale; to make the analysis rigorous, as the next step, one has to control the corrections.

In a similar spirit, we investigate the *scaling regime* of the dynamics of the systems we are considering, that is we fix a finite time t and study what happens, when the size ρ of the perturbation is very small, after $k = \lfloor t/\rho \rfloor$ iterations of the map. We find that, in this regime, the evolution of the slow variable becomes essentially independent of the dynamics on the torus, and is effectively described by the solution of a suitable ordinary differential equation. The differential equation is essentially the continuous limit of the original map, with the function of the fast variables describing the interaction replaced by its average on the torus. More precisely we show that, given an initial condition for the slow variable, and 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. Because of the presence of dissipation, which makes the trajectories to evolve toward the invariant manifold, the probability of such deviations remains small along the full trajectory up to an infinite time. These results are related to the study in refs. [20, 40, 22] on a similar class of models, in which the system described by the slow variable is coupled, through a more general and not necessarily uniformly dissipative perturbation, with an expanding circle map. On the other hand, for dissipative interactions, the presence of both stable and unstable directions for the fast chaotic dynamics on $\mathbb{T}^2$ makes our study more general – see also Subsection 3.2 for a more detailed comparison.

Of course, the fact that the neutral system does not influence the chaotic evolution on $\mathbb{T}^2$ makes the analysis easier. Notwithstanding this simplification, we think that the model we study here contains most of the relevant features to control also the statistical properties over long time scales of models with more general couplings (see Section 4). At the same time, it has the advantage of being well suited for explicit, direct computations, and eliminating details which would introduce technical complications without really adding anything to the underlying physics. Therefore, in our opinion, the model represents a first step toward a full mathematical understanding of the problem, before considering more realistic situations.

Models as those considered above have been widely studied in the literature also as a preliminary step of the more ambitious program of deriving rigorously the heat equation from the microscopic equations of motions. In this perspective, as pointed out in ref. [40], what one would really like to investigate is the case of several chaotic systems coupled with an equal number of neutral systems in a local manner and weakly interacting with each other, and look for results which are uniform in the number of systems. One is ultimately interested in taking the hydrodynamical limit, that is considering infinitely many coupled systems obtained by a suitable scaling limit; a model of this kind, with a different approach with respect to ours, has been studied in ref. [11], where a diffusion equation for the macroscopic energy is derived starting from the microscopic dynamics. For further comments on this line of research we refer to Subsection 5.3.

2 Model and Results

In this section, first, we introduce the basic ingredients that will be used in the rest of the paper: automorphisms of $\mathbb{T}^2$, regularity norms and relative Banach spaces, and correlation functions. Then, we give the formal definition of the model we will study and present our main results, referring to Sections 6 and 7 – and the Appendices – for the proofs.

2.1 Basic Ingredients

Let $\mathbb{T} := \mathbb{R}/2\pi\mathbb{Z}$ and consider an Anosov automorphism A_0 of $\mathbb{T}^2$ [1, 12, 23] 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 $\lambda = \lambda_+$.

Let $\Omega := \mathcal{U} \times \mathbb{T}^2$, where $\mathcal{U}$ is a non-empty closed interval of either $\mathbb{T}$ or $\mathbb{R}$. 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 $\mathbb{T}^2$.

Consider the supremum norm

$$\|f\|_\infty := \sup_{(\varphi,\psi)\in\Omega} |f(\varphi,\psi)|, \tag{2.2}$$

and let $\mathscr{B}(\Omega,\mathbb{R})$ denote the Banach space of the bounded continuous functions f equipped with the norm $\|f\|_\infty$. For $\alpha \in (0, 1]$, consider the Hölder seminorm

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

and the two directional Hölder seminorms

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

that satisfy the inequalities

$$|f|_\alpha^\pm \le |f|_\alpha \le |f|_\alpha^+ + |f|_\alpha^-, \qquad \|f\|_\infty \le c_\alpha |f|_\alpha + \|\langle f\rangle\|_\infty, \tag{2.5}$$

with $c_\alpha := (\pi\sqrt{2})^\alpha$. Introduce also the norms

$$\|f\|_{\alpha_-,\alpha_+} := \|f\|_\infty + \alpha_-|f|_{\alpha_-}^- + \alpha_+|f|_{\alpha_+}^+, \qquad \|f\|_\alpha^- := \|f\|_{\alpha,0}, \qquad \|f\|_\alpha^+ := \|f\|_{0,\alpha}, \tag{2.6}$$

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 functions defined on Ω equipped with the norms $\|\cdot\|_{\alpha_-,\alpha_+}$, $\|\cdot\|^+_\alpha$ and $\|\cdot\|^-_\alpha$, respectively. 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 [6, 18], and this will be exploited in what follows. Finally, observe that $\mathscr{B}_0(\Omega,\mathbb{R}) = \mathscr{B}(\Omega,\mathbb{R})$.

It is easy to see that

$$|fg|_\alpha \le \|f\|_\infty |g|_\alpha + |f|_\alpha \|g\|_\infty, \qquad |fg|^\pm_\alpha \le \|f\|_\infty |g|^\pm_\alpha + |f|^\pm_\alpha \|g\|_\infty, \tag{2.7}$$

and

$$\|fg\|_\alpha \le \|f\|_\alpha \|g\|_\alpha, \qquad \|fg\|_{\alpha_-,\alpha_+} \le \|f\|_{\alpha_-,\alpha_+}\|g\|_{\alpha_-,\alpha_+}. \tag{2.8}$$

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 are α-Hölder continuous in the second one, that is in the variable $\psi\in\mathbb{T}^2$, equipped with the norm

$$|||f|||_{\alpha,k} := \sum_{n=0}^{k-1} \|\partial^n_\varphi f\|_\alpha + \|\partial^k_\varphi f\|_\infty = \sum_{n=0}^{k} \|\partial^n_\varphi f\|_\infty + \alpha \sum_{n=0}^{k-1} |\partial^n_\varphi f|_\alpha\,, \tag{2.9}$$

where ∂_φ denotes the derivative with respect to the first variable. Similarly define the norms $|||\cdot|||^+_{\alpha,k}$ and $|||\cdot|||^-_{\alpha,k}$, as in (2.9) with $|\cdot|_\alpha$ replaced with $|\cdot|^+_\alpha$ and $|\cdot|^-_\alpha$, and call the respective Banach spaces $\mathscr{B}^+_{\alpha,k}(\Omega,\mathbb{R})$ and $\mathscr{B}^-_{\alpha,k}(\Omega,\mathbb{R})$. As in (2.8) we get

$$|||fg|||_{\alpha,k} \le |||f|||_{\alpha,k}|||g|||_{\alpha,k}, \qquad |||fg|||^\pm_{\alpha,k} \le |||f|||^\pm_{\alpha,k}|||g|||^\pm_{\alpha,k}. \tag{2.10}$$

Remark 2.1. In the following we also consider sets of the form

$$\mathcal{A} = \{(\varphi,\psi) : a_-(\psi) \le \varphi \le a_+(\psi)\} \subset \mathbb{R}\times\mathbb{T}^2, \tag{2.11}$$

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 $\mathcal{A}\subset\mathbb{R}\times\mathbb{T}^2$ of the form (2.11). We only need to take the supremum over $(\varphi,\psi)\in\mathcal{A}$ and replace (2.3) and (2.4) with

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

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

respectively. This allows us to define the corresponding Banach spaces in the same way as before, with the set $\mathcal{A}$ 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.

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

Remark 2.2. The behavior of the seminorms (2.3) and (2.4) 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

$$|f \circ A_0^n|_\alpha \le \lambda^{\alpha|n|}|f|_\alpha\,, \qquad |f \circ A_0^n|_\alpha^+ \le \lambda^{\alpha n}|f|_\alpha^+\,, \qquad |f \circ A_0^n|_\alpha^- \le \lambda^{-\alpha n}|f|_\alpha^-\,. \tag{2.12}$$

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

$$\Big|\prod_{i=0}^{n-1} g_{i,+} \circ A_0^{-i}\Big|_\alpha^+ \le \sum_{i=0}^{n-1} \lambda^{-\alpha i}|g_{i,+}|_\alpha^+ \prod_{\substack{j=0\\ j\neq i}}^{n-1} \|g_{i,+}\|_\infty. \tag{2.13}$$

Analogously, one has

$$\Big|\prod_{i=0}^{n-1} g_{i,-} \circ A_0^{i}\Big|_\alpha^- \le \sum_{i=0}^{n-1} \lambda^{-\alpha i}|g_{i,-}|_\alpha^- \prod_{\substack{j=0\\ j\neq i}}^{n-1} \|g_{i,-}\|_\infty, \tag{2.14}$$

if the functions $g_{0,-}, \dots, g_{n-1,-}$ are in $\mathfrak{B}_\alpha^-(\mathbb{T}^2, \mathbb{R})$, for some $\alpha \in (0, 1]$.

A crucial role in our analysis is played by the correlation functions and their decay due to the hyperbolicity of A_0. In Appendix A.3 we prove the following estimate.

Proposition 2.3. *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*

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

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

Remark 2.4. Proposition 2.3 implies that, under the same assumptions, for every $\alpha' < \alpha$ one has

$$|\langle g_+\, g_- \circ A_0^n\rangle - \langle g_+\rangle\langle g_-\rangle| \le C'\lambda^{-\alpha' n}\|\tilde{g}_+\|_\alpha^+\|\tilde{g}_-\|_\alpha^-,$$

with the positive constant C' depending only on α'.

2.2 The model

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

$$\mathscr{S}(\varphi, \psi) = (\mathscr{S}_\varphi(\varphi, \psi), \mathscr{S}_\psi(\varphi, \psi)) = (G(\varphi, \psi), A_0\psi), \qquad G(\varphi, \psi) := \varphi + F(\varphi, \psi)\,, \tag{2.15}$$

with $F \in \mathscr{B}_{\alpha_0,6}(\mathbb{T} \times \mathbb{T}^2, \mathbb{R})$, for some $\alpha_0 \in (0, 1]$, and assume $\mathscr{S}$ to satisfy the following hypotheses.

Remark 2.5. The assumption of Hölder continuity of the map in the variable ψ is very natural, since it cannot be weakened and, at the same time, requiring stronger regularity would not provide stronger results (see also Remarks 2.19 and 6.3 below). On contrast, not only the regularity in the variable φ may not be optimal, but some of the very same results proved along the paper require much less regularity (see for instance Remark 2.15); on the other hand, as far as the variable φ is concerned, higher regularity of the map does yield higher regularity of the long time behavior of the dynamics (see Remark 2.16).

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 $(\varphi, \psi) \in \Omega := \mathcal{U}\times\mathbb{T}^2$.*

Hypothesis 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 $\psi \in \mathbb{T}^2$.*

Hypotheses 1 and 2 imply that for every $\psi \in \mathbb{T}^2$ there is a unique point $S(\psi) \in \mathrm{int}\,\mathcal{U}$ with $F(S(\psi),\psi) = 0$. Set

$$S_m := \inf_{\psi\in\mathbb{T}^2} S(\psi), \qquad S_M := \sup_{\psi\in\mathbb{T}^2} S(\psi), \tag{2.16}$$

and define

$$\Lambda := [S_m, S_M] \times \mathbb{T}^2, \qquad \Gamma := -\sup_{(\varphi,\psi)\in\Lambda} \partial_\varphi F(\varphi,\psi). \tag{2.17}$$

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

Remark 2.6. A simple example of a map satisfying Hypotheses 1–3, with, say, $\mathcal{U} = [-\pi/4, \pi/4]$, is obtained by taking $F(\varphi,\psi) = -\sin(\varphi - g(\psi))$, with $g(\psi) = (\pi/8)\cos(\psi_1 - \psi_2)$. Indeed, in this case we find $S(\psi) = g(\psi)$, so that $\Lambda = [-\pi/8, \pi/8]$ and $\Gamma = \cos(\pi/4)$.

Remark 2.7. Hypothesis 1 yields that $\mathscr{S}$ is injective on Ω and that $\Gamma < 1$. Moreover, as a consequence of both Hypotheses 1 and 2, one has $\mathscr{S}(\Omega) \subsetneq \mathrm{int}\,\Omega$ and $S(\psi) \in (\phi_m, \phi_M)$ for all $\psi \in \mathbb{T}^2$, and hence $[S_m, S_M] \subsetneq (\phi_m, \phi_M)$.

Remark 2.8. 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.

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

Lemma 2.9. *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*

$$N_r \le \frac{\max\{\phi_M - S_M, S_m - \phi_m\}}{\inf\limits_{\Omega\setminus\Lambda_r} |F(\varphi,\psi)|}\,; \tag{2.18}$$

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 2.10. Due to property 6 in Lemma 2.9, without loss of generality, we may and do choose the parametrization in Hypothesis 2 in such a way that $\overline{\varphi} = 0$.

Remark 2.11. By Remark 2.10 we can write

$$F(\varphi,\psi) = \beta(\psi) - \nu(\psi)\,\varphi + \delta(\varphi,\psi)\,\varphi^2, \tag{2.19}$$

where $\langle\beta\rangle = 0$. Conversely, assume that F has the form

$$F(\varphi,\psi) = \beta(\psi) - \nu(\psi)\,\varphi + \delta(\psi)\,\varphi^2, \tag{2.20}$$

with $\langle\beta\rangle = 0$ and $4\|\delta\|_\infty\|\beta\|_\infty < \nu_0^2$, if $\nu_0 := \inf_{\psi\in\mathbb{T}^2}|\nu(\psi)|$. It is easy to see that (2.15) defined by such an F satisfies Hypotheses 1–3 with $\phi_M = \|\beta\|_\infty/2\nu_0$ and $\phi_m = -\phi_M$.

Remark 2.12. Although we defined $\mathscr{S}$ as a map on $\mathbb{T}\times\mathbb{T}^2$, since $\mathscr{S}(\Omega) \subsetneq \Omega$, in the following we will only be interested in the action of $\mathscr{S}$ on Ω and identify Ω as a subset of $\mathbb{R}\times\mathbb{T}^2$. For technical reasons, we will also need to extend the map $\mathscr{S}|_\Omega$ to a map $\mathscr{S}_{\mathrm{ext}}$ on $\Omega_{\mathrm{ext}} = \mathcal{U}_{\mathrm{ext}} \times \mathbb{T}^2$, for some closed interval $\mathcal{U}_{\mathrm{ext}}$ such that $\mathcal{U} \subset \mathcal{U}_{\mathrm{ext}} \subset \mathbb{R}$, in such a way that Hypotheses 1 to 3 remain valid (we refer to Subsection 7.1 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.

2.3 Synchronization

2.3.1 The invariant manifold

Observe that, if in (2.15) we replace the automorphism A_0 with the identity $\mathbb{1}$, that is if we consider the dynamics generated by $\mathscr{S}_{\mathbb{1}}(\varphi,\psi) = (\varphi + F(\varphi,\psi),\psi)$ with F still satisfying Hypotheses 1–3, then $\mathcal{W}_{\mathbb{1}} = \{(S(\psi),\psi) : \psi \in \mathbb{T}^2\}$ is an *invariant manifold* in the sense that $\mathscr{S}_{\mathbb{1}}(\mathcal{W}_{\mathbb{1}}) = \mathcal{W}_{\mathbb{1}}$. 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 $\mathbb{T}^2$. More precisely we say that a manifold $\mathcal{W} = \{(W(\psi),\psi) : \psi \in \mathbb{T}^2\}$ is invariant for $\mathscr{S}$ in (2.15) if we have

$$\mathscr{S}(W(\psi),\psi) = (W(A_0\psi), A_0\psi) \tag{2.21}$$

for every $\psi \in \mathbb{T}^2$. This also means that on $\mathcal{W}$ the dynamics generated by $\mathscr{S}$ is conjugated to the dynamics generated by to the map $\psi \mapsto A_0\psi$.

In Subsection 6.1 we prove the following result.

Theorem 1 (**Synchronization**). *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.15) *satisfying Hypotheses 1–3. There exists a unique invariant manifold $\mathcal{W} = \{(W(\psi),\psi) : \psi \in \mathbb{T}^2\} \subset \Lambda$ for the map $\mathscr{S}$, with $W \in \mathfrak{B}^*_{\alpha_-,\alpha_+}(\mathbb{T}^2,\mathbb{R})$ for suitable $\alpha_-,\alpha_+ \in (0,\alpha_0]$.*

Remark 2.13. The main effect of the hyperbolicity of A_0 is that, in general, the manifold $\mathcal{W}$ is only Hölder continuous in ψ even if we take F very smooth in ψ. Moreover W is, in general, of order 1 in $\|F\|_\infty$ even if F is very close to 0, its maximum size depending mainly on $S(\psi)$ – see also Remark 6.4 below. On the other hand, the existence of the manifold $\mathcal{W}$ (and of the conjugation $\mathscr{H}$ discussed later) remains true if we assume that F is only bounded in ψ.

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

$$\mathscr{S}_0(\eta,\psi) = (\kappa(\psi)\eta, A_0\psi), \qquad \kappa(\psi) := 1 + \partial_\varphi F(W(\psi),\psi), \tag{2.22}$$

where, by Hypothesis 3, one has $\kappa(\psi) \in (0, 1-\Gamma)$ for all $\psi \in \mathbb{T}^2$. This means that we look for a function $\mathscr{H}$ such that

$$\mathscr{H} \circ \mathscr{S} = \mathscr{S}_0 \circ \mathscr{H}. \tag{2.23}$$

A function $\mathscr{H}$ that satisfies (2.23) is called a *conjugating function* – or simply *conjugation.*

Remark 2.14. Since $\mathscr{S}_0$ is linear in η, it is easy to see that if $\mathscr{H}$ is a solution to (2.23) 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 (2.23) and can be written as

$$\mathscr{H}(\varphi,\psi) = (\mathcal{H}(\varphi,\psi),\psi), \tag{2.24}$$

with $\partial_\varphi \mathcal{H}(0,\psi) = 1$ for every $\psi \in \mathbb{T}^2$.

In Subsection 6.3 we prove the following result on the conjugation.

Theorem 2 (**Conjugation**). *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.15) *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* (2.23) *to $\mathscr{S}_0$ given by* (2.22)*. Moreover $\mathscr{H}$ can be written as in* (2.24) *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 2.15. For Theorems 1 and 2 – and for the forthcoming Theorems 3 and 4 as well – to be valid it would be enough to assume $F \in \mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$. Higher regularity will be needed to prove Theorem 5, which in turn plays an essential role in the proof of Theorems 6 and 7.

Remark 2.16. By looking at the proof of Theorem 2 (see Subsection 6.3.3 and Appendix B.1), one may infer that if $F \in \mathscr{B}_{\alpha_0,k}(\Omega,\mathbb{R})$ for some $k \in \mathbb{N}$, then $\mathcal{H} \in \mathscr{B}_{\alpha_*,k-1}(\Omega,\mathbb{R})$. In particular, under the assumption that F is in $\mathscr{B}_{\alpha_0,6}$, the function $\mathcal{H}$ can be proved to be C^5 in the slow variable; however, in the following, we do not need more regularity of the conjugation than that stated in Theorem 2.

Remark 2.17. Theorem 2 implies that, for any initial datum $(\varphi,\psi) \in \Omega$, the evolution generated by $\mathscr{S}$ leads towards the invariant manifold $\mathcal{W}$. Therefore, the invariant manifold 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 6.4.

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 $\mathcal{L} \in \mathscr{B}_{\alpha_*,1}(\Omega_0,\mathbb{R})$.

Remark 2.19. As for W, the conjugation and its inverse are no more than Hölder continuous even if F is very smooth in ψ. From a technical point of view, the analysis might be simplified by assuming slightly stronger regularity conditions for F, such as the strong Hölder condition considered in ref. [5]. Such a condition however would force us to restrict the analysis to systems of the form (2.20) with more stringent bounds on the functions β, ν and δ then those in Remark 2.11.

The proofs of Theorem 2 and of Corollary 2.18 in Subsection 6.3 and in Subsection 6.4, respectively, show that $\mathcal{H}$ and $\mathcal{L}$ can be written as

$$\mathcal{H}(\varphi,\psi) = \varphi - W(\psi) + (\varphi - W(\psi))^2 h(\varphi,\psi), \qquad \mathcal{L}(\eta,\psi) = W(\psi) + \eta + \eta^2 l(\eta,\psi), \tag{2.25}$$

with $h \in \mathscr{B}_{\alpha_*,1}(\Omega,\mathbb{R})$ and $l \in \mathscr{B}_{\alpha_*,1}(\Omega_0,\mathbb{R})$. These representations will be useful in Subsection 2.4.4, when studying the deviations of the conjugation.

2.3.3 The physical measure

Theorems 1 and 2 imply that, given any observable $\mathcal{O} \in \mathscr{B}_{\alpha,0}(\Omega,\mathbb{R})$, one has

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

for almost every $(\varphi_0,\psi_0) \in \Omega$ with respect to the measure

$$\nu_0(\mathrm{d}\varphi\mathrm{d}\psi) := m_0(\mathrm{d}\psi)\frac{\mathrm{d}\varphi}{\phi_M - \phi_m}.$$

Hence, ν_W is the unique *physical measure* for $\mathscr{S}$ on Ω. It is thus interesting to study the mixing property of ν_W with respect to ν_0. The following result is proved in Subsection 6.6.

Theorem 3 (Mixing on the invariant manifold). *Consider the dynamical system described by the map $\mathscr{S}$ in (2.15) satisfying Hypotheses 1–3. For any observables $\mathcal{O}_1$ and $\mathcal{O}_2$ 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*

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

where C depends only on λ_*.

2.4 Averaging

We are interested in the long time evolution generated by the map $\mathscr{S}$ in (2.15), 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

$$\mathscr{S}(\varphi,\psi) = (\varphi + \rho f(\varphi,\psi), A_0\psi) \tag{2.26}$$

when ρ is small. We also define

$$\gamma := -\sup_{(\varphi,\psi)\in\Lambda} \partial_\varphi f(\varphi,\psi), \tag{2.27}$$

so that $\Gamma = \rho\,\gamma$ in (2.17).

Remark 2.20. We may and do assume, without loss of generality, that $|||f|||_{\alpha_0,6} = 1$, and hence, since $\gamma < |||f|||_{\alpha_0,6}$, that $\gamma \in (0,1)$. For any fixed function $f \in \mathscr{B}_{\alpha_0,6}(\mathbb{T}\times\mathbb{T}^2,\mathbb{R})$ such that $|||f|||_{\alpha_0,6} = 1$, both Hypotheses 1 and 3 are automatically satisfied for ρ small enough. On the other hand, Hypothesis 1 requires ρ not to be arbitrarily large because we must have

$$\rho\,\partial_\varphi f(\varphi,\psi) > -1 \qquad \forall(\varphi,\psi)\in\Omega. \tag{2.28}$$

Therefore, when considering a map $\mathscr{S}$ of the form 2.26, we tacitly assume ρ to be smaller than a a suitable value ρ_*, depending on f, such that $\mathscr{S}$ satisfies Hypotheses 1 to 3 for all $\rho \in (0,\rho_*)$.

We investigate the evolution generated by $\mathscr{S}$, as given in (2.26), when $\rho \to 0^+$. To avoid such an evolution to become trivial, for given initial conditions (φ_0,ψ_0), we study the dynamics after a linear *rescaling* of time, that is we consider $\mathscr{S}^k(\varphi_0,\psi_0)$ taking $k = \lfloor t/\rho \rfloor$ for t fixed as $\rho \to 0^+$. We refer to the case with ρ small as the *scaling regime*, even when studying the steady state of $\mathscr{S}$, where k does not explicitly appear.

2.4.1 Heuristic discussion

If ρ is small enough we can take k very large but still much smaller than ρ^{-1}. Expanding $\mathscr{S}^k$ to first order in ρ we write

$$(\mathscr{S}^k)_\varphi(\varphi,\psi) = \varphi + \rho\sum_{i=0}^{k} f(\varphi, A_0^i\psi) + o(k\rho)\,.$$

Since A_0 is strongly mixing, we further obtain that, for most values of ψ_0,

$$(\mathscr{S}^k)_\varphi(\varphi,\psi) \simeq \varphi + k\rho\,\langle f\rangle(\varphi) + o(k\rho)\,,$$

where the right hand side has lost any dependence on ψ, at least at first order in ρ. Calling $k\rho = t$, and writing $\varphi(t) := \mathscr{S}^{\lfloor t/\rho\rfloor}(\varphi,\psi)$ we can read the last expression as

$$\varphi(t) \simeq \varphi + \langle f\rangle(\varphi)\,t + o(t)\,.$$

This propounds that, for ρ small, the evolution $\varphi(t)$ that starts from a given φ and a randomly chosen ψ is essentially independent of ψ and it agrees at first order in t with the solution $\phi(t)$ of the Cauchy problem

$$\begin{cases} \dot\phi = \langle f\rangle(\phi) \\ \phi(0) = \varphi\,. \end{cases} \tag{2.29}$$

To see whether we can get a better agreement, we expand $\mathscr{S}^k$ to second order in ρ and find

$$(\mathscr{S}^k)_\varphi(\varphi,\psi) = \varphi + \rho\sum_{i=0}^{k} f(\varphi,A_0^i\psi) + \rho^2 \sum_{0\le i<j\le k} \partial_\varphi f(\varphi,A_0^j\psi) f(\varphi,A_0^i\psi) + o(k^2\rho^2)\,,$$

so that, for $\varphi(t)$ to remain close to $\phi(t)$ up to corrections $o(t^2)$, we need that

$$\sum_{0\le i<j\le k} \partial_\varphi f(\varphi, A_0^j\psi) f(\varphi_0, A_0^i\psi) \simeq \frac{k^2}{2}\partial_\varphi \langle f\rangle(\varphi)\langle f\rangle(\varphi),$$

that is we need a strong form of decay of correlations for A_0.

This suggests that, on the correct time scale, $(\mathscr{S}^k)_\varphi(\varphi,\psi)$ evolves according to a differential equation involving only the average of f. This is a simple instance of the idea of *averaging* induced by the chaotic behavior of A_0.

Clearly the argument above is only heuristic. In fact, extending the analysis outlined above to all orders is likely to get too tangled and to require very high regularity of the map. Moreover such an analysis is not suitable for dealing with the case of arbitrary t, in particular for deriving results uniform in t. Therefore, the heuristic argument only hints what to look for, but, in order to obtain something rigorous, actually we follow a different approach.

2.4.2 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 (2.26) with ρ small – for the proofs see Subsections 6.2 and 6.5.

Lemma 2.21. *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *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. Then one has $\alpha_+ = O(1)$ and $\alpha_- = O(\rho)$, and, furthermore, both $\|W\|_\infty$ and $|W|^-_{\alpha_-}$ are $O(1)$, while $|W|^+_{\alpha_+} = O(\rho)$.*

Lemma 2.22. *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. Let $\mathscr{H}$ be the conjugation* (2.24) *for the map $\mathscr{S}$ as in Theorem 2, and let h and l be defined as in* (2.25). *Then one has $\alpha_* = O(\rho)$, while both $|||h|||_{\alpha_*,1}$ and $|||l|||_{\alpha_*,1}$ are $O(1)$ in ρ.*

Remark 2.23. Lemma 2.21 shows that, when ρ is small, the manifold $\mathcal{W}$ loses most of the smoothness of f in the stable direction while maintaining it in the unstable one. Moreover the manifold varies slowly in the unstable direction.

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

Corollary 2.24. *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. For any observables $\mathcal{O}_1$ and $\mathcal{O}_2$ in $\mathscr{B}_{\alpha_*,1}$, with α_* as in Theorem 2, and for all $n \in \mathbb{N}$ one has*

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

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

2.4.3 The averaged map

As a intermediate step toward a rigorous justification of the conclusions in Subsection 2.4.1, together with the dynamical system described by (2.26) we consider also the dynamical system given by the *averaged map*

$$\overline{\mathscr{S}}(\varphi,\psi) := (\overline{G}(\varphi), A_0\psi), \qquad \overline{G}(\varphi) := \varphi + \rho\overline{f}(\varphi), \tag{2.30}$$

with

$$\overline{f}(\varphi) := \langle f\rangle(\varphi), \tag{2.31}$$

and its linearization

$$\overline{\mathscr{S}}_0(\varphi,\psi) := (\overline{\mu}\,\varphi, A_0\psi), \qquad \overline{\mu} := \partial_\varphi\overline{G}(0) = 1 + \rho\,\partial_\varphi\overline{f}(0). \tag{2.32}$$

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 the map $\overline{\mathscr{S}}$ satisfies Hypotheses 1–3 if $\mathscr{S}$ does.

Remark 2.25. Since the action of $\overline{\mathscr{S}}$ on each variable is independent of the other one, in fact we have $(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) = \overline{G}^n(\varphi)$. Notwithstanding this, the notation introduced in (2.30) helps clarify the forthcoming discussion.

Remark 2.26. Noting that $\overline{f}(0) = 0$ (see Remark 2.10), 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 $\overline{\mathcal{W}} = \{(0,\psi), \psi\in\mathbb{T}^2\}$.

In the two coming Subsections 2.4.4 and 2.4.5 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 the map $\overline{\mathscr{S}}$, we are able to compare the trajectories of the dynamics generated by $\overline{\mathscr{S}}$ with the solution of (2.29).

First, we proceed as in Theorem 2, and look for an invertible function

$$\overline{\mathscr{H}}(\varphi,\psi) = (\overline{\mathcal{H}}(\varphi),\psi), \qquad\qquad \overline{\mathcal{H}}(\varphi) = \varphi + \varphi^2\overline{h}(\varphi), \tag{2.33}$$

such that

$$\overline{\mathscr{H}}\circ\overline{\mathscr{S}} = \overline{\mathscr{S}}_0\circ\overline{\mathscr{H}}\,, \tag{2.34}$$

that is a function $\overline{\mathscr{H}}$ that conjugates $\overline{\mathscr{S}}$ to $\overline{\mathscr{S}}_0$. Analogously to (2.33), we also write

$$\overline{\mathscr{H}}^{-1}(\eta,\psi) = (\eta + \eta^2\overline{l}(\eta),\psi). \tag{2.35}$$

Then, in Subsection 7.6 we prove the following result.

Lemma 2.27. *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.15) *satisfying Hypotheses 1–3, and define the map $\overline{\mathscr{S}}$ as in* (2.30)*. 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* (2.34) *holds. Moreover, there exists a constant C such that one has $|||\overline{h}|||_{0,3} \leq C$ and $|||\overline{l}|||_{0,3} \leq C$.*

Remark 2.28. Note that $\Omega_0 \neq \mathcal{U}_0\times\mathbb{T}^2$: in fact, one has $\Omega_0 = \mathscr{H}(\Omega)$, while $\mathcal{U}_0 = \overline{\mathcal{H}}(\mathcal{U})$.

We can now introduce the flow Φ generated by (2.29), that is the set of the solutions of

$$\begin{cases} \dfrac{d}{dt}\Phi_t(\varphi) = \overline{f}(\Phi_t(\varphi)), \\ \Phi_0(\varphi) = \varphi, \end{cases} \tag{2.36}$$

when varying $\varphi\in\mathcal{U}$. Because of Hypotheses 1–3, all trajectories $\Phi_t(\varphi)$ of the system (2.36), with $\varphi\in\mathcal{U}$, move towards the origin at exponential rate as t tend to infinity.

Observe that $\Phi_{n\rho} = (\Phi_\rho)^n = \Phi_\rho\circ\ldots\circ\Phi_\rho$ and that $\Phi_\rho(\varphi) - \overline{\mathscr{S}}(\varphi)$ is $O(\rho^2)$. The following lemma, proved in Subsection 7.7, show that the trajectories generated by (2.30) and (2.36) remain close and, in fact, merge asymptotically.

Lemma 2.29. *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3, and define the map $\overline{\mathscr{S}}$ as in* (2.30) *and the flow Φ as in* (2.36)*. For any $\gamma'\in(0,\gamma)$ there exists a positive constant C such that for all $\varphi\in\mathcal{U}$ and all $n>0$ one has*

$$|(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) - \Phi_{n\rho}(\varphi)| \leq C\rho\,(1-\rho\gamma')^n. \tag{2.37}$$

Remark 2.30. With the scaling terminology introduced immediately before Subsection 2.4.1, we can rewrite (2.37) as

$$|(\overline{\mathscr{S}}^{\lfloor t/\rho\rfloor})_\varphi(\varphi,\psi) - \Phi_t(\varphi)| \leq C\rho\, e^{-\gamma' t}.$$

Observe that the decay rate in Lemma 2.29 cannot in general be equal to γ as one could naïvely expect. Indeed, the constant C in Lemma 2.29, as well as in the forthcoming Lemma 2.33 and Theorems 6 and 7, depends on γ' and may diverge as γ' tends to γ.

2.4.4 Oscillations and deviations in the scaling regime

In the next subsection we will show that also $(\mathscr{S}^n)_\varphi(\varphi,\psi)$ and $\Phi_{n\rho}(\varphi)$ remain close in the sense that the first and second moment, with respect to ψ, of their difference are $O(\rho)$ uniformly in n. In this subsection we present several preparatory results that, in our opinion, are also of interest in their own. The proofs of these results form the main technical part of the present work and are reported in Section 7. The main tools used in these proofs are the decay of correlations estimates contained in Propositions 7.6 and 7.30.

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 7.3.

Theorem 4 (**Oscillations of the invariant manifold**). *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *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. Then, there is a constant C such that*

$$|\langle W\rangle| \le C\rho, \qquad \langle W^2\rangle \le C\rho. \tag{2.38}$$

Remark 2.31. From Theorem 4 and Chebyshev inequality we obtain that, for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : |W(\psi)| > \delta\Big\}\Big) \le \frac{C\rho}{\delta^2}\,. \tag{2.39}$$

Therefore, the invariant manifold $\mathcal{W}$ for $\mathscr{S}$ converges in probability to the invariant manifold $\overline{\mathcal{W}}$ of $\overline{\mathscr{S}}$. This also implies that most trajectories of $\mathscr{S}$ starting on $\mathcal{W}$ will spend most of their time very close to $\varphi = 0$ while only rarely venturing away.

Next we compare the linearized maps $\mathscr{S}_0$ and $\overline{\mathscr{S}}_0$, defined in (2.22) and (2.32), respectively. Observe that $\overline{\mathscr{S}}_0^n(\varphi,\psi) = (\overline{\mu}^n\varphi, A_0^n\psi)$, while $\mathscr{S}_0^n(\varphi,\psi) = (\kappa^{(n)}(\psi)\varphi, A_0^n\psi)$, where

$$\kappa^{(n)}(\psi) := \prod_{i=0}^{n-1}\kappa(A_0^i\psi)\,, \tag{2.40}$$

with $\kappa(\psi)$ defined in (2.22).

Remark 2.32. 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. In particular this convention implies that $\kappa^{(0)} = 1$ in (2.40).

The following lemma, proved in Subsection 7.4, shows that the 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 (2.40) is only in $\mathfrak{B}^*_{\alpha_-,\alpha_+}(\mathbb{T}^2,\mathbb{R})$, with $\alpha_- = O(\rho)$ and $|W|^-_{\alpha_-} = O(1)$ in ρ.

Lemma 2.33. *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. Let $\kappa^{(n)}(\psi)$ and $\overline{\mu}$ be defined in* (2.40) *and in* (2.32)*, respectively. Then for any $\gamma' \in (0,\gamma)$ there is a constant C such that, for all $n \in \mathbb{N}$,*

$$\big|\langle\kappa^{(n)} - \overline{\mu}^n\rangle\big| \le C\rho\,(1-\rho\,\gamma')^n, \qquad \big|\big\langle\big(\kappa^{(n)} - \overline{\mu}^n\big)^2\big\rangle\big| \le C\rho\,(1-\rho\gamma')^{2n}. \tag{2.41}$$

Finally we want to estimate the deviation of $\mathcal{H}$ from $\overline{\mathcal{H}}$. To this end, we show that $h(\varphi,\psi)$, defined in (2.25), and $\overline{h}(\varphi,\psi)$, defined in (2.33), are close, more precisely that both $h(\varphi,\psi) - \overline{h}(\varphi,\psi)$ and its derivative are small – always in the sense that their first and second moments are of order ρ. This is ensured by the following result, whose proof is given in Subsection 7.8.

Theorem 5 (**Deviations of the conjugation**). *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. Let h and $\overline{h}$ be defined as in* (2.25) *and in* (2.33), *respectively. Then, there is a constant C such that for all $\varphi \in \mathcal{U}$*

$$\big|\big\langle h(\varphi,\cdot) - \overline{h}(\varphi)\big\rangle\big| \leq C\rho, \qquad \big\langle (h(\varphi,\cdot) - \overline{h}(\varphi))^2\big\rangle \leq C\rho, \tag{2.42a}$$

$$\big|\big\langle \partial_\varphi h(\varphi,\cdot) - \partial_\varphi\overline{h}(\varphi)\big\rangle\big| \leq C\rho, \qquad \big\langle (\partial_\varphi h(\varphi,\cdot) - \partial_\varphi\overline{h}(\varphi))^2\big\rangle \leq C\rho. \tag{2.42b}$$

Bounds analogous to (2.42) hold also for the deviations of the function l, defined in (2.25), from $\bar{l}$, defined in (2.35); we refer to Subsection 7.8 – and to Proposition 7.57 in particular – for a precise statement, which requires introducing a suitable extension of the map $\mathscr{S}$ along the lines considered in Remark 2.12.

Remark 2.34. It is in order to prove the bounds in Theorem 5 – and in the forthcoming Proposition 7.57 – that we need F to be in $\mathscr{B}_{\alpha_0,6}(\Omega,\mathbb{R})$; see also Remark 7.51 in Subsection 7.8.

2.4.5 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 (2.23) and (2.34) we get

$$\big(\mathscr{S}^n\big)_\varphi(\varphi,\psi) - \big(\overline{\mathscr{S}}^n\big)_\varphi(\varphi,\psi) = \mathcal{H}^{-1}\big(\mathscr{S}_0^n(\mathcal{H}(\varphi,\psi),\psi)\big) - \overline{\mathcal{H}}^{-1}\big(\overline{\mathscr{S}}_0^n(\overline{\mathcal{H}}(\varphi),\psi)\big). \tag{2.43}$$

Combining the estimates in Lemma 2.33 with the bounds in Theorem 5 and the analogous bounds for the inverse conjugation in Proposition 7.57, in Subsection 7.10 we prove our main result on the relation between the dynamics generated by $\mathscr{S}$ and the averaged dynamics generated by $\overline{\mathscr{S}}$.

Theorem 6 (**Convergence in square mean**). *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. Then for any $\gamma' \in (0,\gamma)$ there exists a constant C such that, for all $n \in \mathbb{N}$ and all $\varphi \in \mathcal{U}$,*

$$\Big|\Big\langle (\mathscr{S}^n)_\varphi(\varphi,\cdot) - \big((\overline{\mathscr{S}}^n)_\varphi(\varphi,\cdot) + W(A_0^n\cdot)\big)\Big\rangle\Big| \leq C\rho\,(1-\rho\,\gamma')^n, \tag{2.44a}$$

$$\Big\langle \Big((\mathscr{S}^n)_\varphi(\varphi,\cdot) - \big((\overline{\mathscr{S}}^n)_\varphi(\varphi,\cdot) + W(A_0^n\cdot)\big)\Big)^2\Big\rangle \leq C\rho\,(1-\rho\,\gamma')^{2n}. \tag{2.44b}$$

Remark 2.35. The bound (2.44b) for the second moment of the fluctuations and the Cauchy-Schwartz inequality trivially would imply a weaker bound than (2.44a). On the contrary, proving that also the first moment of the fluctuations is of order ρ requires a substantially greater amount of work. A similar comment holds for the results in Theorem 4 and, in fact, it applies 5 to Lemma 2.33 above as well.

The following result is an immediate consequence of Theorems 4 and 6, and Lemma 2.29.

Corollary 2.36. *Under the hypotheses of Theorem 6, there exists a constant C such that, for all $n \in \mathbb{N}$ and all $\varphi \in \mathcal{U}$, one has*

$$\Big|\langle(\mathscr{S}^n)_\varphi(\varphi,\cdot) - \Phi_{n\rho}(\varphi)\rangle\Big| \leq C\rho, \qquad \Big\langle\Big((\mathscr{S}^n)_\varphi(\varphi,\cdot) - \Phi_{n\rho}(\varphi)\Big)^2\Big\rangle \leq C\rho,$$

with the flow Φ defined as in (2.36).

Similarly to Remark 2.31, from Theorem 6, for fixed φ and n, we obtain that, for any $\gamma' \in (0,\gamma)$ and for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}\big|(\mathscr{S}^n)_\varphi(\varphi,\psi) - \big(\Phi_{n\rho}(\varphi) + W(A_0^n\psi)\big)\big| > \delta\Big\}\Big) \leq \frac{C\rho}{\delta^2}. \tag{2.45}$$

Note that the set of angles ψ considered in (2.45) depends on n, even though its measure is bounded independently of the value of n. The following theorem shows that, for most most values of ψ, the difference between $(\mathscr{S}^n)_\varphi(\varphi,\psi)$ and $\Phi_{n\rho}(\varphi) + W(A_0^n\psi)$ is small, exponentially in n.

Theorem 7 (**Convergence in probability, I**). *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. Let the flow Φ be defined as in* (2.36)*. Then for any $\gamma' \in (0,\gamma)$ there exists a constant C such that, for all $\varphi \in \mathcal{U}$,*

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : \sup_{n\geq 0} (1-\rho\,\gamma')^{-n} \big|(\mathscr{S}^n)_\varphi(\varphi,\psi) - \big(\Phi_{n\rho}(\varphi) + W(A_0^n\psi)\big)\big| > \delta\Big\}\Big) \leq \frac{C\rho}{\delta^3}.$$

The proof of Theorem 7 is in Subsection 8.1.

2.4.6 Aftermath: continuous time

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 (2.29) and express them in terms of the rescaled time t. To this aim, for every $\varphi \in \mathbb{T}$ and every $t \in \mathbb{R}^+$, we consider the random variable $X_t \colon \mathbb{T}^2 \to \mathbb{T}$ defined as

$$X_t(\psi) := (\mathscr{S}^{\lfloor t/\rho \rfloor})_\varphi(\varphi,\psi) + (t/\rho - \lfloor t/\rho \rfloor)\left((\mathscr{S}^{\lfloor t/\rho \rfloor+1})_\varphi(\varphi,\psi) - (\mathscr{S}^{\lfloor t/\rho \rfloor})_\varphi(\varphi,\psi)\right). \tag{2.46}$$

Observe that for every ψ, $X_t(\psi)$ is a continuous function of t so that X_t can be seen as a stochastic process with trajectories in $C^0(\mathbb{R}^+,\mathbb{T})$. Similarly we consider the process $\widetilde{X}_t$ with trajectories in $C^0(\mathbb{R}^+,\mathbb{T})$ defined as

$$\widetilde{X}_t(\psi) = X_t(\psi) - W(A_0^{\lfloor t/\rho \rfloor}\psi) - (t/\rho - \lfloor t/\rho \rfloor)\left(W(A_0^{\lfloor t/\rho \rfloor+1}\psi) - W(A_0^{\lfloor t/\rho \rfloor}\psi)\right). \tag{2.47}$$

We want to compare the stochastic processes X_t and $\widetilde{X}_t$ with the flow Φ_t defined in (2.36), seen as a stochastic process on $C^0(\mathbb{R}^+,\mathbb{T})$.

To compare $\widetilde{X}_t$ with Φ_t, for $x \in C^0(\mathbb{R}^+,\mathbb{T})$, we consider the norm

$$\|x\|_{\exp} := \sup_{t\geq 0} e^{\xi t}|x(t)|\,, \tag{2.48}$$

with a suitable $\xi \geq 0$. Combining Lemma 2.29 and Theorem 7 provides the following result (see Subsection 8.2 for the proof).

Theorem 8 (**Convergence in probability, II**). *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. Let $\widetilde{X}_t$ and Φ_t the stochastic process* (2.47) *and the flow defined in* (2.36)*, respectively. Then one has*

$$\lim_{\rho\to 0^+} \widetilde{X}_t = \Phi_t,$$

where the limit is taken in probability in the topology on $C^0(\mathbb{R}^+,\mathbb{T})$ generated by the norm (2.48) *for any $\xi \in [0,\gamma)$.*

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

Corollary 2.37. *Under the hypotheses of Theorem 8, let X_t be the stochastic process defined in* (2.46)*. Then*

$$\lim_{\rho\to 0^+} X_t = \Phi_t,$$

where the limit is taken in probability in the topology of the uniform convergence in $C^0(\mathbb{R}^+,\mathbb{T})$.

2.5 Content of the paper and strategy of the proof

The rest of the paper is mainly devoted to the proof of the results stated above, with the exception of Sections 3 to 5, where the relation with the existing literature is examined.

In Section 3 we report on previous results on similar models, including results by one of the authors where more restrictive hypotheses were assumed, and results for systems where one-dimensional expanding maps are considered instead of Anosov automorphisms, while in Section 4 we discuss a few open problems and possible extensions of our work, also in relation with the kind of problems which are mainly investigated in the literature. Next, in Section 5 we briefly review a few fields of possible application in problems of physical interest, such as the combination of slow and fast motions in describing the effects of weather on climate and the derivation of the heat equation.

The remaining, more technical Sections 6 to 8, which represent the core of the paper, are organized as follows.

Section 6 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 will be needed for dealing with the scaling regime. In particular 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 studied and proved to converge to a function which satisfies the properties stated in Theorem 2. This result then implies the existence of the inverse conjugation as well (Corollary 2.18), which satisfies similar properties, 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 the map

$$\mathscr{S}_1(\theta, \psi) = (\theta + F(\theta + W(\psi), \psi) - F(W(\psi), \psi), A_0\psi),$$

that we call the *translated map*, and the attracting invariant manifold is the flat torus $\mathbb{T}^2$.

In Section 7 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, as a preliminary step, the deviation laws for the invariant manifold (Theorem 4) and the conjugation (Theorem 5), we provide the proof of Theorem 6 on the deviations of the dynamics with respect to that of the averaged system. The analysis is based on delicate correlation inequalities, which 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 2.3. However, the invariant manifold and, hence, the map $\mathscr{S}_1$ are only α-Hölder continuous, with $\alpha = O(\rho)$, so that a naïve generalization of Proposition 2.3 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 the bounds on the variance provided by Theorem 4. Thus, in order to use the bounds on the average in (2.38) we need to study in detail the coefficients of the linear terms in the expansion in W. We achieve this by isolating the contributions which do not depend on the dynamics on the torus, that is on the chaotic variable ψ, and showing that the remaining contributions admit better dimensional bounds. In practice, 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. Then we compare the translated map to the auxiliary map through a series of technical lemmas which aim to extract and study the linear dependence on the function W; this will be treated in Subsection 7.8.4 and in Appendix D. A major issue, from a technical point of view, is that we want that the map $\mathscr{S}_2$ regularizes $\mathscr{S}_1$ and, at the same time, still satisfies Hypotheses 1 to 3: achieving both goals leads to contributions which, albeit depending linearly on 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 need to perform iterated expansions which make the analysis rather intricate: in fact, Subsection 7.8.4 and Appendix D constitute the most technical part of the paper.

Finally, in Section 8, 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 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 7 to obtain more general correlation inequalities. Appendices B, C and – as said above – mainly D contain mostly the proofs of the more technical results presented in Section 7.

3 Relation with previous works

3.1 Synchronization under stronger assumptions

In the scaling regime, dynamical systems described by (2.26) are a generalisation of the systems studied in ref. [29], i.e. continuous systems defined on $\mathbb{T} \times \mathbb{T}^2$ of the form

$$\begin{cases} \dot\varphi = 1 + \varepsilon g(\varphi, \psi, t), \\ \dot\psi = \delta(t) \log A_0 \, \varphi, \end{cases} \tag{3.1}$$

where $\delta(t)$ is the 2π-periodic delta function so defined that its integral from 0 to t equals 1 for any $t > 0$. By integrating the equations (3.1) up to time 2π and using that φ is defined mod 2π, we find the Poincaré map

$$\begin{cases} \varphi(2\pi) = \varphi(0) + \varepsilon \displaystyle\int_0^{2\pi} \mathrm{d}t \, g\Big(\varphi(0) + t + \varepsilon \int_0^t \mathrm{d}s \, g(\varphi(s), A_0\psi, s), A_0\psi, t\Big), \\ \psi(2\pi) = A_0 \, \psi(0). \end{cases} \tag{3.2}$$

If we set $(\varphi, \psi) := (\varphi(0), \psi(0))$ and $\mathscr{S}_c(\varphi, \psi) := (\varphi(2\pi), \psi(2\pi))$, we see that $\mathscr{S}_c$ is of the form (2.15), that is

$$\mathscr{S}_c(\varphi, \psi) = (\varphi + F_c(\varphi, \psi), A_0 \, \psi),$$

with $F_c(\varphi, \psi) = \varepsilon \, F_1(\varphi, \psi) + \varepsilon^2 F_2(\varphi, \psi, \varepsilon)$, where

$$\begin{aligned} F_1(\varphi, \psi) &:= \int_0^{2\pi} \mathrm{d}t \, g(\varphi + t, A_0\psi, t), \\ F_2(\varphi, \psi, \varepsilon) &:= \int_0^{2\pi} \mathrm{d}t \, \partial_\varphi g\big(\varphi^*(t), A_0\psi, t\big) \int_0^t \mathrm{d}s \, g(\varphi(s), A_0\psi, s), \end{aligned}$$

for a suitable function $\varphi^*(t)$. The existence of an invariant manifold for the system (3.1) is proved in ref. [29] under the hypothesis that there exists $\overline{\varphi} \in \mathbb{T}$ such that one has

$$F_1(\overline{\varphi}, \psi) = 0, \qquad \partial_\varphi F_1(\overline{\varphi}, \psi) < 0, \tag{3.3}$$

for all $\psi \in \mathbb{T}^2$. Both F_1 and F_2 are smooth in φ and ψ, and it is easy to see that $\mathscr{S}_c$ satisfies Hypotheses 1–3 with $\Gamma = O(\varepsilon)$. In particular the assumptions (3.3) imply that $F_c(\bar\varphi, \psi) = O(\varepsilon^2)$: this implies that the invariant manifold is such that $W_c(\psi) = O(\varepsilon)$, that is the oscillations of the invariant manifold are small not only in average. Therefore, systems of the form (2.15) extend the class of systems considered in ref. [29]; in fact Theorem 1 provides a positive answer to a question raised in ref. [29], by showing that the assumptions on the functions $\gamma_0(\varphi, \psi) := F_1(\varphi, \psi)$ and $\gamma_1(\varphi, \psi) := \partial_\varphi F_1(\varphi, \psi)$ can be weakened as suggested therein.

3.2 Expanding maps

Our work is also strongly related to the analyses in refs. [21, 22, 14]. In these works a family of one-parameter maps $\mathscr{T}$ on $\mathbb{T} \times \mathbb{T}$ is considered, given by

$$\mathscr{T}(\varphi, \omega) := (\varphi + \rho f(\varphi, \omega), h(\varphi, \omega)), \tag{3.4}$$

where $h(\varphi, \cdot)$ is an expanding circle map for every fixed φ. Systems of the form (3.4), 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 refs. [21, 22] 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 (2.15) but with the baker transformation in place of the automorphism A_0. That is, we consider the dynamical system on $\mathbb{T} \times \mathbb{T}^2$ given by

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

where $\psi = (\psi_1, \psi_2)$ and $\lfloor \cdot \rfloor$ denotes the lower integer part, and assume that $\mathscr{S}_b$ satisfies Hypotheses 1–3. If $f(\varphi, \psi)$ does not depend on ψ_2 then the system $\mathscr{S}_b$ can be seen as an extension of the system $\mathscr{T}_b(\varphi, \psi_1) := (\varphi + \rho f(\varphi, \psi), 2\psi_1 \bmod 2\pi)$ that is part of the family of systems studied in refs. [21, 22]. On the other hand, we expect that most of the results in the present paper apply with minor modifications to $\mathscr{S}_b$; that is, 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 $\mathscr{S}_b(W_b(\psi), \psi) = W_b(B(\psi))$, and $\mathcal{W}_b$ is a global attractor for $\mathscr{S}_b$ and hence $\mathscr{S}_b$ admits a unique physical measure given by

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

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 it is the unique physical measure for $\mathscr{T}_b$. In ref. [14] 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 $\nu_b(\mathrm{d}\varphi\, \mathrm{d}\psi_1) = n_b(\varphi, \psi_1)\mathrm{d}\varphi\, \mathrm{d}\psi_1$, with

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

a well-defined integrable function. We think that, in our case, this follows 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 2.24 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

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

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

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

a result analogous to Corollary 2.37 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 $\mathcal{W}_b$ would allow to control the dynamics for all positive times and hence to obtain stronger results with respect to the models considered in ref. [22] (see in particular ref. [22, Section 3.1]).

4 Extensions and generalizations

As said in the introduction, we consider this work as a first step to fix techniques and strategies to be applied to more general system and/or more refined questions. In this section we present some of these questions and briefly discuss possible strategies to follow in order to solve them by applying the results of this paper.

4.1 More general perturbations: fully coupled systems

The systems considered in this work are usually called *skew products* since the fast variable does not depend on the slow variable. In this subsection we present a path to generalize our results to the case of a fully coupled system.

4.1.1 More general Anosov diffeomorphisms

As a first step we can consider systems of the form

$$\mathscr{S}(\varphi,\psi) = (\varphi + F_\varphi(\varphi,\psi), \mathcal{A}(\psi)), \qquad \mathcal{A}(\psi) := A_0\psi + F_\psi(\psi), \tag{4.1}$$

with $F_\psi : \mathbb{T}^2 \to \mathbb{T}^2$ such that the map $\psi \mapsto \mathcal{A}(\psi)$ describes an Anosov diffeomorphism on $\mathbb{T}^2$, and $F_\varphi(\varphi,\psi) \colon \mathbb{T}\times\mathbb{T}^2 \to \mathbb{T}$ such that $\mathscr{S}$ still satisfies Hypotheses 1–3. However, any Anosov diffeomorphism of the form in (4.1) is conjugated with its linear part A_0 (see also refs. [9, 27, 44] for a more general context). Thus, there exists a Hölder continuous map $\widetilde{\mathcal{H}}_\psi \colon \mathbb{T}^2 \to \mathbb{T}^2$ such that

$$\widetilde{\mathcal{H}}_\psi \circ \mathcal{A} = A_0 \circ \widetilde{\mathcal{H}}_\psi \, .$$

If we decompose $F_\psi(\psi) = F_{\psi,+}(\psi)v_+ + F_{\psi,-}(\psi)v_-$, where v_+ and v_- are the eigenvectors corresponding to the eigenvalues $\lambda > 1$ and λ^{-1} of A_0 (see Subsection 2.1), and write $\widetilde{\mathcal{H}}_\psi(\psi) = \psi + \widetilde{h}_\psi(\psi)$, with $\widetilde{h}_\psi(\psi) = \widetilde{h}_{\psi,+}(\psi)v_+ + \widetilde{h}_{\psi,-}(\psi)v_-$, we find [28]

$$\widetilde{h}_{\psi,+}(\psi) = \sum_{n=0}^{\infty} F_{\psi,+}(\mathcal{A}^n(\psi))\lambda^{-(n+1)}, \qquad \widetilde{h}_{\psi,-}(\psi) = -\sum_{n=1}^{\infty} F_{\psi,+}(\mathcal{A}^{-n}(\psi))\lambda^{-(n+1)}, \tag{4.2}$$

so that, setting $\widetilde{\mathscr{H}}(\varphi,\psi) := (\varphi, \widetilde{\mathcal{H}}_\psi(\psi))$, we may use $\widetilde{\mathscr{H}}$ to conjugate $\mathscr{S}$ in (4.1) with

$$\widetilde{\mathscr{S}}(\varphi,\psi) = (\varphi + \widetilde{F}(\varphi,\psi), A_0\psi), \tag{4.3}$$

with $\widetilde{F}(\varphi, \widetilde{\mathcal{H}}_\psi(\psi)) = F_\varphi(\varphi,\psi)$. Clearly $\widetilde{\mathscr{S}}$ is of the form (2.15) and satisfies Hypotheses 1–3.

In this situation it is still be natural to chose the initial ψ distributed according to the Lebesgue measure m_0 on $\mathbb{T}^2$. This is essentially equivalent to considering the SRB measure $m_{\mathcal{A}}$ associated with $\mathcal{A}$, since $\mathcal{A}^* m_0$ converges exponentially fast to $m_{\mathcal{A}}$ [28]. Thus, to apply the results in Section 2 to $\mathscr{S}$ in (4.1), we need decay of correlations estimates like those in Proposition 2.3 but with $\widetilde{\mathcal{H}}^* m_{\mathcal{A}}$ in place of m_0. Observe that $\widetilde{\mathcal{H}}$ is Hölder continuous and $\widetilde{\mathcal{H}}^* m_{\mathcal{A}}$ is invariant under the action of A_0; thus $m_{\mathcal{A}}$ can be represented as a Gibbs state on the same subshift of finite type used for A_0 in Appendix A.1. We can now extend the proof Proposition 2.3 using the properties of the potential that generates such a Gibbs state as discussed in ref. [28].

4.1.2 Bidirectional perturbations

We can now look at systems with bidirectional perturbations of the form

$$\mathscr{S}(\varphi,\psi) = (\varphi + F_\varphi(\varphi,\psi), \mathcal{A}(\psi;\varphi)), \qquad \mathcal{A}(\psi;\varphi) := A_0\psi + F_\psi(\varphi,\psi)), \tag{4.4}$$

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 $\psi \mapsto \mathcal{A}(\psi;\varphi)$ is an Anosov diffeomorphism for every fixed $\varphi \in [\phi_m,\phi_M]$. Here we briefly sketch how the analysis of the present paper could be adapted to cover such a more general situation

Existence of an invariant manifold for the system (4.4) is proven in ref. [16] in the perturbative regime under the conditions (3.3). Note that, even in the simpler case considered in ref. [16], the invariant manifold must be looked for in the form $\mathcal{W} = \{(W(\psi), H(\psi)) : \psi \in \mathbb{T}^2\}$, with W and H such that $\mathscr{S}(W(\psi), H(\psi)) = (W(A_0\psi), H(A_0\psi))$, since the dynamics of the chaotic variable is no longer trivial.

More generally, we can proceed as in Subsection 4.1.1 , and look for a conjugation $\widetilde{\mathscr{H}}(\varphi,\psi) = (\varphi, \widetilde{\mathcal{H}}_\psi(\varphi,\psi))$ such that

$$\widetilde{\mathscr{H}} \circ \mathscr{S} = \widetilde{\mathscr{S}} \circ \widetilde{\mathscr{H}},$$

with $\widetilde{\mathscr{S}}$ of the form (4.3). This means that

$$\begin{aligned} &\widetilde{\mathcal{H}}_\psi(\varphi + F_\varphi(\varphi,\psi), \mathcal{A}(\psi;\varphi)) = A_0 \widetilde{\mathcal{H}}_\psi(\varphi,\psi), \\ &\widetilde{F}(\varphi, \widetilde{\mathcal{H}}_\psi(\varphi,\psi)) = F(\varphi,\psi). \end{aligned} \tag{4.5}$$

It is possible to write a formal solution to (4.5) along the lines of (4.2) (see also [10] for a similar argument). Applying the conjugation $\widetilde{\mathscr{H}}$ to the system (4.4) allows us to reduce it to a system of the form of (2.15). 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 to extend the results on synchronization to the systems as in (4.4).

Then, as the next step, one must show that the averaging principle proved for the system (2.15) implies, thanks to the existence of $\widetilde{\mathscr{H}}$, an averaging principle for (4.4). More precisely, 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 long time evolution generated by (4.4) for small ρ should be described by the flow $\widetilde{\Phi}$ defined by

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\widetilde{\Phi}_t(\varphi_0) = \widetilde{f}(\widetilde{\Phi}_t(\varphi_0)), \\ \widetilde{\Phi}_0(\varphi) = \varphi_0, \end{cases} \tag{4.6}$$

with

$$\widetilde{f}(\varphi) := \int_{\mathbb{T}^2} f(\varphi,\psi)\, m_\varphi(\mathrm{d}\psi),$$

where the new measure m_φ can be computed from $\widetilde{\mathcal{H}}_\psi$ in (4.2) and, as heuristic arguments suggest, it is expected to be the SRB measure of the Anosov diffeomorphism $\mathcal{A}(\cdot;\varphi)$.

4.1.3 Non-dissipative perturbations

In our present work, and in the generalizations discussed above, the uniform dissipation around $\varphi = 0$ of the map F plays an important role. 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 $\mathscr{S}$. Our Hypothesis 3 requires the map to be strictly contracting in Λ. In refs. [21, 22], where expanding maps are considered instead of Anosov automorphisms, a more general interaction is investigated, since the rate of contraction of the neutral variable is assumed to be non-zero only in average.

In fact, a very interesting case, from a physical point of view, is the conservative one (see also the comments in Subsection 5.3), where one assume that $\langle f(\cdot,\varphi)\rangle$ vanishes for all φ. In such a situation one expects the correct scaling to be $k = \lfloor t/\rho^2 \rfloor$ and to lead in the limit to a stochastic differential equation [17, 26], instead of an ordinary differential equation as in the dissipative case.

4.2 Central Limit Theorem

It would be also of interest to find a more detailed description of the fluctuations of the process X_t in (2.46) around the flow Φ_t in (2.36). Comparing with available results in the literature [35, 25, 26, 19, 20, 21, 22], for systems of the form (3.4), we expect the stochastic process

$$\Delta_t := \frac{1}{\sqrt{\rho}}(X_t - \Phi_t)$$

to converge in distribution, as $\rho \to 0^+$, to the solution of the stochastic differential equation

$$\begin{cases} \mathrm{d}\Delta_t = \partial_\varphi \overline{f}(\Phi_t(\varphi_0))\Delta_t \mathrm{d}t + \sigma(\Phi_t(\varphi_0))\, \mathrm{d}B(t), \\ \Delta_0 = 0, \end{cases} \tag{4.7}$$

where $B(t)$ is a standard Brownian motion and

$$(\sigma(\varphi))^2 := \left\langle \sum_{n=-\infty}^{\infty} f(\varphi, \cdot) f(\varphi, A_0^n \cdot) \right\rangle .$$

Since in the case of conservative interactions the scaling limit is expected to lead to a stochastic differential equation (see the end comments in Subsection 4.1.3), studying how an equation like (4.7) emerges from (2.15) in the scaling regime can be seen as a precursory step before dealing with the more demanding scaling regime needed to study conservative systems.

The fact that the stochastic process Δ_t converges to the solution of equation (4.7) 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

$$\lim_{\rho \to 0^+} m_0 \left(\{ \psi \in \mathbb{T}^2 : W(\psi) \leq \sqrt{\rho} z \} \right) = \frac{1}{\sqrt{2\pi\sigma(0)^2}} \int_{-\infty}^{z} \mathrm{d}y \, e^{-\frac{y^2}{2\sigma(0)^2}} . \tag{4.8}$$

Thus, to start with, as a consistency check we show how to derive (4.8). If we call, using the notation in (2.19) and setting $\mu(\psi) := 1 - \nu(\psi) = 1 + \rho \partial_\varphi f(0, \psi)$,

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

we prove in Subsection 7.3 that $\langle |W - W_0| \rangle = O(\rho)$. This 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. Consider now the new random variable

$$W_{00}(\psi) := \sum_{i=1}^{\infty} \overline{\mu}^{i-1} \beta \circ A_0^{-i} .$$

In Appendix E we present a partial extension of Proposition 7.6 to multi-times correlation functions, which yields that

$$\left\langle (W_0 - W_{00})^2 \right\rangle = O(\rho^2). \tag{4.9}$$

On the other hand, the Central Limit Theorem for Anosov system [15] 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 (4.7), by proceeding along the lines of Subsection 7.8.

4.3 Large Deviations

Finally, the discussion in Subsection 4.2 naturally leads us to consider the validity of a large deviation principle for our model. In particular (4.8) deals with the fluctuation of order $\sqrt{\rho}$ of W around 0. We can then ask if we can describe the fluctuations of W of order 1. In its simplest form we expect that, as $\rho \to 0$,

$$-\lim_{\rho\to 0} \rho \log \left(m_0 \left(\{ \psi \in \mathbb{T}^2 : W(\psi) > x \} \right) \right) = I(x),$$

where I, the *large deviation rate*, should be given by the Legendre transform of the a suitable limit moment generating function for W, that is

$$I(x) = \sup_{\tau>0} \left(\tau x - \lambda(\tau) \right),$$

with

$$\lambda(\tau) = \lim_{\rho\to 0} \rho \log \left(\langle e^{\tau W} \rangle \right).$$

Given the essentially explicit expression for W contained in Subsection 6.1, we think it should be possible to show that I exists and to compute it by using the methods and results of our paper.

It would be more interesting to study the large fluctuation of the full process X_t defined in (2.46). Again, we think that it should be possible to show that, at fixed φ_0, as $\rho \to 0$,

$$-\lim_{\rho\to 0} \rho \log \left(m_0 \left(\left\{ \psi \in \mathbb{T}^2 : \sup_{t\geq 0} e^{\xi t} |\widetilde{X}_t(\psi) - \Phi_t| > x \right\} \right) \right) = J(x),$$

where $\widetilde{X}_t$ is defined in (2.47) and $\xi \in [0, \gamma)$ is as in Theorem 8, while J is a large deviation rate to be related to a suitable moment generating function for the full dynamics.

5 Applications to physical problems

5.1 Periodic orbits: Krylov-Bogolyubov theory

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 problem 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 [37, 8, 34]. 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 [13]. Integrating out the electron coordinates, treated as fast variables, leads to an effective Hamitlonian describing the motion of the two protons. What is found numerically is that, for certain initial conditions of the electron coordinates, the protons are captured in an oscillatory state. This can be seen as a synchronization phenomenon which causes the protons to stabilize on a suitable periodic orbit. On the other hand, the numerical simulations also show that for other initial conditions the motion of the electron becomes chaotic. In this case, apparently no regular pattern emerges for the motion of the protons. It would be interesting to investigate further the chaotic regime, in the light of the increasing results in the literature showing that synchonization may still occur when the dynamics of the fast variable moves from regular to a chaotic; for instance, a behaviour of this kind is observed numerically in electromechanical systems with flexible arms (see ref. [38] and references quoted therein).

5.2 Climate models

The problem of climate change has been increasingly investigated recently, also in the light of its strong relation with society and life on our planet. Earth's climate system is undoubtedly one of the most significant examples of chaotic systems where fast and slow variables interact with each other: while weather processes, such as the atmospheric and ocean dynamics, can be considered as fast motions, what one is ultimately interested in is the slow evolution of Earth's climate [36]. Moreover, one has to take into account also intentional and unintentional human-induced perturbations, such as the global warming due to human activities.

As a consequence of the wide range of processes and 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. To attack the problem analytically, the effects of the small-scale processes are usually taken into account in the equations governing the dynamics on large scale by introducing a suitable parameterization, which may be deterministic in some cases but, more frequently, leads to stochastic differential equations. However, also analytically more accessible models have been studied, both because there are problems which admit a simpler description and because 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. For instance, one can consider 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 timescales, with the deep ocean temperature reacting much more slowly [51].

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 for instance refs. [24, 41, 31] 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 a mathematical abstraction, is justified in consideration of Gallavotti-Cohen chaotic hypothesis [30]. In this regard, we stress that the results obtained by relying on Ruelle response theory [42, 31] exploited the very same assumption.

5.3 Several coupled systems: the heat equation

As we already mentioned, a well established line of research aim to a derivation of the macroscopic law of transport of energy in a crystal, that is the heat equation, starting from the deterministic microscopic dynamics. In this spirit one considers a large number N of microscopic systems – which can be taken equal to each other – organized on a lattice in $\mathbb{R}^3$. Without interaction each microscopic system presents a neutral direction that represents the fact that energy is locally conserved. After a small interaction that couples the neutral directions is introduced, one expect to see the heat equation to emerge as an effective macroscopic equation in the *hydrodynamic limit*, that is the limit in which both the number N of local systems and the (discrete) time k go to infinity in such a way that $k \sim N^2$.

An interesting result in this direction is obtained in ref. [11], where the local systems are assumed to be chaotic maps coupled with a neutral variable which play the role of a local energy, 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 refs. [21, 40], 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 (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.

At the moment, we are not aware of substantial progress in respect to the second step. As far as the first step is concerned, one wishes the dynamics to be well understood in the absence of interaction. On the other hand, integrable systems have to be excluded because they are too special and are expected to display a non-typical behaviour. For these reasons, the local systems are usually assumed to be chaotic, as in [11] – for instance expanding maps or, as in our paper, Anosov maps. A further simplifying hypothesis is to consider an interaction which makes non-conservative the evolution of the neutral variable, so that local attractors appear: this is required in order to control the dynamics 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 (see also Subsection 4.2 for more comments on this point).

The present paper deals with the first step of the strategy outlined above, in the strictly dissipative case. The next step would be considering a finite region $\Lambda \subset \mathbb{Z}^d$ and a set of variable $(\underline{\varphi}, \underline{\psi}) \in \mathbb{T}^\Lambda \times (\mathbb{T}^2)^\Lambda$, such that, for $i \in \Lambda$, the dynamics is given by

$$\mathscr{S}_i(\underline{\varphi}, \underline{\psi}) = (\varphi_i + \rho f_i(\underline{\varphi}, \underline{\psi}), A_0 \psi_i),$$

with $f_i(\underline{\varphi}, \underline{\psi})$ depending only on the variables φ_j and ψ_j with $j \in \Lambda$ close to i, for instance the first neighbours. The results of the paper should extend to the case of a finite number of systems, while extending the analysis to an arbitrarily large region Λ requires substantial additional work in order to obtain bounds uniform in the size of the region. Of course, because of the dissipation, one does not expect to obtain the heat equation when the hydrodynamic limit is taken; nevertheless, studying the limit of infinitely many systems in a simpler case could shed light on the more realistic models.

As the last comment suggests, another non-trivial extension would be removing the dissipation hypothesis on the dynamics of the neutral variable. However, as stressed in ref. [40] this is a much harder problem with respect to the dissipative case, already in the case of a single system.

6 Mapping to a simpler model

In this section we prove Theorems 1 and 2 by explicitly solving (2.21) and (2.23). The first theorem is obtained by relying on Banach Fixed-Point Theorem, while the second one exploits the dynamics being uniformly contracting around the invariant manifold. Finally, the two results together are showed to yield immediately Theorem 3.

6.1 The invariant manifold: proof of Theorem 1

From (2.21) we get

$$W(A_0\psi) = G(W(\psi), \psi) = W(\psi) + F(W(\psi), \psi) \,. \tag{6.1}$$

To show that a solution of (6.1) exists we define the map

$$\mathscr{E}[W](\psi) := G(W(A_0^{-1}\psi), A_0^{-1}\psi) \tag{6.2}$$

so that the invariant manifold is the solution of the fixed point equation $\mathscr{E}[W] = W$.

If we define

$$B := \{W \in \mathfrak{B}(\mathbb{T}^2, \mathbb{R}) : S_m \le W \le S_M\}, \tag{6.3}$$

then $\mathscr{E}(B) \subset B$ by Hypotheses 1 and 2. Moreover

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

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 (6.1).

To discuss the regularity of W we observe that, since F is α_0-Hölder continuous, we get, for $\alpha_+ \in (0, \alpha_0]$,

$$\begin{aligned}|\mathscr{E}[W]|^+_{\alpha_+} &= \lambda^{-\alpha_+} \sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+} \left|G(W(\psi + xv_+), \psi + xv_+) - G(W(\psi), \psi)\right| \\ &\le \lambda^{-\alpha_+} \sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+} \Big(|G(W(\psi + xv_+), \psi + xv_+) - G(W(\psi), \psi + xv_+)| \\ &\qquad + |G(W(\psi), \psi + xv_+) - G(W(\psi), \psi)|\Big) \\ &\le \lambda^{-\alpha_+} \Big((1-\Gamma)|W|^+_{\alpha_+} + |F|_{\alpha_+}\Big),\end{aligned} \tag{6.4}$$

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

$$|\mathscr{E}[W]|^-_{\alpha_-} \le \lambda^{\alpha_-} \Big((1-\Gamma)|W|^-_{\alpha_-} + |F|_{\alpha_-}\Big) . \tag{6.5}$$

This implies that the set

$$B_{\alpha_-,\alpha_+} := \left\{W \in \mathfrak{B}_{\alpha_-,\alpha_+}(\mathbb{T}^2, \mathbb{R}) : |W|^+_{\alpha_+} \le \frac{|F|_{\alpha_+}}{\lambda^{\alpha_+} - (1-\Gamma)} \ , \quad |W|^-_{\alpha_-} \le \frac{|F|_{\alpha_-}}{\lambda^{-\alpha_-} - (1-\Gamma)}\right\} \tag{6.6}$$

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

$$\begin{aligned}|\mathscr{E}[W_1] - \mathscr{E}[W_2]|^+_{\alpha_+} &= \lambda^{-\alpha_+} \sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+} \Big|G(W_1(\psi), \psi) - G(W_2(\psi), \psi) \\ &\qquad - G(W_1(\psi + xv_+), \psi + xv_+) + G(W_2(\psi + xv_+), \psi + xv_+)\Big| \\ &\le \lambda^{-\alpha_+} \sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+} \Big|G(W_1(\psi), \psi) - G(W_2(\psi), \psi) \\ &\qquad - G(W_1(\psi + xv_+), \psi) + G(W_2(\psi + xv_+), \psi)\Big| \\ &+ \lambda^{-\alpha_+} \sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}} |x|^{-\alpha_+} \Big|G(W_1(\psi + xv_+), \psi) - G(W_2(\psi + xv_+), \psi) \\ &\qquad - G(W_1(\psi + xv_+), \psi + xv_+) + G(W_2(\psi + xv_+), \psi + xv_+)\Big| .\end{aligned} \tag{6.7}$$

Writing

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

we bound, in the contribution from the fifth and sixth lines in (6.7),

$$\begin{aligned}\sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}}|x|^{-\alpha_+}\Big|&G(W_1(\psi+xv_+),\psi)-G(W_2(\psi+xv_+),\psi)\\ &-G(W_1(\psi+xv_+),\psi+xv_+)+G(W_2(\psi+xv_+),\psi+xv_+)\Big|\le|\partial_\varphi G|_{\alpha_+}\|W_1-W_2\|_\infty.\end{aligned}\tag{6.8}$$

For any C^2 function h we have

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

so that we can bound, in the contribution from the third and fourth lines of (6.7),

$$\begin{aligned}\sup_{\psi\in\mathbb{T}^2}\sup_{x\in\mathbb{R}}|x|^{-\alpha_+}\Big|&G(W_1(\psi),\psi)-G(W_2(\psi),\psi)-G(W_1(\psi+xv_+),\psi)\\ &+G(W_2(\psi+xv_+),\psi)\Big|\le(1-\Gamma)|W_1-W_2|^+_{\alpha_+}+\|\partial^2_\varphi G\|_\infty\max_{i=1,2}|W_i|^+_{\alpha_+}\|W_1-W_2\|_\infty.\end{aligned}\tag{6.9}$$

Collecting the bounds (6.8) and (6.9), we obtain

$$\begin{aligned}&|\mathscr{E}[W_1]-\mathscr{E}[W_2]|^+_{\alpha_+}\\ &\quad\le\lambda^{-\alpha_+}\Big((1-\Gamma)|W_1-W_2|^+_{\alpha_+}+\big(\|\partial^2_\varphi F\|_\infty\max_{i=1,2}|W_i|^+_{\alpha_+}+|\partial_\varphi F|^+_{\alpha_+}\big)\|W_1-W_2\|_\infty\Big)\end{aligned}\tag{6.10}$$

and, analogously,

$$\begin{aligned}&|\mathscr{E}[W_1]-\mathscr{E}[W_2]|^-_{\alpha_-}\\ &\quad\le\lambda^{\alpha_-}\Big((1-\Gamma)|W_1-W_2|^-_{\alpha_-}+\big(\|\partial^2_\varphi F\|_\infty\max_{i}|W_i|^-_{\alpha_-}+|\partial_\varphi F|^-_{\alpha_-}\big)\|W_1-W_2\|_\infty\Big).\end{aligned}\tag{6.11}$$

Let now $\alpha_-,\alpha_+\in(0,\alpha_0]$ be such that

$$\lambda^{\alpha_-}(1-\Gamma)<1\,,\tag{6.12a}$$

$$\alpha_+\lambda^{-\alpha_+}\left(\frac{\|\partial^2_\varphi F\|_\infty|F|_{\alpha_+}}{\lambda^{\alpha_+}-(1-\Gamma)}+|\partial_\varphi F|^+_{\alpha_+}\right)+\alpha_-\lambda^{\alpha_-}\left(\frac{\lambda^{\alpha_-}\|\partial^2_\varphi F\|_\infty|F|_{\alpha_-}}{1-\lambda^{\alpha_-}(1-\Gamma)}+|\partial_\varphi F|^-_{\alpha_-}\right)<\Gamma\,.\tag{6.12b}$$

For such α_- and α_+, combining the bounds (6.4), (6.5), (6.10) and (6.11), the map $\mathscr{E}$ turns out to be a contraction on B_{α_-,α_+}. This concludes the proof of Theorem 1.

Remark 6.1. Recalling that $S_m<0<S_M$, by Remark 2.10, the discussion above implies that the sequence $\mathscr{E}^n[0]$ converges to W in $\mathfrak{B}(\mathbb{T}^2,\mathbb{R})$.

Remark 6.2. The inequalities (6.12) are satisfied for α_-,α_+ small enough. If ρ is small and $|||F|||_{\alpha_0,2}=O(\rho)$, then (6.12a) requires $\alpha_-=O(\rho)$, which inserted into (6.12b) shows that α_+ is allowed to be $O(1)$ in ρ.

Remark 6.3. Assuming the map $\mathscr{S}$ to satisfy much stronger regularity properties in the fast variable (such as smoothness or even analyticity) does not really improve the regularity of the manifold. At best, we might obtain a regularity somewhat stronger than the Hölder continuity, such as the strong Hölder condition considered in ref. [5], but only at the price of assuming stronger conditions on the map, in particular on the variation of the function μ – see also Remark 2.19.

6.2 The invariant manifold in the scaling regime: proof of Lemma 2.21

In the scaling regime, where F, together with all its derivatives, is proportional to ρ, the best we can say about the invariant manifold is that $W \in B$, with B as in (6.3), and hence $|W| \le \max\{|S_m|, S_M\}$, so that we still have $\|W\|_\infty = O(1)$. Furthermore, since $W \in B_{\alpha_-,\alpha_+}$, with B_{α_-,α_+} as in (6.6), so that $\alpha_+ = O(1)$ and $\alpha_- = O(\rho)$ by Remark 6.2, we obtain $|W|^-_{\alpha_-} = O(1)$, while $|W|^+_{\alpha_+} = O(\rho)$.

Remark 6.4. If ψ_0 is a fixed point of A_0, that is $A_0\psi_0 = \psi_0$, 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 $\lim_{\rho\to 0^+} W(\psi) \ne 0$.

6.3 The conjugation: proof of Theorem 2

6.3.1 The translated map

We will construct the conjugation $\mathscr{H}$ by first subtracting the steady state and then linearizing the resulting dynamics around 0. More precisely, we write $\varphi = \theta + W(\psi)$ so that, in terms of the variables (θ, ψ), the dynamics is described by the map

$$\mathscr{S}_1(\theta,\psi) = ((\mathscr{S}_1)_\theta(\theta,\psi), (\mathscr{S}_1)_\psi(\theta,\psi)) := (G_1(\theta,\psi), A_0\psi), \qquad G_1(\theta,\psi) := \theta + F_1(\theta,\psi), \tag{6.13}$$

with

$$F_1(\theta,\psi) := F(\theta + W(\psi),\psi) - F(W(\psi),\psi) = F(\theta + W(\psi),\psi) + W(\psi) - W(A_0\psi). \tag{6.14}$$

We call $\mathscr{S}_1$ the *translated map*. Defining

$$\Omega_1 := \{(\theta,\psi) \in \mathbb{T}\times\mathbb{T}^2 \,:\, \phi_m - W(\psi) \le \theta \le \phi_M - W(\psi)\}, \tag{6.15}$$

we have that $\mathscr{S}_1$ is injective from Ω_1 into itself.

Remark 6.5. 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 2.26).

Remark 6.6. The iterations of the maps $\mathscr{S}$ and $\mathscr{S}_1$ are such that

$$(\mathscr{S}^n)_\varphi(\varphi,\psi) = (\mathscr{S}_1^n)_\theta(\varphi - W(\psi),\psi) + W(A_0^n\psi), \tag{6.16}$$

while $(\mathscr{S}^n)_\psi(\varphi,\psi) = (\mathscr{S}_1^n)_\psi(\varphi - W(\psi),\psi) = A_0^n\psi$. Conversely we have

$$(\mathscr{S}_1^n)_\theta(\theta,\psi) = (\mathscr{S}^n)_\varphi(\theta + W(\psi),\psi) - W(A_0^n\psi).$$

In Subsections 6.3.2 and 6.3.3 we will study the conjugation relation

$$\mathscr{H}_1 \circ \mathscr{S}_1 = \mathscr{S}_0 \circ \mathscr{H}_1. \tag{6.17}$$

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

$$\mathscr{H}(\varphi,\psi) = (\mathcal{H}_1(\varphi - W(\psi),\psi),\psi), \tag{6.18}$$

with $\mathscr{H}$ satisfying the conjugation relation (2.23). 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, $\mathscr{H}$ and $\mathscr{H}_1$ have the same image Ω_0 defined in Theorem 2. In analogy with (6.18), we write $\mathscr{H}_1^{-1}(\eta,\psi) = (\mathcal{L}_1(\eta,\psi),\psi)$.

Considering also Remark 2.14, we look for functions $\mathcal{H}_1 : \Omega_1 \to \mathbb{R}$ and $\mathcal{L}_1 : \Omega_0 \to \mathbb{R}$ of the form

$$\mathcal{H}_1(\theta,\psi) = \theta + \theta^2 h_1(\theta,\psi), \qquad \mathcal{L}_1(\eta,\psi) = \eta + \eta^2 l_1(\eta,\psi). \tag{6.19}$$

If we show that functions $\mathcal{H}_1$ and $\mathcal{L}_1$ of the form (6.19) 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 (2.25), with

$$h(\varphi, \psi) = h_1(\varphi - W(\psi), \psi), \qquad l(\eta, \psi) = l_1(\eta, \psi). \tag{6.20}$$

Remark 6.7. 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 $\{(\varphi, \psi) \in \mathbb{T} \times \mathbb{T}^2 : \varphi - W(\psi) \in \Theta\} \subset \Omega$.

6.3.2 Dynamics toward the steady state

From (6.17) we see that $\mathcal{H}_1$ satisfies the equation

$$\kappa(\psi)\mathcal{H}_1(\theta, \psi) = \mathcal{H}_1(\mathscr{S}_1(\theta, \psi)) . \tag{6.21}$$

According to (6.19), we get

$$h_1(\theta, \psi) = \frac{1}{\kappa(\psi)\theta^2}\Big(G_1(\theta, \psi) + G_1(\theta, \psi)^2 h_1(\mathscr{S}_1(\theta, \psi)) - \kappa(\psi)\,\theta\Big),$$

with $G_1(\theta, \psi)$ as in (6.13), so that, setting

$$p_1(\theta, \psi) := \frac{(G_1(\theta, \psi))^2}{\theta^2\kappa(\psi)} = \frac{1}{\kappa(\psi)}\left(1 + \int_0^1 \mathrm{d}t\, \partial_\varphi F(t\theta + W(\psi), \psi)\right)^2, \tag{6.22a}$$

$$q_1(\theta, \psi) := \frac{G_1(\theta, \psi) - \theta\kappa(\psi)}{\theta^2\kappa(\psi)} = \frac{1}{\kappa(\psi)}\int_0^1 \mathrm{d}t\,(1-t)\,\partial_\varphi^2 F(t\theta + W(\psi), \psi)\,, \tag{6.22b}$$

we obtain

$$h_1(\theta, \psi) = q_1(\theta, \psi) + p_1(\theta, \psi)\, h_1(\mathscr{S}_1(\theta, \psi))\,, \tag{6.23}$$

whose solution can be formally written as

$$h_1(\theta, \psi) = \sum_{n=1}^{\infty} p_1^{(n)}(\theta, \psi)\, q_1(\mathscr{S}_1^n(\theta, \psi)), \tag{6.24}$$

with

$$p_1^{(n)}(\theta, \psi) := \prod_{i=0}^{n-1} p_1(\mathscr{S}_1^i(\theta, \psi)) \tag{6.25}$$

where, according to our conventions (see Remark 2.32), $p_1^{(0)}(\theta, \psi) = 1$.

Remark 6.8. Both $p_1(\theta, \psi)$ and $q_1(\theta, \psi)$ are well defined at $\theta = 0$: using (6.1) and the definition of $\kappa(\psi)$ after (2.22), we find $\kappa(\psi)\, q_1(0, \psi) = \partial_\varphi^2 F(0, \psi)$ and $p_1(0, \psi) = \kappa(\psi)$.

We start with the regularity properties of the functions p_1 and q_1. The following result is an easy consequence of the representation in (6.22) together with Lemma 2.9.

Lemma 6.9. *Assume the map $\mathscr{S}$ in (2.15) to satisfy Hypotheses 1–3. For any $\Gamma' \in (0, \Gamma)$ and any $\alpha \in (0, \min\{\alpha_-, \alpha_+\})$, with α_- and α_+ as in Theorem 1, 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*

$$\theta_1 := \frac{\Gamma - \Gamma'}{2\|\kappa^{-1}\|_\infty(1 + \|\partial_\varphi F\|_\infty)\|\partial_\varphi^2 F\|_\infty}. \tag{6.26}$$

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 (6.22a) and fix θ_1 as in (6.26), 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. ☐

Remark 6.10. In the scaling regime, where $\Gamma = \rho\,\gamma$, one has $|||p_1|||_{\alpha,5} = 1 + O(\rho)$, $|||q_1|||_{\alpha,5} = O(\rho)$, while $|||\partial_\theta p_1|||_{0,4} = O(\rho)$ and $|p_1|_\alpha = O(\rho)$. Moreover, if one takes $\Gamma' = \rho\,\gamma'$, with $\gamma' \in (0,\gamma)$, then (6.26) in Lemma 6.9 implies that $\theta_1 = O(1)$.

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

Lemma 6.11. *Assume the map $\mathscr{S}$ in (2.15) to satisfy Hypotheses 1–3. For any $\Gamma' \in (0,\Gamma)$ and $\alpha \in (0, \min\{\alpha_-,\alpha_+\})$, with α_- and α_+ as in Theorem 1, and for all $n \geq 0$, one has $(\mathscr{S}_1^n)_\theta \in \mathscr{B}_{\alpha,2}(\Omega_1,\mathbb{R})$ and*

$$|||\partial_\theta(\mathscr{S}_1^n)_\theta|||_{0,1} \leq C_0\,(1-\Gamma')^n\,, \qquad |(\mathscr{S}_1^n)_\theta|_\alpha \leq C_0\lambda^{\alpha n}, \qquad |\partial_\theta(\mathscr{S}_1^n)_\theta|_\alpha \leq C_0\lambda^{\alpha n},$$

with the constant C_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 2.9, and observe that $N_r = O((\Gamma')^{-1})$ in such a case. We have $\|\partial_\theta G_1\|_\infty \leq 1 + \|\partial_\varphi F\|_\infty$, while for $n \geq N_r$ we may bound

$$\|(\partial_\theta G_1)\circ\mathscr{S}_1^n\|_\infty \leq \|(1+\partial_\varphi F)\circ\mathscr{S}^n\|_\infty \leq (1-\Gamma').$$

Thus, noting that $(\mathscr{S}_1^n)_\theta = G_1\circ\mathscr{S}_1^{n-1}$, we get

$$\|\partial_\theta(\mathscr{S}_1^n)_\theta\|_\infty = \|\partial_\theta(G_1\circ\mathscr{S}_1^{n-1})\|_\infty \leq \prod_{i=0}^{n-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^i\|_\infty \leq C_1\,(1-\Gamma')^n\,, \tag{6.27}$$

with

$$C_1 := \left(\frac{1+\|\partial_\varphi F\|_\infty}{1-\Gamma'}\right)^{N_r}. \tag{6.28}$$

Moreover, for $n \geq 2$, by (2.7) and (2.12), we have

$$|G_1\circ\mathscr{S}_1^{n-1}|_\alpha \leq \|(\partial_\theta G_1)\circ\mathscr{S}_1^{n-1}\|_\infty |G_1\circ\mathscr{S}_1^{n-2}|_\alpha + \lambda^{\alpha(n-1)}|G_1|_\alpha, \tag{6.29}$$

where $|G_1|_\alpha \leq \|\partial_\varphi F\|_\infty |W|_\alpha + |F|_\alpha$, and hence, iterating, we get

$$\begin{aligned} |(\mathscr{S}_1^n)_\theta|_\alpha = |G_1\circ\mathscr{S}_1^{n-1}|_\alpha &\leq |G_1|_\alpha \sum_{i=1}^{n} \lambda^{\alpha(n-i)} \prod_{j=1}^{i-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^{n-j}\|_\infty \\ &\leq C_1|G_1|_\alpha\,\lambda^{\alpha n} \sum_{i=1}^{n} (1-\Gamma')^{i-1}\,\lambda^{-\alpha i} \leq C_2\lambda^{\alpha n}, \end{aligned} \tag{6.30}$$

with C_1 as in (6.28) and

$$C_2 := C_1|G_1|_\alpha\,(\Gamma')^{-1}. \tag{6.31}$$

We have also

$$\|\partial_\theta^2(\mathscr{S}_1^n)_\theta\|_\infty \leq \sum_{i=0}^{n-1} \left\|(\partial_\theta^2 G_1)\circ\mathscr{S}_1^i\right\|_\infty \left\|\partial_\theta(\mathscr{S}_1^i)_\theta\right\|_\infty \prod_{\substack{j=0\\ j\neq i}}^{n-1} \|(\partial_\theta G_1)\circ\mathscr{S}_1^j\|_\infty,$$

so that, by exploiting (6.27), we find that

$$\|\partial_\theta^2(\mathscr{S}_1^n)_\theta\|_\infty \leq C_3\,(1-\Gamma')^n\,,$$

with

$$C_3 := C_1\|\partial_\varphi^2 F\|_\infty\,(\Gamma')^{-1}. \tag{6.32}$$

Finally, noting that

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

and proceeding as done to get the bound (6.30), we obtain

$$|\partial_\theta(\mathscr{S}_1^n)_\theta|_\alpha \leq C_4\lambda^{\alpha n},$$

with

$$C_4 := \left(C_2\|\partial_\varphi^2 F\|_\infty + |\partial_\theta G_1|_\alpha\right)(\Gamma')^{-1}. \tag{6.33}$$

Then the bounds follow with $C_0 = \max\{C_1, C_2, C_3, C_4\}$. □

Remark 6.12. In the scaling regime, where $\Gamma = \rho\,\gamma$, if we choose $\Gamma' = \rho\gamma'$, with $\gamma' \in (0,\gamma)$, we find that the constant C_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, in (6.28), one has $N_r = O(\rho^{-1})$ by (2.18).

Remark 6.13. Using $(\mathscr{S}_1^n)_\theta(0,\psi) = 0$ allows us to bound also $|(\mathscr{S}_1^n)_\theta(\theta,\psi)| \leq \|\partial_\theta(\mathscr{S}_1^n)_\theta\|_\infty|\theta|$ and hence

$$\|(\mathscr{S}_1^n)_\theta\|_\infty \leq C_1\Big(\max\{|\phi_m|,\phi_M\} + \max\{|S_m|, S_M\}\Big)(1-\Gamma')^n.$$

6.3.3 Existence and regularity of the conjugation

We now have all the ingredients to complete the proof of Theorem 2. Fix $\Gamma' \in (0,\Gamma)$ and set $r = r(\Gamma')$, with the notation of Lemma 2.9. Fix θ_1 be as in Lemma 6.9, and set

$$S_r := \max\{S_M, |S_m|\} + r, \qquad M_r := \max\{2\,(\Gamma')^{-1}\log(S_r/\theta_1), 0\}.$$

If $(\theta + W(\psi),\psi) \in \Lambda_r$ and $n > M_r$, then $|(\mathscr{S}_1^n(\theta,\psi))_\theta| \leq \theta_1$. Combining the latter with property 3 in Lemma 2.9, for every $(\theta,\psi) \in \Omega_1$ we find that $|(\mathscr{S}_1^n(\theta,\psi))_\theta| \leq \theta_1$ for $n \geq M_r' := M_r + N_r$. This implies that, for $n \geq M_r'$,

$$\big\|p_1^{(n)}\big\|_\infty \leq \|p_1\|_\infty^{M_r'}\,(1-\Gamma')^{n-M_r'}\,, \tag{6.34}$$

with $p_1^{(n)}$ as in (6.25), and hence, using the bound (6.34) in (6.24), we obtain

$$\|h_1\|_\infty \leq D_1\sum_{n=0}^{\infty}(1-\Gamma')^n\,\|q_1\|_\infty \leq D_1(\Gamma')^{-1}\|q_1\|_\infty, \qquad D_1 := \left(\frac{\|p_1\|_\infty}{1-\Gamma'}\right)^{M_r'}. \tag{6.35}$$

Furthermore, using Lemma 6.11, we see that

$$\begin{aligned}\big\|\partial_\theta p_1^{(n)}\big\|_\infty &\leq \sum_{i=0}^{n-1}\big\|\partial_\theta\big(p_1\circ\mathscr{S}_1^i\big)\big\|_\infty\prod_{\substack{j=0\\ j\neq i}}^{n-1}\big\|p_1\circ\mathscr{S}_1^j\big\|_\infty \\ &\leq C_0D_1\|\partial_\theta p_1\|_\infty\,(1-\Gamma')^{n-1}\sum_{i=0}^{n-1}(1-\Gamma')^i \leq C_0D_1(\Gamma')^{-1}\|\partial_\theta p_1\|_\infty\,(1-\Gamma')^n\;,\end{aligned} \tag{6.36}$$

with D_1 as in (6.35), so that, summing over n, we get

$$\|\partial_\theta h_1\|_\infty \leq \sum_{n=1}^{\infty}\Big(\big\|\partial_\theta p_1^{(n)}\big\|_\infty\|q_1\|_\infty + \big\|p_1^{(n)}\big\|_\infty\|\partial_\theta(q_1\circ\mathscr{S}_1^n)\|_\infty\Big) \leq D_2\,(\Gamma')^{-1}|||q|||_{0,1}\,, \tag{6.37}$$

for a suitable constant D_2. By studying along the same lines the second derivative $\partial_\theta^2 h_1$, we obtain (see Appendix B.1 for details)

$$\|\partial_\theta^2 h_1\|_\infty \le D_3\,(\Gamma')^{-1}|||q|||_{0,2}\,, \tag{6.38}$$

for some other constant D_3.

Finally, for any $\alpha \in (0, \min\{\alpha_-, \alpha_+\}]$, we get, again relying on Lemma 6.11,

$$\begin{aligned} \left|p_1^{(n)}\right|_\alpha &\le \sum_{i=0}^{n-1} \left|p_1 \circ \mathscr{S}_1^i\right|_\alpha \prod_{\substack{j=0\\ j\neq i}}^{n-1} \left\|p_1 \circ \mathscr{S}_1^j\right\|_\infty \\ &\le C_0 D_1 (|p_1|_\alpha + \|\partial_\theta p_1\|_\infty)\,(1-\Gamma')^{n-1}\,\frac{\lambda^{\alpha n}-1}{\lambda^\alpha - 1} \le D_4\,\lambda^{\alpha n}\,(1-\Gamma')^n\ , \end{aligned} \tag{6.39}$$

for a suitable constant D_4 proportional to D_1. To sum over n we take $\alpha = \alpha_*$ in Lemma 6.11, with α_* such that $\lambda^{\alpha_*}(1-\Gamma') < 1 - \Gamma''$, with $\Gamma'' < \Gamma'$, so as to obtain

$$|h_1|_{\alpha_*} \le \sum_{n=1}^{\infty} \left(\left|p_1^{(n)}\right|_{\alpha_*} \|q\|_\infty + \left\|p_1^{(n)}\right\|_\infty |q_1 \circ \mathscr{S}_1^n|_{\alpha_*} \right) \le D_4\,(\Gamma'')^{-1}|||q_1|||_{\alpha_*,1}, \tag{6.40}$$

for a suitable constant D_4 proportional to D_1. Once more we bound $|\partial_\theta h_1|_{\alpha_*}$ by reasoning in a similar way (again we refer to Appendix B.1 for details) and find

$$|\partial_\theta h_1|_{\alpha_*} \le D_5\,(\Gamma'')^{-1}|||q_1|||_{\alpha_*,1}, \tag{6.41}$$

with D_5 proportional to D_2.

By collecting together the bounds (6.35), (6.37), (6.38), (6.40) and (6.41), we find that $\|h_1\|_{\alpha_*,2}$ is bounded. Therefore, by recalling the first relation in (6.20), Theorem 2 is proved, with $\Omega_0 = \mathscr{H}_1(\Omega_1) = \mathscr{H}(\Omega)$.

Remark 6.14. The argument above show that, essentially, it is enough to prove the existence of the conjugation inside Λ. Indeed, once the conjugation has been defined in Λ, it can be easily extended to the whole Ω by using the fact that all trajectories fall inside a neighborhood of the attracting invariant manifold in a finite time.

6.4 The inverse conjugation: proof of Corollary 2.18

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 $|\theta| < \theta_*$. For every $(\theta,\psi) \in \Omega_1$ and every $M \in \mathbb{N}$, we have

$$\mathcal{H}_1(\theta,\psi) = \frac{\mathcal{H}_1(\mathscr{S}_1^M(\theta,\psi))}{\kappa^{(M)}(\psi)}\,.$$

Reasoning like in the derivation of (6.34) we see that there exists M_2 such that $|(\mathscr{S}_1^{M_2}(\theta,\psi))_\theta| \le \theta_*$. Thus, using Lemma 2.9 we get, for all $(\theta,\psi) \in \Omega_1$,

$$\partial_\theta \mathcal{H}_1(\theta,\psi) \ge \|\kappa\|_\infty^{-M_2} \inf_{|\theta|\le\theta^*} \partial_\theta \mathcal{H}_1(\theta,\psi) \left(\inf_{(\theta,\psi)\in\Omega_1} \partial_\theta G_1(\theta,\psi) \right)^{M_2} =: \tau_1 > 0\,, \tag{6.42}$$

where we used Hypothesis 1. From the Inverse Function Theorem it follows easily that there exist $l_1 \in \mathscr{B}_{0,1}(\mathscr{H}_1(\Omega_1, \mathbb{R}))$ such that

$$(\mathscr{H}_1^{-1})_\eta(\eta,\psi) = \eta + \eta^2 l_1(\eta,\psi)\,.$$

Finally, 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

$$\frac{|\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi'),\psi') - \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi')|}{|\psi'-\psi|^{\alpha_*}} = \frac{|\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi') - \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)|}{|\psi'-\psi|^{\alpha_*}},$$

which gives

$$\left(\inf_{(\theta,\psi)\in\Omega_1} \partial_\theta \mathcal{H}_1(\theta,\psi)\right) |\mathcal{H}_1^{-1}|_{\alpha_*} \le |\mathcal{H}_1|_{\alpha_*},$$

and hence $|\mathcal{H}_1^{-1}|_{\alpha_*} \le \tau_1^{-1}|\mathcal{H}_1|_{\alpha_*}$. Therefore, the fact that $h_1 \in \mathscr{B}_{\alpha_*,1}(\Omega_1,\mathbb{R})$ implies also that $l_1 \in \mathscr{B}_{\alpha_*,1}(\mathscr{H}_1(\Omega_1),\mathbb{R})$, and, by the second relation of (6.20), Corollorary 2.18 follows.

Remark 6.15. By reasoning as in the derivation of (6.42), 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 7.9).

6.5 The conjugation in the scaling regime: proof of Lemma 2.22

Recall that, with the notation of Lemma 2.9, for $r = r(\Gamma')$ we have $N_r = O((\Gamma')^{-1})$. Thus, from Remark 6.12 it is easy to see that in the scaling regime the constant C_0 in Lemma 6.11 is $O(1)$ in ρ.

Furthermore, in Subsection 6.3.3 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 6.10), 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 (6.35) 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.

6.6 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 6.13, we have

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

so that, if ν_W and ν_0 are defined as in Subsection 2.3.3, it follows that

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

Since $W \in \mathfrak{B}^-_{\alpha_-}(\mathbb{T}^2,\mathbb{R})$, the result follows from Remark 2.4.

7 Averaging and deviations

The study of the convergence of the dynamics (2.26) to the deterministic dynamics (2.29) will be structured in several steps: we start with the first and second moments of the invariant manifold W (Subsections 7.2 and 7.3); then we consider the moments of the functions h and l and their derivatives (Subsections 7.4 to 7.9); eventually we draw the conclusions about the deviations of the dynamics with respect to the averaged system (Subsection 7.10).

Remark 7.1. Throughout this section, as well as in the Appendices we refer to for the proofs (actually from Appendix B on), we assume ρ to be such that Hypotheses 1 to 3 hold (see Remark 2.20), and we call C any constant independent of ρ whose numerical value is not relevant.

7.1 The extended map

Even though the set Ω is positively invariant for the map $\mathscr{S}$ in (2.26), 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 $x \leq 0$ and $\chi_{\rm ext}(x) = 0$ for $x \geq 1$ while $\partial_x \chi_{\rm ext}(x) \leq 0$ for every $x \in \mathbb{R}$;
- set $s_M := \min\{1, \nu_M/2\}$, where $\nu_M := \inf_\psi |f(\varphi_M, \psi)|$;
- 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 $\varphi_m \leq \varphi \leq \varphi_M$ while

  $$f_{\rm ext}(\varphi, \psi) = \left(\sum_{i=0}^{6} \partial_\varphi^i f(\varphi_M, \psi)(\varphi - \varphi_M)^i \right) \chi_{\rm ext}\left(\frac{\varphi - \varphi_M}{s_M} \right) - \frac{\nu_M}{2}\left(1 - \chi_{\rm ext}\left(\frac{\varphi - \varphi_M}{s_M} \right) \right),$$

  for $\varphi > \varphi_M$, and an analogous expression for $\varphi < \varphi_m$.

It follows that, for any $r > 0$, $f_{\rm ext}$ satisfies the bounds

$$|||f_{\rm ext}|||_{\alpha_0,6} \leq s_M^{-6} |||\chi_{\rm ext}|||_{0,6} F|||_{\alpha_0,6}, \qquad \inf_{\mathbb{R}\setminus\Lambda_r} |f_{\rm ext}(\varphi, \psi)| \geq \frac{1}{2} \inf_{\Omega\setminus\Lambda_r} |f(\varphi, \psi)|\,.$$

Moreover the map

$$\mathscr{S}_{\rm ext}(\varphi, \psi) := (\varphi + \rho f_{\rm ext}(\varphi, \psi), A_0\psi) \tag{7.1}$$

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 $\Omega_{\rm ext}$ in place of Ω.

Remark 7.2. 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 $(\mathbb{T} \times \mathbb{T}^2)\setminus\Omega$, is that we want to compare $\mathscr{S}$ with other maps, constructed starting from $\mathscr{S}$, which, albeit being closely related to $\mathscr{S}$, not only may fail to admit Ω as an invariant set (see Remark 7.11), but also are not necessarily defined in the whole Ω (see the beginning of Section 5). Thus, in order to avoid discussing 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 the original maps. We stress here that all the functions appearing in Theorems 4 to 8 depend only on $\mathscr{S}|_\Omega$. The errors introduced in estimating these functions, as an effect of the arbitrariness of the extension, are under control and are proved to be of order ρ (see Remark 7.19).

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 6.3, with the only difference that $\mathscr{S}$ has to be replaced with $\mathscr{S}_{\rm ext}$ everywhere. Extending $\mathscr{S}$ on $\Omega_{\rm ext}$ naturally defines an extended map $\mathscr{S}_{1,\rm ext}$ such that

$$\mathscr{S}_{1,\rm ext}(\theta, \psi) := \mathscr{S}_{\rm ext}(\theta + W(\psi), \psi) - W(A_0\psi)$$

is defined on $\Omega_{1,\rm ext} := \{(\theta, \psi) \in \mathbb{R} \times \mathbb{T}^2 : (\theta + W(\psi), \psi) \in \Omega_{\rm ext}\}$. In particular, in the following discussion we need to choose $\Omega_{\rm ext}$ in such a way that $\Omega \subset \Omega_{1,\rm ext}$ (see Subsection 7.8).

Remark 7.3. The actual set $\Omega_{\rm ext}$ that we need depends on the details of the proofs in the remaining of this section. We will specify how to construct it in Remark 7.11 and in (7.40) below.

7.2 A correlation inequality

This subsection is dedicated to a generalization of Proposition 2.3 that plays a central role in the proof of Theorem 4 and Lemma 2.33. Considering its importance, we first discuss a very simple example in Subsection 7.2.1 before stating the result in its full generality in Subsection 7.2.2.

7.2.1 A simple example

In Subsection 7.3 we will estimate, among more complex ones, expressions of the form

$$a_n := \left\langle \left(\prod_{j=0}^{n-1} \mu\circ A_0^j \right) b\circ A_0^n \right\rangle = \int_{\mathbb{T}^2} \left(\prod_{j=0}^{n-1} \mu(A_0^j\psi) \right) b(A_0^n\psi)\, m_0(\mathrm{d}\psi),$$

where $\langle b\rangle = 0$ and, writing $\mu(\psi) = \langle\mu\rangle + \rho\tilde{v}(\psi)$, with $\langle\tilde{v}\rangle = 0$, the functions μ and $\tilde{v}$ are such that $0 < \langle\mu\rangle \le \|\mu\|_\infty \le 1-\rho\,\gamma$ and $\|\tilde{v}\|_\infty \le C$.

A naïve estimate immediately gives

$$|a_n| \le (1-\rho\,\gamma)^n \|b\|_\infty\,. \tag{7.2}$$

We now show that, assuming $\tilde{v}$ and b to be α-Hölder continuous for some $\alpha \in (0,1)$ and $\langle b\rangle$ to vanish, such an estimate can be improved. Indeed we can write

$$\prod_{j=0}^{n-1} \mu(A_0^j\psi) = \rho \left(\prod_{j=0}^{n-2} \mu(A_0^j\psi) \right) \tilde{v}(A^{n-1}\psi) + \left(\prod_{j=0}^{n-2} \mu(A_0^j\psi) \right) \langle\mu\rangle$$

and, after iterations, we get

$$a_n = \langle\mu\rangle^n\langle b\rangle + \rho\sum_{k=1}^{n} \langle\mu\rangle^{n-k} \left\langle \left(\prod_{j=0}^{k-2} \mu\circ A_0^j \right) \tilde{v}\circ A_0^{k-1}\; b\circ A_0^n \right\rangle,$$

where we are following the convention in Remark 2.32 for sums and products, so that, by applying Proposition 2.3, we find

$$|a_n| \le C\rho \sum_{k=1}^{n} (1-\rho\,\gamma)^{n-k}(1+\alpha(n-k))\lambda^{-\alpha(n-k)} \left\| \left(\prod_{j=1}^{k-1} \mu\circ A_0^{-j} \right) \tilde{v} \right\|_\alpha^+ \|b\|_\alpha.$$

We can use (2.8) and bound

$$\left\| \left(\prod_{j=1}^{k-1} \mu\circ A_0^{-j} \right) \tilde{v} \right\|_\alpha^+ \le \|\tilde{v}\|_\alpha \left\| \prod_{j=1}^{k-1} \mu\circ A_0^{-j} \right\|_\alpha^+,$$

while from (2.13) we get

$$\left\| \prod_{j=1}^{k-1} \mu\circ A_0^{-j} \right\|_\alpha^+ \le \|\mu\|_\infty^{k-2}\|\mu\|_\alpha^+ \sum_{j=1}^{k-1} \lambda^{-\alpha j}. \tag{7.3}$$

Finally, we obtain

$$|a_n| \le C\rho(1-\rho\,\gamma)^n\, \frac{\lambda^\alpha+\alpha-1}{(1-\rho\,\gamma)^2(1-\lambda^\alpha)^2}\, \|\tilde{v}\|_\alpha\|\mu\|_\alpha\|b\|_\alpha. \tag{7.4}$$

Comparing (7.4) with (7.2) we see that, for $\alpha = \alpha_0$, we have gained a factor ρ at the cost of a possibly worse constant. Note that, on contrast, if $\alpha = \alpha_*$, the factor $(\lambda^\alpha+\alpha-1)/(1-\lambda^\alpha)^2$ is $O(1/\rho)$, so that there is no gain with respect to the bound (7.2) in such a case.

7.2.2 The general inequality

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 $i = 0,\dots,n-k$,

$$\mathfrak{u}_i^{(k)}(\psi) := \mathfrak{u}_i^{(k-1)}(\psi)\,\mathfrak{u}_{i+k-1}(A_0^{k-1}\psi) = \prod_{j=0}^{k-1} \mathfrak{u}_{i+j}(A_0^j\psi), \qquad \mathfrak{v}_i^{(k)}(\psi) := \prod_{j=0}^{k-1} \mathfrak{v}_{i+j}(A_0^j\psi), \tag{7.5}$$

where, according to Remark 2.32, $\mathfrak{u}_i^{(0)}(\psi) = \mathfrak{v}_i^{(0)}(\psi) = 1$. Finally, set $\mathfrak{u}^{(k)} := \mathfrak{u}_0^{(k)}$ and $\mathfrak{v}^{(k)} := \mathfrak{v}_0^{(k)}$.

Remark 7.4. If we define

$$\mathfrak{u}^{(-k)}(\psi) := \prod_{j=1}^{k} \mathfrak{u}_{k-j}(A_0^{-j}\psi),$$

then one has $\mathfrak{u}^{(-k)} = \mathfrak{u}^{(k)} \circ A_0^{-k}$.

Remark 7.5. We can write

$$\mathfrak{u}_i^{(k)} - \mathfrak{v}_i^{(k)} = \rho\, \mathfrak{u}_i^{(k-1)}\, \mathfrak{w}_{i+k-1} \circ A_0^{k-1} + \big(\mathfrak{u}_i^{(k-1)} - \mathfrak{v}_i^{(k-1)}\big)\, \mathfrak{v}_{i+k-1} \circ A_0^{k-1},$$

so that, after iterating, we get

$$\begin{aligned} \mathfrak{u}_i^{(k)} &= \mathfrak{v}_i^{(k)} + \rho \sum_{j=0}^{k-1} \mathfrak{u}_i^{(j)}\, \mathfrak{w}_{i+j} \circ A_0^{j}\, \mathfrak{v}_{i+j+1}^{(k-j-1)} \circ A_0^{j+1} \\ &= \mathfrak{v}_i^{(k)} + \rho \sum_{j=0}^{k-1} \mathfrak{v}_i^{(j)}\, \mathfrak{w}_{i+j} \circ A_0^{j}\, \mathfrak{v}_{i+j+1}^{(k-j-1)} \circ A_0^{j+1} \\ &\quad + \rho^2 \sum_{j=1}^{k-1} \sum_{j'=0}^{j-1} \mathfrak{u}_i^{(j')}\, \mathfrak{w}_{i+j'} \circ A_0^{j'}\, \mathfrak{v}_{i+j'+1}^{(j-j'-1)} \circ A_0^{j'+1}\, \mathfrak{w}_{i+j} \circ A_0^{j}\, \mathfrak{v}_{i+j+1}^{(k-j-1)} \circ A_0^{j+1}. \end{aligned} \tag{7.6}$$

where the second equality follows applying the first equality to the factor $\mathfrak{u}_i^{(j)}$ in the first line. According to the conventions established in Remark 2.32, the sum in the second line 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

$$\mathfrak{u}_i^{(k)} = \mathfrak{v}_i^{(k)} + \rho \sum_{j=0}^{k-1} \mathfrak{v}_i^{(j)}\, \mathfrak{w}_{i+j} \circ A_0^{j}\, \mathfrak{u}_{i+j+1}^{(k-j-1)} \circ A_0^{j+1}, \tag{7.7}$$

where we avoided writing the equivalent of the second expansion in (7.6) since we will not need it.

The following result plays a important role in the forthcoming analysis. The proof is based on Remark 7.5 and extends the reasoning of Subsection 7.2.1.

Proposition 7.6. *Let $\mathfrak{u}_0, \ldots, \mathfrak{u}_{n-1}$ 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 $i = 0, \ldots, n-1$,*
2. *$\tilde{\mathfrak{u}}_i = O(\rho)$ for all $i = 0, \ldots, n-1$.*

Given $g_+ \in \mathfrak{B}_\alpha^+(\mathbb{T}^2, \mathbb{R})$ and $g_- \in \mathfrak{B}_\alpha^-(\mathbb{T}^2, \mathbb{R})$, one has

$$\begin{aligned} \Big| \langle g_+\, \mathfrak{u}^{(n)} \circ A_0\, g_- \circ A_0^{n+1} \rangle - \langle \mathfrak{u} \rangle^{(n)} \langle g_+ \rangle \langle g_- \rangle \Big| &\le C(1 - \rho\,\gamma)^n \Big((1 + \alpha n) \lambda^{-\alpha n} \|\tilde{g}_+\|_\alpha^+ \|\tilde{g}_-\|_\alpha^- \\ &\quad + \rho \big(\|\tilde{g}_+\|_\alpha^+ \|g_-\|_\alpha^- + \|g_+\|_\alpha^+ \|\tilde{g}_-\|_\alpha^- \big) + \rho^2 n |\langle g_+ \rangle| |\langle g_- \rangle| \Big), \end{aligned} \tag{7.8}$$

where $\langle \mathfrak{u} \rangle^{(n)} := \langle \mathfrak{u}_0 \rangle \langle \mathfrak{u}_1 \rangle \ldots \langle \mathfrak{u}_{n-1} \rangle$.

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

Then we write

$$\langle g_+\, \mathfrak{u}^{(n)} \circ A_0\, g_- \circ A_0^{n+1} \rangle = \langle g_+\, \mathfrak{u}^{(n)} \circ A_0\, \tilde{g}_- \circ A_0^{n+1} \rangle + \langle \tilde{g}_+\, \mathfrak{u}^{(n)} \circ A_0 \rangle \langle g_- \rangle + \langle g_+ \rangle \langle \mathfrak{u}^{(n)} \rangle \langle g_- \rangle, \tag{7.9}$$

so that, using the first line of (7.6) in Remark 7.5, with $\mathfrak{u}^{(n)}\circ A_0$ instead of $\mathfrak{u}^{(n)}$, we rewrite the first term in (7.9) as

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

Using (2.7) and the first bound in Remark 2.2, we get

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

so that, from (2.12) and Proposition 2.3, we obtain in (7.9)

$$\begin{aligned}&\big|\langle g_+\,\mathfrak{u}^{(n)}\circ A_0\,\tilde g_-\circ A_0^{n+1}\rangle\big|\\ &\quad\le C(1-\rho\,\gamma)^n\lambda^{-\alpha n}(1+\alpha n)\|\tilde g_+\|_\alpha^+\|\tilde g_-\|_\alpha^- + C\rho(1-\rho\,\gamma)^n\|g_+\|_\alpha^+\|\tilde g_-\|_\alpha^-\,.\end{aligned}\tag{7.10}$$

Analogously to (7.10), using (7.7), with $\mathfrak{u}^{(n)}\circ A_0$ instead of $\mathfrak{u}^{(n)}$, and Proposition 2.3, we get

$$\big|\langle\tilde g_+\,\mathfrak{u}^{(n)}\circ A_0\rangle\big| \le \rho\sum_{k=0}^{n-1}\langle\mathfrak{u}\rangle_0^k\big|\langle\tilde g_+\,\tilde{\mathfrak{u}}_k\circ A_0^k\,\mathfrak{u}^{(n-k-1)}\circ A_0^{k+1}\rangle\,\big| \le C\rho(1-\rho\,\gamma)^n\|\tilde g_+\|_\alpha^+\,.\tag{7.11}$$

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 (7.6), once more with $\mathfrak{u}^{(n)}\circ A_0$ instead of $\mathfrak{u}^{(n)}$, we have

$$\langle\,\mathfrak{u}^{(n)}\rangle-\langle\mathfrak{u}\rangle^n = \rho^2\sum_{k=1}^{n-1}\sum_{j=0}^{k-1}\langle\mathfrak{u}\rangle_{j+1}^{k-j-1}\langle\mathfrak{u}\rangle_{k+1}^{n-k-1}\langle\mathfrak{u}^{(j)}\circ A_0^{-j}\,\tilde{\mathfrak{u}}_j\,\tilde{\mathfrak{u}}_k\circ A_0^{k-j}\rangle$$

and hence, by Proposition 2.3 and Remark 7.4,

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

so that we obtain

$$\big|\langle\,\mathfrak{u}^{(n)}\rangle-\langle\mathfrak{u}\rangle_0^n\big| \le C\rho^2 n(1-\rho\,\gamma)^n\,.\tag{7.12}$$

Collecting all contributions (7.10) to (7.12) gives the thesis. □

Remark 7.7. The factor $\rho\big(\|\tilde g_+\|_\alpha^+\|g_-\|_\alpha^- + \|g_+\|_\alpha^+\|\tilde g_-\|_\alpha^-\big)$ in the estimate (7.8) may be replaced more pragmatically with $2\rho\,\|g_+\|_\alpha^+\|g_-\|_\alpha^-$. However the more explicit bound (7.8) can be useful in some cases. In particular, for $g_+=g_-=1$, we obtain that $|\langle\mathfrak{u}^{(n)}\rangle-\langle\mathfrak{u}\rangle^{(n)}|\le C\rho^2 n(1-\rho\,\gamma)^n$ 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'.

7.3 Oscillations of the invariant manifold: proof of Theorem 4

Following (2.19), we write

$$f(\varphi,\psi)=b(\psi)+v(\psi)\,\varphi+d(\varphi,\psi)\varphi^2,\tag{7.13}$$

where $b(\psi) := f(0,\psi)$ and $v(\psi) := \partial_\varphi f(0,\psi)$, with $\|v\|_\infty \le C$ by Remark 2.20. We also define

$$\mu(\psi) := 1 + \rho v(\psi) = 1 + \rho \partial_\varphi f(0,\psi) \tag{7.14}$$

and, for $i \ge 0$, according to (7.5) and Remark 7.4,

$$\mu^{(i)}(\psi) = \prod_{j=0}^{i-1} \mu(A_0^j \psi), \qquad \mu^{(-i)}(\psi) = \prod_{j=1}^{i} \mu(A_0^{-j}\psi) = \mu^{(i)}(A^{-i}\psi). \tag{7.15}$$

Remark 7.8. Decomposing $\mu(\psi) = \langle \mu \rangle + \tilde{\mu}(\psi)$ according to (2.1), we note that $\langle \mu \rangle = \overline{\mu}$, with $\overline{\mu}$ as in (2.32), and $\tilde{\mu}(\psi) = \rho \tilde{v}(\psi)$. We get $\langle \mu \rangle \le \|\mu\|_\infty \le 1 - \rho\gamma$; indeed, one has $\partial_\varphi f(\varphi,\psi) \le -\gamma$ for all $\varphi \in [S_m, S_M]$ and $\bar{\varphi} = 0 \in [S_m, S_M]$ (see (2.27) and Remark 2.10). Moreover one has $|\mu|^+_{\alpha_0} = O(\rho)$, as we see directly from (7.14).

In Section 6 we proved (see Remark 6.1) that, uniformly on $\mathbb{T}^2$,

$$W = \lim_{n\to\infty} \mathscr{E}^n[0],$$

with $\mathscr{E}$ defined in (6.2). Thus, to prove (2.38), we will show that for all $n \in \mathbb{N}$

$$\big|\langle \mathscr{E}^n[0] \rangle\big| \le C\rho, \qquad \langle (\mathscr{E}^n[0])^2 \rangle \le C\rho,$$

for some constant C independent of n. To this end, analogously to (6.2), we set

$$\mathscr{E}_0[W](\psi) := G_0(W(A_0^{-1}\psi), A_0^{-1}\psi),$$

where

$$G_0(\varphi,\psi) := \varphi + \rho f_0(\varphi,\psi), \qquad f_0(\varphi,\psi) := b(\psi) + v(\psi)\varphi\,. \tag{7.16}$$

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

$$\mathscr{E}_0^n[h] = \rho \sum_{i=1}^{n} \mu^{(-i+1)} b \circ A_0^{-i} + \mu^{(-n)} h \circ A_0^{-n}\,. \tag{7.17}$$

Remark 7.9. Note that $\mathscr{E}_0$ is a contraction on the space of bounded continuous functions defined from $\mathbb{T}^2$ to $\mathbb{R}$. Thus the fixed point equation $\mathscr{E}_0[W] = W$ admits a unique solution, that we call W_0, and $\mathscr{E}_0^n[0]$ converges uniformly to W_0. In fact we can write

$$W_0 = \rho \sum_{i=1}^{n} \mu^{(-i+1)} b \circ A_0^{-i} + \mu^{(-n)} W_0 \circ A_0^{-n} = \rho \sum_{i=1}^{\infty} \mu^{(-i+1)} b \circ A_0^{-i}\,, \tag{7.18}$$

from which we get $\|W_0\|_\infty \le \gamma^{-1}\|b\|_\infty$, while $|W_0|^+_{\alpha_0} = O(\rho)$ by (2.13).

Remark 7.10. To prove Theorem 4 we first show, via an essentially explicit computation, that $|\langle W_0 \rangle| = O(\rho)$ and $\langle W_0^2 \rangle = O(\rho)$. We then show that $\langle |W - W_0| \rangle = O(\rho)$ by comparing the iterates of $\mathscr{E}_0$ and those of $\mathscr{E}$ and using their contractive properties.

From Remark 7.4, 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 7.6, with $n = i-1$, $\mathfrak{u}_k = \mu$ for $k = 0, \ldots, n-1$, $g_+ = b$ and $g_- = 1$, we find

$$|\langle \mathscr{E}_0^n[0] \rangle| \le C\rho^2 \sum_{i=0}^{n-1} (1 - \rho\gamma)^i \|b\|_{\alpha_0} \le C\rho\,. \tag{7.19}$$

Moreover we have

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

where

$$
\begin{aligned}
&\Big\langle (\mu^{(-i+1)})^2 \, b \circ A_0^{-i} \, \mu \circ A_0^{-i} \, \mu^{(-j+i+1)} \circ A_0^{-i} \, b \circ A_0^{-j} \Big\rangle \\
&\qquad = \Big\langle b \, \mu^{(j-i-1)} \circ A_0 (\mu^{(i-1)})^2 \circ A_0^{j-i+1} \, b \circ A_0^{j-i} \, \mu \circ A_0^{j-i} \Big\rangle,
\end{aligned}
$$

so that, first using (2.12) and the second bound in Remark 2.2 to estimate

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

then applying once more Proposition 7.6, 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

$$
\begin{aligned}
\langle \mathscr{E}_0^n[0]^2 \rangle &\le C\rho \|b\|_\infty^2 + C\rho^3 \sum_{1\le i<j\le n} (1-\rho\,\gamma)^{j-i} \|b\|_{\alpha_0} \|(\mu^{(i-1)})^2 \circ A_0 \, b\,\mu\|_{\alpha_0}^- \\
&\le C\rho + C\rho^3 \sum_{1\le i<j\le n}^{n} (1-\rho\,\gamma)^{j+i} \le C\rho.
\end{aligned}
\tag{7.20}
$$

Thus, if W_0 is defined as in Remark 7.9, we obtain

$$
\langle W_0 \rangle = \lim_{n\to\infty} \langle \mathscr{E}_0^n[0] \rangle = O(\rho), \qquad \langle W_0^2 \rangle = \lim_{n\to\infty} \langle \mathscr{E}_0^n[0]^2 \rangle = O(\rho). \tag{7.21}
$$

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

$$
\begin{aligned}
\mathscr{E}^n[0] - \mathscr{E}_0^n[0] &= \sum_{i=0}^{n-1} \mathscr{E}^{n-i}[\mathscr{E}_0^i[0]] - \mathscr{E}^{n-i-1}[\mathscr{E}_0^{i+1}[0]] \\
&= \sum_{i=0}^{n-1} \mathscr{E}^{n-i-1}[\mathscr{E}[\mathscr{E}_0^i[0]]] - \mathscr{E}^{n-i-1}[\mathscr{E}_0[\mathscr{E}_0^i[0]]] \,.
\end{aligned}
\tag{7.22}
$$

Remark 7.11. 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 $n \ge 0$ and all $\psi \in \mathbb{T}^2$. Thus for the r.h.s. of (7.22) to be well defined, we compute $\mathscr{E}$ by replacing the map $\mathscr{S}$ with the extension $\mathscr{S}_{\rm ext}$ introduced in Subsection 7.1 such that $\Omega_{\rm ext} \supset \Omega'$ (see also Remark 7.2).

By taking $\mathscr{S}_{\rm ext}$ according to Remark 7.11 and reasoning as in Lemma 2.9, we find that for any $r > 0$ there exists

$$
N_r' \le \frac{\max\{\phi_M' - S_M, S_m - \phi_m'\}}{\rho \inf\limits_{\Omega' \setminus \Lambda_r} |f_{\rm ext}(\varphi,\psi)|} \,,
$$

such that $\mathscr{S}_{\rm ext}^{N_r'}(\Omega') \subset \Lambda_r$, 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

$$
|\mathscr{E}^{N_r'}[W_1] - \mathscr{E}^{N_r'}[W_2]| \le (1 + \rho \|\partial_\varphi f\|_\infty)^{N_r'} |W_1 - W_2|.
$$

Thus, for any $\gamma' \in (0, \gamma)$, with the notation in property 4 of Lemma 2.9, 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 $k \ge 0$, we may bound $|\mathscr{E}^k[W_1] - \mathscr{E}^k[W_2]| \le (1-\rho\,\gamma')^k |W_1 - W_2|$.

Summing up, (7.22) gives

$$
\begin{aligned}
&|\mathscr{E}^n[0](\psi) - \mathscr{E}_0^n[0](\psi)| \\
&\quad \le \left(\frac{1+\rho\|\partial_\varphi f\|_\infty}{1-\rho\gamma'}\right)^{N_r'} \rho \sum_{i=0}^{n-1} (1-\rho\gamma')^{n-i-1} \left|d(\mathscr{E}_0^i[0](A_0^{-1}\psi), A_0^{-1}\psi)\right| \, (\mathscr{E}_0^i[0](A_0^{-1}\psi))^2,
\end{aligned}
\tag{7.23}
$$

where $d(\varphi,\psi)$ is defined in (7.13). Integrating over ψ and using (7.20), 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

$$
\langle |W - W_0| \rangle \le C\rho \tag{7.24}
$$

and, because of (7.21),

$$
|\langle W\rangle| \le |\langle W - W_0\rangle| + |\langle W_0\rangle| \le \langle |W - W_0|\rangle + |\langle W_0\rangle| \le C\rho\,,
$$

that is the first bound in (2.38).

It is now easy to see that

$$
|\langle W^2 - W_0^2\rangle| \le \langle |W - W_0| \, |W + W_0|\rangle \le C\rho
$$

because of (7.24), and hence, thanks to (7.21), $\langle W^2\rangle \le |\langle W^2 - W_0^2\rangle| + \langle W_0^2\rangle \le C\rho$, which provides the second bound in (2.38).

7.4 Fluctuations of the linearized dynamics: proof of Lemma 2.33

Having studied the oscillations of the steady state 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

$$
\kappa(\psi) = \mu(\psi) + \rho\,\xi(\psi), \qquad \xi(\psi) := d_1(\psi)\, W(\psi) + d_2(\psi)\, W^2(\psi), \qquad d_1(\psi) := \partial_\varphi^2 f(0,\psi), \tag{7.25}
$$

which implicitly defines the function $d_2 \in \mathfrak{B}^*_{\alpha_-,\alpha_+}(\mathbb{T}^2,\mathbb{R})$. From Proposition 7.6, with $g_+ = g_- = 1$ and either $\mathfrak{u}_i = \mu$ or $\mathfrak{u}_i = \mu^2$ for all i, and Remark 7.7, it follows that

$$
\left|\langle \mu^{(n)} - \overline{\mu}^n\rangle\right| \le C\rho^2 n(1-\rho\gamma)^n, \qquad \left\langle (\mu^{(n)} - \overline{\mu}^n)^2\right\rangle \le C\rho^2 n(1-\rho\gamma)^{2n},
$$

and hence, in order to prove (2.41), it is enough to show that

$$
\left|\langle \kappa^{(n)} - \mu^{(n)}\rangle\right| \le C\rho(1-\rho\gamma)^n \sum_{k=0}^{2} \rho^k n^k, \qquad \left\langle (\kappa^{(n)} - \mu^{(n)})^2\right\rangle \le C\rho(1-\rho\gamma)^{2n} \sum_{k=0}^{2} \rho^k n^k. \tag{7.26}
$$

Remark 7.12. The proof of (7.26) represents a first instance of the strategy outlined at the end of Section 2.5: in order to bypass the low regularity of the function 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 averages of the quadratic contributions and the bound (7.24) to reduce the analysis of the averages of the linear contributions to the more manageable function W_0. Finally we use the explicit expression for W_0 in (7.18) to obtain the desired estimates.

Thus, we proceed with the proof of (7.26). By Remark 7.5 we can write

$$
\begin{aligned}
\kappa^{(n)} - \mu^{(n)} &= \rho \sum_{j=0}^{n-1} \kappa^{(j)}\, \xi \circ A_0^j\, \mu^{(n-j-1)} \circ A^{j+1} \\
&= \rho \sum_{j=0}^{n-1} \mu^{(j)}\, \xi \circ A_0^j\, \mu^{(n-j-1)} \circ A^{j+1} \\
&+ \rho^2 \sum_{j=1}^{n-1} \sum_{k=0}^{j-1} \kappa^{(k)}\, \xi \circ A_0^k\, \mu^{(j-k-1)} \circ A_0^{k+1}\, \xi \circ A_0^j\, \mu^{(n-j-1)} \circ A_0^{j+1}.
\end{aligned} \tag{7.27}
$$

Observing that

$$
\begin{aligned}
&\left| \left\langle \kappa^{(k)}\, \xi \circ A_0^k\, \mu^{(j-k-1)} \circ A_0^{k+1}\, \xi \circ A_0^j\, \mu^{(n-j-1)} \circ A_0^{j+1} \right\rangle \right| \\
&\qquad \le (1-\rho\gamma)^{n-2} \left\langle \left| \xi \circ A_0^k\, \xi \circ A_0^j \right| \right\rangle \le (1-\rho\gamma)^{n-2} \langle \xi^2 \rangle
\end{aligned}
$$

and that $\langle \xi^2 \rangle \le C\rho$, by (7.25) and (2.38), eventually we get, for $n \ge 2$,

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

On the other hand we have, again by (2.38),

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

so that we just need to estimate

$$
\left\langle \mu^{(j)}\, (d_1 W) \circ A_0^j\, \mu^{(n-j-1)} \circ A_0^{j+1} \right\rangle = \left\langle \mu^{(j)} \circ A_0^{-j}\, d_1 W\, \mu^{(n-j-1)} \circ A_0 \right\rangle.
$$

By Proposition 7.6, with $\alpha = \alpha_+$, n replaced with $n-j-1$, $\mathfrak{u}_i = \mu$ for $i = 0, \ldots, n-j-2$, $g_+ = \mu^{(j)} \circ A_0^{-j}\, d_1 W$ and $g_- = 1$, we obtain

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

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 (2.13), respectively. Then, by (7.24) and Remark 7.9, 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, by Proposition 7.6, with $n = j$, $\mathfrak{u}_i = \mu^2$ for $i = 0, \ldots, j-1$, $g_+ = W_0 \circ A_0^{-1}$ and $g_- = d_1$,

$$
\left| \left\langle W_0\, (\mu^{(j)})^2\, d_1 \circ A_0^j \right\rangle \right| \le C\, (1-\rho\gamma)^{2j} \big((1+\alpha_0 j) \lambda^{-\alpha_0 j} + \rho + \rho^3 j \big),
$$

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 7.6, 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 (7.27), we obtain the first of (7.26).

For the second of (2.41), we use the first line of (7.27) to write

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

Thus, the second bound in (7.26) follows using again the second bound in (2.38).

7.5 Iterated products

We introduce here some 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

$$\mathscr{S}(\varphi,\psi) = (\mathscr{S}_\varphi(\varphi,\psi), \mathscr{S}_\psi(\varphi,\psi)) := (G(\varphi,\psi), A_0\psi)\,, \tag{7.28}$$

and let $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ be any set of functions defined in $\mathcal{U}\times\mathbb{T}^2$. Define, for $k=0,\dots,n$ and $i=0,\dots,n-k$,

$$\mathfrak{p}_i^{(k)}(\mathscr{S};\varphi,\psi) := \prod_{j=0}^{k-1} \mathfrak{p}_{i+j}(\mathscr{S}^j(\varphi,\psi)). \tag{7.29}$$

Remark 7.13. The map $\mathscr{S}$ in (7.28) is not necessarily the map (2.15) which defines our model. In particular, in what follows, we shall use the notation (7.29) for several maps, including the translated map $\mathscr{S}_1$ and the auxiliary map $\mathscr{S}_2$ which will be introduced in Subsection 7.8.1.

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 (7.29) we may consider

$$\mathfrak{p}_i^{(k)}(\mathscr{S};\varphi) := \prod_{j=0}^{k-1} \mathfrak{p}_{i+j}(G^j(\varphi))\,, \tag{7.30}$$

with $G(\varphi) := \mathscr{S}_\varphi(\varphi,\psi)$ and G^j denoting is the composition of G with itself j times. In particular, for $\mathscr{S}(\varphi,\psi)=\overline{\mathscr{S}}(\varphi,\psi)$, where $\overline{\mathscr{S}}(\varphi,\psi)=(\overline{G}(\varphi),A_0\psi)$ is the averaged map (2.30), then, given any functions $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ on $\mathcal{U}\times\mathbb{T}^2$, we have

$$\langle\mathfrak{p}\rangle_i^{(k)}(\overline{\mathscr{S}};\varphi) = \prod_{j=0}^{k-1} \langle\mathfrak{p}_{i+j}\rangle(\overline{G}^j(\varphi)). \tag{7.31}$$

According to the convention established in Remark 2.32, we set $\mathfrak{p}_i^{(-1)}(\mathscr{S};\varphi,\psi)=\mathfrak{p}_i^{(0)}(\mathscr{S};\varphi,\psi)=1$.

7.6 Conjugation of the averaged dynamics: proof of Lemma 2.27

Proceeding as in Subsection 6.3.2 we see that the function $\overline{h}\colon \mathcal{U} \to \mathcal{U}_0$ introduced in (2.33) satisfies the equation

$$\overline{h}(\varphi) = \overline{q}(\varphi) + \overline{p}(\varphi)\,\overline{h}(\overline{G}(\varphi)), \tag{7.32}$$

with $\overline{G}(\varphi) = \overline{\mathscr{S}}_\varphi(\varphi, \psi)$ defined in (2.30) and

$$\overline{p}(\varphi) := \frac{(\overline{G}(\varphi))^2}{\varphi^2\,\overline{\mu}} = \frac{1}{1+\rho\,\partial_\varphi \overline{f}(0)} \left(\frac{\varphi + \rho \overline{f}(\varphi)}{\varphi}\right)^2, \tag{7.33a}$$

$$\overline{q}(\varphi) = \frac{\overline{G}(\varphi) - \varphi\,\overline{\mu}}{\varphi^2\,\overline{\mu}} = \frac{\rho}{1+\rho\,\partial_\varphi \overline{f}(0)} \frac{\overline{f}(\varphi) - \partial_\varphi \overline{f}(0)\varphi}{\varphi^2}. \tag{7.33b}$$

Thus we can write (see (7.29) and (7.30) for the notation)

$$\overline{h}(\varphi) = \sum_{n=1}^{\infty} \overline{p}^{(n)}(\varphi)\,\overline{q}(\overline{G}^n(\varphi)), \tag{7.34}$$

with

$$\overline{p}^{(n)}(\varphi) = \overline{p}^{(n)}(\overline{\mathscr{S}};\varphi) = \prod_{i=0}^{n-1} \overline{p}(\overline{G}^i(\varphi)). \tag{7.35}$$

Analogously to Lemma 6.9 the following result holds.

Lemma 7.14. *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 7.15. By reasoning as in the proof of Lemma 6.11 (and hence, actually, Lemma 2.9), we find that, for any $\gamma' \in (0,\gamma)$, there exist $r > 0$ and $\overline{N} = O(\rho^{-1})$ such that $\overline{\mathscr{S}}^k(\mathcal{U}) \subset [S_m - r, S_M + r]$ and $(1+\rho\partial_\varphi \overline{f})\circ\overline{\mathscr{S}}^k \le (1-\rho\,\gamma')$ for all $k \ge \overline{N}$.

The following result is proved in Appendix B.2.

Lemma 7.16. *Given $\gamma' \in (0,\gamma)$ and $\bar{\theta}$ as in Lemma 7.14, consider any functions $\mathfrak{p}_0, \ldots, \mathfrak{p}_{n-1}$ in $\mathscr{B}_{0,3}(\Omega, \mathbb{R})$, independent of ψ, such that*

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

and define $\mathfrak{p}^{(n)}(\varphi) = \mathfrak{p}_0^{(n)}(\overline{\mathscr{S}};\varphi)$, with $\mathfrak{p}_0^{(n)}(\overline{\mathscr{S}};\varphi)$ as in (7.30) with $\mathscr{S} = \overline{\mathscr{S}}$. Then one has

$$|||\mathfrak{p}^{(n)}|||_{0,3} \le C(1-\rho\,\gamma')^n. \tag{7.36}$$

where the constant C does not depend on n. In particular one finds

$$\|\partial_\varphi^k \overline{G}^n\|_\infty \le C(1-\rho\,\gamma')^n\,, \qquad k = 1,2,3. \tag{7.37}$$

Remark 7.17. An immediate consequence of the functions $\mathfrak{p}_0, \ldots, \mathfrak{p}_{n-1}$ satisfying condition 1 in Lemma 7.16 is that $|||\partial_\varphi \mathfrak{p}_i|||_{0,2} = O(\rho)$ for all $i = 0, \ldots, n-1$.

Remark 7.18. In the following we have to consider also cases in which property 2 of Lemma 7.16 holds for all $i = 0, \ldots, n-1$ except at most n_* values, for some $n_* < n$ independent of n. However such a case is easily reduced to Lemma 7.16. Indeed if $\mathfrak{p}_{i_1}, \ldots, \mathfrak{p}_{i_{n_*}}$ are the functions which do not satisfy the bound in property 2, then setting

$$I_* = \{i_1, \ldots i_{n_*}\}, \qquad K_* = \max\{\|\mathfrak{p}_i\|_\infty : i \in I_*\},$$

it may be convenient to 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 as to obtain that $|\tilde{\mathfrak{p}}_i(\varphi)| \le (1-\rho\,\gamma')^{-1}$ for all $i = 1, \dots, n-1$. Moreover, if $\|\mathfrak{p}_i - 1\|_{0,3} = O(\rho)$ then also $\|\tilde{\mathfrak{p}}_i - 1\|_{0,3} = O(\rho)$. Hence the functions $\tilde{\mathfrak{p}}_0, \dots, \tilde{\mathfrak{p}}_{n-1}$ satisfy all the hypotheses of Lemma 7.16. 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.

The bounds for $\overline{h}$ now follow easily considering (7.34), reasoning like in Subsection 6.3.3 and using the bound (7.36) 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 6.4), and the bounds for the function $\overline{l}$ are easily obtained using the Inverse Function Theorem.

7.7 The averaged and the continuous time system: proof of Lemma 2.29

Let $\Phi_t(\varphi)$ be the solution of (2.36). Observe first that

$$\Phi_\rho(\varphi) = \varphi + \rho\overline{f}(\varphi) + \rho^2 \int_0^1 (1-t)\,\partial_\varphi \overline{f}(\Phi_{t\rho}(\varphi))\,\overline{f}(\Phi_{t\rho}(\varphi))\,\mathrm{d}t \tag{7.38}$$

and that for any $\gamma'' \in (0,\gamma)$ there exists a constant C independent of ρ such that, for all $n \ge 0$,

$$|\Phi_{n\rho}(\varphi)| \le C(1-\rho\,\gamma'')^n, \qquad |(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi)| \le C(1-\rho\,\gamma'')^n. \tag{7.39}$$

Then, using (2.30) and (7.38), we obtain, for some φ_n between $(\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi)$ and $\Phi_{n\rho}(\varphi)$,

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

where we have used (7.39) and the fact that $\overline{f}(0) = 0$ in order to bound $|\overline{f}(\Phi_{t\rho}(\varphi))| \le C|\Phi_{t\rho}(\varphi))|$. Iterating we get, for suitable $\varphi_1, \dots, \varphi_{n-1}$,

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

Let $\overline{N}$ be defined as in Remark 7.15. By the first of (7.39), there exists $\overline{M} \ge N$ 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 $k \ge M$, with r such that $|1+\rho\,\partial_\varphi\overline{f}(\varphi)| \le C(1-\rho\gamma'')$ for all $\varphi \in [S_m - r, S_M + r]$. Thus, we get

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

from which the thesis follows immediately, by choosing $\gamma'' \in (0,\gamma)$ and taking $\gamma' \in (0,\gamma'')$ such that $n\rho(1-\rho\,\gamma'')^n \le (1-\rho\,\gamma')^n$. □

7.8 Deviations of the conjugation: proof of Theorem 5

If we aim to compare $\mathscr{S}^n(\varphi,\psi)$ with $\overline{\mathscr{S}}^n(\varphi,\psi)$, according to (2.43) we need to control the deviations of $h(\varphi,\psi)$ with respect to $\overline{h}(\varphi)$ and of $l(\eta,\psi)$ with respect to $\overline{l}(\eta)$.

Since $\overline{G}(0) = 0$, the map $\overline{\mathscr{S}}$ admits the invariant manifold $\overline{\mathcal{W}} := \{(0,\psi) \,|\, \psi \in \mathbb{T}^2\}$, that is the same invariant manifold as $\mathscr{S}_1$ (see Remarks 2.26 and 6.5). Thus, in order to complete the program outlined at the beginning of Subsection 2.4.3, it is more convenient to compare first $h_1(\varphi,\psi)$ with $\overline{h}(\varphi)$ and show that, in average, they are close, and then show that the same happens when comparing $h(\varphi,\psi)$ with $h_1(\varphi,\psi)$.

However, the domain Ω_1 of the map $\mathscr{S}_1(\varphi,\psi)$ – and hence of $h_1(\varphi,\psi)$ – is of the form (2.11), with $a_+(\psi) = \phi_M - W(\psi)$ and $a_-(\psi) = \phi_m - W(\psi)$. For $\langle h_1(\varphi,\cdot)\rangle$ to be 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 6.7). Since eventually 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 7.1 – to functions $F_{\rm ext}$ and $\mathcal{H}_{\rm ext}$ defined on a larger domain $\Omega_{\rm ext}$, with the set $\Omega_{\rm ext}$ such that the extended map $\mathscr{S}_{1,{\rm ext}}(\theta,\psi) := \mathscr{S}_{\rm ext}(\theta + W(\psi),\psi) - W(A_0\psi)$ is defined for all $(\theta,\psi) \in \Omega_{\rm ext}$. To this end we set $\Omega_{\rm ext} = \mathcal{U}_{\rm ext} \times \mathbb{T}^2$ with

$$\mathcal{U}_{\rm ext} = \Big[\phi_m + \min_{\psi\in\mathbb{T}^2} W(\psi), \phi_M + \max_{\psi\in\mathbb{T}^2} W(\psi)\Big] \tag{7.40}$$

so that we have $\Omega_{\rm ext} \supset \{(\varphi,\psi) \in \mathbb{R}\times\mathbb{T}^2 : \phi_m + \min\{0, W(\psi)\} \le \varphi \le \phi_M + \max\{0, W(\psi)\}\}$.

Then, when considering $h(\varphi,\psi) - \overline{h}(\varphi)$, we may write, for $\varphi \in \mathcal{U}$,

$$\begin{aligned} h(\varphi,\psi) &= h_{\rm ext}(\varphi,\psi) = h_{\rm ext}(\varphi + W(\psi),\psi) - \big(h_{\rm ext}(\varphi + W(\psi),\psi) - h_{\rm ext}(\varphi,\psi)\big) \\ &= h_{1,{\rm ext}}(\varphi,\psi) - W(\psi)\,\partial_\varphi h_{\rm ext}(\varphi,\psi) - (W(\psi))^2 \int_0^1 {\rm d}t\,(1-t)\,\partial_\varphi^2 h_{\rm ext}(\varphi - tW(\psi),\psi) \end{aligned} \tag{7.41}$$

and start by studying the average of $h_{1,{\rm ext}}(\varphi,\psi) - \overline{h}(\varphi)$. The next step will be to show that the average of the other terms appearing in (7.41) produce corrections of order ρ. This will follow from the control on the first and second moments of W, ensured by Theorem 4.

Remark 7.19. It is important to stress that, although we use the extended maps along the proof of Theorem 6, the final result does not depend on the extension that we have used – and that in principle is quite arbitrary. For instance, another way to proceed could be to set $h_1(\varphi,\psi) = 0$ for $(\varphi,\psi) \in \Omega\setminus\Omega_1$. The reason why the exact form of the extension is not relevant is that the difference between Ω_1 and Ω has measure of order ρ. As a consequence, any extended map we may consider produces corrections which are at most of order ρ. The advantage of taking $h_{1,{\rm ext}}$ as defined before (7.41) is that it has the same regularity of the original h_1.

Remark 7.20. 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 $\overline{h}$, we will 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 7.21. *Let h_1 and $\overline{h}$ be defined as in* (6.19) *and in* (2.33)*, respectively. Then, for all $\theta \in \mathcal{U}$, one has*

$$\big|\langle h_1(\theta,\cdot) - \overline{h}(\theta)\rangle\big| \le C\rho, \qquad \big|\big\langle (h_1(\theta,\cdot) - \overline{h}(\theta))^2\big\rangle\big| \le C\rho, \tag{7.42a}$$

$$\big|\langle \partial_\theta h_1(\theta,\cdot) - \partial_\theta\overline{h}(\theta)\rangle\big| \le C\rho, \qquad \big|\big\langle (\partial_\theta h_1(\theta,\cdot) - \partial_\theta\overline{h}(\theta))^2\big\rangle\big| \le C\rho. \tag{7.42b}$$

After proving Proposition 7.21, to complete the proof of Theorem 5 we need to reexpress h_1 in terms of h. This will be done in the last two Subsections 7.8.5 and 7.8.6.

7.8.1 The auxiliary map

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

$$\mathscr{S}_2(\theta,\psi):=(G_2(\theta,\psi),A_0\psi),\qquad G_2(\theta,\psi):=\theta+\rho f_2(\theta,\psi),\tag{7.43}$$

where

$$f_2(\theta,\psi):=f(\theta,\psi)-\chi(\theta)f(0,\psi).$$

Observe that, setting $F_1(\theta,\psi)=\rho f_1(\theta,\psi)$ in (6.14), with

$$f_1(\theta,\psi):=f(\theta+W(\psi),\psi)-f(W(\psi),\psi),$$

we can write

$$f_1(\theta,\psi)=f_2(\theta,\psi)+\zeta(\theta,\psi),\tag{7.44}$$

where

$$c_0(\theta):=\chi(\theta)-1,\tag{7.45a}$$

$$\zeta(\theta,\psi):=f(\theta+W(\psi),\psi)-f(\theta,\psi)-\chi(\theta)\big(f(W(\psi),\psi)-f(0,\psi)\big)+c_0(\theta)\,f(W(\psi),\psi).\tag{7.45b}$$

Remark 7.22. By (6.1), with $F=\rho f$, in (7.45b) we can write

$$\mathfrak{f}(\psi):=f(W(\psi),\psi)=\rho^{-1}(W(A_0\psi)-W(\psi)),\tag{7.46}$$

an identity which will be used at length in the following.

Remark 7.23. From the definition of the function χ it follows that $\mathscr{S}_2$ satisfies Hypotheses 1–3, provided ρ is such that

$$\rho\,\partial_\theta f_2(\theta,\psi)>-1\qquad\forall(\theta,\psi)\in\Omega.\tag{7.47}$$

Thus if we wish to use the results in Subsections 2.3 and 6.3 with the map $\mathscr{S}_2$ in place of the map $\mathscr{S}$ we need to restrict ρ to an interval $(0,\rho_0)$ with $\rho_0>0$ possibly smaller than ρ_*, as defined in Remark 2.20. 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 7.26).

Remark 7.24. 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, (7.44) shows that $\mathscr{S}_2$ can be seen as a regularization of $\mathscr{S}_1$. In particular $\mathscr{S}_2(0,\psi)=(0,A_0\psi)$, 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 Remark 2.26 and 6.5).

Remark 7.25. Instead of $\mathscr{S}_2$ one might like to consider the simpler map

$$\mathscr{S}_3(\theta,\psi):=(G_3(\theta,\psi),A_0\psi),\quad G_3(\theta,\psi):=\theta+\rho f_3(\theta,\psi),\quad f_3(\theta,\psi):=f(\theta,\psi)-f(0,\psi).\tag{7.48}$$

However, even though one has $f_3(0,\psi)=0$, it may happen that $f_3(\theta,\psi)=0$ also for some $\theta\neq0$, so that $\mathscr{S}_3$ does not satisfy Hypothesis 2. It follows that $\overline{\mathcal{W}}$ may fail to be a global attractor for $(\Omega,\mathscr{S}_3)$ and hence, in general, $\mathscr{S}_3$ cannot 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. Of course, the same goal would be achieved by assuming stronger hypotheses on the map $\mathscr{S}$, for instance by requiring the map to be uniformly contracting along the direction of the slow variable on the whole Ω; on the other hand, this would introduce too restrictive and unnecessary conditions for the results to hold.

From Theorem 2, with $\mathscr{S}_2$ instead of $\mathscr{S}$, and Remark 7.24 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

$$\mathscr{H}_2(\theta, \psi) := (\mathcal{H}_2(\theta, \psi), \psi), \qquad \mathcal{H}_2(\theta, \psi) := \theta + \theta^2 h_2(\theta, \psi), \tag{7.49}$$

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 $\mu(\psi)$ as in (7.14).

Introducing the functions, analogous to the function $p_1(\theta, \psi)$ and $q_1(\theta, \psi)$ defined in Subsection 6.3.2,

$$p_2(\theta, \psi) := \frac{1}{1 + \rho\partial_\theta f_2(0, \psi)} \left(\frac{\theta + \rho f_2(\theta, \psi)}{\theta} \right)^2, \tag{7.50a}$$

$$q_2(\theta, \psi) := \frac{\rho}{1 + \rho\,\partial_\theta f_2(0, \psi)} \frac{f_2(\theta, \psi) - \partial_\theta f_2(0, \psi)\theta}{\theta^2}, \tag{7.50b}$$

and setting

$$p_2^{(n)}(\theta, \psi) := \prod_{i=0}^{n-1} p_2(\mathscr{S}_2^i(\theta, \psi)),$$

we get

$$h_2(\theta, \psi) := \sum_{n=0}^{\infty} p_2^{(n)}(\theta, \psi)\, q_2(\mathscr{S}_2^n(\theta, \psi)) q_2(\mathscr{S}_2^n(\theta, \psi)). \tag{7.51}$$

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

$$h_1(\theta, \psi) - \overline{h}(\theta) = (h_1(\theta, \psi) - h_2(\theta, \psi)) + \big(h_2(\theta, \psi) - \overline{h}(\theta)\big), \tag{7.52a}$$

$$\partial_\theta h_1(\theta, \psi) - \partial_\theta \overline{h}(\theta) = (\partial_\theta h_1(\theta, \psi) - \partial_\theta h_2(\theta, \psi)) + \big(\partial_\theta h_2(\theta, \psi) - \partial_\theta \overline{h}(\theta)\big), \tag{7.52b}$$

and study separately the two contributions in both (7.52a) and (7.52b). This is the content of Propositions 7.32, 7.33, 7.49, and 7.50 below, which combined immediately imply Proposition 7.21. For both maps $\mathscr{S}_1$ and $\mathscr{S}_2$, the θ-component vanishes at $\theta = 0$, i.e. one has $G_1(0, \psi) = G_2(0, \psi) = 0$ for all $\psi \in \mathbb{T}^2$. However, while $\mathscr{S}_1$ depends on W and hence inherits the low regularity of the invariant manifold, the map $\mathscr{S}_2$ has the same regularity as the map $\mathscr{S}$. Therefore, through the splitting (7.52), we aim at controlling first the deviations of h_2 from $\overline{h}$ (see Subsections 7.8.2 and 7.8.3) 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 (7.45b) and the bounds in Theorem 4 (see Subsection 7.8.4). Finally, we study the deviations of h from h_1 (see Subsections 7.8.5 and 7.8.6), in order to complete the proof of Theorem 5.

Remark 7.26. For $\mathscr{S}_2$ to satisfy Hypotheses 1–3 we need to restrict the maximum value allowed for ρ to ρ_0, as defined in Remark 7.22, since condition (7.47) is more stringent than condition (2.28). However the bounds in Theorem 5 are trivially satisfied for any fixed ρ, by possibly taking a large enough constant C. Thus we may and do take for granted that the bounds hold for $\rho \geq \rho_0$. For this reason in what follows we confine ourselves to consider $\rho \in (0, \rho_0)$ and hence assume that both $\mathscr{S}$ and $\mathscr{S}_2$ satisfy Hypotheses 1–3.

7.8.2 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 (7.51), 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 (7.59) below). This comparison is similar to the comparison in Proposition 7.6, the main difference being that the analogues of g_+, $\mathfrak{u}$ and g_- now depend also on θ. Thus we need a new correlation inequality, generalizing the previous one to this new case.

The following preliminary result, proved in Appendix B.3, shows that bounds analogous to those in Lemma 7.16, 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 7.27. *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. *$|\mathfrak{p}_i(\theta)| \le 1 - \rho\gamma'$ for $|\theta| \le \theta'$ for some θ' independent of ρ and for all $i = 0, \ldots, n-1$,*

and set $\mathfrak{p}^{(n)}(\theta, \psi) := \mathfrak{p}_0^{(n)}(\mathscr{S}_2; \theta, \psi)$, with $\mathfrak{p}_0^{(n)}(\mathscr{S}_2; \theta, \psi)$ defined according to (7.29). Then one has

$$|||\mathfrak{p}^{(n)}|||^-_{\alpha_0,3} \le C\,(1 - \rho\gamma')^n, \tag{7.53}$$

where the constant C does not depend on n. From this it follows that

$$|||\partial_\theta(\mathscr{S}_2^n)_\theta|||^-_{\alpha_0,4} \le C\,(1 - \rho\gamma')^n, \tag{7.54}$$

with C independent of n.

Remark 7.28. Condition 1 in Lemma 7.27 implies that

$$|\mathfrak{p}_i|_{\alpha_0} = O(\rho), \qquad |||\partial_\theta \mathfrak{p}_i|||_{\alpha_0,2} = O(\rho), \qquad \|\mathfrak{p}_i - \langle \mathfrak{p}_i \rangle\|_\infty = O(\rho).$$

Remark 7.29. A comment analogous to Remark 7.18 applies also to Lemma 7.27 and the forthcoming Proposition 7.45: if we assume that property 2 in Lemma 7.27 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 state the new correlation inequality, which is proved in Appendix C.2.

Proposition 7.30. *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 7.27. Then, for any $g_+ \in \mathscr{B}^+_{\alpha_0}(\Omega, \mathbb{R})$ and $g_- \in \mathscr{B}^-_{\alpha_0,3}(\Omega, \mathbb{R})$, one has*

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

where $\mathfrak{p}^{(n)}(\theta, \psi) = \mathfrak{p}_0^{(n)}(\mathscr{S}_2; \theta, \psi)$ and $\langle \mathfrak{p} \rangle^{(n)}(\theta) := \langle \mathfrak{p} \rangle_0^{(n)}(\overline{\mathscr{S}}; \theta)$.

Remark 7.31. Proposition 7.30 can be seen as a generalization of Proposition 7.6 to functions which also depend smoothly on the slow variable. This will be exploited in Subsection 7.8.3 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 7.8.4, will be to compare the auxiliary map with the translated map, where the dependence on the fast variable is only α_*-Hölder continuous, with $\alpha_* = O(\rho)$.

7.8.3 Deviations of the conjugation of the auxiliary map

Recall that we are working with the extension of the map $\mathscr{S}$, although not explicitly indicated (see Remark 7.20). 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 (7.51) and (7.34), we may write

$$\begin{aligned} &\langle p_2 \rangle^{(n)} \langle q_2 \rangle \circ \overline{G}^n - \overline{p}^{(n)}\, \overline{q} \circ \overline{G}^n \\ &\quad = \langle q_2 \rangle \circ \overline{G}^n \sum_{i=0}^{n-1} \langle p_2 \rangle^{(i)} \left(\langle p_2 \rangle - \overline{p} \right) \circ \overline{G}^i\, \overline{p}^{(n-i-2)} \circ \overline{G}^{i+1} + \overline{p}^{(n)} \left(\langle q_2 \rangle \circ \overline{G}^n - \overline{q} \circ \overline{G}^n \right). \end{aligned} \tag{7.55}$$

We can now prove the following result.

Proposition 7.32. *Let h_2 and $\overline{h}$ be defined according to* (7.49) *and* (2.33)*, respectively. Then, for all $\theta \in \mathcal{U}$, one has*

$$\left|\left\langle h_2(\theta,\cdot) - \overline{h}(\theta)\right\rangle\right| \le C\rho, \qquad \left\langle (h_2(\theta,\cdot) - \overline{h}(\theta))^2\right\rangle \le C\rho\,. \tag{7.56}$$

Proof. As discussed in Remark 7.22, we can assume, without loss of generality, that $\rho \le \rho_0$. Observe that $p_2(\theta,\psi) = 1 + O(\rho)$, so that (compare with Remark 7.28)

$$\|p_2 - \langle p_2\rangle\|_\infty \le C\rho, \qquad |p_2|_{\alpha_0} \le C\rho, \qquad |||\partial_\theta p_2|||_{\alpha_0,2} \le C\rho\,. \tag{7.57}$$

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

$$|||q_2|||_{\alpha_0,3} \le C\rho, \tag{7.58}$$

so that we can apply Proposition 7.30 and obtain

$$\left|\langle p_2^{(n)} q_2 \circ \mathscr{S}_2^n\rangle - \langle p_2\rangle^{(n)} \langle q_2\rangle \circ \overline{G}^n\right| \le C(1-\rho\,\gamma')^n \rho^2\,(1+\rho n), \tag{7.59}$$

where $\langle p_2\rangle^{(n)}(\theta)$ is given by (7.31), with $\mathfrak{p}_i = p_2$ $\forall i = 0,\dots,n-1$. Since

$$|||\langle p_2\rangle - \overline{p}|||_{0,1} \le C\rho^2, \qquad |||\langle q_2\rangle - \overline{q}|||_{0,1} \le C\rho^2, \tag{7.60}$$

we have also, by (7.55),

$$\begin{aligned} &\left|\langle p_2\rangle^{(n)} \langle q_2\rangle \circ \overline{G}^n - \overline{p}^{(n)}\,\overline{q} \circ \overline{G}^n\right| \\ &\qquad \le C(1-\rho\,\gamma')^n n\rho^2 \|q_2\|_\infty + C(1-\rho\,\gamma')^n\,\|\langle q_2\rangle - \overline{q}\|_\infty \le C(1-\rho\,\gamma')^n \rho^2 (1+\rho n), \end{aligned} \tag{7.61}$$

so that

$$\left|\langle p_2^{(n)} q_2 \circ \mathscr{S}_2^n\rangle - \overline{p}^{(n)}\,\overline{q} \circ \overline{G}^n\right| \le C(1-\rho\,\gamma')^n \rho^2 (1+\rho n)\,.$$

Summing over n we get the first bound in (7.56).

For the second bound, we start considering

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

and observe that, thanks to (7.59), 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 contribution 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

$$p_3(\theta,\psi) := (p_2(\theta,\psi))^2\,, \qquad q_{3,n}(\theta,\psi) := q_2(\theta,\psi) p_2^{(n)}(\theta,\psi)\, q_2(\mathscr{S}_2^n(\theta,\psi)).$$

From Proposition 7.30 we get

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

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 7.27, of bound (7.58) and of inequality (2.10). Thus we are left with studying

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

Using again Proposition 7.30 and (7.58), we obtain

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

while the first bound in (7.57) yields $|\langle p_3\rangle - \langle p_2\rangle^2| \le C\rho^2$, which in turn gives

$$\begin{aligned}
&\Big|\langle p_3\rangle^{(n_1)}\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1} - \langle p_2\rangle^{(n_1)}\langle p_2\rangle^{(n_2)}\Big| = \Big|\Big(\langle p_3\rangle^{(n_1)} - \langle p_2\rangle^{(n_1)}\langle p_2\rangle^{(n_1)}\Big)\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1}\Big|\\
&\quad = \Big|\sum_{i=0}^{n_1}\langle p_3\rangle^{(i)}\left(\langle p_3\rangle - \langle p_2\rangle^2\right)\circ\overline{G}^{i}\,\langle p_2\rangle^{(n_2-n_1)}\circ\overline{G}^{n_1}\Big| \le C\,(1-\rho\,\gamma')^{n_1+n_2}n_1\rho^2.
\end{aligned}$$

Combining all the estimates together, we get

$$\Big\langle\Big(\sum_{n=0}^{\infty}\Big(p_2^{(n)}\,q_2\circ\mathscr{S}_2^n - \langle p_2\rangle^{(n)}\langle q_2\rangle\circ\overline{G}^n\Big)\Big)^2\Big\rangle \le C\rho. \tag{7.62}$$

Finally, proceeding like in (7.55), we get

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

which, together with (7.62), provides the second bound in (7.56). □

The following result extends the analysis above to the first derivatives of the functions h_2 and $\overline{h}$; the proof, based on the same ideas used for Proposition 7.32 up to technical intricacies, is deferred to Appendix D.1.

Proposition 7.33. *Let h_2 and $\overline{h}$ be defined as in* (7.51) *and in* (7.34)*, respectively. Then, for all $\theta \in \mathcal{U}$, one has*

$$\big|\big\langle\partial_\theta h_2(\theta,\cdot) - \partial_\theta\overline{h}(\theta)\big\rangle\big| \le C\rho, \qquad \big\langle(\partial_\theta h_2(\theta,\cdot) - \partial_\theta\overline{h}(\theta))^2\big\rangle \le C\rho\,. \tag{7.63}$$

Remark 7.34. To prove Proposition 7.33 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 will be required to prove the forthcoming Proposition 7.50 (see Remark 7.51).

7.8.4 Comparison between the translated map and the auxiliary map

In order to complete the proof of Proposition 7.21, we are left to study the contributions $h_1 - h_2$ and $\partial_\theta h_1 - \partial_\theta h_2$ in (7.52). In the light of (6.24) we have

$$h_1 - h_2 = \sum_{n=0}^{\infty}\big(p_1^{(n)}\,q_1\circ\mathscr{S}_1^n - p_2^{(n)}\,q_2\circ\mathscr{S}_2^n\big). \tag{7.64}$$

Remark 7.35. For $\mathfrak{p}_1, \mathfrak{p}_2 \in \mathscr{B}_{0,3}(\Omega, \mathbb{R})$, let $\mathfrak{p}_1^{(n)}$ and $\mathfrak{p}_2^{(n)}$ be defined as in (7.76). Reasoning as in Subsection 7.4, one may write

$$\mathfrak{p}_1^{(n)} - \mathfrak{p}_2^{(n)} = \sum_{k=0}^{n-1} \mathfrak{p}_1^{(k)} \big(\mathfrak{p}_1 \circ \mathscr{S}_1^k - \mathfrak{p}_2 \circ \mathscr{S}_2^k\big)\, \mathfrak{p}_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1}, \tag{7.65a}$$

$$\mathfrak{p}_1^{(n)} \mathfrak{q}_1 \circ \mathscr{S}_1^n - \mathfrak{p}_2^{(n)}\, \mathfrak{q}_2 \circ \mathscr{S}_2^n = \big(\mathfrak{p}_1^{(n)} - \mathfrak{p}_2^{(n)}\big)\, \mathfrak{q}_2 \circ \mathscr{S}_2^n + \mathfrak{p}_1^{(n)} \left(\mathfrak{q}_1 \circ \mathscr{S}_1^n - \mathfrak{q}_2 \circ \mathscr{S}_2^n\right), \tag{7.65b}$$

and, in a similar way, one finds

$$\mathfrak{p}_2^{(n)} \circ \mathscr{S}_1^i - \mathfrak{p}_2^{(n)} \circ \mathscr{S}_2^i = \sum_{k=0}^{n-1} \mathfrak{p}_2^{(k)} \circ \mathscr{S}_1^i \big(\mathfrak{p}_2 \circ \mathscr{S}_2^k \circ \mathscr{S}_1^i - \mathfrak{p}_2 \circ \mathscr{S}_2^k \circ \mathscr{S}_2^i\big)\, \mathfrak{p}_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1+i} \tag{7.66a}$$

$$\mathfrak{p}_1^{(n)} \circ \mathscr{S}_1^i - \mathfrak{p}_2^{(n)} \circ \mathscr{S}_2^i = \sum_{k=0}^{n-1} \mathfrak{p}_1^{(k)} \circ \mathscr{S}_1^i \big(\mathfrak{p}_1 \circ \mathscr{S}_1^{k+i} - \mathfrak{p}_2 \circ \mathscr{S}_2^{k+i}\big)\, \mathfrak{p}_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1+i}, \tag{7.66b}$$

with the latter reducing to (7.65a) for $i = 0$.

Taking into account the expansions in Remark 7.35, we may rewrite (7.64) as

$$\begin{aligned} h_1 - h_2 = & \sum_{n=0}^{\infty} p_2^{(n)} \big(q_1 \circ \mathscr{S}_1^n - q_2 \circ \mathscr{S}_2^n\big) + \sum_{n=0}^{\infty} \big(p_1^{(n)} - p_2^{(n)}\big)\, q_2 \circ \mathscr{S}_2^n \\ & + \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} p_1^{(k)} \left(p_1 \circ \mathscr{S}_1^k - p_2 \circ \mathscr{S}_2^k\right) p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1} (q_1 \circ \mathscr{S}_1^n - q_2 \circ \mathscr{S}_2^n). \end{aligned} \tag{7.67}$$

In this subsection we will estimate $h_1 - h_2$ by studying the differences that appear as summands in (7.67). To do this we will first prove a series a technical lemmas based on the structure of the difference $f_1 - f_2$. We will come back to (7.67) in Lemma 7.46 below.

To start with, we write (7.45b) as

$$\begin{aligned} \zeta(\theta, \psi) = f_1(\theta, \psi) - f_2(\theta, \psi) &= c_0(\theta)\, f(W(\psi), \psi) + \zeta_0(\theta, \psi) \\ &= c_0(\theta)\, f(W(\psi), \psi) + c_1(\theta, \psi)\, W(\psi) + c_2(\theta, \psi)\, W(\psi)^2, \end{aligned} \tag{7.68}$$

with

$$\begin{aligned} c_0(\theta) &:= \chi(\theta) - 1, \\ c_1(\theta, \psi) &:= \left(\partial_\theta f(\theta, \psi) - \chi(\theta)\, \partial_\theta f(0, \psi)\right), \\ c_2(\theta, \psi) &:= \int_0^1 \mathrm{d}t\, (1-t) \left(\partial_\theta^2 f(\theta + tW(\psi), \psi) - \chi(\theta)\, \partial_\theta^2 f(tW(\psi), \psi)\right), \end{aligned}$$

so that $|||\zeta|||_{0,2} \le C$, while we have $\langle \zeta_0^2 \rangle \le C\rho$ and $\langle (\partial_\theta \zeta_0)^2 \rangle \le C\rho$ by (2.38) in Theorem 4.

Remark 7.36. A key observation in the argument used below and in the related appendices is the following. According to (7.68) the function ζ can be written as sum of three terms. While we expect the last two terms, which depend linearly and quadratically on W, to be controlled with by relying on Theorem 4, the first one depends on $W(\psi)$ through the function $\mathfrak{f}(\psi) = f(W(\psi), \psi)$. One can write $\mathfrak{f}$ in terms of the difference $W \circ A_0 - W$ (and hence linearly in W) by using (7.46), but, in doing so, a factor ρ is lost. However, in order to compare $\mathscr{S}_1$ with $\mathscr{S}_2$, one has to deal 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 (7.93) and (7.95) in the proof of Lemma 7.46 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 7.41 and 7.43), 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, given in Appendix D.2, is based on the idea illustrated in Remark 7.36 (see the beginning of the appendix for more details).

Lemma 7.37. *For any* $\mathfrak{p}$ *in* $\mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$ *one has*

$$\mathfrak{p}\circ\mathscr{S}_1^n - \mathfrak{p}\circ\mathscr{S}_2^n = \rho\sum_{i=0}^{n}\mathfrak{C}_{\mathfrak{p},n,n-i}\circ\mathscr{S}_2^i\, W\circ A_0^i + \rho\,\mathfrak{R}_{\mathfrak{p},n}, \tag{7.69}$$

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, for* ρ *to be such that the map* $\mathscr{S}_2$ *satisfies Hypotheses 1–3,*

$$|||\mathfrak{C}_{\mathfrak{p},n,k}|||^-_{\alpha_0,2} \le C\,(1-\rho\,\gamma')^k\,|||\partial_\theta\mathfrak{p}|||_{\alpha_0,2}, \qquad k=1,\ldots,n-1, \tag{7.70a}$$

$$\sum_{k=0}^{n}|||\mathfrak{C}_{\mathfrak{p},n,k}|||^-_{\alpha_0,2} \le C\,\rho^{-1}|||\partial_\theta\mathfrak{p}|||_{\alpha_0,2}, \tag{7.70b}$$

$$\rho\|\mathfrak{R}_{\mathfrak{p},n}\|_\infty + \langle|\mathfrak{R}_{\mathfrak{p},n}|\rangle \le C\,|||\partial_\theta\mathfrak{p}|||_{\alpha_0,2}. \tag{7.70c}$$

The two next results are immediate consequences of Lemma 7.37.

Lemma 7.38. *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*

$$(\mathfrak{p}_1-\mathfrak{p}_2)(\theta,\psi) = \rho\,\mathfrak{c}_1(\theta,\psi)W(\psi) + \rho\,\mathfrak{c}_2(\theta,\psi) + \rho\,\mathfrak{c}_3(\theta)\,\mathfrak{f}(\psi), \tag{7.71}$$

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* (7.46)*, one has*

$$\mathfrak{p}_1\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n = \rho\sum_{i=0}^{n}\mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,n,n-i}\circ\mathscr{S}_2^i\,W\circ A_0^i + \rho\,\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n} + \rho\,\mathfrak{c}_3\circ\mathscr{S}_1^n\,\mathfrak{f}\circ A_0^n, \tag{7.72}$$

where

$$\mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,n,k} := \mathfrak{C}_{\mathfrak{p}_2,n,k} + \delta_{k,0}\mathfrak{c}_1, \qquad k=0,\ldots,n, \tag{7.73a}$$

$$\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n} := \mathfrak{R}_{\mathfrak{p}_2,n} + \mathfrak{c}_2\circ\mathscr{S}_1^n + (\mathfrak{c}_1\circ\mathscr{S}_1^n - \mathfrak{c}_1\circ\mathscr{S}_2^n)\,W\circ A_0^n. \tag{7.73b}$$

Proof. Write

$$\mathfrak{p}_1\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n = \mathfrak{p}_2\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n + (\mathfrak{p}_1-\mathfrak{p}_2)\circ\mathscr{S}_1^n,$$
$$\mathfrak{c}_1\circ\mathscr{S}_1^n = \mathfrak{c}_1\circ\mathscr{S}_1^n - \mathfrak{c}_1\circ\mathscr{S}_2^n + \mathfrak{c}_1\circ\mathscr{S}_2^n,$$

and use Lemma 7.37 to deal with the contribution $\mathfrak{p}_2\circ\mathscr{S}_1^n - \mathfrak{p}_2\circ\mathscr{S}_2^n$. □

Corollary 7.39. *For any* $\mathfrak{p}_2\in\mathscr{B}_{\alpha_0,3}(\Omega,\mathbb{R})$ *such that* $|||\partial_\theta\mathfrak{p}_2|||_{\alpha_0,2}\le C\rho$*, and any* $\mathfrak{p}_1\in\mathscr{B}_{0,3}(\Omega,\mathbb{R})$ *such that* (7.71) *holds, with*

$$|||\mathfrak{c}_1|||^-_{\alpha_0,2}\le C, \qquad \rho\|\mathfrak{c}_2\|_\infty + \langle|\mathfrak{c}_2|\rangle \le C\rho, \tag{7.74}$$

then the functions (7.73) *satisfy the bounds*

$$\sum_{k=0}^{n}|||\mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,n,k}|||_{\alpha_0,2}\le C, \qquad \rho\|\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n}\|_\infty + \langle|\mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n}|\rangle \le C\rho. \tag{7.75}$$

Proof. The assertion follows immediately from the bounds (7.70) and from Theorem 4. □

Remark 7.40. 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 (6.22) and p_2 and q_2 as in (7.50), the hypotheses of Lemma 7.38 are verified, with the functions $\mathfrak{c}_1$ and $\mathfrak{c}_2$ satisfying the estimates (7.74) in Corollary (7.39) in both cases. In particular, 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 one has $\mathfrak{c}_3\in C^\infty(\mathcal{U},\mathbb{R})$ and $\|\mathfrak{c}_3\|_\infty\le C$ in both cases. Note that if one restricts $\mathscr{S}$ to Λ_{2r_2}, according to Remark 7.25, c_0 is replaced with 0 and both functions $\mathfrak{a}_2$ and $\mathfrak{a}_3$ vanish. In that case, the contributions with $\mathfrak{c}_3$ in (7.71) and, as a consequence, in (7.72), disappear. In particular the coming Lemmas 7.41 and 7.43 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 all the proofs.

Corollary 7.39 allows us to deal with the first two contributions in the r.h.s. of (7.72). In order to deal with the last contribution we need also the following two results.

Lemma 7.41. *For* $\mathfrak{p}_1, \mathfrak{p}_2 \in \mathscr{B}_{0,3}(\Omega, \mathbb{R})$*, define*

$$\mathfrak{p}_1^{(n)} := \prod_{i=0}^{n-1} \mathfrak{p}_1 \circ \mathscr{S}_1^i, \qquad \mathfrak{p}_2^{(n)} := \prod_{i=0}^{n-1} \mathfrak{p}_2 \circ \mathscr{S}_2^i. \tag{7.76}$$

Assume that $\mathfrak{p}_1$ *and* $\mathfrak{p}_2$ *are such that* (7.71) *is satisfied and* $\|\mathfrak{p}_1 - 1\|_\infty \le C\rho$*. Then, if* ρ *is such that the map* $\mathscr{S}_2$ *satisfies Hypotheses 1–3, 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*

$$\begin{aligned} &\sum_{k=0}^{n-1} \mathfrak{p}_1^{(k)} \circ \mathscr{S}_1^j \, \mathfrak{a} \circ \mathscr{S}_1^{k+j} \mathfrak{p}_2^{(n-1-k)} \circ \mathscr{S}_2^{k+1+j} \big(W \circ A_0^{k+1+j} - W \circ A_0^{k+j}\big) \\ &\qquad = \sum_{k=0}^{n} \mathfrak{p}_1^{(k-1)} \circ \mathscr{S}_1^j \, \mathfrak{D}_{\mathfrak{a},n,k,j} \circ \mathscr{S}_1^{k+j} \, \mathfrak{p}_2^{(n-1-k)} \circ \mathscr{S}_2^{k+1+j} \, W \circ A_0^{k+j}, \end{aligned} \tag{7.77}$$

where

$$\mathfrak{D}_{\mathfrak{a},n,k,j} := \begin{cases} -\mathfrak{a}, & k = 0, \\ \mathfrak{a} \circ \mathscr{S}_1^{-1} \, \mathfrak{p}_2 \circ \mathscr{S}_2^{k+j} \circ \mathscr{S}_1^{-(k+j)} - \mathfrak{p}_1 \circ \mathscr{S}_1^{-1} \, \mathfrak{a}, & k = 1, \dots, n-1, \\ \mathfrak{a} \circ \mathscr{S}_1^{-1}, & k = n, \end{cases} \tag{7.78}$$

is such that

$$\|\mathfrak{D}_{\mathfrak{a},n,k,j}\|_\infty \le C\rho \, |||\mathfrak{a}|||_{0,1}, \qquad k = 1, \dots, n-1. \tag{7.79}$$

Proof. The identity (7.77) 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 (7.72) in Lemma 7.38 and (7.75) in Corollary 7.39. □

Remark 7.42. The coefficients $\mathfrak{D}_{\mathfrak{a},n,k,j}$ with $k = 0, \dots, n-1$ 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,

$$\mathfrak{D}_{\mathfrak{a},k,j} := \mathfrak{a} \circ \mathscr{S}_1^{-1} \, \mathfrak{p}_2 \circ \mathscr{S}_2^{k+j} \circ \mathscr{S}_1^{-(k+j)} - \mathfrak{p}_1 \circ \mathscr{S}_1^{-1} \, \mathfrak{a}. \tag{7.80}$$

Lemma 7.43. *Let* $\mathfrak{p}_1, \mathfrak{p}_2 \in \mathscr{B}_{0,3}(\Omega, \mathbb{R})$ *be as in Lemma 7.41, and let* $\mathfrak{p}_1^{(n)}$ *and* $\mathfrak{p}_2^{(n)}$ *be defined as in* (7.76)*. Then, for any function* $\mathfrak{a}\colon \mathcal{U} \to \mathbb{R}$ *of class* C^3*, one has*

$$\sum_{n=0}^{\infty} \mathfrak{p}_1^{(n)} \, \mathfrak{a} \circ \mathscr{S}_1^n \big(W \circ A_0^{n+1} - W \circ A_0^n\big) = \sum_{n=0}^{\infty} \mathfrak{p}_1^{(n-1)} \, \mathfrak{D}_{0,\mathfrak{a},n} \circ \mathscr{S}_1^n \, W \circ A_0^n, \tag{7.81}$$

with

$$\mathfrak{D}_{0,\mathfrak{a},n} := \begin{cases} -\mathfrak{a}, & n = 0, \\ \mathfrak{a} \circ \mathscr{S}_1^{-1} - \mathfrak{p}_1 \circ \mathscr{S}_1^{-1} \, \mathfrak{a}, & n \ge 1, \end{cases} \tag{7.82}$$

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

$$\sum_{n=0}^{\infty} \mathfrak{p}_2^{(n)}\, \mathfrak{a} \circ \mathscr{S}_r^n \big(W \circ A_0^{n+1} - W \circ A_0^n \big) = \sum_{n=0}^{\infty} \mathfrak{p}_2^{(n-1)}\, \mathfrak{D}_{r,\mathfrak{a},n} \circ \mathscr{S}_r^n\, W \circ A_0^n, \tag{7.83}$$

with

$$\mathfrak{D}_{r,\mathfrak{a},n} := \begin{cases} -\mathfrak{a}, & n = 0, \\ \mathfrak{a} \circ \mathscr{S}_r^{-1} - \mathfrak{p}_2 \circ \mathscr{S}_2^{n-1} \circ \mathscr{S}_r^{-n}\, \mathfrak{a}, & n \geq 1. \end{cases} \tag{7.84}$$

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

$$\|\mathfrak{D}_{r,\mathfrak{a},n}\|_\infty \leq C\rho\, |||\mathfrak{a}|||_{0,1}, \qquad n \geq 1, \quad r = 0, 1, 2, \tag{7.85a}$$

$$|||\mathfrak{D}_{2,\mathfrak{a},n}|||_{0,2} \leq C\rho\, |||\mathfrak{a}|||_{0,3}, \qquad n \geq 1. \tag{7.85b}$$

Proof. After checking (7.81) and (7.83) by direct computation, the bounds (7.85a) are easily obtained by writing in (7.84)

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

and using that

$$\|\mathfrak{a} \circ \mathscr{S}_r^{-1} - \mathfrak{p}_r \circ \mathscr{S}_r^{-1}\, \mathfrak{a}\|_\infty \leq \|\mathfrak{a} \circ \mathscr{S}_r^{-1} - \mathfrak{a}\|_\infty + \|(1 - \mathfrak{p}_r \circ \mathscr{S}_r^{-1})\|_\infty \|\mathfrak{a}\|_\infty \leq C\rho\, |||\mathfrak{a}|||_{0,1}, \quad r = 1, 2,$$
$$\|\mathfrak{p}_1 \circ \mathscr{S}_1^{n-1} - \mathfrak{p}_2 \circ \mathscr{S}_2^{n-1}\|_\infty \leq C\rho,$$

with the second bound holding by Lemma 7.38.

The bound (7.85b) 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 7.27. $\square$

In order to simplify the notation, it is useful to set

$$\Delta_{\mathfrak{p},W,n} := \rho \sum_{i=0}^{n} \mathfrak{C}_{\mathfrak{p},n,n-i} \circ \mathscr{S}_2^i\, W \circ A_0^i + \rho\, \mathfrak{R}_{\mathfrak{p},n}, \tag{7.86a}$$

$$\Delta_{\mathfrak{p}_1,\mathfrak{p}_2,W,n} := \rho \sum_{i=0}^{n} \mathfrak{C}_{\mathfrak{p}_1,\mathfrak{p}_2,n,n-i} \circ \mathscr{S}_2^i\, W \circ A_0^i + \rho\, \mathfrak{R}_{\mathfrak{p}_1,\mathfrak{p}_2,n}. \tag{7.86b}$$

Then we can rewrite (7.69) and (7.72), respectively, as

$$\mathfrak{p} \circ \mathscr{S}_1^n - \mathfrak{p} \circ \mathscr{S}_2^n = \Delta_{\mathfrak{p},W,n}, \tag{7.87a}$$

$$\mathfrak{p}_1 \circ \mathscr{S}_1^n - \mathfrak{p}_2 \circ \mathscr{S}_2^n = \Delta_{\mathfrak{p}_1,\mathfrak{p}_2,W,n} + \rho\, \mathfrak{c}_3 \circ \mathscr{S}_1^n\, \mathfrak{f} \circ A_0^n. \tag{7.87b}$$

In particular, by Remark 7.40 and Theorem 4, we find

$$p_1 \circ \mathscr{S}_1^n - p_2 \circ \mathscr{S}_2^n = \Delta_{p_1,p_2,W,n} + \rho\, \mathfrak{a}_1 \circ \mathscr{S}_1^n\, \mathfrak{f} \circ A_0^n, \tag{7.88a}$$

$$q_1 \circ \mathscr{S}_1^n - q_2 \circ \mathscr{S}_2^n = \Delta_{q_1,q_2,W,n} + \rho\, \mathfrak{a}_2 \circ \mathscr{S}_1^n\, \mathfrak{f} \circ A_0^n, \tag{7.88b}$$

with $\mathfrak{a}_1, \mathfrak{a}_2 \in C^\infty(\mathcal{U}, \mathbb{R})$, while $\Delta_{p_1,p_2,W,n}$ and $\Delta_{q_1,q_2,W,n}$ are such that, for any $n \geq 0$,

$$|||\Delta_{p_1,p_2,W,n}|||^-_{\alpha_0,2} \leq C\rho, \qquad |||\Delta_{q_1,q_2,W,n}|||^-_{\alpha_0,2} \leq C\rho, \tag{7.89}$$

Remark 7.44. Relations analogous to (7.89) hold also for $\Delta_{p_2,W,n}$ and $\Delta_{q_2,W,n}$. Indeed, for any $n \geq 0$, one has

$$|||\Delta_{p_2,W,n}|||^-_{\alpha_0,2} \leq C\rho, \qquad |||\Delta_{q_2,W,n}|||^-_{\alpha_0,2} \leq C\rho.$$

Moreover, by using Theorem 4, one obtains, for any $n, k \geq 0$,

$$\big\langle|\Delta_{p_2,W,n}\, W\circ A_0^k|\big\rangle \leq C\rho^2, \qquad \big\langle|\Delta_{q_2,W,n}\, W\circ A_0^k|\big\rangle \leq C\rho^2, \tag{7.90a}$$

$$\big\langle|\Delta_{p_1,p_2,W,n}\, W\circ A_0^k|\big\rangle \leq C\rho^2, \qquad \big\langle|\Delta_{p_1,p_2,W,n}\,\Delta_{q_1,q_2,W,k}|\big\rangle \leq C\rho^3. \tag{7.90b}$$

Also the next result, essentially based on Proposition 7.30, is used at length in what follows.

Proposition 7.45. *Let $\mathfrak{p}_0, \dots, \mathfrak{p}_k \in \mathscr{B}_{\alpha_0,3}(\Omega, \mathbb{R})$ satisfy the properties 1 and 2 in Lemma 7.27, and let $\mathfrak{p}^{(k)}(\theta,\psi) := \mathfrak{p}_0^{(k)}(\mathscr{S}_2;\theta,\psi)$ be defined according to (7.29). Then, 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 $i = 0,\dots,k$, 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 $i = 0,\dots,k$, let $\mathfrak{p}_i^{(k)}(\theta,\psi) := \mathfrak{p}_i^{(k)}(\mathscr{S}_2;\theta,\psi)$ be as in (7.29) with $\mathscr{S} = \mathscr{S}_2$. We have

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

by Remarks 7.4 and 7.9, so that its average can be bounded as

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

where we have used (7.24) to obtain the first term in the second line, and applied Proposition 7.30 twice, first with $g_+ = W_0$, $g_- = \mathfrak{p}_i^{(k-i)}\mathfrak{C}\,\mathfrak{u}\circ\mathscr{S}_2^{k+1-i}$ and $\mathfrak{p}^{(n)}$ replaced with $(\mathfrak{p}\mu)^{(i)}$, to obtain the second term in the second line, and next, after writing $\mathfrak{p}_{i-j}^{(k-i+j)} = \mathfrak{p}_{i-j}\,\mathfrak{p}_{i-j+1}^{(j-1)}\circ\mathscr{S}_2\,\mathfrak{p}_i^{(k-i)}\circ\mathscr{S}_2^j$, with $g_+ = 1$, $g_- = b\,\mathfrak{p}_{i-j}(\mathfrak{p}\,\mu)^{(j-1)}_{i-j+1}\circ\mathscr{S}_2\,(\mathfrak{p}_i^{(k-i)}\,\mathfrak{C})\circ\mathscr{S}_2^j\,\mathfrak{u}\circ\mathscr{S}_2^{k+1-i+j}$ and $\mathfrak{p}^{(n)}$ replaced with $\mathfrak{p}^{(i-j)}$, to obtain the last two lines.

Using once more Proposition 7.30, with $g_+ = b$, $g_- = (\mathfrak{p}_i^{(k-i)}\,\mathfrak{C})\,\mathfrak{u}\circ\mathscr{S}_2^{k+1-i}$ 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

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

Collecting all the bounds together, we obtain the assertion. $\square$

We can now come back to the study of $h_1 - h_2$ as represented in (7.67). We study separately the averages of the three contributions in (7.67), starting from the last one.

Lemma 7.46. *For p_1 and q_1 as in* (6.22)*, and p_2 and q_2 as in* (7.50)*, let $p_1^{(n)}$ and $p_2^{(n)}$ be defined as in* (6.25) *and in* (7.51)*, respectively. Then, one has*

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

Proof. Using (7.46) and the notation (7.86), we get

$$\begin{aligned}
&p_1^{(k)}(p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k)\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n)\\
&\quad = p_1^{(k)}\Delta_{p_1,p_2,W,k}\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\\
&\quad + p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1} - W\circ A_0^k)\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\\
&\quad + p_1^{(k)}\Delta_{p_1,p_2,W,k}\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^n\,(W\circ A_0^{n+1} - W\circ A_0^n)\\
&\quad + p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1} - W\circ A_0^k)\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^n\,(W\circ A_0^{n+1} - W\circ A_0^n).
\end{aligned} \tag{7.91}$$

The bounds (7.90) give immediately

$$\left|\left\langle p_1^{(k)}\Delta_{p_1,p_2,W,k}\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\right\rangle\right| \le C\rho^3\, n\,(1-\rho\,\gamma')^n. \tag{7.92}$$

Furthermore, by Lemma 7.41, we can write

$$\begin{aligned}
&\sum_{k=0}^{n-1} p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1} - W\circ A_0^k)\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\\
&\qquad = \sum_{k=0}^{n} p_1^{(k-1)}\,\mathfrak{D}_{\mathfrak{a}_1,n,k,0}\circ\mathscr{S}_1^k\, W\circ A_0^k\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\Delta_{q_1,q_2,W,n},
\end{aligned} \tag{7.93}$$

so that, by using the bounds (7.90) and (7.79), we obtain

$$\left|\sum_{k=0}^{n-1}\left\langle p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^k\,(W\circ A_0^{k+1} - W\circ A_0^k)\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n}\right\rangle\right| \le C\rho^2\, n\,(1-\rho\,\gamma')^n. \tag{7.94}$$

Analogously, using Lemma 7.43 and the expansion (7.66a) in Remark 7.35, together with the notation

(7.87a), in order to write

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{k=0}^{n-1} p_1^{(k)}\Delta_{p_1,p_2,W,k}\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\,(W\circ A_0^{n+1}-W\circ A_0^{n})\\
&=\sum_{k=0}^{\infty} p_1^{(k)}\Delta_{p_1,p_2,W,k}\sum_{n=k+1}^{\infty} p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\,(W\circ A_0^{n+1}-W\circ A_0^{n})\\
&=\sum_{k=0}^{\infty} p_1^{(k)}\Delta_{p_1,p_2,W,k}\sum_{n=0}^{\infty}\big(p_2^{(n)}\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\,(W\circ A_0^{n+1}-W\circ A_0^{n})\big)\circ\mathscr{S}_1^{k+1}\\
&+\sum_{k=0}^{\infty} p_1^{(k)}\Delta_{p_1,p_2,W,k}\sum_{n=0}^{\infty}\big(p_2^{(n)}\circ\mathscr{S}_2^{k+1}-p_2^{(n)}\circ\mathscr{S}_1^{k+1}\big)\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\,\mathfrak{f}\circ A_0^{n+k+1}\\
&=\sum_{k=0}^{\infty} p_1^{(k)}\Delta_{p_1,p_2,W,k}\sum_{n=0}^{\infty}\big(p_2^{(n-1)}\,\mathfrak{D}_{1,\mathfrak{a}_2,n}\circ\mathscr{S}_1^{n}\,W\circ A_0^{n}\big)\circ\mathscr{S}_1^{k+1}-\sum_{k=0}^{\infty} p_1^{(k)}\Delta_{p_1,p_2,W,k}\\
&-\sum_{k=0}^{\infty} p_1^{(k)}\Delta_{p_1,p_2,W,k}\sum_{n=0}^{\infty}\sum_{i=0}^{n-1} p_2^{(i)}\circ\mathscr{S}_1^{k+1}\,\Delta_{p_2\circ\mathscr{S}_2^{i},W,k+1}\, p_2^{(n-1-i)}\circ\mathscr{S}_2^{k+2+i}\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\,\mathfrak{f}\circ A_0^{n+k+1},
\end{aligned}\tag{7.95}$$

we obtain, again thanks to (7.90) and (7.85a),

$$\left|\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\Big\langle p_1^{(k)}\Delta_{p_1,p_2,W,k}\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\,(W\circ A_0^{n+1}-W\circ A_0^{n})\Big\rangle\right|\le C\,\rho. \tag{7.96}$$

Finally, using first Lemma 7.41 and Remark 7.80, hence Lemma 7.43 and thence the expansion (7.65a), we obtain

$$\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{k=0}^{n-1} p_1^{(k)}\,\mathfrak{a}_1\circ\mathscr{S}_1^{k}\,(W\circ A_0^{k+1}-W\circ A_0^{k})\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\,(W\circ A_0^{n+1}-W\circ A_0^{n})\\
&=\sum_{n=0}^{\infty}\sum_{k=0}^{n} p_1^{(k-1)}\,\mathfrak{D}_{\mathfrak{a}_1,n,k,0}\circ\mathscr{S}_1^{k}\,W\circ A_0^{k}\, p_2^{(n-1-k)}\circ\mathscr{S}_2^{k+1}\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\,(W\circ A_0^{n+1}-W\circ A_0^{n})\\
&=\sum_{k=0}^{\infty} p_1^{(k-1)}\,\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^{k}\,W\circ A_0^{k}\sum_{n=0}^{\infty} p_2^{(n)}\circ\mathscr{S}_2^{k+1}\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\,(W\circ A_0^{n+k+2}-W\circ A_0^{n+k+1})\\
&\qquad+\sum_{n=0}^{\infty} p_1^{(n-1)}\,\mathfrak{D}_{\mathfrak{a}_1,n,n,0}\circ\mathscr{S}_1^{n}\,W\circ A_0^{n}\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\,(W\circ A_0^{n+1}-W\circ A_0^{n})\\
&=\sum_{k=0}^{\infty} p_1^{(k-1)}\,\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^{k}\,W\circ A_0^{k}\sum_{n=0}^{\infty}\big(p_2^{(n-1)}\,\mathfrak{D}_{1,\mathfrak{a}_2,n}\circ\mathscr{S}_1^{n}\,W\circ A_0^{n}\big)\circ\mathscr{S}_1^{k+1}\\
&-\sum_{k=0}^{\infty} p_1^{(k-1)}\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^{k}W\circ A_0^{k}\sum_{n=0}^{\infty}\sum_{i=0}^{n-1} p_2^{(i)}\circ\mathscr{S}_1^{k+1}\Delta_{p_2\circ\mathscr{S}_2^{i},W,k+1}p_2^{(n-1-i)}\circ\mathscr{S}_2^{k+2+i}\rho\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\mathfrak{f}\circ A_0^{n+k+1};\\
&\qquad+\sum_{n=1}^{\infty} p_1^{(n-1)}\,\mathfrak{a}_1\circ\mathscr{S}_1^{n-1}\,W\circ A_0^{n}\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n}\mathfrak{f}\circ A_0^{n},
\end{aligned}$$

where, observing that

$$W\circ A_0^{n}\,\rho\,\mathfrak{f}(A_0^{n})=\frac{1}{2}\Big((W\circ A_0^{n+1})^2-(W\circ A_0^{n})^2-\big(\rho\,\mathfrak{f}\circ A_0^{n}\big)^2\Big) \tag{7.97}$$

the contribution in the last line can be written as

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

as it is easily checked proceeding as in the proof of Lemma 7.41.

Thus, relying once more on (7.90), (7.79) and (7.85a), we get the assertion. □

The first application of Proposition 7.45 is to estimate the other two contributions in (7.67). This leads to the two following lemmas, whose proof makes also use of the argument given in Remark 7.36 (in particular of Lemmas 7.37, 7.41 and 7.43, which are based on the latter).

Lemma 7.47. *For q_1 as in* (6.22)*, and p_2 and q_2 as in* (7.50)*, let $p_2^{(n)}$ be defined as in* (7.51)*. Then, one has*

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

Proof. As in the proof of Lemma 7.46 we confine ourselves to the case $\rho\in(0,\rho_0)$. Using Lemma 7.38 and Remark 7.40, we write

$$\begin{aligned}\left|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\,\big(q_1\circ\mathscr{S}_1^n-q_2\circ\mathscr{S}_2^n\big)\Big\rangle\right|&\\ \le\left|\rho\sum_{n=0}^{\infty}\sum_{k=0}^{n}\Big\langle p_2^{(n)}\mathfrak{C}_{q_1,q_2,n,n-k}\circ\mathscr{S}_2^k\,W\circ A_0^k\Big\rangle\right|&+\left|\rho\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}R_{q_1,q_2,n}\Big\rangle\right|\\ +\left|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\mathfrak{a}_2\circ\mathscr{S}_2^n\big(W\circ A_0^{n+1}-W\circ A_0^n\big)\Big\rangle\right|&+\left|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\,\rho\big(\mathfrak{a}_2\circ\mathscr{S}_1^n-\mathfrak{a}_2\circ\mathscr{S}_2^n\big)\mathfrak{f}\circ A_0^n\Big\rangle\right|.\end{aligned}\tag{7.98}$$

We use Proposition 7.45, with k and i replaced with n and k, 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 (7.70) in Lemma 7.37 to bound the second line in (7.98). In the first contribution of the third line, using Lemma 7.43, we write, for $n\ge1$,

$$\left|\sum_{n=0}^{\infty}\Big\langle p_2^{(n)}\mathfrak{a}_2\circ\mathscr{S}_2^n\big(W\circ A_0^{n+1}-W\circ A_0^n\big)\Big\rangle\right|=\left|\sum_{n=0}^{\infty}\Big\langle p_2^{(n-1)}\mathfrak{D}_{2,\mathfrak{a}_2,n}\circ\mathscr{S}_2^n\,W\circ A_0^n\Big\rangle\right|\le C\rho,\tag{7.99}$$

where the last bound follows from Proposition 7.45, with $k=n-1$, $i=n$, $\mathfrak{p}_j=p_2$ for $j=0,\dots,n-2$ (see Remark 7.29), $\mathfrak{C}=\mathfrak{D}_{2,\mathfrak{a}_2,n}$, 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 third line of (7.98), setting

$$\mathfrak{f}_0(\psi):=f(0,\psi),\qquad \mathfrak{f}\circ A_0^n-\mathfrak{f}_0\circ A_0^n=\mathfrak{g}_n\,W\circ A_0^n,\qquad \mathfrak{g}_n(\psi):=\int_0^1\partial_\theta\mathfrak{f}(tW(A_0^n\psi),A_0\psi),\tag{7.100}$$

and using Lemma 7.37, we rewrite

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

so that, by exploiting Proposition 7.45, 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 7.37, we obtain that, for any $\gamma' \in (0, \gamma)$

$$\begin{aligned}
&\left| \sum_{n=0}^{\infty} \sum_{k=0}^{n} \rho^2 \left\langle p_2^{(n)} \mathfrak{C}_{\mathfrak{a}_2,n,n-k} \circ \mathscr{S}_2^k \, W \circ A_0^k \, \mathfrak{f}_0 \circ A_0^n \right\rangle \right| \\
&\qquad = \left| \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \rho^2 \left\langle p_2^{(n+k)} \mathfrak{C}_{\mathfrak{a}_2,n+k,n} \circ \mathscr{S}_2^k \, W \circ A_0^k \, \mathfrak{f}_0 \circ A_0^{n+k} \right\rangle \right| \\
&\qquad \le \sum_{k=0}^{\infty} \rho^2 (1 - \rho \gamma')^k \big((1 + \alpha_0 k) \, \lambda^{-\alpha_0 k} + \rho + k\rho^2 + k\rho^3 \big) \sum_{n=0}^{\infty} |||\mathfrak{C}_{\mathfrak{a}_2,n+k,n}|||_{\alpha_0,2}^{-} \le C\rho, \\
&\left| \sum_{n=0}^{\infty} \rho^2 \left\langle p_2^{(n)} \, R_{\mathfrak{a}_2,n} \, \mathfrak{f}_0 \circ A_0^n \right\rangle \right| \le C \sum_{n=0}^{\infty} \rho^2 (1 - \rho \gamma')^n \big\langle \big|\mathfrak{R}_{\mathfrak{a}_2,n}\big| \big\rangle \le C\rho, \\
&\left| \sum_{n=0}^{\infty} \rho^2 \left\langle p_2^{(n)} \mathfrak{g}_n \left(\sum_{k=0}^{n} \mathfrak{C}_{\mathfrak{a}_2,n,n-k} \circ \mathscr{S}_2^k \, W \circ A_0^k + R_{\mathfrak{a}_2,n} \right) W \circ A_0^n \right\rangle \right| \le C\rho,
\end{aligned}$$

which imply the desired bound. □

Lemma 7.48. *For p_1 and q_1 as in (6.22), and p_2 and q_2 as in (7.50), let $p_1^{(n)}$ and $p_2^{(n)}$ be defined as in (6.25) and in (7.51), respectively. Then, one has*

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

Proof. As in the proof of Lemma 7.46 we confine ourselves to the case $\rho \in (0, \rho_0)$. We expand $p_1^{(n)} - p_2^{(n)}$ as in Remark 7.35, hence split $p_1^{(k)} = p_1^{(k)} - p_2^{(k)} + p_2^{(k)}$ and thence expand again $p_1^{(k)} - p_2^{(k)}$, so as to obtain

$$\begin{aligned}
&\sum_{n=0}^{\infty} \big(p_1^{(n)} - p_2^{(n)} \big) \, q_2 \circ \mathscr{S}_2^n \\
&\quad = \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} \sum_{j=0}^{k-1} p_1^{(j)} \big(p_1 \circ \mathscr{S}_1^j - p_2 \circ \mathscr{S}_2^j \big) p_2^{(k-1-j)} \circ \mathscr{S}_2^{j+1} \, \big(p_1 \circ \mathscr{S}_1^k - p_2 \circ \mathscr{S}_2^k \big) \\
&\qquad \times p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1} \, q_2 \circ \mathscr{S}_2^n \\
&\quad + \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} p_2^{(k)} \big(p_1 \circ \mathscr{S}_1^k - p_2 \circ \mathscr{S}_2^k \big) p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1} \, q_2 \circ \mathscr{S}_2^n .
\end{aligned} \tag{7.101}$$

Thus, if 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 (7.91), with n, k, q_1 and q_2 replaced with k, j, p_1 and p_2, respectively, we reason as in the proof of Lemma 7.46 to study the first sum in the r.h.s. of (7.101). In particular, by relying on Lemmas 7.41

and 7.43, and using (7.97), we can rewrite the sum as

$$
\begin{aligned}
&\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\sum_{j=0}^{k-1} 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}\\
&\quad+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\sum_{j=0}^{k} 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}\\
&\qquad+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1} 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}\\
&\quad+\sum_{j=0}^{\infty} 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}_2^{j+1}\sum_{n=0}^{\infty}\big(p_2^{(n)}\, q_2\circ\mathscr{S}_2^{n}\big)\circ\mathscr{S}_2^{k+1}\\
&\qquad-\sum_{j=0}^{\infty} 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,W,j+i+1}\, p_2^{(k-1-i)}\circ\mathscr{S}_2^{j+i+1}\\
&\qquad\qquad\times\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{k+j+1}\,\mathfrak{f}\circ A_0^{k+j+1}\sum_{n=0}^{\infty}\big(p_2^{(n)}\, q_2\circ\mathscr{S}_2^{n}\big)\circ\mathscr{S}_2^{k+1}\\
&\quad+\sum_{j=0}^{\infty} 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}\sum_{n=0}^{\infty}\big(p_2^{(n)}\, q_2\circ\mathscr{S}_2^{n}\big)\circ\mathscr{S}_2^{k+j+2}\\
&\qquad-\sum_{j=0}^{\infty} 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)}\,\mathscr{S}_1^{j+1}\,\Delta_{p_2,W,i+j+1}\, p_2^{(k-1-i)}\circ\mathscr{S}_2^{k+j+1}\\
&\qquad\qquad\times\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{k+j+1}\,\mathfrak{f}\circ A_0^{k+j+1}\sum_{n=0}^{\infty}\Big(p_2^{(n)}\, q_2\circ\mathscr{S}_2^{n}\Big)\circ\mathscr{S}_2^{k+j+2}\\
&\quad+\frac{1}{2}\sum_{k=0}^{\infty} 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\sum_{n=0}^{\infty}\big(p_2^{(n)}\, q_2\circ\mathscr{S}_2^{n}\big)\circ\mathscr{S}_2^{k+1}\\
&\qquad-\frac{1}{2}\sum_{k=0}^{\infty} 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\sum_{n=0}^{\infty}\big(p_2^{(n)}\, q_2\circ\mathscr{S}_2^{n}\big)\circ\mathscr{S}_2^{k+1},
\end{aligned}
$$

so that the average of all contributions is found to be bounded as $C\rho$.

The last sum of (7.101) is dealt with by writing once more $p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k$ according to Remark 7.40, so that we obtain

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

The average of the second line is bounded by using Proposition 7.45, with $\mathfrak{p}_i = p_2\ \forall i=1,\ldots,k-1$, $\mathfrak{C}=\mathfrak{C}_{p_1,p_2,k,k-i}$ and $\mathfrak{u}=p_2^{(n-k-1)}\, q_2\circ\mathscr{S}^{n-k-1}$. Finally, the average of the third line is bounded by splitting $\mathfrak{a}_2\circ\mathscr{S}_1^k = \mathfrak{a}_2\circ\mathscr{S}_2^k + (\mathfrak{a}_2\circ\mathscr{S}_1^k - \mathfrak{a}_2\circ\mathscr{S}_2^k)$ and reasoning as done for the last line of (7.98): the contribution with $\mathfrak{a}_2\circ\mathscr{S}_2^k$ is dealt with as (7.99), the only difference being that, when applying

Proposition 7.45, one sets $\mathfrak{u} = p_2^{(n-k-1)} \circ \mathscr{S}_2\, q_2 \circ \mathscr{S}_2^{n-k}$, while in the second contribution we expand $\mathfrak{a}_2 \circ \mathscr{S}_1^k - \mathfrak{a}_2 \circ \mathscr{S}_2^k$ as in (7.69) and write $\mathfrak{f} \circ A_0^k$ as in (7.100), so as to obtain the three contributions

$$\sum_{i=0}^{\infty} \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \rho\, p_2^{(k+i)}\, \rho\, \mathfrak{C}_{\mathfrak{a}_2,k+i,k} \circ \mathscr{S}_2^i\, W \circ A_0^i\, \mathfrak{f}_0 \circ A_0^{k+1}\, p_2^{(n-1)} \circ \mathscr{S}_2^{k+i+1}\, q_2 \circ \mathscr{S}_2^{n+k+i},$$

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} \rho\, p_2^{(k)}\, \rho\, \mathfrak{R}_{\mathfrak{a}_2,k}\, \mathfrak{f}_0 \circ A_0^{k+1}\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1}\, q_2 \circ \mathscr{S}_2^{n},$$

$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} \rho\, p_2^{(k)} \left(\rho \sum_{i=0}^{k} \mathfrak{C}_{\mathfrak{a}_2,k,k-i} \circ \mathscr{S}_2^i\, W \circ A_0^i + \rho\, \mathfrak{R}_{\mathfrak{a}_2,k} \right) \mathfrak{g}_k\, W \circ A_0^k\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1}\, q_2 \circ \mathscr{S}_2^{ni},$$

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 7.50, it completes the first step in order to prove Proposition 7.33, as outlined at the beginning of the present subsection.

Proposition 7.49. *Let h_2 and h_1 be defined according to* (7.51) *and* (6.18)*, respectively. Then, for all $\theta \in \mathcal{U}$, one has*

$$|\langle h_1(\theta,\cdot) - h_2(\theta,\cdot)\rangle| \le C\rho, \qquad \left\langle (h_1(\theta,\cdot) - h_2(\theta,\cdot))^2 \right\rangle \le C\rho\,. \tag{7.102}$$

Proof. According to (7.64) – and recalling the definitions (6.22) of p_1 and q_1, and (7.50) of p_2 and q_2, and the notation in (6.25) and in (7.51) – to obtain the first bound in (7.102) it is enough to prove that

$$\left| \sum_{n=0}^{\infty} \left\langle p_1^{(n)}\, q_1 \circ \mathscr{S}_1^n - p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n \right\rangle \right| \le C\rho, \tag{7.103}$$

for $\rho \in (0, \rho_0)$, with ρ_0 as in Remark 7.26. On the other hand, the bound (7.103) follows immediately from (7.67) and from the estimates in Lemma 7.46, 7.47 and 7.48.

To obtain the second bound in (7.102), we expand

$$\begin{aligned}
(h_1 - h_2)^2 = & \sum_{n_1,n_2=0}^{\infty} p_1^{(n_1)} \left(q_1 \circ \mathscr{S}_1^{n_1} - q_2 \circ \mathscr{S}_2^{n_1} \right) p_1^{(n_2)} \left(q_1 \circ \mathscr{S}_1^{n_2} - q_2 \circ \mathscr{S}_2^{n_2} \right) \\
& + \sum_{n_1,n_2=0}^{\infty} p_1^{(n_1)} \left(q_1 \circ \mathscr{S}_1^{n_1} - q_2 \circ \mathscr{S}_2^{n_1} \right) \left(p_1^{(n_2)} - p_2^{(n_2)} \right) q_2 \circ \mathscr{S}_2^{n_2} \\
& + \sum_{n_1,n_2=0}^{\infty} \left(p_1^{(n_1)} - p_2^{(n_1)} \right) q_2 \circ \mathscr{S}_2^{n_1}\, p_1^{(n_2)} \left(q_1 \circ \mathscr{S}_1^{n_2} - q_2 \circ \mathscr{S}_2^{n_2} \right) \\
& + \sum_{n_1,n_2=0}^{\infty} \left(p_1^{(n_1)} - p_2^{(n_1)} \right) q_2 \circ \mathscr{S}_2^{n_1} \left(p_1^{(n_2)} - p_2^{(n_2)} \right) q_2 \circ \mathscr{S}_2^{n_2},
\end{aligned}$$

where, writing, for $i = 1, 2$,

$$\begin{aligned}
q_1 \circ \mathscr{S}_1^{n_i} - q_2 \circ \mathscr{S}_2^{n_i} &= \Delta_{q_1,q_2,W,n_i} + \rho\, \mathfrak{a}_2 \circ \mathscr{S}_1^{n_i}\, \mathfrak{f} \circ A_0^{n_i}, \\
p_1^{(n_i)} - p_2^{(n_i)} &= \sum_{k_i=0}^{n-1} \mathfrak{p}_1^{(k_i)} \left(\Delta_{p_1,p_2,W,n_i} + \rho\, \mathfrak{a}_2 \circ \mathscr{S}_1^{n_i}\, \mathfrak{f} \circ A_0^{n_i} \right) \mathfrak{p}_2^{(n_i-k_i-1)} \circ \mathscr{S}_2^{k_i+1},
\end{aligned}$$

according to (7.88) and (7.65a), we obtais a sum of contributions which can be dealt with as the contributions in (7.91). 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 7.49 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 7.49, is given in Appendix D.3.

Proposition 7.50. *Let h_1 and h_2 be defined as in* (6.18) *and in* (7.51)*, respectively. Then, for all $\theta \in \mathcal{U}$, one has*

$$|\langle \partial_\theta h_1(\theta,\cdot) - \partial_\theta h_2(\theta,\cdot)\rangle| \le C\rho, \qquad \big\langle (\partial_\theta h_1(\theta,\cdot) - \partial_\theta h_2(\theta,\cdot))^2 \big\rangle \le C\rho\,. \tag{7.104}$$

Remark 7.51. By looking at the proof of Proposition 7.50 in in Appendix D.3, we see that $F = \rho f$ has to be required to belong to $\mathscr{B}_{\alpha_0,5}$, because we need to apply Proposition 7.45, where both one u and $\mathfrak{p}^{(n)}$ may contain a factor $\partial_\theta p_2$, which in turn is bounded in terms of $\partial_\varphi^2 f$ (see in particular (7.50a)). If we wanted to control the deviations only of the difference $h_1 - \overline{h}$, and not of its derivative too, we could require less regularity on F: it would be enough to have $F \in \mathscr{B}_{\alpha_0,4}$. The further condition required after (2.15) that F be in $\mathscr{B}_{\alpha_0,6}$ will be needed in order to control the deviations of the second derivative of the conjugation, which in turn will be used to estimate the deviations of the first derivative of the inverse conjugation (see Remark 7.53 below).

7.8.5 Deviations of the conjugation I: proof of the bounds (2.42a)

We can now come back to $h(\varphi,\psi) - \overline{h}(\varphi)$, with $h(\varphi,\psi)$ written according to (7.41), and estimate the whole average. Recall that we are working with the extended maps (see Remark 7.20). After rewriting $\langle h(\varphi,\cdot) - \overline{h}(\varphi)\rangle$ as

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

we bound $|\langle h_1(\varphi,\cdot) - \overline{h}(\varphi)\rangle| \le C\rho$ by Proposition 7.21, 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 7.21, together again with Theorem 4, to obtain

$$\big|\big\langle W(\cdot)\,\partial_\varphi\left(h_1(\varphi,\cdot) - \overline{h}(\varphi)\right)\big\rangle\big| \le \Big(\big\langle (W(\cdot))^2\big\rangle \big\langle \left(\partial_\varphi h_1(\varphi,\cdot) - \partial_\varphi \overline{h}(\varphi)\right)^2\big\rangle\Big)^{\frac{1}{2}} \le C\rho. \tag{7.105}$$

This concludes the proof of the first bound in (2.42a).

The second bound in (2.42a) is easily obtained by writing

$$\begin{aligned} h(\varphi,\psi) - \overline{h}(\varphi) &= h(\varphi - W(\psi),\psi) - \overline{h}(\varphi) - W(\psi)\int_0^1 \mathrm{d}t\,\partial_\varphi h(\varphi - tW(\psi),\psi) \\ &= h_1(\varphi,\psi) - \overline{h}(\varphi) - W(\psi)\int_0^1 \mathrm{d}t\,\partial_\varphi h(\varphi - tW(\psi),\psi), \end{aligned}$$

which implies

$$\begin{aligned} \left(h(\varphi,\psi) - \overline{h}(\varphi)\right)^2 &\le 2\Big(\left(h_1(\varphi,\psi) - \overline{h}(\varphi)\right)^2 + (W(\psi))^2\,\|\partial_\varphi h\|^2\Big) \\ &\le 4\,(h_1(\varphi,\psi) - h_2(\varphi,\psi))^2 + 4\left(h_2(\varphi,\psi) - \overline{h}(\varphi)\right)^2 + 2(W(\psi))^2\,\|\partial_\varphi h\|^2 \end{aligned}$$

which in turn is bounded using once more Proposition 7.21 and Theorem 4.

7.8.6 Deviations of the conjugation II: proof of the bounds (2.42b)

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

$$\begin{aligned}\partial_\varphi h(\varphi,\psi) &= \partial_\varphi h_1(\varphi,\psi) - W(\psi)\int_0^1 \mathrm{d}t\, \partial_\varphi^2 h(\varphi - tW(\psi),\psi) \\ &= \partial_\varphi h_1(\varphi,\psi) - W(\psi)\left(\partial_\varphi^2 h_1(\varphi,\psi) - \partial_\varphi^2 \overline{h}(\varphi)\right) \\ &\qquad - W(\psi)\,\partial_\varphi^2\overline{h}(\varphi) + (W(\psi))^2 \int_0^1 \mathrm{d}t\,\partial_\varphi^3 h(\varphi - tW(\psi),\psi).\end{aligned} \tag{7.106}$$

Therefore, using the second expansion in (7.106), we can bound

$$\begin{aligned}\left|\left\langle \partial_\varphi h(\varphi,\cdot) - \partial_\varphi\overline{h}(\varphi)\right\rangle\right| &\le \left|\left\langle \partial_\varphi h_1(\varphi,\cdot) - \partial_\varphi \overline{h}(\varphi)\right\rangle\right| \\ &\quad + \left|\left\langle W(\cdot)\left(\partial_\varphi^2 h_1(\varphi,\cdot) - \partial_\varphi^2\overline{h}(\varphi)\right)\right\rangle\right| + \langle W(\psi)\rangle \partial_\varphi^2 \overline{h}(\varphi) + \left\langle (W(\psi))^2\right\rangle \|\partial_\varphi^3 h\|_\infty .\end{aligned} \tag{7.107}$$

The term in the first line and the last two terms in the second line are bounded by (2.42a) and by Theorem 4, respectively. To bound the first term in the second line of (7.107) we need the following result, which is proved in Appendix D.4.

Lemma 7.52. *Let h_1 and $\overline{h}$ be defined as in* (6.18) *and in* (7.34)*, respectively. Then, for all $\theta \in \mathcal{U}$, one has*

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

Remark 7.53. As mentioned in Remark 7.51, the condition $F \in \mathscr{B}_{\alpha_0,6}$ is required to obtain the bounds in Lemma 7.52. A bound like $|\langle \partial_\theta^2 h_1(\theta,\cdot) - \partial_\theta^2\overline{h}(\theta)\rangle| \le C\rho$ 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 using the Cauchy-Schwarz inequality and applying Lemma 7.52, we get

$$\left|\left\langle W(\cdot)\left(\partial_\varphi^2 h_1(\varphi,\cdot) - \partial_\varphi^2 \overline{h}(\varphi)\right)\right\rangle\right| \le C\rho,$$

which inserted into (7.107) yields the first bound in (2.42b).

The second bound in (2.42b) is obtained by using the first expansion in (7.106), which allows us to write

$$\left|\partial_\varphi h(\varphi,\psi) - \partial_\varphi\overline{h}(\varphi)\right| \le \left|\partial_\varphi h_1(\varphi,\psi) - \partial_\varphi \overline{h}(\varphi)\right| + |W(\psi)|\,\|\partial_\varphi^2 h\|_\infty$$

and hence

$$\left\langle \left(\partial_\varphi h(\varphi,\cdot) - \partial_\varphi\overline{h}(\varphi)\right)^2\right\rangle \le 2\left\langle\left(\partial_\varphi h_1(\varphi,\cdot) - \partial_\varphi\overline{h}(\varphi)\right)^2\right\rangle + 2\|\partial_\varphi^2 h\|_\infty^2 \langle (W(\cdot))^2\rangle,$$

which immediately implies the second bound in (2.42b) by (7.63) and by Theorem 4.

7.9 Deviations of the inverse conjugation

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

$$\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) = \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta)) = \eta, \tag{7.108}$$

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

Lemma 7.54. *Let l_1 and $\bar{l}$ be defined as in* (6.19) *and in* (2.35)*, respectively. Then, for all $\eta \in \mathcal{U}_0$, one has*

$$|\langle l_1(\eta,\cdot) - \bar{l}(\eta)\rangle| \le C\rho, \qquad \langle (l_1(\eta,\cdot) - \bar{l}(\eta))^2\rangle \le C\rho\,. \tag{7.109}$$

Proof. By using (7.108) we get

$$\mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) - \mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi) = \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta)) - \mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi). \tag{7.110}$$

If we write in (7.110)

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

and use the lower bound (6.42), we obtain

$$\gamma_1^2 \big\langle \big(\mathcal{H}_1^{-1} - \overline{\mathcal{H}}^{-1}\big)^2 \big\rangle(\eta) \le \big\langle \big(\mathcal{H}_1 - \overline{\mathcal{H}}\big)^2 \big\rangle \big(\overline{\mathcal{H}}^{-1}(\eta)\big). \tag{7.111}$$

Analogously, if we write in (7.110)

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

then we find

$$\begin{aligned} \overline{\gamma}\big|\big\langle \mathcal{H}_1^{-1} - \overline{\mathcal{H}}^{-1}\big\rangle(\eta)\big| &\le \partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\big|\big\langle \mathcal{H}_1^{-1} - \overline{\mathcal{H}}^{-1}\big\rangle(\eta)\big| \\ &\le \big|\big\langle \mathcal{H}_1 - \overline{\mathcal{H}}\big\rangle\big(\overline{\mathcal{H}}^{-1}(\eta)\big)\big| \\ &+ \|\partial_\theta^2 \mathcal{H}_1\|_\infty \big\langle \big(\mathcal{H}_1^{-1} - \overline{\mathcal{H}}^{-1}\big)^2\big\rangle(\eta) \\ &+ \Big(\big\langle \big(\mathcal{H}_1^{-1} - \overline{\mathcal{H}}^{-1}\big)^2\big\rangle(\eta) \big\langle \big(\partial_\theta \mathcal{H}_1 - \partial_\theta \overline{\mathcal{H}}\big)^2\big\rangle\big(\overline{\mathcal{H}}^{-1}(\eta)\big)\Big)^{1/2}, \end{aligned} \tag{7.112}$$

where $\overline{\gamma} = O(1)$ is a lower bound on $\partial_\theta \overline{\mathcal{H}}$ (see Remark 6.15).

Thus, using that

$$\mathcal{H}_1(\theta,\psi) - \overline{\mathcal{H}}(\theta) = \theta^2\big(h_1(\theta,\psi) - \overline{h}(\theta)\big), \qquad \mathcal{H}_1^{-1}(\eta,\psi) - \overline{\mathcal{H}}^{-1}(\eta) = \eta^2\big(l_1(\eta,\psi) - \overline{l}(\eta)\big)$$

and that there exists a constant d_0 such that $d_0^{-1} \le \theta/\eta \le d_0$ for $\theta = \overline{\mathcal{H}}^{-1}(\eta)$, as discussed in Section 6.4, the bound (7.111), together with the second bound in (7.42a), gives the second bound in (7.109), which, in turn, inserted into (7.112), together with the first bound in (7.42a) and the second bound in (7.42b), yields the first bound in (7.109). □

Remark 7.55. As the proof of Lemma 7.54 shows, in order to control the deviations of $l_1 - \overline{l}$, we need an estimate on the squared deviations of $\partial_\theta h_1 - \partial_\theta \overline{h}$, so that, even if we confined ourselves to the deviations of the conjugations and the inverse conjugations, in order to deal with the inverse conjugations we would need to study the derivatives of h, and hence we should require F to be in $\mathscr{B}_{\alpha_0,5}$ and not only in $\mathscr{B}_{\alpha_0,4}$. In the same way, in order to study the deviations of the derivatives of the inverse conjugation, we will need to control the squared deviations of $\partial_\theta^2 h_1 - \partial_\theta^2 \overline{h}$ through Lemma 7.52 and hence to require F to be in $\mathscr{B}_{\alpha_0,6}$ (see Remark 7.53).

Lemma 7.56. *Let l_1 and $\overline{l}$ be defined as in (6.19) and in (2.35), respectively. Then, for all $\eta \in \mathcal{U}_0$, one has*

$$|\langle \partial_\eta l_1(\eta,\cdot) - \partial_\eta \overline{l}(\eta)\rangle| \le C\rho, \qquad \langle (\partial_\eta l_1(\eta,\cdot) - \partial_\eta \overline{l}(\eta))^2\rangle \le C\rho\,. \tag{7.113}$$

Proof. Differentiating (7.108) with respect to η, we obtain

$$\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)\,\partial_\eta \mathcal{H}_1^{-1}(\eta,\psi) = \partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\,\partial_\eta \overline{\mathcal{H}}^{-1}(\eta) = 1, \tag{7.114}$$

that implies

$$\partial_\eta \mathcal{H}_1^{-1}(\eta,\psi) - \partial_\eta \overline{\mathcal{H}}^{-1}(\eta) = -\frac{\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) - \partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))}{\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)\,\partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))}. \tag{7.115}$$

Taking the square, we find

$$\begin{aligned}&\big(\partial_\eta \mathcal{H}_1^{-1}(\eta,\psi) - \partial_\eta \overline{\mathcal{H}}^{-1}(\eta)\big)^2\\ &\qquad \le \frac{2}{\tau_1^2\overline{\tau}^2}\Big(\|\partial_\theta^2\mathcal{H}_1\|_\infty^2\big(\mathcal{H}_1^{-1}(\eta,\psi) - \overline{\mathcal{H}}^{-1}(\eta)\big)^2 + \big(\partial_\theta \mathcal{H}_1(\overline{\mathcal{H}}^{-1}(\eta),\psi) - \partial_\theta\overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\big)^2\Big),\end{aligned}$$

where γ_1 and $\overline{\gamma}$ are on lower bounds, respectively, on $\partial_\theta\mathcal{H}_1$ and $\partial_\theta\overline{\mathcal{H}}$ (see Subsection 6.4), so that

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

which, together with the second bound in (7.42b) and the second bound in (7.109), implies the second of (7.113). Rewriting (7.115) as

$$\begin{aligned}&\partial_\eta \mathcal{H}_1^{-1}(\eta,\psi),\psi) - \partial_\eta \overline{\mathcal{H}}^{-1}(\eta)\\ &\quad = -\frac{\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) - \partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))}{\big(\partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\big)^2} + \frac{\big(\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi) - \partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\big)^2}{\partial_\theta \mathcal{H}_1(\mathcal{H}_1^{-1}(\eta,\psi),\psi)\,\big(\partial_\theta \overline{\mathcal{H}}(\overline{\mathcal{H}}^{-1}(\eta))\big)^2},\end{aligned}$$

and expanding

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

we obtain eventually

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

so that the first of (7.113) follows by using the previous estimates of Propositions 7.21 and Lemmas 7.52 and 7.56. □

Recalling that $l = l_1$, by (6.20), Lemmas 7.54 and 7.56 immediately imply the following result, which is the analogue of Theorem 5 for the inverse conjugation.

Proposition 7.57. *Let l be defined as in* (2.25)*. Then there is a constant C such that, for all $\eta \in \mathcal{U}_0$,*

$$\big|\big\langle l(\eta,\cdot) - \bar{l}(\eta)\big\rangle\big| \le C\rho, \qquad \big\langle(l(\eta,\cdot) - \overline{h}(\eta))^2\big\rangle \le C\rho \tag{7.116a}$$

$$\big|\big\langle\partial_\eta l(\eta,\cdot) - \partial_\eta\bar{l}(\eta)\big\rangle\big| \le C\rho, \qquad \big\langle(\partial_\eta l(\eta,\cdot) - \partial_\eta\bar{l}(\eta))^2\big\rangle \le C\rho\,. \tag{7.116b}$$

7.10 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 7.58. *For any $\gamma' \in (0,\gamma)$, there exists a constant C such that, for all $\theta \in \mathcal{U}$ and all $n \in \mathbb{N}$, one has*

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

Proof. From (2.34), (6.17) and (6.19), we find

$$\begin{aligned}
(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta) &= \mathcal{H}_1^{-1}(\mathscr{S}_0^n(\mathcal{H}_1(\theta,\psi),\psi)) - \overline{\mathcal{H}}^{-1}(\overline{\mathscr{S}}_0^n(\overline{\mathcal{H}}(\theta),\psi)) \\
&= \kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\overline{\mathcal{H}}(\theta) \\
&+ \Big(\big(\kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi)\big)^2 - \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2\Big)\,\bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big) \\
&+ \Big(\big(\kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi)\big)^2 - \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2\Big)\Big(l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big) \\
&+ \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2\Big(l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big),
\end{aligned} \tag{7.117}$$

where, after further expanding, in the third and fourth line,

$$\begin{aligned}
&\big(\kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi)\big)^2 - \big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2 \\
&\qquad = \Big(\kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\Big)^2 + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\,\Big(\kappa^{(n)}(\psi)\,\mathcal{H}_1(\theta,\psi)\big) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\Big),
\end{aligned} \tag{7.118}$$

and, in the fourth and fifth line,

$$\begin{aligned}
&l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big) \\
&\quad = \partial_\eta\bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\,\Big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\Big) \\
&\quad + \big(\partial_\eta l_1\big(\overline{\mu}^n\overline{\mathcal{H}}(\theta), A_0^n\psi\big) - \partial_\eta\bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\,\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big) \\
&\quad + \big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)^2 \\
&\qquad\qquad \times\int_0^1 \mathrm{d}t\,(1-t)\,\partial_\eta^2 l_1\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta) + t\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big), A_0^n\psi\big) \\
&\quad + l_1\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta), A_0^n\psi\big) - \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big),
\end{aligned} \tag{7.119}$$

we write

$$\begin{aligned}
&\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta) \\
&\quad = \big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\big(\mathcal{H}_1(\theta,\psi) - \overline{\mathcal{H}}(\theta)\big) + \big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\overline{\mathcal{H}}(\theta) + \overline{\mu}^n\big(\mathcal{H}_1(\theta,\psi) - \overline{\mathcal{H}}(\theta)\big) \\
&\quad = \theta^2\big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\big(h_1(\theta,\psi) - \bar{h}(\theta)\big) + \big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\overline{\mathcal{H}}(\theta) + \theta^2\overline{\mu}^n\big(h_1(\theta,\psi) - \bar{h}(\theta)\big).
\end{aligned} \tag{7.120}$$

Thus, using Lemma 2.33, Propositions 7.21 and 7.56, and the Cauchy-Schwarz inequality, the first bound follows.

The second bound is obtained in a similar way. Shortening, for notational simplicity,

$$\begin{aligned}
\mathcal{A}_n(\theta,\psi) &:= \kappa^{(n)}\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}, \\
\lambda_1(\theta,\psi) &:= l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big), \qquad \lambda_2(\theta,\psi) := \partial_\eta l_1\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi), A_0^n\psi\big), \\
\bar{\lambda}(\theta) &:= \bar{l}\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big)\Big), \qquad \bar{\lambda}_1(\theta,\psi) := l_1\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta), A_0^n\psi\big)\Big), \\
I(\theta,\psi) &:= \int_0^1 \mathrm{d}t\,(1-t)\,\partial_\eta^2 l_1\big(\overline{\mu}^n\,\overline{\mathcal{H}}(\theta) + t\big(\kappa^{(n)}(\psi)\mathcal{H}_1(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\big), A_0^n\psi\big)
\end{aligned}$$

and using (7.118), we can rewrite (7.117) as

$$(\mathscr{S}_1^n)_\theta - (\overline{\mathscr{S}}^n)_\theta = \mathcal{A}_n + \big(\mathcal{A}_n + 2\,\overline{\mu}^n\,\overline{\mathcal{H}}\,\mathcal{A}_n\big)^2\,\lambda_1 + \big(\overline{\mu}^n\,\overline{\mathcal{H}}\big)^2\big(\lambda_1 - \bar{\lambda}\big),$$

and (7.119) as

$$\lambda_1 - \bar{\lambda} = \lambda_2\mathcal{A}_n + I_n\,\mathcal{A}_n^2 + \big(\bar{\lambda}_1 - \bar{\lambda}\big),$$

so that, bounding

$$\big((\mathscr{S}_1^n)_\theta - (\overline{\mathscr{S}}^n)_\theta\big)^2 \le 3\left(\mathcal{A}_n^2 + 2\,\mathcal{A}_n^2\big(\mathcal{A}_n^2 + 4\,(\overline{\mu}^n\,\overline{\mathcal{H}})^2\big)\lambda_1^2 + \big(\overline{\mu}^n\,\overline{\mathcal{H}}\big)^4\big(\lambda_1 - \bar{\lambda}\big)^2\right),$$

$$\big(\lambda_1 - \bar{\lambda}\big)^2 \le 3\left(\lambda_2^2\mathcal{A}_n^2 + I_n^2\,\mathcal{A}_n^4 + \big(\bar{\lambda}_1 - \bar{\lambda}\big)^2\right),$$

and, thanks to (7.120),

$$\mathcal{A}_n^2 \le 2\left(\big(\kappa^{(n)} - \overline{\mu}^n\big)^2\overline{\mathcal{H}} + \theta^4\overline{\mu}^{2n}\big(h_1 - \bar{h}\big)^2\right),$$

we obtain

$$\begin{aligned}&\big\langle\big((\mathscr{S}_1^n)_\theta(\theta,\cdot) - (\overline{\mathscr{S}}^n)_\theta(\theta)\big)^2\big\rangle\\ &\quad\le C\Big(\big(\big\langle\kappa^{(n)}(\cdot) - \overline{\mu}^n\big)^2\big\rangle + \overline{\mu}^{2n}\big\langle\big(h_1(\theta,\cdot) - \bar{h}(\theta)\big)^2\big\rangle + \overline{\mu}^{4n}\big\langle\big(l_1(\overline{\mu}^n\overline{\mathcal{H}}(\theta),\cdot) - \bar{l}(\overline{\mu}^n\overline{\mathcal{H}}(\theta)\big)^2\big\rangle\Big).\end{aligned}$$

Then, using Lemma 2.33 and Propositions 7.21 and 7.56 gives immediately the second bound. □

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 7.58. Thus, we decompose $(\mathscr{S}^n)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi)$ as done for $(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta)$, relying on (2.23) and (2.25) instead of (6.17) and (6.19), 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

$$\begin{aligned}&\kappa^{(n)}(\psi)\,\mathcal{H}(\theta,\psi) - \overline{\mu}^n\,\overline{\mathcal{H}}(\theta)\\ &\quad= \theta^2\big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\big(h(\theta,\psi) - \bar{h}(\theta)\big) + \big(\kappa^{(n)}(\psi) - \overline{\mu}^n\big)\,\overline{\mathcal{H}}(\theta) + \theta^2\overline{\mu}^n\big(h(\theta,\psi), A_0^n\psi) - \bar{h}(\theta)\big).\end{aligned}$$

Then, using the bounds of Lemma 2.33, Theorem 4, Proposition 7.57, and the Cauchy-Schwarz inequality, we obtain the bound (2.44a) of Theorem 6. Again, the bound (2.44b) is obtained by reasoning in a similar way.

8 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 2.4.5. Below, as in Subsections 7.8 to 7.10, for notational simplicity, $\mathscr{S}_1$ denotes the extended map $\mathscr{S}_{1,\mathrm{ext}}$.

8.1 The probability of deviations: proof of Theorem 7

From Theorem 1 and a direct application of Chebyshev inequality we obtain that, for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : |W(\psi)| > \delta\Big\}\Big) \le \frac{C\rho}{\delta^2}, \tag{8.1}$$

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 7.58, we obtain that, if $\gamma' \in (0,\gamma)$, then, for fixed θ and n, and for any $\delta > 0$,

$$m_0\Big(\Big\{\psi \in \mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \delta\Big\}\Big) \le \frac{C(\gamma')\,\rho}{\delta^2}, \tag{8.2}$$

for a suitable positive constant $C(\gamma')$ depending on γ' but neither on θ nor n.

Note that the set of angles ψ considered in (8.2) depends on n, even though its measure is bounded independently of the value of n. The following result provides a uniform version of the estimate above, 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 8.1. *Consider the dynamical system described by the map $\mathscr{S}$ in* (2.26) *satisfying Hypotheses 1–3. Let $\mathscr{S}_1$ be defined as in* (6.13)*. 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. Using that $|||f|||_{\alpha_0,6} = 1$, for any $\theta, \theta' \in \mathcal{U}$, we get

$$|(\mathscr{S}_1)_\theta(\theta,\psi) - (\overline{\mathscr{S}})_\theta(\theta',\psi)| \ge (1-\rho)|\theta-\theta'| - \rho|f_1(\theta,\psi) - \bar{f}(\theta)| \ge (1-\rho)|\theta-\theta'| - C\rho|\theta|\,,$$

where we have also used that $f_1(0,\psi) = \overline{f}(0) = 0$, as it follows from the definition of f_1 in (7.44) and from Remark 2.10. Iterating we obtain, for any $\theta, \theta' \in \mathcal{U}$ and for any $k \ge 0$,

$$|(\mathscr{S}_1^k)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^k_\theta(\theta',\psi)| \ge (1-\rho)^k|\theta-\theta'| - Ck(1-\rho\,\gamma')^k\rho|\theta|,$$

where we have used that $(1-\rho) < (1-\rho\,\gamma')$, since $\gamma' < \gamma < 1$ (see Remark 2.20), and that, by Remark 6.12, $|(\mathscr{S}_1^k)_\theta(\theta,\psi)| \le C(1-\rho\,\gamma')^k|\theta|$. Assuming that for some $n > 0$ we have

$$(1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)| > \delta,$$

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, for all $k \le M$,

$$Ck\rho \le \frac{\delta}{4}, \qquad \text{and} \qquad \left(\frac{1-\rho}{1-\rho\,\gamma'}\right)^k \ge \frac{1}{2}.$$

This means that for $k \leq M$ we have

$$(1-\rho\,\gamma')^{-(n+k)}|(\mathscr{S}_1^{n+k})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n+k})_\theta(\theta,\psi)| \geq \frac{\delta}{4}.$$

We thus obtain that

$$\begin{aligned}&\sum_{n=0}^{\infty} m_0\left(\left\{\psi\in\mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^{n})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n})_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right)\\ &\qquad\geq \frac{M_0\delta}{\rho}\, m_0\Big(\Big\{\psi\in\mathbb{T}^2 : \sup_{n\geq 0}(1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^{n})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n})_\theta(\theta,\psi)| > \delta\Big\}\Big).\end{aligned} \tag{8.3}$$

On the other hand, one has

$$\sum_{n=0}^{\infty} m_0\left(\left\{\psi\in\mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^{n})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n})_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right) \leq \frac{C_0(\gamma')}{\delta^2}, \tag{8.4}$$

for a suitable constant $C_0(\gamma')$ depending on γ'. This can be seen in the following way. For any $\gamma',\gamma''\in(0,\gamma)$ we find, by using (8.2),

$$m_0\left(\left\{\psi\in\mathbb{T}^2 : (1-\rho\,\gamma'')^{-n}|(\mathscr{S}_1^{n})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n})_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right) \leq \left(\frac{1-\rho\,\gamma''}{1-\rho\,\gamma'}\right)^{2n}\frac{16\,C(\gamma')\,\rho}{\delta^2},$$

with $C(\gamma')$ as in (8.2). If we invert the roles of γ' and γ'', we obtain

$$m_0\left(\left\{\psi\in\mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}_1^{n})_\theta(\theta,\psi) - (\overline{\mathscr{S}}^{n})_\theta(\theta,\psi)| > \frac{\delta}{4}\right\}\right) \leq \left(\frac{1-\rho\,\gamma'}{1-\rho\,\gamma''}\right)^{2n}\frac{16\,C(\gamma'')\,\rho}{\delta^2},$$

so that, fixing $\gamma'\in(0,\gamma)$ and taking $\gamma''=\gamma'/2$, then (8.4) follows with

$$C_0(\gamma') = 16\,C(\gamma'/2)\,\rho\sum_{n=0}^{\infty}\left(\frac{1-\rho\,\gamma'}{1-\rho\,\gamma'/2}\right)^{2n}.$$

Inserting (8.4) into (8.3) delivers the thesis. □

Remark 8.2. As a consequence of Lemma 8.1, for the overwhelming majority of initial ψ, the dynamics generated by $\mathscr{S}_1$ is very well described by the averaged dynamics generated by $\overline{\mathscr{S}}$ uniformly in n.

To deal with the map $\mathscr{S}$, 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 $\delta>0$,

$$m_0\Big(\Big\{\psi\in\mathbb{T}^2 : (1-\rho\,\gamma')^{-n}|(\mathscr{S}^{n})_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^{n})_\theta(\varphi,\psi) - W(A_0^n\psi)| > \delta\Big\}\Big) \leq \frac{C(\gamma')\,\rho}{\delta^2},$$

with the constant $C(\gamma')$ independent of both θ and n. Next, observe that we can write

$$\begin{aligned}&\big|\mathscr{S}_\varphi(\varphi,\psi) - \overline{\mathscr{S}}_\varphi(\varphi',\psi) - W(A_0\psi)\big| = \big|\mathscr{S}_\varphi(\varphi,\psi) - \overline{\mathscr{S}}_\varphi(\varphi',\psi) - W(\psi) - \rho\, f(W(\psi),\psi)\big|\\ &\quad= \big|\varphi - W(\psi) - \varphi' + \rho\big(f(\varphi,\psi) - \rho\, f(W(\psi),\psi) - \rho\,\overline{f}(\varphi - W(\psi)\big) + \rho\big(\overline{f}(\varphi - W(\psi)) - \overline{f}(\varphi')\big)\big|\\ &\quad\geq (1-\rho)\big|\varphi - W(\psi) - \varphi'\big| - C\rho\big|\varphi - W(\psi)\big|,\end{aligned}$$

so that, iterating and using that

$$\big|(\mathscr{S}^i)_\varphi(\varphi,\psi) - W(A_0^i\psi)\big| = \big|(\mathscr{S}_1^i)_\theta(\theta,\varphi)\big| \leq C(1-\rho\,\gamma')^i|\theta|,$$

where $\theta = \varphi - W(\psi)$, we find, once again thanks to Lemma 6.11,

$$\big|(\mathscr{S}^k)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^k)_\varphi(\varphi',\psi) - W(A_0^n\psi)\big| \geq (1-\rho)^k\big|\varphi - W(\psi) - \varphi'\big| - Ck(1-\rho\gamma')^k\rho\big|\varphi - W(\psi)\big|.$$

Therefore, assuming that for some $n > 0$ we have

$$(1-\rho\gamma')^{-n}|(\mathscr{S}^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi) - W(A_0^n\psi)| > \delta$$

and proceeding as done for $\mathscr{S}_1$ leads to

$$\begin{aligned}\frac{C_1(\gamma')}{\delta^2} &\geq \sum_{n=0}^{\infty} m_0\left(\left\{\psi\in\mathbb{T}^2 : (1-\rho\gamma')^{-n}|(\mathscr{S}^n)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) - W(A_0^n\psi)| > \frac{\delta}{4}\right\}\right)\\ &\geq \frac{M_0\delta}{\rho}\, m_0\Big(\Big\{\psi\in\mathbb{T}^2 : \sup_{n\geq 0}(1-\rho\gamma')^{-n}|(\mathscr{S}^n)_\varphi(\varphi,\psi) - (\overline{\mathscr{S}}^n)_\varphi(\varphi,\psi) - W(A_0^n\psi)| > \delta\Big\}\Big).\end{aligned}$$

for a suitable constant $C_1(\gamma')$. This completes the proof of Theorem 7.

8.2 Continuous time: proof of Theorem 8

From (6.16) we get

$$\begin{aligned}\widetilde{X}_t(\psi) &= (\mathscr{S}_1^{\lfloor t/\rho\rfloor})_\theta(\varphi_0 - W(\psi),\psi)\\ &\quad + (t/\rho - \lfloor t/\rho\rfloor)\left((\mathscr{S}_1^{\lfloor t/\rho\rfloor+1})_\theta(\varphi_0 - W(\psi),\psi) - (\mathscr{S}_1^{\lfloor t/\rho\rfloor})_\theta(\varphi_0 - W(\psi),\psi)\right).\end{aligned}$$

Since, by Lemma 6.11,

$$\big|(\mathscr{S}_1^k)_\theta(\varphi_0 - W(\psi),\psi) - (\mathscr{S}_1^k)_\theta(\varphi_0,\psi)\big| \leq C(1-\rho\gamma')^k|W(\psi)|, \tag{8.5}$$

by combining the bound (8.5) with Lemmas 7.58 and 2.29, we get, for $\xi\in(0,\gamma)$,

$$\sup_{t\geq 0} e^{\xi t}|\widetilde{X}_t(\psi) - \Phi_t(\varphi_0)| \leq \sup_{n\geq 0}(1-\rho\xi)^{-n}\left|(\mathscr{S}_1^n)_\theta(\theta,\psi) - (\overline{\mathscr{S}}^n)_\theta(\theta,\psi)\right| + C\left(|W(\psi)| + \rho\right),$$

so that, for every $\delta > 0$, we have

$$\lim_{\rho\to 0^+} m_0\Big(\Big\{\psi\in\mathbb{T}^2 : \sup_{t\geq 0} e^{\xi t}|\widetilde{X}_t(\psi) - \Phi_t(\varphi_0)| \geq \delta\Big\}\Big) = 0,$$

because of (2.39) and Lemma 8.1.

A Decay of correlations for Hölder continuous functions

A.1 Symbolic dynamics

Consider a Markov partition $\mathcal{Q} = \{Q_1, Q_2, \ldots, Q_s\}$ for the Anosov automorphism A_0 in (2.15), and call T be the corresponding compatibility matrix. For any Anosov automorphism A_0 on $\mathbb{T}^n$, the largest eigenvalue λ_1 of the compatibility matrix is simple (by Perron-Frobenius Theorem) and equals the largest eigenvalue $\lambda_+ = \lambda$ of A_0 [28, Ch. 6]; a stronger result holds for $n = 2$ (see Remark A.2 below). 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).

Set $\mathcal{N} = \{1, 2, \ldots, s\}$. 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. Given two sequences $\underline{\sigma}$ and $\underline{\sigma}'$, we 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

$$|\psi(\underline{\sigma}) - \psi(\underline{\sigma}')| \leq C_s\, d(\underline{\sigma},\underline{\sigma}'),$$

where C_s is the maximum diameter of the sets $Q_1, \ldots, Q_s$.

We call $C_{\underline{\sigma}}^J$ the T-compatible cylinder with base $J = (j_1, \ldots, j_q) \subset \mathbb{Z}$ and specification $\underline{\sigma} = (\sigma_{j_1}, \ldots, \sigma_{j_q}) \in \mathcal{N}_T^q$, i.e. the set of sequences $\underline{\sigma'} \in \mathcal{N}_T^{\mathbb{Z}}$ such that $\sigma'_{j_k} = \sigma_{j_k}$ for $k = 1, \ldots, q$. In the following we mainly consider sets J such that $j_1, \ldots, j_q$ are consecutive, i.e. $j_{k+1} = j_k + 1$ for $k = 1, \ldots, q-1$; in that case we also write $J = [j_1, j_q]$ and $\sigma_{j_k} = \underline{\sigma}(j_k)$ for $k = 1, \ldots, q$, and we say that the subsequence $(\sigma_{j_1}, \ldots, \sigma_{j_q})$ is T-compatible on J if $T_{\underline{\sigma}(j_k)\underline{\sigma}(j_{k+1})} > 0$ for $k = 1, \ldots, q-1$.

Finally call m the Gibbs measure associated with the Lebesgue measure on $\mathbb{T}^2$ [28, Ch. 5]. If $\Sigma(M,\sigma,\sigma')$ denotes the set of T-compatible sequences $\underline{\sigma}$ on $[-M-1, M+1]$ starting with σ and ending with σ', i.e. such that $\underline{\sigma}(-M-1) = \sigma$ and $\underline{\sigma}(M+1) = \sigma'$, one has

$$m(C_{\underline{\sigma}}^J) = \lim_{M\to\infty} \frac{|\Sigma(M,\sigma,\sigma') \cap C_{\underline{\sigma}}^J|}{|\Sigma(M,\sigma,\sigma')|}, \tag{A.1}$$

where the limit does not depend on σ and σ' (which can be fixed arbitrarily, say equal to 1).

Lemma A.1. *Let $n_1, n_2, n_1', n_2' \in \mathbb{Z}$ such that $n_1 < n_2$, $n_1' < n_2'$ and $n_1' - n_2 > a$. Consider two cylinders $C_{\underline{\sigma}_1}^{J_1}$ and $C_{\underline{\sigma}_2}^{J_2}$, with $J_1 = [n_1, n_2]$ and $J_2 = [n_1', n_2']$. Then one has*

$$1 - K_0\lambda_0^{-(n_1'-n_2)} \leq \frac{m(C_{\underline{\sigma}_1}^{J_1} \cap C_{\underline{\sigma}_2}^{J_2})}{m(C_{\underline{\sigma}_1}^{J_1})\, m(C_{\underline{\sigma}_2}^{J_2})} \leq 1 + K_0\lambda_0^{-(n_1'-n_2)},$$

with $\lambda_0 := \lambda/|\lambda_2|$, where λ_2 is the second largest eigenvalue of T in absolute value, and K_0 is a suitable constant independent of n_1, n_2, n_1', n_2'.

Proof. Calling $\#(M,\sigma,\sigma')$ the number of T-compatible sequences of length $M+1$ starting with σ end ending with σ', i.e. such that

- $\underline{\sigma} = (\underline{\sigma}(j_1), \ldots, \underline{\sigma}(j_{M+1}))$,
- $T_{\underline{\sigma}(k)\underline{\sigma}(k+1)} = 1$ for $k = 1, \ldots, M$,
- $\underline{\sigma}(j_1) = \sigma$ and $\underline{\sigma}(j_{M+1}) = \sigma'$,

then (A.1) gives

$$\begin{aligned}
m(C_{\underline{\sigma}_1}^{J_1}) &= \lim_{M\to\infty} \frac{\#(M-n_1, 1, \underline{\sigma}_1(n_1))\, \#(M-n_2, \underline{\sigma}_1(n_2), 1)}{\#(2M,1,1)}, \\
m(C_{\underline{\sigma}_2}^{J_2}) &= \lim_{M\to\infty} \frac{\#(M-n_1', 1, \underline{\sigma}_2(n_1'))\, \#(M-n_2', \underline{\sigma}_2(n_2'), 1)}{\#(2M,1,1)}, \\
m(C_{\underline{\sigma}_1}^{J_1} \cap C_{\underline{\sigma}_2}^{J_2}) &= \lim_{M\to\infty} \frac{\#(M-n_1, 1, \underline{\sigma}_1(n_1))\#(n_1'-n_2, \underline{\sigma}_1(n_2), \underline{\sigma}_2(n_1'))\#(M-n_2', \underline{\sigma}_2(n_2'), 1)}{\#(2M,1,1)},
\end{aligned} \tag{A.2}$$

where the boundary conditions on the sites $\pm(M+1)$ have been fixed equal to 1. Clearly one has

$$\#(M,\sigma,\sigma') = (T^{M+1})_{\sigma,\sigma'}\,. \tag{A.3}$$

Moreover T has s eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_s$, with $\lambda_1 = \lambda > |\lambda_k|$ for $k = 2, \ldots, s$. If $v_1, \ldots, v_s$ are the eigenvectors of T associated with the eigenvalues $\lambda_1, \ldots, \lambda_s$, respectively, and V denotes the matrix with entries $V_{ij} = (v_i)_j$, then one has

$$(T^n)_{\sigma\sigma'} = \sum_{i=1}^{s} \lambda_i^n (V^{-1})_{i\sigma} V_{i\sigma'}, \tag{A.4}$$

Replacing (A.4) in (A.3), inserting (A.3) in (A.2) and using that $\lambda_0 > 1$ delivers the thesis. □

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 [50]. Therefore we have $\lambda_0 \geq \lambda$ in Lemma A.1.

We can now extend Lemma A.1 to all bounded measurable functions as follows. 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 $J \subset [k, \infty)$.

Corollary A.3. *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. Then one has*

$$|\langle g_- g_+ \rangle - \langle g_- \rangle \langle g_+ \rangle| \leq K_1 \lambda_0^{-(k'-k)} \|g_-\|_\infty \|g_+\|_\infty \, . \tag{A.5}$$

for a suitable constant $K_1 \geq K_0$, with K_0 s in Lemma A.1.

Proof. Assume first that $k' - k > a$. Let B be a cylinder in $\mathcal{F}^{\geq k'}$ and let $\mathcal{G}_-$ be the smallest subset of $\mathcal{F}^{\leq k}$ containing all the sets A such that

$$1 - K_0 \lambda_0^{-(n_1' - n_2)} \leq \frac{m(A \cap B)}{m(A)\, m(B)} \leq 1 + K_0 \lambda_0^{-(n_1' - n_2)}. \tag{A.6}$$

Observe now that all cylinders of $\mathcal{F}^{\leq k}$ are in $\mathcal{G}_-$. If $C^{J_1}_{\underline{\sigma}_1}$ and $C^{J_2}_{\underline{\sigma}_2}$ are two cylinders in $\mathcal{F}^{\leq k}$, then $C^{J_1}_{\underline{\sigma}_1} \cap C^{J_2}_{\underline{\sigma}_2}$ is also a cylinder, while $C^{J_1}_{\underline{\sigma}_1} \cup C^{J_2}_{\underline{\sigma}_2}$ can be written as the union of all cylinders $C^{J_1 \cup J_2}_{\underline{\sigma}'}$ such that $\underline{\sigma}'$ coincides with $\underline{\sigma}_1$ on J_1 or with $\underline{\sigma}_2$ on J_2. In other words, the union of cylinders can be decomposed into a possibly much larger, disjoint union of cylinders.

Finally, if $A = \bigcup_{i=1}^k A_k$, where $A_1, \ldots, A_k$ are disjoint measurable sets and B is a measurable set, then

$$\frac{m(A \cap B)}{m(A) m(B)} = \frac{\displaystyle\sum_{i=1}^{k} m(A_i \cap B)}{\displaystyle\sum_{i=1}^{k} m(A_i) m(B)},$$

so that

$$\min_{i=1,\ldots,k} \frac{m(A_i \cap B)}{m(A_i)\, m(B)} \leq \frac{m(A \cap B)}{m(A)\, m(B)} \leq \max_{i=1,\ldots,k} \frac{m(A_i \cap B)}{m(A_i)\, m(B)} \, .$$

It follows that $\mathcal{G}_-$ contains the algebra generated by the cylinders.

On the other hand, if $\{A_i\}_{i=0}^\infty$ is a decreasing sequence of sets in $\mathcal{G}_-$, then continuity of the measure implies that also $\bigcap_{i=0}^\infty A_i \in \mathcal{G}_-$. The analogous statement for a increasing sequence follows similarly since m is a finite measure. In conclusion, $\mathcal{G}_-$ is a monotone class that contains the algebra generated by the cylinders and thus it contains $\mathcal{F}^{\leq k}$.

A similar argument for B shows that (A.6) holds for all $A \in \mathcal{F}^{\leq k}$ and $B \in \mathcal{F}^{\geq k'}$.

Now, let g_- and g_+ be simple functions, that is

$$g_-(\underline{\sigma}) = \sum_i a_i \chi^{J_i}_{A_i}(\underline{\sigma}), \qquad g_+(\underline{\sigma}) = \sum_{i'} b_{i'} \chi_{B_{i'}}(\underline{\sigma}),$$

where χ_A is the characteristic function of the set A, the sets A_i are disjoint $\mathcal{F}^{\leq k}$-measurable sets and the sets $B_{i'}$ are disjoint $\mathcal{F}^{\geq k'}$-measurable sets. We get

$$
\begin{aligned}
|\langle g_- g_+\rangle - \langle g_-\rangle\langle g_+\rangle| &= \sum_{i,i'} |a_k||b_{i'}| m(A_i \cap B_{i'}) - m(A_i)m(B_{i'})| \\
&\leq K_0 \lambda_0^{-(k'-k)} \sum_{i,i'} |a_i||b_{i'}| m(A_i)m(B_{i'}) \leq K_0 \lambda_0^{-(k'-k)} \max_i |a_i| \max_{i'} |b_{i'}| \,.
\end{aligned}
$$

We can approximate any bounded $\mathcal{F}^{\leq k}$-measurable function g_- with a sequence $g_{-,k}$ of simple functions with $\|g_{-,k}\|_\infty \leq \|g_-\|_\infty$ and $g_{-,k} \to g_-$ pointwise m-almost everywhere, and similarly for g_+. We thus find that (A.5) holds for every g_- bounded $\mathcal{F}^{\leq k}$-measurable function and g_+ bounded $\mathcal{F}^{\geq k'}$-measurable function. Finally we can remove the condition $k' - k > a$ by replacing K_0 with a suitable constant $K_1 \geq K_0$. $\square$

Remark A.4. 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 $\mathcal{F}^{\geq k}$, 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 $i \geq k'$.

A.2 Correlation functions

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 $|x| \leq C_s \lambda^{-n}$. 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 $\psi(\underline{\sigma}') = \psi(\underline{\sigma}) + xv_+$ 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 $\|\psi(\tau^{2n}\underline{\sigma}') - \psi(\tau^{2n}\underline{\sigma})\| \leq C_s \lambda^{-n}$, so that, by the previous argument, we find again that $|x| \leq C_s \lambda^{-n}$. 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 $\psi(\underline{\sigma}') = \psi(\underline{\sigma}) + xv_-$, with $|x| \leq C_s \lambda^n$.

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 $\underline{\sigma}_{(-\infty,n]}$ [28, 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

$$
\hat{g}^{(\leq n)}(\underline{\sigma}) := \int \hat{g}(\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)})\, m(\mathrm{d}\underline{\sigma}'_{(n,\infty)}|\underline{\sigma}_{(-\infty,n]}),
$$

By construction $\hat{g}^{(\leq n)}(\underline{\sigma})$ is $\mathcal{F}^{(\leq n)}$-measurable and $\langle \hat{g}^{(\leq n)}\rangle = \langle \hat{g}\rangle$, where, with a slight abuse of notation, we use $\langle\cdot\rangle$ also for the average with respect to m. Analogously we define

$$
\hat{g}^{(\geq n)}(\underline{\sigma}) := \int \hat{g}(\underline{\sigma}'_{(-\infty,n)} \vee \underline{\sigma}_{[n,\infty)})\, m(\mathrm{d}\underline{\sigma}'_{(-\infty,n)}|\underline{\sigma}_{[n,\infty)}),
$$

and considerations similar to those above still hold. Observe that $\hat{g}^{(\leq n)}(\underline{\sigma})$ is the average of $g(\psi)$ over a segment of the unstable manifold of $\psi(\underline{\sigma})$ of length $O(\lambda^{-n})$. Thus if $n \gg 1$ we expect $\hat{g}^{(\leq n)}(\underline{\sigma})$ to be very close to $\hat{g}(\underline{\sigma})$.

Finally, define

$$
\begin{aligned}
\hat{g}^{(n,+)}(\underline{\sigma}) &:= \hat{g}^{(\leq \lfloor n/\alpha_+ \rfloor)}(\underline{\sigma}) - g^{(\leq \lfloor (n-1)/\alpha_+ \rfloor)}(\underline{\sigma}), \\
\hat{g}^{(n,-)}(\underline{\sigma}) &:= \hat{g}^{(\geq \lfloor n/\alpha_- \rfloor)}(\underline{\sigma}) - g^{(\geq \lfloor (n+1)/\alpha_- \rfloor))}(\underline{\sigma})\,.
\end{aligned}
\tag{A.7}
$$

From the argument above it follows that if g is Hölder continuous along the unstable manifold we expect $\hat{g}^{(n,+)}$ and $\hat{g}^{(-n,-)}$ to be small for $n \gg 1$.

Remark A.5. Definition A.7 allows us to decompose

$$\hat{g}(\underline{\sigma}) = \hat{g}^{(\le 0)}(\underline{\sigma}) + \sum_{k=1}^{\infty} \hat{g}^{(k,+)}(\underline{\sigma}),$$

where $\hat{g}^{(k,+)}(\underline{\sigma})$ depends only on $\underline{\sigma}_{(-\infty,\lfloor k/\alpha_+\rfloor)}$.

The above discussion is made precise in the following lemma.

Lemma A.6. *If g is (α_-,α_+)-Hölder continuous on $\mathbb{T}^2$, then one has*

$$|\hat{g}^{(\le n)}(\underline{\sigma}) - g(\underline{\sigma})| \le K_2\lambda^{-\alpha_+ n}|g|^+_{\alpha_+}\,, \qquad |\hat{g}^{(\ge n)}(\underline{\sigma}) - g(\underline{\sigma})| \le K_2\lambda^{\alpha_- n}|g|^-_{\alpha_-}\,, \tag{A.8a}$$

$$|\hat{g}^{(n,+)}(\underline{\sigma})| \le K_2\lambda^{-n}|g|^+_{\alpha_+}, \qquad |\hat{g}^{(n,-)}(\underline{\sigma})| \le K_2\lambda^{n}|g|^-_{\alpha_-}, \tag{A.8b}$$

for a suitable 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, then $(\underline{\sigma})_i = (\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)})_i$ for $i \le n$ and, from the discussion at the beginning of this subsection, we get $\psi(\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)}) = \psi(\sigma) + xv_+$ with $|x| \le C_s\lambda^{-n}$, so that

$$|\hat{g}(\underline{\sigma}_{(-\infty,n]} \vee \underline{\sigma}'_{(n,\infty)}) - \hat{g}(\underline{\sigma})| \le C_s\lambda^{-n\alpha_+}|g|^+_{\alpha_+}\,.$$

Integrating over $m(\mathrm{d}\underline{\sigma}'_{(n,\infty)}|\underline{\sigma}_{(-\infty,n]})$ gives the first inequality in (A.8a). The second inequality can be 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 we get

$$\begin{aligned}|\hat{g}(\underline{\sigma}_{(-\infty,\lfloor(n-1)/\alpha_+\rfloor]} \vee \underline{\sigma}'_{(\lfloor(n-1)/\alpha_+\rfloor,\infty)}) - \hat{g}(\underline{\omega}_{\lfloor(n-1)/\alpha_+\rfloor}(\underline{\sigma}))| &\le C_s\lambda^{-\lfloor(n-1)/\alpha_+\rfloor\alpha_+}|g|^+_{\alpha_+}\\ &\le C_s\lambda^{\alpha_+}\lambda^{-(n-1)}|g|^+_{\alpha_+}\end{aligned}$$

and, analogously,

$$|\hat{g}(\underline{\sigma}_{(-\infty,\lfloor n/\alpha_+\rfloor]} \vee \underline{\sigma}'_{(\lfloor n/\alpha_+\rfloor],\infty)}) - \hat{g}(\underline{\omega}_{\lfloor(n-1)/\alpha_+\rfloor}(\underline{\sigma}))| \le C_s\lambda^{\alpha_+}\lambda^{-(n-1)}|g|^+_{\alpha_+}\,.$$

Integrating over $m(\mathrm{d}\underline{\sigma}'_{(\lfloor(n-1)/\alpha_+\rfloor,\infty)}|\underline{\sigma}_{(-\infty,\lfloor(n-1)/\alpha_+\rfloor]})$ and $m(\mathrm{d}\underline{\sigma}'_{(\lfloor n/\alpha_+\rfloor,\infty)}|\underline{\sigma}_{(-\infty,\lfloor n/\alpha_+\rfloor]})$, respectively, we get

$$|\hat{g}^{(\le\lfloor(n-1)/\alpha_+\rfloor)}(\underline{\sigma}) - \hat{g}(\underline{\omega}_{\lfloor(n-1)/\alpha_+\rfloor}(\underline{\sigma}))| \le C_s\lambda^{\alpha_+}\lambda^{-(n-1)}|g|^+_{\alpha_+},$$
$$|\hat{g}^{(\le\lfloor n/\alpha_+\rfloor)}(\underline{\sigma}) - \hat{g}(\underline{\omega}_{\lfloor(n-1)/\alpha_+\rfloor}(\underline{\sigma}))| \le C_s\lambda^{\alpha_+}\lambda^{-(n-1)}|g|^+_{\alpha_+},$$

from which the first inequality of (A.8b) follows with $K_2 = 2C_s\lambda^{1+\alpha_+}$. The second inequality is obtained in a similar way. □

A.3 A correlation inequality: proof of Proposition 2.3

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 notations of (2.1), we can assume that $\langle g_+\rangle = \langle g_-\rangle = 0$. Then, calling $\tilde{n} := \lfloor \alpha n\rfloor$, we write

$$\hat{g}_+(\underline{\sigma}) = \hat{g}_+^{(\le 0)}(\underline{\sigma}) + \sum_{k=1}^{\tilde{n}} \hat{g}_+^{(k,+)}(\underline{\sigma}) + (\hat{g}_+(\underline{\sigma}) - \hat{g}_+^{(\le\lfloor\tilde{n}/\alpha\rfloor)}(\underline{\sigma})) =: \sum_{k=0}^{\tilde{n}+1} \breve{g}_+^{(k)}(\underline{\sigma}),$$

$$\hat{g}_-(\underline{\sigma}) = \hat{g}_-^{(\ge 0)}(\underline{\sigma}) + \sum_{k'=1}^{\tilde{n}} \hat{g}_-^{(-k',-)}(\underline{\sigma}) + (\hat{g}_-(\underline{\sigma}) - \hat{g}_-^{(\ge-\lfloor\tilde{n}/\alpha\rfloor)}(\underline{\sigma})) =: \sum_{k'=0}^{\tilde{n}+1} \breve{g}_-^{(k')}(\underline{\sigma}),$$

with obvious meaning of the symbols. By Lemma A.6 we get $\|\check{g}_+^{(k)}\|_\infty \le K_2\lambda^{-k}\|g_+\|_\alpha^+$ and, analogously, $\|\check{g}_-^{(k)}\|_\infty \le K_2\lambda^{-k}\|g_-\|_\alpha^-$.

Observe that $(\tau^n\underline{\sigma})_{(-k',\infty)} = \underline{\sigma}_{(n-k',\infty)}$, so that $\check{g}_-^{(k')}(\tau^n\underline{\sigma})$ depends only on $\underline{\sigma}_{(n-k',\infty)}$. Thus if $k+k'<\tilde{n}$, by Corollary A.3 and Lemma A.6 we have

$$\left|\left\langle \check{g}_+^{(k)}\, \check{g}_-^{(k')}\circ\tau^n\right\rangle\right| \le K_1K_2\|g_+\|_\alpha^+\|g_-\|_\alpha^-\lambda^{-(k+k')}\lambda_0^{-(n-\lfloor k/\alpha\rfloor-\lfloor k'/\alpha\rfloor)},$$

while for $k+k'\ge\tilde{n}$ we have, by Lemma A.6,

$$\left|\left\langle \check{g}_+^{(k)}\, \check{g}_-^{(k')}\circ\tau^n\right\rangle\right| \le K_2\|g_+\|_\alpha^+\|g_-\|_\alpha^-\lambda^{-(k+k)}\,.$$

Summing over k and k' we get

$$|\langle g_+\, g_-\circ A_0^n\rangle| \le K_2\lambda^{-\tilde{n}}\|g_+\|_\alpha^+\|g_-\|_\alpha^-\left(\sum_{\substack{k,k'=0,\dots,\tilde{n}+1\\ k+k'<\tilde{n}}} K_1\lambda^3\lambda^{-(\tilde{n}-k-k')\frac{1-\alpha}{\alpha}} + \sum_{\substack{k,k'=0,\dots,\tilde{n}+1\\ k+k'\ge\tilde{n}}}\lambda^{\tilde{n}-k-k'}\right),$$

where we have used that $(n-\lfloor k/\alpha\rfloor-\lfloor k'/\alpha\rfloor)\ge(\tilde{n}-k-k')/\alpha-3$, so that

$$(k+k')+(n-\lfloor k/\alpha\rfloor-\lfloor k'/\alpha\rfloor)\ge\tilde{n}/\alpha-(k+k')(1-1/\alpha)-3\ge\tilde{n}+\frac{1-\alpha}{\alpha}(\tilde{n}-k-k')-3\,.$$

Thus we obtain

$$|\langle g_+\, g_-\circ A_0^n\rangle| \le \max\{1,K_1\}K_2\lambda^3\|g_+\|_\alpha^+\|g_-\|_\alpha^-\lambda^{-\tilde{n}}\left(\tilde{n}\sum_{q=0}^{\tilde{n}}\lambda^{-q\frac{1-\alpha}{\alpha}}+\sum_{q=0}^{\tilde{n}+2}(\tilde{n}+q+1)\lambda^{-q}\right)$$

from which the thesis follows observing that $\lambda^{-\tilde{n}}\le\lambda^\alpha\lambda^{-\alpha n}\le\lambda\,\lambda^{-\alpha n}$.

B Bounds on the norms of the iterated products

The following lemma is easily checked by direct computation.

Lemma B.1. *Let $\mathscr{S}$ be a map on $\mathcal{U}\times\mathbb{T}^2$ of the form (7.28), 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 $\mathcal{U}\times\mathbb{T}^2$ of class C^3 in the first variable φ. Set $\mathfrak{p}^{(n)}(\varphi,\psi):=$*

$\mathfrak{p}_0^{(n)}(\mathscr{S};\psi)$, *with* $\mathfrak{p}_0^{(n)}(\mathscr{S};\varphi,\psi)$ *defined in* (7.29). *Then one has*

$$\partial_\varphi \mathfrak{p}^{(n)} = \sum_{i=0}^{n-1} \big(\partial_\varphi \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi(\mathscr{S}^i)_\varphi\big) \prod_{\substack{j=0\\ j\neq i}}^{n-1} \big(\mathfrak{p}_j \circ \mathscr{S}^j\big), \tag{B.1a}$$

$$\begin{aligned}\partial_\varphi^2 \mathfrak{p}^{(n)} = & \sum_{i=0}^{n-1} \Big(\big(\partial_\varphi^2 \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi(\mathscr{S}^i)_\varphi\big)^2 + \big(\partial_\varphi \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi^2(\mathscr{S}^i)_\varphi\big)\Big) \prod_{\substack{j=0\\ j\neq i}}^{n-1} \big(\mathfrak{p}_j \circ \mathscr{S}^j\big) \\ & + \sum_{\substack{i,j=0\\ i\neq j}}^{n-1} \big(\partial_\varphi \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi(\mathscr{S}^i)_\varphi\big)\big(\partial_\varphi \mathfrak{p}_j \circ \mathscr{S}^j\big)\big(\partial_\varphi(\mathscr{S}^j)_\varphi\big) \prod_{\substack{k=0\\ k\neq i,j}}^{n-1} \big(\mathfrak{p}_k \circ \mathscr{S}^k\big),\end{aligned} \tag{B.1b}$$

$$\begin{aligned}\partial_\varphi^3 \mathfrak{p}^{(n)} = & \sum_{i=0}^{n-1} \Big(\big(\partial_\varphi^3 \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi(\mathscr{S}^i)_\varphi\big)^3 + 3\big(\partial_\varphi^2 \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi(\mathscr{S}^i)_\varphi\big)\big(\partial_\varphi^2(\mathscr{S}^i)_\varphi\big) \\ & + \big(\partial_\varphi \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi^3(\mathscr{S}^i)_\varphi\big)\Big) \prod_{\substack{j=0\\ j\neq i}}^{n-1} \big(\mathfrak{p}_j \circ \mathscr{S}^j\big) + 2 \sum_{\substack{i,j=0\\ i\neq j}}^{n-1} \Big(\big(\partial_\varphi^2 \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi(\mathscr{S}^i)_\varphi\big)^2 \\ & + \big(\partial_\varphi \mathfrak{p}_i \circ \mathscr{S}^i\big)\big(\partial_\varphi^2(\mathscr{S}^i)_\varphi\big)\Big)\big(\partial_\varphi \mathfrak{p}_j \circ \mathscr{S}^j\big)\big(\partial_\varphi(\mathscr{S}^j)_\varphi\big) \prod_{\substack{k=0\\ k\neq i,j}}^{n-1} \big(\mathfrak{p}_k \circ \mathscr{S}^k\big) + \sum_{\substack{i,j,k=0\\ i\neq j\neq k\neq i}}^{n-1} \big(\partial_\varphi \mathfrak{p}_i \circ \mathscr{S}^i\big) \\ & \times \big(\partial_\varphi(\mathscr{S}^i)_\varphi\big)\big(\partial_\varphi \mathfrak{p}_j \circ \mathscr{S}^j\big)\big(\partial_\varphi(\mathscr{S}^j)_\varphi\big)\big(\partial_\varphi \mathfrak{p}_k \circ \mathscr{S}^k\big)\big((\partial_\varphi(\mathscr{S}^k)_\varphi\big) \prod_{\substack{h=0\\ h\neq i,j,k}}^{n-1} \big(\mathfrak{p}_h \circ \mathscr{S}^h\big).\end{aligned} \tag{B.1c}$$

Remark B.2. 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 $\mathfrak{p}^{(n)}(\varphi) = \mathfrak{p}_0^{(n)}(\mathscr{S};\varphi)$, equations (B.1) still hold, with $\mathscr{S}(\varphi,\psi)$ and $\mathfrak{p}_i(\mathscr{S}^i(\varphi\psi))$ replaced, respectively, with $G(\varphi)$ and $\mathfrak{p}_i(G^i(\varphi))$.

B.1 A first application: completion of the proof of Theorem 2

Differentiating twice the function h_1 in (6.24) gives

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

where $\big\|\partial_\theta p_1^{(n)}\big\|_\infty$ can be bounded as in 6.36. Thus, noting that

$$\partial_\theta(q_1\circ\mathscr{S}_1^n) = (\partial_\theta q_1\circ\mathscr{S}_1^n)\,\partial_\theta\mathscr{S}_1^n, \qquad \partial_\theta^2(q_1\circ\mathscr{S}_1^n) = (\partial_\theta^2 q_1\circ\mathscr{S}_1^n)\,(\partial_\theta\mathscr{S}_1^n)^2 + (\partial_\theta q_1\circ\mathscr{S}_1^n)\,\partial_\theta^2\mathscr{S}_1^n,$$

and writing $\partial_\theta^2 p_1^{(n)}$ according to (B.1b), we can use Lemma 6.11 to obtain the bound (6.38), for a suitable constant D_3.

Finally, since

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

we find

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

where the factors which are not differentiated with respect to θ are bounded as in done to obtain (6.40), while the differentiated factors are bounded once more by relaying on (B.1a) and the bounds in Lemma 6.11.

B.2 Products independent of the fast variable: proof of Lemma 7.16

Let $\overline{N} = O(\rho^{-1})$ be defined as in Remark 7.15. Then, using that $|\overline{G}^n(\varphi)| \leq \overline{\theta}$ for $n \geq \overline{N}$ and that

$$\partial_\varphi \overline{G}^n(\varphi) = \prod_{i=0}^{n-1} \partial_\varphi \overline{G}(\overline{G}^i(\varphi)), \tag{B.2}$$

we get

$$\|\partial_\varphi \overline{G}^n\|_\infty \leq C(1-\rho\,\gamma')^n \,. \tag{B.3}$$

Then, combining (B.1a) and (7.30) together with the bound (B.3), and reasoning like in (6.34) and (6.36), we obtain

$$\left|\prod_{i=0}^{n-1} \mathfrak{p}_i(\overline{G}^i(\varphi)\right| \leq C(1-\rho\,\gamma')^n \tag{B.4a}$$

$$\left|\partial_\varphi\left(\prod_{i=0}^{n-1} \mathfrak{p}_i(\overline{G}^i(\varphi))\right)\right| \leq C\rho(1-\rho\,\gamma')^{n-1}\sum_{i=0}^{n-1}(1-\rho\,\gamma')^i \leq C(1-\rho\,\gamma')^n, \tag{B.4b}$$

so that (B.4a) gives

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

while (B.4b) gives

$$|||\mathfrak{p}^{(n)}|||_{0,1} \leq C(1-\rho\,\gamma')^n, \qquad \|\partial_\varphi^2 \overline{G}^n\|_\infty \leq C(1-\rho\,\gamma')^n, \tag{B.5}$$

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 (B.5) in (B.1b) (again see (7.30) for notations) and reasoning like in (B.4b), we get

$$|||\mathfrak{p}^{(n)}|||_{0,2} \leq C(1-\rho\,\gamma')^n, \qquad \|\partial_\varphi^3 \overline{G}^n\|_\infty \leq C(1-\rho\,\gamma')^n. \tag{B.6}$$

Using (B.6) in (B.1c) delivers the inequality (7.36).

Finally, from (B.5), (B.6) and the bound (B.3) one obtains (7.37).

B.3 Products depending on the fast variable: proof of Lemma 7.27

Assume the hypotheses of Lemma 7.27 to be satisfied. Reasoning as in the proof of Lemma 7.16, we find

$$|||\mathfrak{p}^{(n)}|||_{0,3} \leq C(1-\rho\,\gamma')^n. \tag{B.7}$$

Since $\partial_\theta(\mathscr{S}_2^n)_\theta = (\partial_\theta G_2)^{(n)}$, where $(\partial_\theta G_2)^{(n)}$ is given by (7.29) 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

$$\|\partial_\theta^k(\mathscr{S}_2^n)_\theta\|_\infty \leq C(1-\rho\,\gamma')^n, \qquad k = 0,1,2,3. \tag{B.8}$$

On the other hand writing

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

we have, for $n \geq 2$,

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

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

$$\begin{aligned} |(\mathscr{S}_2^n)_\theta|_{\alpha_0}^- &\leq |\!|\!| G_2 |\!|\!|_{\alpha_0,1} \sum_{i=0}^{n-1} (1-\rho\,\gamma')^{n-1-i}\lambda^{-\alpha_0(n-1-i)} \prod_{j=1}^{i} \|(\partial_\theta G_2)\circ \mathscr{S}_2^{n-j}\|_\infty \\ &\leq C |\!|\!| G_2 |\!|\!|_{\alpha_0,1} (1-\rho\,\gamma')^{n-1} \sum_{i=0}^{n-1} \lambda^{-\alpha_0(n-1-i)} \leq C(1-\rho\,\gamma')^n\,, \end{aligned} \tag{B.9}$$

which holds true also for $n = 1$. We can now write, by (2.7),

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

where

$$|\mathfrak{p}_i \circ \mathscr{S}_2^i|_{\alpha_0}^- \leq \|\partial_\theta \mathfrak{p}_i\|_\infty |(\mathscr{S}_2^i)_\theta|_{\alpha_0}^- + \lambda^{-\alpha_0 i}|\mathfrak{p}_i|_{\alpha_0},$$

with the first term missing if $i = 0$, so that we get

$$|\mathfrak{p}^{(n)}|_{\alpha_0}^- \leq C(1-\rho\,\gamma')^{n-1}\left(\rho\sum_{i=1}^{n-1}(1-\rho\,\gamma')^i + \rho\sum_{i=0}^{n-1}\lambda^{-\alpha_0 i}\right) \leq C(1-\rho\,\gamma')^n, \tag{B.10}$$

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

$$|\partial_\theta(\mathscr{S}_2^n)_\theta|_{\alpha_0}^- \leq C(1-\rho\,\gamma')^n. \tag{B.11}$$

Analogously, using the expression (B.1b) for $\partial_\theta^2 \mathfrak{p}^{(n)}(\theta)$ and (B.8), we get

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

that, by the same argument used in (B.10), delivers

$$|\partial_\theta \mathfrak{p}^{(n)}|_{\alpha_0}^- \leq C\,(1-\rho\,\gamma')^n, \qquad |\partial_\theta^2(\mathscr{S}_2^n)_\theta|_{\alpha_0}^- \leq C\,(1-\rho\,\gamma')^n. \tag{B.12}$$

It is now easy, by using the expression (B.1c) for $\partial_\theta^3 p^{(n)}(\theta)$ and the bounds (B.8)–(B.12), and reasoning once more as in in (B.10), to obtain

$$|\partial_\theta^2 \mathfrak{p}^{(n)}|_{\alpha_0}^- \leq C\,(1-\rho\,\gamma')^n, \tag{B.13}$$

so as to complete the proof of (7.53).

Finally the bounds (7.54) follow collecting together the bounds (B.8)–(B.12).

C A correlation inequality involving the slow variable

In this and the following Appendix D, the maps $\mathscr{S}$, $\mathscr{S}_1$ and $\mathscr{S}_2$ are meant as the extended maps which are obtained by following the procedure described in Subsection 7.1, and, analogously, the domains are meant as the extended domains where the extended maps are defined.

C.1 Some preliminary rewriting

Define $\mathfrak{p}_i^{(k)} := \mathfrak{p}_i^{(k)}(\mathscr{S}_2;\theta,\psi)$ and $\langle\mathfrak{p}\rangle_i^{(k)} := \langle\mathfrak{p}\rangle_i^{(k)}(\overline{\mathscr{S}};\theta)$, for $k=0,\dots,n-1$ and $i=0,\dots,n-k$. according to (7.29) and (7.31), respectively, and set $\mathfrak{p}^{(n)} = \mathfrak{p}_0^{(n)}$ and $\langle\mathfrak{p}\rangle^{(n)} = \langle\mathfrak{p}\rangle_0^{(n)}$. Observe that

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

Moreover we have

$$
\mathfrak{p}_i^{(n-i)}(\overline{\mathscr{S}}^i(\theta,\psi))\, g_-(\mathscr{S}_2^{n-i}(\overline{\mathscr{S}}^i(\theta,\psi))) = \mathfrak{p}_i^{(n-i)}(\overline{G}^i(\theta), A_0^i\psi)\, g_-(\mathscr{S}_2^{n-i}(\overline{G}^i(\theta), A_0^i\psi)),
\tag{C.2}
$$

that is $\langle\mathfrak{p}\rangle^{(i)}(\theta)\, \mathfrak{p}_i^{(n-i)}(\overline{\mathscr{S}}^i(\theta,\psi))\, g_-(\mathscr{S}_2^{n-i}(\overline{\mathscr{S}}^i(\theta,\psi)))$ depends on ψ only through $\psi' := A_0^i\psi$; a similar consideration holds for the second term inside the summation in (C.1). Defining

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

we get

$$
\mathfrak{p}^{(n)}(\theta,\psi)\, g_-(\mathscr{S}_2^n(\theta,\psi)) - \langle\mathfrak{p}\rangle^{(n)}(\theta)\, g_-(\overline{\mathscr{S}}^n(\theta,\psi)) = \sum_{i=0}^{n-1} \langle\mathfrak{p}\rangle^{(i)}(\theta)\, \Delta_i^{(n-i)}(\overline{\mathscr{S}}^i(\theta,\psi)).
\tag{C.3}
$$

C.2 The new correlation inequality: proof of Proposition 7.30

We start with a particular case, by assuming the function g_+ in Proposition 7.30 to have zero average. Eventually we extend the result to any $g_+ \in \mathscr{B}^+_{\alpha_0}(\Omega,\mathbb{R})$.

Lemma C.1. *Assume ρ to be such that the map $\mathscr{S}_2$ satisfies Hypotheses 1–3. Let $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ 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 $i = 0,\dots,n-1$,*
2. *$|\mathfrak{p}_i(\theta)| \le 1 - \rho\gamma'$ for $|\theta| \le \theta'$ for some $\theta' = O(1)$ in ρ and for all $i = 0,\dots,n-1$,*

and set $\mathfrak{p}^{(n)}(\theta,\psi) := \mathfrak{p}_0^{(n)}(\mathscr{S}_2;\theta,\psi)$, with $\mathfrak{p}_0^{(n)}(\mathscr{S}_2;\theta,\psi)$ defined according to (7.29). Then, for any $g_+ \in \mathscr{B}^+_{\alpha_0}(\Omega,\mathbb{R})$with $\langle g_+\rangle(\theta) \equiv 0$, and $g_- \in \mathscr{B}^-_{\alpha_0,2}(\Omega,\mathbb{R})$, 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. Note that $\Delta_i^{(n-i)}(\overline{\mathscr{S}}^i(\theta,\psi)) = \Delta_i^{(n-i)}(\overline{G}^i(\theta), A_0\psi))$ in (C.3) depends on ψ only through $\psi' := A_0^i\psi$. Thus, using Proposition 2.3, we bound

$$
\begin{aligned}
\left|\left\langle g_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle\right| &\le \left|\left\langle g_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle - \langle\mathfrak{p}\rangle^{(n)}\left\langle g_+ g_- \circ \overline{\mathscr{S}}^n \right\rangle\right| + \left|\langle\mathfrak{p}\rangle^{(n)}\right| \left|\left\langle g_+\, g_- \circ \overline{\mathscr{S}}^n \right\rangle\right| \\
&\le C\sum_{i=0}^{n-1} \|\mathfrak{p}^{(i)}\|_\infty (1+\alpha_0 i)\, \lambda^{-\alpha_0 i}\|g_+\|^+_{\alpha_0}\|\Delta_i^{(n-i)}\|^-_{\alpha_0} \\
&\quad + C\|\langle\mathfrak{p}\rangle^{(n)}\|_\infty (1+\alpha_0 n)\, \lambda^{-\alpha_0 n}\|g_+\|^+_{\alpha_0}\|\tilde{g}_-\|^-_{\alpha_0}.
\end{aligned}
\tag{C.4}
$$

Observe finally that $\mathscr{S}_2(\theta,\psi) = (\overline{G}(\theta) + \rho\Delta f(\theta,\psi), A_0\psi)$, with

$$\Delta f(\theta,\psi) := \tilde{f}(\theta,\psi) - \chi(\theta)\, f(0,\psi), \tag{C.5}$$

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

$$\begin{aligned}&\mathfrak{p}_{i+1}^{(k-1)}(\overline{G}(\theta) + \rho\Delta f(\theta,\psi), A_0\psi) - \mathfrak{p}_{i+1}^{(k-1)}(\overline{G}(\theta), A_0\psi) \\ &\qquad = \rho\Delta f(\theta,\psi)\int_0^1 \mathrm{d}t\, \partial_\theta \mathfrak{p}_{i+1}^{(k-1)}(\overline{G}(\theta) + t\rho\Delta f(\theta,\psi), A_0\psi),\end{aligned} \tag{C.6}$$

and

$$\begin{aligned}&g_-(\mathscr{S}_2^{k-1}(\overline{G}(\theta) + \rho\Delta f(\theta,\psi)) - g_-(\mathscr{S}_2^{k-1}(\overline{G}(\theta), A_0\psi)) = \rho\Delta f(\theta,\psi)\int_0^1 \mathrm{d}t \\ &\quad \times \left(\big(\partial_\theta g_-\circ\mathscr{S}_2^{k-1}(\overline{G}(\theta) + t\rho\Delta f(\theta,\psi), A_0\psi)\big)\,\big(\partial_\theta(\mathscr{S}_2^{k-1})_\theta(\overline{G}(\theta) + t\rho\Delta f(\theta,\psi), A_0\psi)\big)\right),\end{aligned} \tag{C.7}$$

and using that

$$\big|\partial_\theta g_-\circ\mathscr{S}_2^k\big|_{\alpha_0}^- \le \|\partial_\theta^2 g_-\circ\mathscr{S}_2^k\|_\infty \big|(\mathscr{S}_2^k)_\theta\big|_{\alpha_0}^- + \lambda^{-\alpha_0 k}\big|\partial_\theta g_-\big|_{\alpha_0}^-,$$

then the assumptions on the functions $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ and the bounds provided in Lemma 7.27 and Remark 7.28 give

$$\|\Delta_i^{(k)}\|_{\alpha_0}^- \le C\rho(1-\rho\,\gamma')^{k-1}|||g_-|||_{\alpha_0,2}^-,$$

which, inserted into (C.4), completes the proof. □

Remark C.2. Note that, using (C.6) and (C.7) and still relying on the bounds of Lemma 7.27 and assuming g_- to be in $\mathscr{B}_{\alpha_0,3}^-(\Omega,\mathbb{R})$, in the same way we get the bound on $\|\Delta^{(k)}\|_{\alpha_0}^-$ we may prove more generally the bound

$$|||\Delta^{(k)}|||_{\alpha_0,2}^- \le C\rho(1-\rho\,\gamma')^{k-1}|||g_-|||_{\alpha_0,3}^-.$$

Lemma C.3. *Assume ρ to be such that the map $\mathscr{S}_2$ satisfies Hypotheses 1–3. Let $\mathfrak{p}_0,\dots,\mathfrak{p}_{n-1}$ be any functions in $\mathscr{B}_{0,3}(\Omega,\mathbb{R})$ as in Lemma C.1, and let $\langle\mathfrak{p}\rangle^{(n)}$ defined as in (7.29). Then, for any $g_-\in\mathscr{B}_{\alpha_0,2}^-(\Omega,\mathbb{R})$, if one defines*

$$\Xi^{(k)}(\theta,\psi) := \langle\mathfrak{p}\rangle^{(k)}(\theta)\, g_-(\overline{\mathscr{S}}^k(\theta,\psi)),$$

one has

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

Proof. Writing

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

with $\Delta f(\theta,\psi)$ as in (C.5), 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 7.16 and C.1,

$$\left|\left\langle \Delta f\, \partial_\theta \Xi^{(k)} \circ \overline{\mathscr{S}} \right\rangle\right| \le C(1+\alpha_0 k)\, \lambda^{-\alpha_0 k} \|\Delta f\|_{\alpha_0} (1-\rho\,\gamma')^k |||g_-|||^-_{\alpha_0,2},$$

while, using (2.10), we find, again by Lemma 7.16,

$$\|\partial_\theta^2 \Xi^{(k)}\|_\infty \le C|||\langle \mathfrak{p}\rangle^{(k)}|||_{0,2}\, |||g_- \circ \overline{\mathscr{S}}^{(k)}|||_{0,2} \le C(1-\rho\,\gamma')^k |||g_-|||_{0,2}\,.$$

Combining the above estimates we obtain the desired bound. □

Now we can prove Proposition 7.30. Assume ρ to be such that the map $\mathscr{S}_2$ satisfies Hypotheses 1–3, and observe that

$$\left\langle g_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle = \langle g_+\rangle \left\langle \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle + \left\langle \tilde{g}_+ \mathfrak{p}^{(n)} g_- \circ \mathscr{S}_2^n \right\rangle,$$

where the second term can be bounded using Lemma C.1. If we write the first term using (C.3), we see we need to estimate $\langle \Delta_i^{(n-i)}\rangle$. We can write, for any $k \ge 1$,

$$\begin{aligned} \Delta_i^{(k)}(\theta,\psi) &= (\mathfrak{p}_i(\theta,\psi) - \langle \mathfrak{p}_i\rangle(\theta)) \\ &\times \Big(\mathfrak{p}_{i+1}^{(k-1)}(\mathscr{S}_2(\theta,\psi))\, g_-(\mathscr{S}_2^k(\theta,\psi)) - \mathfrak{p}_{i+1}^{(k-1)}(\overline{\mathscr{S}}(\theta,\psi))\, g_-(\mathscr{S}_2^{k-1}(\overline{\mathscr{S}}(\theta,\psi)))\Big) \\ &+ (\mathfrak{p}_i(\theta,\psi) - \langle \mathfrak{p}_i\rangle(\theta))\, \mathfrak{p}_{i+1}^{(k-1)}(\overline{\mathscr{S}}(\theta,\psi))\, g_-(\mathscr{S}_2^{k-1}(\overline{\mathscr{S}}(\theta,\psi))) \\ &+ \langle \mathfrak{p}_i\rangle(\theta) \Big(\mathfrak{p}_{i+1}^{(k-1)}(\mathscr{S}_2(\theta,\psi))\, g_-(\mathscr{S}_2^k(\theta,\psi)) - \mathfrak{p}_{i+1}^{(k-1)}(\overline{\mathscr{S}}(\theta,\psi))\, g_-(\mathscr{S}_2^{k-1}(\overline{\mathscr{S}}(\theta,\psi)))\Big). \end{aligned} \tag{C.8}$$

We have $\|(\mathfrak{p}_i(\theta,\psi) - \langle \mathfrak{p}_i\rangle(\theta))\|_\infty \le C\rho$, while, reasoning like when studying (C.6) and (C.7), we get

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

so that the average of the first contribution in the r.h.s. of (C.8) is bounded by $C\rho^2(1-\rho\,\gamma')^k |||g_-|||_{0,1}$. A similar bound for the contribution in the third line of (C.8) is obtained as in the proof of Lemma C.1, by using (C.6) and (C.7), with the gain of a further factor ρ because of the factor $\mathfrak{p}_i - \langle \mathfrak{p}_i\rangle$. As to the contribution in the last line, we use (C.3) twice and write

$$\begin{aligned} &\mathfrak{p}_{i+1}^{(k-1)}(\mathscr{S}_2(\theta,\psi))\, g_-(\mathscr{S}_2^k(\theta,\psi)) - \mathfrak{p}_{i+1}^{(k-1)}(\overline{\mathscr{S}}(\theta,\psi))\, g_-(\mathscr{S}_2^{k-1}(\overline{\mathscr{S}}(\theta,\psi))) \\ &= \sum_{j=0}^{k-2} \Big(\langle \mathfrak{p}\rangle_{i+1}^{(j)}(G_2(\theta,\psi)) \Delta_j^{(k-1-j)}(\overline{\mathscr{S}}^j(\mathscr{S}_2(\theta,\psi))) - \langle \mathfrak{p}\rangle_{i+1}^{(j)}(\overline{G}(\theta)) \Delta_j^{(k-1-j)}(\overline{\mathscr{S}}^{j+1}(\theta,\psi)) \Big) \\ &+ \langle \mathfrak{p}\rangle_{i+1}^{(k-1)}(G_2(\theta,\psi))\, g_-(\overline{\mathscr{S}}^{k-1}(\mathscr{S}_2(\theta,\psi))) - \langle \mathfrak{p}\rangle_{i+1}^{(k-1)}(\overline{G}(\theta))\, g_-(\overline{\mathscr{S}}^k(\theta,\psi)). \end{aligned} \tag{C.9}$$

Thus, we apply Lemma C.3 first to the last line, which gives

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

and then to the second line, with $g_- = \Delta_j^{(k-1-j)}$, so as to obtain

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

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 C.2. This concludes the proof of Proposition 7.30.

D Proof of some technical results

Recall that, as mentioned at the beginning of Appendix C.2, $\mathscr{S}$, $\mathscr{S}_1$ and $\mathscr{S}_2$ are a shortened notation for the corresponding extended maps.

D.1 Derivative of the auxiliary map: proof of Proposition 7.33

In the light of Remark 7.26, we may and so assume that $\rho < \rho_0$. First of all we note that

$$\partial_\theta\,(p_2^{(n)}\,q_2{\circ}\mathscr{S}_2^n) = \partial_\theta p_2^{(n)}\,q_2{\circ}\mathscr{S}_2^n + p_2^{(n)}\,\partial_\theta(q_2{\circ}\mathscr{S}_2^n), \tag{D.1a}$$

$$\partial_\theta\,(\overline{p}^{(n)}\,\overline{q}{\circ}\overline{G}^n) = \partial_\theta \overline{p}^{(n)}\,\overline{q}{\circ}\overline{G}^n + \overline{p}^{(n)}\,\partial_\theta(\overline{q}{\circ}\overline{G}^n), \tag{D.1b}$$

where both $\partial_\theta p_2^{(n)}$ and $\partial_\theta \overline{p}^{(n)}$ can be written according to (B.1a), so that we can bound separately

$$\langle \partial_\theta p_2^{(n)}\,q_2{\circ}\mathscr{S}_2^n\rangle - \partial_\theta\overline{p}^{(n)}\,\overline{q}{\circ}\overline{G}^n, \qquad \langle p_2^{(n)}\,\partial_\theta(q_2{\circ}\mathscr{S}_2^n)\rangle - \overline{p}^{(n)}\,\partial_\theta(\overline{q}{\circ}\overline{G}^n). \tag{D.2}$$

Since $\partial_\theta(\mathscr{S}_2)_\theta^n = (\partial_\theta G_2)^{(n)}$ and $\partial_\theta\overline{G}^n = (\partial_\theta\overline{G})^{(n)}$, we bound the first contribution in (D.2) as

$$\begin{aligned}
\big|\langle \partial_\theta p_2^{(n)}\,q_2{\circ}\mathscr{S}_2^n\rangle - \partial_\theta\overline{p}^{(n)}\,\overline{q}{\circ}\overline{G}^n\big| \le \sum_{k=0}^{n-1}\Big|&\langle (p_2\,\partial_\theta G_2)^{(k)}\big(\partial_\theta p_2\,p_2^{(n-1-k)}{\circ}\mathscr{S}_2\,q_2{\circ}\mathscr{S}_2^{n-k}\big){\circ}\mathscr{S}_2^k\rangle \\
&- \langle p_2\,\partial_\theta G_2\rangle^{(k)}\langle \partial_\theta p_2\,p_2^{(n-1-k)}{\circ}\mathscr{S}_2\,q_2{\circ}\mathscr{S}_2^{n-k}\rangle{\circ}\overline{G}^k\Big| \\
+\sum_{k=0}^{n-1}\big|\langle p_2\,\partial_\theta G_2\rangle^{(k)}\big|\big|&\big(\langle\partial_\theta p_2\,p_2^{(n-1-k)}{\circ}\mathscr{S}_2\,q_2{\circ}\mathscr{S}_2^{n-k}\rangle{\circ}\overline{G}^k \\
&- \langle\partial_\theta p_2\rangle\langle p_2^{(n-1-k)}{\circ}\mathscr{S}_2\rangle\langle q_2\rangle{\circ}\overline{G}^{n-k}\big){\circ}\overline{G}^k\big| \\
+\sum_{k=0}^{n-1}\big|&\langle p_2\partial_\theta G_2\rangle^{(k)}\big(\langle\partial_\theta p_2\rangle\langle p_2^{(n-1-k)}{\circ}\mathscr{S}_2\rangle\langle q_2\rangle{\circ}\overline{G}^{n-k}\big){\circ}\overline{G}^k \\
&- (\overline{p}\,\partial_\theta\overline{G})^{(k)}\big(\partial_\theta\overline{p}\,\overline{p}^{(n-1-k)}{\circ}\overline{G}\,\overline{q}{\circ}\overline{G}^{n-k}\big){\circ}\overline{G}^k\big|.
\end{aligned} \tag{D.3}$$

Next, we use Proposition 7.30, 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 contributions in the first two lines of (D.3) with

$$\sum_{k=1}^{n} C(1-\rho\,\gamma')^k\big(\rho + k\rho^2\big)\,|||\partial_\theta p_2\,p_2^{(n-1-k)}{\circ}\mathscr{S}_2\,q_2{\circ}\mathscr{S}_2^{n-k}|||_{\alpha_0,3}^-, \tag{D.4}$$

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 contributions in the third and fourth line of (D.3) with

$$\sum_{k=0}^{n-1} C(1-\rho\,\gamma')^k\Big((n-1-k)\lambda^{-\alpha_0(n-1-k)} + \big(\rho + (n-1-k)\rho^2\big)\,|||\partial_\theta p_2|||_{\alpha_0}^+\,|||q_2|||_{\alpha_0,3}^-\Big), \tag{D.5}$$

and, after writing

$$\begin{aligned}
&\big|\langle (p_2\,\partial_\theta G_2)^{(k)}\rangle\big(\langle\partial_\theta p_2\rangle\langle p_2^{(n-1-k)}{\circ}\mathscr{S}_2\rangle\langle q_2\rangle{\circ}\overline{G}^{n-k}\big){\circ}\overline{G}^k - (\overline{p}\,\partial_\theta\overline{G})^{(k)}\big(\partial_\theta\overline{p}\;\overline{p}^{(n-1-k)}{\circ}\overline{G}\,\overline{q}{\circ}\overline{G}^{n-k}\big){\circ}\overline{G}^k\big| \\
&\quad\le \big|\langle p_2\,\partial_\theta G_2\rangle^{(k)} - (\overline{p}\,\partial_\theta\overline{G})^{(k)}\big|\big|\big(\langle\partial_\theta p_2\rangle\langle p_2^{(n-1-k)}{\circ}\mathscr{S}_2\rangle\langle q_2\rangle{\circ}\overline{G}^{n-k}\big){\circ}\overline{G}^k\big| \\
&\qquad + \big|(\overline{p}\,\partial_\theta\overline{G})^{(k)}\big|\big|(\langle\partial_\theta p_2\rangle - \partial_\theta\overline{p}){\circ}\overline{G}^k\big|\big|\big(\langle p_2^{(n-1-k)}{\circ}\mathscr{S}_2\rangle\langle q_2\rangle{\circ}\overline{G}^{n-k}\big){\circ}\overline{G}^k\big| \\
&\qquad + \big|(\overline{p}\,\partial_\theta\overline{G})^{(k)}\,\partial_\theta\overline{p}{\circ}\overline{G}^k\big|\big|\big(\langle p_2^{(n-1-k)}{\circ}\mathscr{S}_2\rangle - \overline{p}^{(n-1-k)}{\circ}\overline{G}\big){\circ}\overline{G}^k\big|\big|\langle q_2\rangle{\circ}\overline{G}^n\big| \\
&\qquad + \big|(\overline{p}\,\partial_\theta\overline{G})^{(k)}\,\big(\partial_\theta\overline{p}\;\overline{p}^{(n-1-k)}{\circ}\overline{G}\big){\circ}\overline{G}^k\big|\big|\langle q_2\rangle{\circ}\overline{G}^n - \overline{q}{\circ}\overline{G}^n\big|,
\end{aligned}$$

we rely once more Proposition 7.30 to bound the last two lines of (D.3) with

$$\sum_{k=0}^{n-1} C(1-\rho\gamma')^k\Big((\rho+k\rho^2)\|\partial_\theta p_2\|_\infty\|p_2^{(n-1-k)}\|_\infty\|q_2\|_\infty+\rho\,\|\langle\partial_\theta p_2\rangle-\partial_\theta\overline{p}\|_\infty\|p_2^{(n-1-k)}\|_\infty\|q_2\|_\infty \\ +(\rho+(n-k-1)\,\rho^2)\,(1-\rho\,\gamma')^{n-k-1}\|q_2\|_\infty+\rho\,(1-\rho\,\gamma')^{n-1-k}\|\langle q_2\rangle-\overline{q}\|_\infty\Big). \tag{D.6}$$

Therefore, thanks to (7.57) and (7.60), which yield

$$|||\partial_\theta p_2|||_{\alpha_0,2}\le C\rho,\qquad |||\langle p_2\rangle-\overline{p}|||_{0,1}\le C\rho^2,\qquad \|q_2\|_\infty\le C\rho,\qquad \|\langle q_2\rangle-\overline{q}\|_\infty\le C\rho^2,$$

by collecting together the bounds (D.4) to (D.6) and using (2.10), we obtain

$$\big|\langle\partial_\theta p_2^{(n)}\,q_2\circ\mathscr{S}_2^n\rangle-\partial_\theta\overline{p}^{(n)}\,\overline{q}\circ\overline{G}^n\big|\le\sum_{k=0}^{n-1}C(1-\rho\,\gamma')^n\rho^3\left(1+k\rho+k^2\rho^2\right). \tag{D.7}$$

Next, we consider the second contribution in (D.2), that we rewrite as

$$\langle p_2^{(n)}\,\partial_\theta(q_2\circ\mathscr{S}_2^n)\rangle-\overline{p}^{(n)}\,\partial_\theta(\overline{q}\circ\overline{G}^n) \\ =\Big(\langle p_2^{(n)}\,\partial_\theta(q_2\circ\mathscr{S}_2^n)\rangle-\langle p_2\rangle^{(n)}\langle\partial_\theta(q_2\circ\overline{G}^n)\rangle\Big)+\Big(\langle p_2\rangle^{(n)}\langle\partial_\theta(q_2\circ\overline{G}^n\rangle-\overline{p}^{(n)}\,\partial_\theta(\overline{q}\circ\overline{G}^n)\Big),$$

and, using that

$$\partial_\theta(q_2\circ\mathscr{S}_2^n)=(\partial_\theta(\mathscr{S}_2)_\theta)^{(n)}(\partial_\theta q_2)\circ\mathscr{S}_2^n, \\ \partial_\theta(q_2\circ\overline{G}^n)=(\partial_\theta\overline{G})^{(n)}\partial_\theta q_2\circ\overline{G}^n, \\ \partial_\theta(\overline{q}\circ\overline{G}^n)=(\partial_\theta\overline{G})^{(n)}(\partial_\theta\overline{q})\circ\overline{G}^n,$$

we apply first Proposition 7.30, 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 reason as in (7.55) to get

$$\Big|\langle p_2\partial_\theta(\mathscr{S}_2)_\theta\rangle^{(n)}\langle\partial_\theta q_2\rangle\circ\overline{G}^n-(\overline{p}\,\partial_\theta\overline{G})^{(n)}\,(\partial_\theta\overline{q})\circ\overline{G}^n\Big| \\ \le\langle\partial_\theta q_2\rangle\circ\overline{G}^n\sum_{i=0}^{n-1}\langle p_2\,\partial_\theta(\mathscr{S}_2)_\theta\rangle^{(i)}\left(\langle p_2\partial_\theta(\mathscr{S}_2)_\theta\rangle-\overline{p}\,\partial_\theta\overline{G}\right)\circ\overline{G}^i\,(\overline{p}\,\partial_\theta\overline{G})^{(n-i-2)}\circ\overline{G}^{i+1} \\ +(\overline{p}\,\partial_\theta\overline{G})^{(n)}\left(\langle\partial_\theta q_2\rangle\circ\overline{G}^n-\partial_\theta\overline{q}\circ\overline{G}^n\right) \\ \le C(1-\rho\,\gamma')^n n\,\|p_2\partial_\theta(\mathscr{S}_2)_\theta-\overline{p}\,\partial_\theta\overline{G}\|_\infty\,\|\partial q_2\|_\infty+C(1-\rho\,\gamma')^n\,\|\langle\partial_\theta q_2\rangle-\partial_\theta\overline{q}\|_\infty \\ \le C(1-\rho\,\gamma')^n n\rho^2,$$

so that eventually we obtain

$$\big|\langle p_2^{(n)}\,\partial_\theta(q_2\circ\mathscr{S}_2^n)-\overline{p}^{(n)}\,\partial_\theta(\overline{q}\circ\overline{G}^n)\rangle\big|\le C\rho. \tag{D.8}$$

and hence, summing together (D.7) and (D.8), the first bound in (7.63) follows.

The second bound in (7.63) is obtained in a similar way by reasoning as in the second part of the proof of Proposition 7.32 and using (D.7) and (D.8) instead of (7.59).

D.2 Comparing the translated and auxiliary maps I: proof of Lemma 7.37

As in Appendix D.1, we consider explicitly only the case $\rho < \rho_0$. Note that all the proofs from here on, until the end of this appendix, undergo major simplifications if $\mathscr{S}$ is restricted to the set Λ_{2r_2} as discussed in Remark 7.40; in particular all term which involve the function c_0 vanish identically.

We follow the strategy we have outlined in Remark 7.36, by rearranging sums where differences $W\circ A_0^{i+1} - W\circ A_0^i$ appear (see the equations (D.10) to (D.14) below) in such a way that the new summands contain differences of more regular functions (see Lemmas D.1 and D.4 below).

Thus, if we define, for $k \geq 0$,

$$D_{\mathfrak{p},1,k}(\theta,\psi) := \int_0^1 \mathrm{d}t\, \partial_\theta(\mathfrak{p}\circ\mathscr{S}_1^k)\,(G_2(\theta,\psi) + t\,\rho\,\zeta(\theta,\psi), A_0\psi), \tag{D.9a}$$

$$D_{\mathfrak{p},2,k}(\theta,\psi) := \int_0^1 \mathrm{d}t\,(1-t)\,\partial_\theta^2(\mathfrak{p}\circ\mathscr{S}_1^k)\,(G_2(\theta,\psi) + t\,\rho\,\zeta(\theta,\psi), A_0\psi). \tag{D.9b}$$

$$R_{\mathfrak{p},1,k}(\theta,\psi) := \sum_{i=0}^{k} \big(\partial_\theta\big((D_{\mathfrak{p},1,k-i}\,\zeta)\circ\mathscr{S}_2^i\big)\big)\,(\theta,\psi), \tag{D.9c}$$

$$R_{\mathfrak{p},2,k}(\theta,\psi) := \sum_{i=0}^{k} \big(\big(D_{\mathfrak{p},2,k-i}\,\zeta^2\big)\circ\mathscr{S}_2^i\big)\,(\theta,\psi), \tag{D.9d}$$

we can write, for $n \geq 1$,

$$\begin{aligned}
\mathfrak{p}\circ\mathscr{S}_1^n - \mathfrak{p}\circ\mathscr{S}_2^n &= \sum_{i=0}^{n-1}\big(\mathfrak{p}\circ\mathscr{S}_1^{n-i}\circ\mathscr{S}_2^i - \mathfrak{p}\circ\mathscr{S}_1^{n-1-i}\circ\mathscr{S}_2^{i+1}\big)\\
&= \rho\sum_{i=0}^{n-1}\big(D_{\mathfrak{p},1,n-1-i}\,\zeta\big)\circ\mathscr{S}_2^i\\
&= \rho\sum_{i=0}^{n-1}\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_1^{n-1-i})\circ\mathscr{S}_2\,\zeta\big)\circ\mathscr{S}_2^i + \rho^2 R_{\mathfrak{p},2,n-1},
\end{aligned} \tag{D.10}$$

and, using the second line of (D.10), with $n-i-1$ instead of n, to write

$$\partial_\theta(\mathfrak{p}\circ\mathscr{S}_1^{n-1-i}) = \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1-i}) + \rho\sum_{i=0}^{n-1-i}\partial_\theta\big(\big(D_{\mathfrak{p},1,n-i-1-j}\,\zeta\big)\circ\mathscr{S}_2^j\big)$$

in the last line, we obtain

$$\mathfrak{p}\circ\mathscr{S}_1^n - \mathfrak{p}\circ\mathscr{S}_2^n = \rho\sum_{i=0}^{n-1}\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1-i})\circ\mathscr{S}_2\,\zeta\big)\circ\mathscr{S}_2^i + \rho^2\sum_{i=0}^{n-1}\left(R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta\right)\circ\mathscr{S}_2^i + \rho^2 R_{\mathfrak{p},2,n-1}. \tag{D.11}$$

Note that, by Lemma 6.9 and Remark 6.13, for all $n \geq 1$ one has

$$\|R_{\mathfrak{p},2,n-1}\|_\infty \leq C, \qquad \langle|R_{\mathfrak{p},2,n-1}|\rangle \leq C. \tag{D.12}$$

Next, using (7.68) and recalling (7.46), if we set

$$E_{\mathfrak{p},r,k} := \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^k)\circ\mathscr{S}_2\,c_r, \qquad r = 0,1,2, \tag{D.13}$$

then in the first sum on the r.h.s. of (D.11) we can write

$$\begin{aligned}
&\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-i-1})\circ\mathscr{S}_2\,\zeta\big)\circ\mathscr{S}_2^i\\
&\quad = \rho^{-1}E_{\mathfrak{p},0,n-i-1}\circ\mathscr{S}_2^i\,(W\circ A_0^{i+1} - W\circ A_0^i) + \sum_{r=1}^{2} E_{\mathfrak{p},r,n-i-1}\circ\mathscr{S}_2^i\,\left(W\circ A_0^i\right)^r.
\end{aligned} \tag{D.14}$$

The following result is easily checked.

Lemma D.1. *Let $\{E_k\}_{k=0}^{n-1}$ be a set of functions $E_k : \Omega \to \Omega$, with $n \geq 1$. For $m \in \mathbb{N}$ such that $1 \leq m \leq n$, define*

$$\Delta E_{m,0} = E_0 \circ \mathscr{S}_2^{-1}, \quad \Delta E_{m,k} = E_k \circ \mathscr{S}_2^{-1} - E_{k-1},\ k = 1, \dots, m-1, \quad \Delta E_{m,m} = -E_{m-1}, \tag{D.15}$$

where $\Delta E_{m,k}$ with $0 < k < m$ are meaningful only if $m \geq 2$. Then, for any m as above, one has

$$\sum_{i=0}^{m-1} E_{m-1-i} \circ \mathscr{S}_2^i \left(W \circ A_0^{i+1} - W \circ A_0^i\right) = \sum_{i=0}^{m} \Delta E_{m,m-i} \circ \mathscr{S}_2^i\, W \circ A_0^i. \tag{D.16}$$

Remark D.2. If the functions E_k in Lemma D.1 are such that

$$|||E_k|||_{\alpha_0,3}^{-} \leq (1 - \rho\,\gamma')^k\, \Gamma_0, \qquad k = 0, \dots, n-1, \tag{D.17a}$$

$$|||E_k \circ \mathscr{S}_2^{-1} - E_{k-1}|||_{\alpha_0,3}^{-} \leq \rho\,(1 - \rho\,\gamma')^k\, \Gamma_0, \qquad k = 1, \dots, n-1, \tag{D.17b}$$

for some constant Γ_0, then one finds

$$\sum_{i=0}^{n} |||\Delta E_{n,n-i}|||_{\alpha_0,3}^{-} \leq C\,\Gamma_0. \tag{D.18}$$

Analogously, if

$$\|E_k\|_\infty \leq (1 - \rho\,\gamma')^k\, \Gamma_0, \qquad k = 0, \dots, n-1, \tag{D.19a}$$

$$\|E_k \circ \mathscr{S}_2^{-1} - E_{k-1}\|_\infty \leq \rho\,(1 - \rho\,\gamma')^k\, \Gamma_0, \qquad k = 1, \dots, n-1, \tag{D.19b}$$

one has

$$\sum_{i=0}^{n} \|\Delta E_{n,n-i}\|_\infty \leq C\,\Gamma_0. \tag{D.20}$$

Remark D.3. The functions $E_{\mathfrak{p},0,k}$ introduced in (D.13) satisfy the conditions (D.17) with $\Gamma_0 = C|||\partial_\theta \mathfrak{p}|||_{\alpha_0,3}$. Indeed, using that $\partial_\theta(\mathfrak{p} \circ \mathscr{S}_2^k) = \partial_\theta \mathfrak{p} \circ \mathscr{S}_2^k \partial_\theta(\mathscr{S}_2^k)_\theta$, one has

$$\begin{aligned}
&|||E_{\mathfrak{p},0,k} \circ \mathscr{S}_2^{-1} - E_{\mathfrak{p},0,k-1}|||_{\alpha_0,3}^{-} \\
&\quad \leq |||\partial_\theta(\mathfrak{p} \circ \mathscr{S}_2^k)|||_{\alpha_0,3}^{-} |||c_0 \circ \mathscr{S}_2^{-1} - c_0|||_{\alpha_0,3} + |||\partial_\theta(\mathfrak{p} \circ \mathscr{S}_2^k) \circ \mathscr{S}_2^{-1} - \partial_\theta(\mathfrak{p} \circ \mathscr{S}_2^{k-1})|||_{\alpha_0,3}^{-} |||c_0|||_{\alpha_0,3} \\
&\quad \leq |||\partial_\theta \mathfrak{p}|||_{\alpha_0,3} |||\partial_\theta(\mathscr{S}_2^k)_\theta|||_{\alpha_0,3}^{-} |||c_0 \circ \mathscr{S}_2^{-1} - c_0|||_{\alpha_0,3} \\
&\quad + |||\partial_\theta \mathfrak{p}|||_{\alpha_0,3} |||\partial_\theta(\mathscr{S}_2^{k-1})_\theta|||_{\alpha_0,3}^{-} |||\partial_\theta(\mathscr{S}_2)_\theta - \mathbb{1}|||_{\alpha_0,3}^{-} |||c_0|||_{\alpha_0,3} \\
&\quad \leq C \rho\,(1 - \rho\,\gamma')^k\, |||\partial_\theta \mathfrak{p}|||_{\alpha_0,3},
\end{aligned}$$

as a consequence of the bounds (7.54) in Lemma 7.27.

Lemma D.1 is extended immediately as follows.

Lemma D.4. *Let $\{E_k\}_{k=0}^{n-1}$ and, for $1 \leq m \leq n$, $\{\Delta E_{m,k}\}_{k=1}^{m}$ be as in Lemma D.1, and let $\{F_k\}_{k=0}^{n}$ be a set of functions $F_k : \Omega \to \Omega$. Then, for any m as above, one has*

$$\begin{aligned}
&\sum_{i=0}^{m-1} E_{m-1-i} \circ \mathscr{S}_2^i\, F_i (W \circ A_0^{i+1} - W \circ A_0^i) \\
&\qquad = \sum_{i=0}^{m} \Delta E_{m,m-i} \circ \mathscr{S}_2^i\, F_i W \circ A_0^i + \sum_{i=1}^{m} E_{m-i} \circ \mathscr{S}_2^{i-1} \big(F_{i-1} - F_i\big) W \circ A_0^i.
\end{aligned} \tag{D.21}$$

Remark D.5. It is easy to check that

$$\begin{aligned}
&\|\partial_\theta(\mathscr{S}_2^k)_\theta\, c_r\circ\mathscr{S}_2^{-1} - \partial_\theta(\mathscr{S}_2^k)_\theta\circ\mathscr{S}_2\, c_r\|_\infty \\
&\quad\le \|\partial_\theta(\mathscr{S}_2^k)_\theta\|_\infty\|c_r\circ(\mathscr{S}_2^{-1})_\theta - c_r\|_\infty + \|\partial_\theta^2(\mathscr{S}_2^k)_\theta\|_\infty\|(\mathscr{S}_2)_\theta - \mathbb{1}\|_\infty\|\,c_r\|_\infty \\
&\quad\le C\rho(1-\rho\,\gamma')^k,
\end{aligned}$$

and, analogously,

$$\begin{aligned}
&\|\partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big)\circ\mathscr{S}_2^{-1} - \partial_\theta\big(D_{\mathfrak{p},1,k-1}\,c_r\big)\|_\infty \\
&\quad\le \|\partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big)\|_\infty\|\circ\mathscr{S}_2^{-1} - \mathbb{1}\|_\infty + \|\partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big) - \partial_\theta\big(D_{\mathfrak{p},1,k-1}\,c_r\big)\|_\infty \\
&\quad\le C\rho\,(1-\rho\,\gamma')^{k+j}|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}.
\end{aligned}$$

Therefore the functions $Q_{r,k}$ and $Q_{\mathfrak{p},r,k}$, for $r=0,1,2$, satisfy the conditions (D.19), with $\Gamma_0 = C$ and $\Gamma_0 = C|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}$ respectively, and hence, by Remark D.2, we can bound

$$\|\Delta Q_{r,m,k}\|_\infty \le C\rho(1-\rho\,\gamma')^k, \qquad \|\Delta Q_{\mathfrak{p},r,m,k}\|_\infty \le C\rho(1-\rho\,\gamma')^k\,|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}, \tag{D.22}$$

for $r=0,1$, $m\ge 2$ and $k=1,\dots,m-1$.

By Lemma D.1, after inserting (D.14) into (D.11), we obtain

$$\sum_{i=0}^{n-1} E_{\mathfrak{p},0,n-1-i}\circ\mathscr{S}_2^i\,(W\circ A_0^{i+1} - W\circ A_0^i) = \sum_{i=0}^{n} \Delta E_{\mathfrak{p},0,n,n-i}\circ\mathscr{S}_2^i\,W\circ A_0^i, \tag{D.23}$$

where the functions $\Delta E_{\mathfrak{p},0,m,k}$ are defined as in (D.15) with $E_k = E_{\mathfrak{p},0,k}$, so that,

$$\sum_{i=0}^{n} |||\Delta E_{\mathfrak{p},0,n,n-i}|||_{\alpha_0,3}^{-} \le C\,|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}, \tag{D.24}$$

by Remarks D.2 and D.3. Furthermore, in the second sum on the r.h.s. of (D.11), if we set

$$Q_{\mathfrak{p},r,k} := \partial_\theta\big(D_{\mathfrak{p},1,k}\,c_r\big), \qquad Q_{r,k} := \partial_\theta(\mathscr{S}_2^k)_\theta\circ\mathscr{S}_2\,c_r, \qquad r=0,1,2, \tag{D.25}$$

and define $\Delta Q_{\mathfrak{p},r,m,k}$ and $\Delta Q_{r,m,k}$ according to (D.15) with $E_k = Q_{\mathfrak{p},r,k}$ and $E_k = Q_{r,k}$, respectively, we obtain

$$
\begin{aligned}
\sum_{i=0}^{n-1} (R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta)\circ\mathscr{S}_2^i &= \sum_{i=0}^{n-1}\sum_{j=0}^{i}\Big(\partial_\theta\left(\left(D_{\mathfrak{p},1,n-1-i}\,\zeta\right)\circ\mathscr{S}_2^{i-j}\right)\circ\mathscr{S}_2\,\zeta\Big)\circ\mathscr{S}_2^j \\
&= \sum_{i=0}^{n-1}\sum_{j=0}^{i}\sum_{r,s=1}^{2} Q_{\mathfrak{p},r,n-1-i}\circ\mathscr{S}_2^{i+1}Q_{s,i-j}\circ\mathscr{S}_2^j\left(W\circ A_0^{i+1}\right)^r\left(W\circ A_0^j\right)^s \\
&+ \rho^{-1}\sum_{i=0}^{n-1}\sum_{j=0}^{i+1}\sum_{r=1}^{2} Q_{\mathfrak{p},r,n-1-i}\circ\mathscr{S}_2^{i+1}\Delta Q_{0,i+1,i+1-j}\circ\mathscr{S}_2^j\left(W\circ A_0^{i+1}\right)^r W\circ A_0^j \\
&+ \rho^{-1}\Bigg(\sum_{j=0}^{n-1}\sum_{i=0}^{n-j}\sum_{s=1}^{2}\Big(\Delta Q_{\mathfrak{p},0,n-j,n-j-i}\circ\mathscr{S}_2^{i+1}Q_{s,i}\,W\circ A_0^{i+1}\Big)\circ\mathscr{S}_2^j\left(W\circ A_0^j\right)^s \\
&\quad+\sum_{j=0}^{n-1}\sum_{i=1}^{n-j}\sum_{s=1}^{2}\Big(Q_{\mathfrak{p},0,n-j-i}\circ\mathscr{S}_2^i\left(Q_{s,i-1}-Q_{s,i}\right)W\circ A_0^{i+1}\Big)\circ\mathscr{S}_2^j\left(W\circ A_0^j\right)^s\Bigg) \\
&+ \rho^{-2}\Bigg(\sum_{j=0}^{n-1}\sum_{i=0}^{n-j}\Big(\Delta Q_{\mathfrak{p},0,n-j,n-j-i}\circ\mathscr{S}_2^{i+1}\Delta Q_{0,i+1,i+1}W\circ A_0^{i+1}\Big)\circ\mathscr{S}_2^j\;W\circ A_0^j \\
&\quad+\sum_{j=0}^{n-1}\sum_{i=1}^{n-j}\Big(Q_{\mathfrak{p},0,n-j-i}\circ\mathscr{S}_2^i\big(\Delta Q_{0,i,i}-\Delta Q_{0,i+1,i+1}\big)\Big)\circ\mathscr{S}_2^j\,W\circ A_0^{i+j+1}\,W\circ A_0^j \\
&\quad+\sum_{i=0}^{n-1} Q_{\mathfrak{p},0,n-1-i}\circ\mathscr{S}_2^{i+1}\,\Delta Q_{0,i+1,0}\circ\mathscr{S}_2^{i+1}\left(W\circ A_0^{i+2}-W\circ A_0^{i+1}\right)W\circ A_0^{i+1}\Bigg),
\end{aligned}
\tag{D.26}
$$

where we have used Lemma D.1 to obtain the third line, Lemma D.4 to obtain the fourth and fifth lines, and first Lemma D.1 and then Lemma D.4 to obtain the last three lines.

Remark D.6. To bound the second to last line of (D.26) we use that, for $m > k \geq 0$, we have

$$
\begin{aligned}
\Delta Q_{0,m,k}-\Delta Q_{0,m+1,k+1} &= \left(Q_{0,k}\circ\mathscr{S}_2^{-1}-Q_{0,k-1}\right)-\left(Q_{0,k+1}\circ\mathscr{S}_2^{-1}-Q_{0,k}\right) \\
&= \partial_\theta(\mathscr{S}_2^{k-1})_\theta\left(\partial_\theta(\mathscr{S}_2)_\theta\,c_0\circ\mathscr{S}_2^{-1}-c_0\right)-\partial_\theta(\mathscr{S}_2^{k})_\theta\left(\partial_\theta(\mathscr{S}_2)_\theta\,c_0\circ\mathscr{S}_2^{-1}-c_0\right) \\
&= \left(\partial_\theta(\mathscr{S}_2^{k-1})_\theta-\partial_\theta(\mathscr{S}_2^{k})_\theta\right)\left(\partial_\theta(\mathscr{S}_2)_\theta\,c_0\circ\mathscr{S}_2^{-1}-c_0\right) \\
&= \partial_\theta(\mathscr{S}_2^{k-1})_\theta\left(\mathbb{1}-\partial_\theta(\mathscr{S}_2)_\theta\right)\left(\left(\partial_\theta(\mathscr{S}_2)_\theta-\mathbb{1}\right)c_0\circ\mathscr{S}_2^{-1}+\left(c_0\circ\mathscr{S}_2^{-1}-c_0\right)\right),
\end{aligned}
$$

so that $\|\Delta Q_{0,m,k}-\Delta Q_{0,m,k+1}\|_\infty \leq C\rho^2(1-\rho\,\gamma')^k$.

The bounds (7.54) in Lemma 7.27 and the second bound of (2.38) in Theorem 4, together with the bounds in Remarks D.5 and D.6, yield that

$$
\left\|\sum_{i=0}^{n-1}(R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta)\circ\mathscr{S}_2^i\right\|_\infty \leq C\,\rho^{-2}|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}, \tag{D.27a}
$$

$$
\left\langle\left|\sum_{i=0}^{n-1}(R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta)\circ\mathscr{S}_2^i\right|\right\rangle \leq C\rho^{-1}|||\partial_\theta\mathfrak{p}|||_{\alpha_0,3}. \tag{D.27b}
$$

In conclusion, we obtain (7.69) with $C_{\mathfrak{p},n,k}=E_{\mathfrak{p},1,k-1}+\rho^{-1}\Delta E_{\mathfrak{p},0,n,k}$ for $k=1,\ldots,n$ and $C_{\mathfrak{p},n,0}=$

$\Delta E_{\mathfrak{p},1,n,0}$, i.e.

$$\begin{aligned} C_{\mathfrak{p},n,0} &:= \rho^{-1}\partial_\theta \mathfrak{p}\, c_0 \circ \mathscr{S}_2^{-1}, \\ C_{\mathfrak{p},n,k} &:= \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{k-1})\circ\mathscr{S}_2\, c_1 \\ &\quad + \rho^{-1}\big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{k})\, c_0\circ\mathscr{S}_2^{-1} - \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{k-1})\circ\mathscr{S}_2\, c_0\big), \quad k=1,\dots,n-1, \\ C_{\mathfrak{p},n,n} &:= \partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1})\circ\mathscr{S}_2\, \mathfrak{c}_1 - \rho^{-1}\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1})\, c_0\circ\mathscr{S}_2^{-1} \end{aligned} \tag{D.28}$$

so that the bound (7.70b) follows from the estimates (7.54) and (D.24), while

$$R_{\mathfrak{p},n} := \sum_{i=0}^{n-1}\Big(\partial_\theta(\mathfrak{p}\circ\mathscr{S}_2^{n-1-i})\circ\mathscr{S}_2\, c_2\Big)\circ\mathscr{S}_2^i\, W^2\circ A_0^i + \rho\sum_{i=0}^{n-1}(R_{\mathfrak{p},1,n-1-i}\circ\mathscr{S}_2\,\zeta)\circ\mathscr{S}_2^i + \rho\, R_{\mathfrak{p},2,n-1} \tag{D.29}$$

satisfies the bounds in the statement because of the bounds (D.12) and (D.27).

Remark D.7. While the functions $C_{\mathfrak{p},n,0},\dots,C_{\mathfrak{p},n,n}$ depend on $\mathscr{S}_2$ alone, the function $R_{\mathfrak{p},n}$ involves both $\mathscr{S}_1$ and $\mathscr{S}_2$ through the contributions $R_{\mathfrak{p},2,n-1}$ and $R_{\mathfrak{p},1,k}$.

D.3 Comparing the translated and auxiliary maps II: proof of Proposition 7.50

Once more, we can confine ourselves to the case $\rho < \rho_0$. We have

$$\partial_\theta\big(p_1^{(n)} q_1\circ\mathscr{S}_1^n\big) = \partial_\theta p_1^{(n)}\, q_1\circ\mathscr{S}_1^n + p_1^{(n)}\,\partial_\theta(q_1\circ\mathscr{S}_1^n), \tag{D.30}$$

which, together with (D.1a), allows us to write the first contribution in (7.52b) as

$$\begin{aligned} \partial_\theta h_1 - \partial_\theta h_2 &= \sum_{n=0}^{\infty}\partial_\theta\,(p_1^{(n)} q_1\circ\mathscr{S}_1^n) - \partial_\theta\,(p_2^{(n)} q_2\circ\mathscr{S}_2^n) \\ &= \sum_{n=0}^{\infty}(\partial_\theta p_1^{(n)} - \partial_\theta p_2^{(n)})\,(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n) + \sum_{n=0}^{\infty}(\partial_\theta p_1^{(n)} - \partial_\theta p_2^{(n)})\, q_2\circ\mathscr{S}_2^n \\ &\quad + \sum_{n=0}^{\infty}\partial_\theta p_2^{(n)}\,(q_1\circ\mathscr{S}_1^n - q_2\circ\mathscr{S}_2^n) + \sum_{n=0}^{\infty}(p_1^{(n)} - p_2^{(n)})\,(\partial_\theta q_1\circ\mathscr{S}_1^n - \partial_\theta q_2\circ\mathscr{S}_2^n) \\ &\quad + \sum_{n=0}^{\infty}(p_1^{(n)} - p_2^{(n)})\,\partial_\theta q_2\circ\mathscr{S}_2^n + \sum_{n=0}^{\infty} p_2^{(n)}\,(\partial_\theta q_1\circ\mathscr{S}_1^n - \partial_\theta q_2\circ\mathscr{S}_2^n). \end{aligned} \tag{D.31}$$

In (D.31) we can expand

$$p_1^{(n)} - p_2^{(n)} = \sum_{k=0}^{n-1} p_1^{(k)}\big(p_1\circ\mathscr{S}_1^k - p_2\circ\mathscr{S}_2^k\big)\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}, \tag{D.32}$$

according to (7.65a), and

$$\begin{aligned} \partial_\theta p_1^{(n)} - \partial_\theta p_2^{(n)} &= \sum_{k=0}^{n-1}\big(\pi_1^{(k)} - \pi_2^{(k)}\big)\,\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\ &\quad + \sum_{k=0}^{n-1}\pi_1^{(k)}\big(\partial_\theta p_1\circ\mathscr{S}_1^k - \partial_\theta p_2\circ\mathscr{S}_2^k\big) p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1} \\ &\quad + \sum_{k=0}^{n-1}\pi_1^{(k)}\,\partial_\theta p_1\circ\mathscr{S}_1^k\big(p_1^{(n-k-1)}\circ\mathscr{S}_1^{k+1} - p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\big) \end{aligned} \tag{D.33}$$

with $\pi_1 := p_1\partial_\theta G_1$ and $\pi_2 := p_2\partial_\theta G_2$, first taking into account (B.1a) and then proceeding as in deriving (7.65a). Next, we apply (7.87b), (7.66b) and, once more, (7.65a) to write in the first line of (D.33)

$$\pi_1^{(k)} - \pi_2^{(k)} = \sum_{i=0}^{k-1} \pi_1^{(i)} \big(\Delta_{\pi_1,\pi_2,W,i} + \rho\, \mathfrak{a}_3 \circ \mathscr{S}_1^i\, \mathfrak{f} \circ A_0^i\big)\, \pi_2^{(k-i-1)} \circ \mathscr{S}_2^{i+1}, \tag{D.34}$$

with $\mathfrak{a}_3(\theta) := \mathfrak{a}_1(\theta) + \partial_\theta c_0(\theta)$, in the second line

$$\partial_\theta p_1 \circ \mathscr{S}_1^k - \partial_\theta p_2 \circ \mathscr{S}_2^k = \Delta_{\partial_\theta p_1, \partial_\theta p_2, W, k} + \rho\, \mathfrak{a}_4 \circ \mathscr{S}_1^k\, \mathfrak{f} \circ A_0^k, \tag{D.35}$$

with $\mathfrak{a}_4(\theta) := \partial_\theta \mathfrak{a}_1(\theta)$, and in the third line

$$\begin{aligned} &p_1^{(n-k-1)} \circ \mathscr{S}_1^{k+1} - p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1} \\ &\quad = \sum_{i=0}^{n-k-2} p_1^{(i)} \circ \mathscr{S}_1^{k+1} \big(\Delta_{p_1,p_2,W,i+k+1} + \rho\, \mathfrak{a}_1 \circ \mathscr{S}_1^{i+k+1} \mathfrak{f} \circ A_0^{i+k} + 1\big)\, p_2^{(n-k-2-i)} \circ \mathscr{S}_2^{i+k+2}. \end{aligned} \tag{D.36}$$

Moreover in (D.31), (D.32) and (D.33) we express $p_1 \circ \mathscr{S}_1^n - p_2 \circ \mathscr{S}_2^n$ and $q_1 \circ \mathscr{S}_1^n - q_2 \circ \mathscr{S}_2^n$ according to (7.88) and, analogously to (D.35), we write

$$\partial_\theta q_1 \circ \mathscr{S}_1^n - \partial_\theta q_2 \circ \mathscr{S}_2^n = \Delta_{\partial_\theta q_1, \partial_\theta q_2, W, n} + \rho\, \mathfrak{a}_5 \circ \mathscr{S}_1^n\, \mathfrak{f} \circ A_0^n, \tag{D.37}$$

with $\mathfrak{a}_5(\theta) = \partial_\theta \mathfrak{a}_2(\theta)$.

Remark D.8. Bounds analogous to (7.90) hold also for the functions $\Delta_{\pi_1,\pi_2,W,n}$, $\Delta_{\partial_\theta p_1,\partial_\theta p_2,W,n}$ 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

$$|||\Delta|||_{\alpha_0,3}^{-} \leq C\rho, \qquad \langle|\Delta\, W \circ A_0^k|\rangle \leq C\rho^2, \qquad \langle|\Delta\Delta'|\rangle \leq C\rho^3.$$

Eventually, by collecting together the expansions from (D.31) to (D.37), we can write (D.31) as a sum of several contributions in the form

$$\partial_\theta h_1 - \partial_\theta h_2 = \mathcal{A}_1 + \mathcal{A}_2 + \mathcal{A}_3 + \mathcal{A}_4 + \mathcal{A}_5 + \mathcal{A}_6 + \mathcal{A}_7 + \mathcal{A}_8, \tag{D.38}$$

where the functions $\mathcal{A}_1, \ldots, \mathcal{A}_8$ are defined and bounded in the following way.

- The contribution

$$\begin{aligned} \mathcal{A}_1 := &\sum_{n=0}^{\infty} \sum_{k=0}^{n-1} \sum_{i=0}^{k-1} \pi_1^{(i)}\, \Delta_{\pi_1,\pi_2,W,i}\, \pi_2^{(k-i-1)} \circ \mathscr{S}_2^{i+1}\, \partial_\theta p_2 \circ \mathscr{S}_2^k\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1}\, \Delta_{q_1,q_2,W,n} \\ &+ \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} \pi_1^{(k)}\, \Delta_{\partial_\theta p_1,\partial_\theta p_2,W,k}\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1}\, \Delta_{q_1,q_2,W,n} \\ &+ \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} \sum_{i=0}^{n-k-2} \pi_1^{(k)}\, \partial_\theta p_1 \circ \mathscr{S}_1^k\, p_1^{(i)} \circ \mathscr{S}_1^{k+1}\, \Delta_{p_1,p_2,W,i+k+1}\, p_2^{(n-k-2-i)} \circ \mathscr{S}_2^{i+k+2}\, \Delta_{q_1,q_2,W,n} \\ &+ \sum_{n=0}^{\infty} \sum_{k=0}^{n-1} p_1^{(k)}\, \Delta_{p_1,p_2,W,k}\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1}\, \Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n} \end{aligned}$$

  is such that its average is bounded as $|\langle \mathcal{A}_1 \rangle| \leq C\rho$ because of Remark D.8 and the fact that $\|\partial_\theta p_2\|_\infty \leq C\rho$.

- The average of

$$
\begin{aligned}
\mathcal{A}_2 := \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} & \pi_1^{(i-1)}\mathfrak{D}_{\mathfrak{a}_3,i,0}\circ\mathscr{S}_1^i\, W\circ A_0^i\,(\pi_2^{(k)}\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n)}\circ\mathscr{S}_2^{k+1})\circ\mathscr{S}_2^{i+1}\Delta_{q_1,q_2,W,n+k+i+2} \\
& + \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\pi_1^{(k-1)}\mathfrak{a}_3\circ\mathscr{S}_1^{k-1}\, W\circ A_0^k\,\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n+k+1} \\
+ \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} & \pi_1^{(k-1)}\mathfrak{D}_{\mathfrak{a}_4,k,0}\circ\mathscr{S}_1^k\, W\circ A_0^k\, p_2^{(n)}\circ\mathscr{S}_2^{k+1}\Delta_{q_1,q_2,W,n+k+1} \\
& + \sum_{n=0}^{\infty}\pi_1^{(n-1)}\mathfrak{a}_4\circ\mathscr{S}_1^{n-1}\, W\circ A_0^n\,\Delta_{q_1,q_2,W,n} \\
+ \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} & \pi_1^{(k)}\partial_\theta p_1\circ\mathscr{S}_1^k\big(p_1^{(i-1)}\Delta_{\mathfrak{a}_1,i,0}\circ\mathscr{S}_1^i W\circ A_0^i\big)\circ\mathscr{S}_1^{k+1}\, p_2^{(n)}\circ\mathscr{S}_2^{i+k+2}\Delta_{q_1,q_2,W,n+i+k+2} \\
& + \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\pi_1^{(k)}\partial_\theta p_1\circ\mathscr{S}_1^k\big(p_1^{(i-1)}\mathfrak{a}_1\circ\mathscr{S}_1^{n-1}W\circ A_0^n\big)\circ\mathscr{S}_1^{k+1}\,\Delta_{q_1,q_2,W,n+k+1} \\
+ \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} & p_1^{(k-1)}\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^k\, W\circ A_0^k\, p_2^{(n)}\circ\mathscr{S}_2^{k+1}\,\Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n+k+1} \\
& + \sum_{n=0}^{\infty} p_1^{(n-1)}\mathfrak{a}_1\circ\mathscr{S}_1^{n-1}\, W\circ A_0^n\,\Delta_{\partial_\theta q_1,\partial_\theta q_2,W,n},
\end{aligned}
$$

where we have used Lemma 7.41, is bounded as $|\langle\mathcal{A}_2\rangle| \le C\rho$ because of Remark D.8, the bound (7.79) and the fact that $\|\partial_\theta p_2\|_\infty \le C\rho$.

- The contribution

$$
\begin{aligned}
\mathcal{A}_3 := & \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\pi_1^{(i)}\,\Delta_{\pi_1,\pi_2,W,i}\,(\pi_2^{(k)}\,\partial_\theta p_2\circ\mathscr{S}_2^k\, p_2^{(n)}\circ\mathscr{S}_2^{k+1})\circ\mathscr{S}_2^{i+1}\,\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+i+2}\,\mathfrak{f}\circ A_0^{n+k+i+2} \\
& + \sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\pi_1^{(k)}\,\Delta_{\partial_\theta p_1,\partial_\theta p_2,W,k}\, p_2^{(n)}\circ\mathscr{S}_2^{k+1}\,\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\,\mathfrak{f}\circ A_0^{n+k+1} \\
& + \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\pi_1^{(k)}\partial_\theta p_1\circ\mathscr{S}_1^k p_1^{(i)}\circ\mathscr{S}_1^{k+1}\Delta_{p_1,p_2,W,i+k+1}p_2^{(n)}\circ\mathscr{S}_2^{i+k+2}\rho\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+i+2}\mathfrak{f}\circ A_0^{n+k+i+2} \\
& + \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} p_1^{(k)}\,\Delta_{p_1,p_2,W,k}\, p_2^{(n)}\circ\mathscr{S}_2^{k+1}\,\rho\,\mathfrak{a}_5\circ\mathscr{S}_1^{n+k+1}\,\mathfrak{f}\circ A_0^{n+k+1}
\end{aligned}
$$

is dealt with first writing

$$
\begin{aligned}
p_2^{(n)}\circ\mathscr{S}_2^{k+1} &= p_2^{(n)}\circ\mathscr{S}_1^{k+1} + (p_2^{(n)}\circ\mathscr{S}_2^{k+1} - p_2^{(n)}\circ\mathscr{S}_1^{k+1}), \\
p_2^{(n)}\circ\mathscr{S}_2^{i+k+2} &= p_2^{(n)}\circ\mathscr{S}_1^{i+k+2} + (p_2^{(n)}\circ\mathscr{S}_2^{i+k+2} - p_2^{(n)}\circ\mathscr{S}_1^{i+k+1}),
\end{aligned}
$$

then expanding $p_2^{(n)}\circ\mathscr{S}_2^{k+1} - p_2^{(n)}\circ\mathscr{S}_1^{k+1}$ and $p_2^{(n)}\circ\mathscr{S}_2^{i+k+2} - p_2^{(n)}\circ\mathscr{S}_1^{i+k+2}$ according to (7.66a), so as to apply Lemma 7.37, while using Lemma 7.43 for the contributions containing the remaining factors $p_2^{(n)}\circ\mathscr{S}_1^{k+1}$ and $p_2^{(n)}\circ\mathscr{S}_1^{k+i+1}$ in order to obtain a factor either $\mathfrak{D}_{1,\mathfrak{a}_2,n}\circ\mathscr{S}_1^n$ or $\mathfrak{D}_{1,\mathfrak{a}_5,n}\circ\mathscr{S}_1^n$, so that eventually one finds $|\langle\mathcal{A}_3\rangle| \le C\rho$ because of Remark D.8, the bound (7.79) and the fact that $\|\partial_\theta p_2\|_\infty \le C\rho$.

- The contribution

$$
\begin{aligned}
\mathcal{A}_4 := \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}&\pi_1^{(i-1)}\mathfrak{D}_{\mathfrak{a}_3,i,0}\circ\mathscr{S}_1^{i}W\circ A_0^{i}\\
&\times(\pi_2^{(k)}\partial_\theta p_2\circ\mathscr{S}_2^{k}p_2^{(n)}\circ\mathscr{S}_2^{k+1})\circ\mathscr{S}_2^{i+1}\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+i+2}\mathfrak{f}\circ A_0^{n+k+i+2}\\
&+\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\pi_1^{(k-1)}\mathfrak{a}_3\circ\mathscr{S}_1^{k-1}W\circ A_0^{k}\partial_\theta p_2\circ\mathscr{S}_2^{k}p_2^{(n)}\circ\mathscr{S}_2^{k+1}\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\mathfrak{f}\circ A_0^{n+k+1}\\
+\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}&\pi_1^{(k-1)}\,\mathfrak{D}_{\mathfrak{a}_4,k,0}\circ\mathscr{S}_1^{k}\,W\circ A_0^{k}\,p_2^{(n)}\circ\mathscr{S}_2^{k+1}\,\rho\,\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\,\mathfrak{f}\circ A_0^{n+k+1}\\
&+\sum_{n=0}^{\infty}\pi_1^{(n-1)}\,\mathfrak{a}_4\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}\\
+\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}&\pi_1^{(k)}\partial_\theta p_1\circ\mathscr{S}_1^{k}p_1^{(i-1)}\circ\mathscr{S}_1^{k+1}\\
&\times\mathfrak{D}_{\mathfrak{a}_1,i,0}\circ\mathscr{S}_1^{i+k+1}\,W\circ A_0^{i+k+1}p_2^{(n)}\circ\mathscr{S}_2^{i+k+2}\rho\mathfrak{a}_2\circ\mathscr{S}_1^{n+i+k+2}\mathfrak{f}\circ A_0^{n+i+k+2}\\
&+\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}\pi_1^{(k)}\partial_\theta p_1\circ\mathscr{S}_1^{k}p_1^{(n-1)}\circ\mathscr{S}_1^{k+1}\mathfrak{a}_1\circ\mathscr{S}_1^{n-1}\,W\circ A_0^{n}\rho\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}\mathfrak{f}\circ A_0^{n+k+1}\\
+\sum_{k=0}^{\infty}\sum_{n=0}^{\infty}&p_1^{(k-1)}\,\mathfrak{D}_{\mathfrak{a}_1,k,0}\circ\mathscr{S}_1^{k}\,W\circ A_0^{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}\\
&+\sum_{n=0}^{\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}
\end{aligned}
$$

where we have used Lemma 7.41, can be studied by decomposing,

$$\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}=\mathfrak{a}_2\circ\mathscr{S}_2^{n+k+1}+(\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+1}-\mathfrak{a}_2\circ\mathscr{S}_2^{n+k+1}),$$

so as to apply Lemma 7.43 to obtain a factor $\mathfrak{D}_{2,\mathfrak{a}_2,n+k+1}$ from the terms containing the first factor and Lemma 7.37 to extract a further function W from the terms with the difference, which allows us to use the bounds in Remark D.8, and then proceeding in a similar way for the factors $\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+i+2}$, $\mathfrak{a}_2\circ\mathscr{S}_1^{n+k+i+2}$ and $\mathfrak{a}_5\circ\mathscr{S}1_1^{n+k+1}$ appearing in the other contributions, so that eventually, after using also (7.97) for the contributions where only the sum over n appears, a bound $|\langle\mathcal{A}_4\rangle|\le C\rho$ is obtained.

- The average of

$$
\begin{aligned}
\mathcal{A}_5 :=& \sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\sum_{i=0}^{k-1}\pi_1^{(i)}\,\Delta_{\pi_1,\pi_2,W,i}\,\pi_2^{(k-i-1)}\circ\mathscr{S}_2^{i+1}\,\partial_\theta p_2\circ\mathscr{S}_2^{k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^{n}\\
&+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\pi_1^{(k)}\,\Delta_{\partial_\theta p_1,\partial_\theta p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,q_2\circ\mathscr{S}_2^{n}\\
&+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}\sum_{i=0}^{n-k-2}\pi_1^{(k)}\,\partial_\theta p_1\circ\mathscr{S}_1^{k}p_1^{(i)}\circ\mathscr{S}_1^{k+1}\,\Delta_{p_1,p_2,W,i+k+1}\,p_2^{(n-k-2-i)}\circ\mathscr{S}_2^{i+k+2}\,q_2\circ\mathscr{S}_2^{n}\\
&+\sum_{n=0}^{\infty}\sum_{k=0}^{n-1}p_1^{(k)}\,\Delta_{p_1,p_2,W,k}\,p_2^{(n-k-1)}\circ\mathscr{S}_2^{k+1}\,\partial_\theta q_2\circ\mathscr{S}_2^{n}
\end{aligned}
$$

after writing

$$\pi_1^{(i)} = \pi_2^{(i)} + (\pi_1^{(i)} - \pi_2^{(i)}), \qquad \pi_1^{(k)} = \pi_2^{(k)} + (\pi_1^{(k)} - \pi_2^{(k)}), \qquad p_1^{(i)} = p_2^{(i)} + (p_1^{(i)} - p_2^{(i)}),$$
$$\pi_1^{(k)}\,\partial_\theta p_1 \circ \mathscr{S}_1^k p_1^{(i)} \circ \mathscr{S}_1^{k+1} = \pi_2^{(k)} \partial_\theta p_2 \circ \mathscr{S}_2^k p_2^{(i)} \circ \mathscr{S}_2^{k+1} + (\pi_1^{(k)} \partial_\theta p_1 \circ \mathscr{S}_1^k p_1^{(i)} \circ \mathscr{S}_1^{k+1} - \pi_2^{(k)} \partial_\theta p_2 \circ \mathscr{S}_2^k p_2^{(i)} \circ \mathscr{S}_2^{k+1}),$$

is found to be bounded as $|\langle \mathcal{A}_5 \rangle| \le C\rho$ by using Proposition 7.45 for the contributions with the factors involving the smooth terms $\pi_2^{(i)}$, $\pi_2^{(k)}$, $p_2^{(i)}$ and $\pi_2^{(k)} \partial_\theta p_2 \circ \mathscr{S}_2^k p_2^{(i)} \circ \mathscr{S}_2^{k+1}$, and by decomposing the differences $\pi_1^{(i)} - \pi_2^{(i)}$, $\pi_1^{(k)} - \pi_2^{(k)}$ and $p_1^{(i)} - p_2^{(i)}$, according to (7.65a), and, similarly, the last difference as

$$\begin{aligned}
&\pi_1^{(k)} \partial_\theta p_1 \circ \mathscr{S}_1^k p_1^{(i)} \circ \mathscr{S}_1^{k+1} - \pi_2^{(k)} \partial_\theta p_2 \circ \mathscr{S}_2^k p_2^{(i)} \circ \mathscr{S}_2^{k+1} \\
&\quad = \sum_{j=0}^{k-1} \pi_1^{(j)} \big(\pi_1 \circ \mathscr{S}_1^j - \pi_2 \circ \mathscr{S}_2^j\big) \pi_2^{(k-1-j)} \circ \mathscr{S}_2^{j+1}\, \partial_\theta p_2 \circ \mathscr{S}_2^k\, p_2^{(i)} \circ \mathscr{S}_2^{k+1} \\
&\quad + \pi_1^{(k)} \big(\partial_\theta \pi_1 \circ \mathscr{S}_1^k - \partial_\theta \pi_2 \circ \mathscr{S}_2^k\big)\, p_2^{(i)} \circ \mathscr{S}_2^{k+1} \\
&\quad + \sum_{j=k+1}^{k+1+i} \pi_1^{(k)}\, \partial_\theta p_1 \circ \mathscr{S}_1^k p_1^{(j-k-1)} \circ \mathscr{S}_1^k \big(p_1 \circ \mathscr{S}_1^j - p_2 \circ \mathscr{S}_2^j\big)\, p_2^{(i-j+k+1)} \circ \mathscr{S}_2^{k+1+j},
\end{aligned}$$

so as to use (7.72) and the bounds in Remark D.8.

- The average of
$$\mathcal{A}_6 := \sum_{n=0}^{\infty} \partial_\theta p_2^{(n)}\, \Delta_{q_1,q_2,W,n} + \sum_{n=0}^{\infty} p_2^{(n)}\, \Delta_{\partial_\theta q_1, \partial_\theta q_2, W, n}$$
is bounded as $|\langle \mathcal{A}_6 \rangle| \le C\rho$, as it follows from the definition (7.86) and Lemma 7.37.
- The contribution

$$\begin{aligned}
\mathcal{A}_7 := &\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \pi_1^{(i-1)} \mathfrak{D}_{\mathfrak{a}_3,i,0} \circ \mathscr{S}_1^i\, W \circ A_0^i\, (\pi_2^{(k)} \partial_\theta p_2 \circ \mathscr{S}_2^k\, p_2^{(n)} \circ \mathscr{S}_2^{k+1}) \circ \mathscr{S}_2^{i+1}\, q_2 \circ \mathscr{S}_2^{n+k+i+2} \\
&\qquad + \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \pi_1^{(k-1)} \mathfrak{a}_3 \circ \mathscr{S}_1^{k-1}\, W \circ A_0^k\, \partial_\theta p_2 \circ \mathscr{S}_2^k\, p_2^{(n)} \circ \mathscr{S}_2^{k+1}\, q_2 \circ \mathscr{S}_2^{n+k+1} \\
&+ \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \pi_1^{(k-1)} \mathfrak{D}_{\mathfrak{a}_4,k,0} \circ \mathscr{S}_1^k\, W \circ A_0^k\, p_2^{(n)} \circ \mathscr{S}_2^{k+1}\, q_2 \circ \mathscr{S}_2^{n+k+1} \\
&\qquad + \sum_{n=0}^{\infty} \pi_1^{(n-1)} \mathfrak{a}_4 \circ \mathscr{S}_1^{n-1}\, W \circ A_0^n\, q_2 \circ \mathscr{S}_2^n \\
&+ \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \pi_1^{(k)} \partial_\theta p_1 \circ \mathscr{S}_1^k \big(p_1^{(i-1)} \Delta_{\mathfrak{a}_1,i,0} \circ \mathscr{S}_1^i W \circ A_0^i\big) \circ \mathscr{S}_1^{k+1}\, p_2^{(n)} \circ \mathscr{S}_2^{i+k+2} q_2 \circ \mathscr{S}_2^{n+i+k+1} \\
&\qquad + \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} \pi_1^{(k)} \partial_\theta p_1 \circ \mathscr{S}_1^k \big(p_1^{(i-1)} \mathfrak{a}_1 \circ \mathscr{S}_1^{n-1} W \circ A_0^n\big) \circ \mathscr{S}_1^{k+1}\, q_2 \circ \mathscr{S}_2^{n+k+1} \\
&+ \sum_{k=0}^{\infty}\sum_{n=0}^{\infty} p_1^{(k-1)} \mathfrak{D}_{\mathfrak{a}_1,k,0} \circ \mathscr{S}_1^k\, W \circ A_0^k\, p_2^{(n)} \circ \mathscr{S}_2^{k+1}\, \partial_\theta q_2 \circ \mathscr{S}_2^{n+k+1} \\
&\qquad + \sum_{n=0}^{\infty} p_1^{(n-1)} \mathfrak{a}_1 \circ \mathscr{S}_1^{n-1}\, W \circ A_0^n\, \partial_\theta q_2 \circ \mathscr{S}_2^n,
\end{aligned}$$

where we have used Lemma 7.41, can be dealt with as follows. Consider for instance the sum in the fourth line and, after writing $\pi_1^{(n-1)} = \pi_2^{(n-1)} + (\pi_1^{(n-1)} - \pi_2^{(n-1)})$, expand $\pi_1^{(n-1)} - \pi_2^{(n-1)}$ according to (D.34), so as to extract either a factor $\Delta_{\pi_1,\pi_2,W,n-1}$ or a factor $\rho\,\mathfrak{f}\circ A_0^{n-1}$, while for the remaining term with $\pi_2^{(i-1)}$, after writing $\mathfrak{a}_4\circ\mathscr{S}_1^{n-1} = \mathfrak{a}_4\circ\mathscr{S}_2^{n-1} + (\mathfrak{a}_4\circ\mathscr{S}_1^{n-1} - \mathfrak{a}_4\circ\mathscr{S}_2^{n-1})$, expand $\mathfrak{a}_4\circ\mathscr{S}_1^{n-1} - \mathfrak{a}_4\circ\mathscr{S}_2^{n-1}$ according to (7.69), again with the aim of extracting either a sum of terms containing a function W or a term whose averaged absolute value is of order ρ. In that way we obtain two contributions with a further sum, but also with a further factor which is of order ρ and, if the average of its absolute value is not of order ρ, contains a further function W, so that the overall average is bounded proportionally to ρ. The average of the remaining term

$$\sum_{n=0}^{\infty} \pi_2^{(n-1)}\mathfrak{a}_4\circ\mathscr{S}_2^{n-1}\, W\circ A_0^n\, q_2\circ\mathscr{S}_2^n$$

is controlled through Proposition 7.45, which ensures an overall bound of order ρ. The contributions in the other lines are discussed analogously, and eventually the bound $|\langle\mathcal{A}_7\rangle| \le C\rho$ follows.

- The contribution

$$\mathcal{A}_8 := \sum_{n=0}^{\infty} \partial_\theta p_2^{(n)}\, \rho\, \mathfrak{a}_2\circ\mathscr{S}_1^n\, \mathfrak{f}\circ A_0^n + \sum_{n=0}^{\infty} p_2^{(n-1)}\, \rho\, \mathfrak{a}_5\circ\mathscr{S}_1^n\, \mathfrak{f}\circ A_0^n$$

can be studied as follows. First of all, write

$$\mathfrak{a}_2\circ\mathscr{S}_1^n = \mathfrak{a}_2\circ\mathscr{S}_2^n + (\mathfrak{a}_2\circ\mathscr{S}_1^n - \mathfrak{a}_2\circ\mathscr{S}_2^n), \qquad \mathfrak{a}_5\circ\mathscr{S}_1^n = \mathfrak{a}_5\circ\mathscr{S}_2^n + (\mathfrak{a}_5\circ\mathscr{S}_1^n - \mathfrak{a}_5\circ\mathscr{S}_2^n),$$

so that, when considering the terms $\mathfrak{a}_2\circ\mathscr{S}_2^n$ and $\mathfrak{a}_5\circ\mathscr{S}_2^n$, we write $\rho\,\mathfrak{f}\circ A_0^n = W\circ A_0^{n+1} - W\circ A_0^n$, in order to apply first Lemma 7.43 to obtain

$$\begin{aligned} &\sum_{n=0}^{\infty} \partial_\theta p_2^{(n-1)}\, \mathfrak{D}_{\mathfrak{a}_2,n,0}\circ\mathscr{S}_2^n\, W\circ A_0^n + \sum_{n=0}^{\infty} p_2^{(n-2)}\mathfrak{D}_{\mathfrak{a}_5,n-1,0}\circ\mathscr{S}_2^n\, W\circ A_0^n \\ &= -\mathfrak{a}_2\, W + \sum_{n=1}^{\infty} \partial_\theta p_2^{(n-1)}\mathfrak{D}_{\mathfrak{a}_2,n,0}\circ\mathscr{S}_2^n W\circ A_0^n - \mathfrak{a}_5\, W + \sum_{n=1}^{\infty} p_2^{(n-2)}\mathfrak{D}_{\mathfrak{a}_5,n-1,0}\circ\mathscr{S}_2^n W\circ A_0^n, \end{aligned} \tag{D.39}$$

and hence Proposition 7.45 to bound the average as

$$C\rho + C\sum_{n=1}^{\infty}(1-\rho\gamma')^n\big((1+\alpha_0 i)\lambda^{-\alpha_0 n} + \rho + n\rho + n^2\rho^3\big)\big(|||\mathfrak{D}_{\mathfrak{a}_2,n,0}|||_{\alpha_0,2}^{-} + |||\mathfrak{D}_{\mathfrak{a}_5,n,0}|||_{\alpha_0,2}^{-}\big) \le C\rho,$$

while, when considering the terms with the differences, i.e.

$$\sum_{n=0}^{\infty} \partial_\theta p_2^{(n)}\, \rho\, \big(\mathfrak{a}_2\circ\mathscr{S}_1^n - \mathfrak{a}_2\circ\mathscr{S}_2^n\big)\, \mathfrak{f}\circ A_0^n + \sum_{n=0}^{\infty} p_2^{(n-1)}\, \rho\, \big(\mathfrak{a}_5\circ\mathscr{S}_1^n - \mathfrak{a}_5\circ\mathscr{S}_2^n\big)\, \mathfrak{f}\circ A_0^n, \tag{D.40}$$

we reason as done when bounding the second contribution in the third line of (7.98) in the proof of Lemma 7.47, in order to obtain a bound proportional to ρ, which, together with the previous bound, yields that $|\langle\mathcal{A}_8\rangle| \le C\rho$.

To obtain the second bound in (7.104), we consider

$$\Big(\partial_\theta(p_1^{(n_1)}q_2\circ\mathscr{S}_1^{n_1}) - \partial_\theta(p_2^{(n_2)}q_2\circ\mathscr{S}_2^{n_2})\Big)\,\Big(\partial_\theta(p_1^{(n_1)}q_2\circ\mathscr{S}_1^{n_1}) - \partial_\theta(p_2^{(n_2)}q_2\circ\mathscr{S}_2^{n_2})\Big),$$

where both factors can be written as in (D.38). Then observe that all the contributions $\mathcal{A}_1$ to $\mathcal{A}_7$ are $O(1)$ in ρ and contains at least a function W, while the contribution $\mathcal{A}_8$ can be dealt with as when discussing the first bound in (7.104) and written as the sum of the two contributions (D.39) and (D.40), which both contains a function W. This yields that

$$\left|\langle \mathcal{A}_i \mathcal{A}_j \rangle\right| \leq C\rho, \qquad i,j = 1,\ldots,8,$$

by Theorem 4. Then the second bound in (7.104) follows as well.

D.4 Second derivative of the conjugation: proof of Lemma 7.52

Since the proof of Lemma 7.52 follows very closely the same scheme as Propositions 7.49 and 7.50, we confine ourselves to discuss briefly how to proceed without entering into the details.

Once more, as in proving Lemma 7.37 and Proposition 7.50 (see Appendices D.2 and D.3, and recall Remark 7.26), also to prove Lemma 7.52 we discuss explicitly the case $\rho < \rho_0$, since the statement trivially holds true for $\rho \geq \rho_0$.

We want to show that

$$|\langle(\partial_\theta^2 h_2(\theta,\cdot) - \partial_\theta^2 \overline{h}(\theta))^2\rangle| \leq C\rho, \qquad \langle(\partial_\theta^2 h_1(\theta,\cdot) - \partial_\theta^2 h_2(\theta,\cdot))^2\rangle \leq C\rho\,, \tag{D.41}$$

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 D.1 and D.3, 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 (B.1b),

$$\begin{aligned}
\partial_\theta^2 p_2^{(n)} &= \sum_{k=0}^{n-1} (p_2\,(\partial_\theta G_2)^2)^{(k)}\, \partial_\theta^2 p_2 \circ \mathscr{S}_2^k\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1} \\
&+ \sum_{k=0}^{n-1}\sum_{i=0}^{k-1} (p_2\,(\partial_\theta G_2)^2)^{(i)}\, (\partial_\theta p_2 \partial_\theta G_2) \circ \mathscr{S}_2^i\, (p_2 \partial_\theta G_2)^{(k-i-1)} \circ \mathscr{S}_2^{i+1} \partial_\theta p_2 \circ \mathscr{S}_2^k\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1} \\
&+ \sum_{k=0}^{n-1}\sum_{i=0}^{k-1} (p_2\,(\partial_\theta G_2)^2)^{(i)}\, (p_2\, \partial_\theta^2 G_2) \circ \mathscr{S}_2^i\, (p_2 \partial_\theta G_2)^{(k-i-1)} \circ \mathscr{S}_2^{i+1} \partial_\theta p_2 \circ \mathscr{S}_2^k\, p_2^{(n-k-1)} \circ \mathscr{S}_2^{k+1} \\
&+ \sum_{k=0}^{n-1}\sum_{i=0}^{n-k-2} (p_2\, \partial_\theta G_2)^{(k)}\, \partial_\theta p_2 \circ \mathscr{S}_2^k\, (p_2 \partial_\theta G_2)^{(i)} \circ \mathscr{S}_2^{k+1} \partial_\theta p_2 \circ \mathscr{S}_2^{k+1+i}\, p_2^{(n-k-i-2)} \circ \mathscr{S}_2^{k+2+i},
\end{aligned} \tag{D.42}$$

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

$$\partial_\theta^2 p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n - \partial_\theta^2 \overline{p}^{(n)}\, \overline{q} \circ \overline{G}^n, \qquad \partial_\theta^2 p_1^{(n)}\, q_1 \circ \mathscr{S}_1^n - \partial_\theta^2 p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n,$$

by following the same scheme as in Appendices D.1 and D.3, respectively.

In that way, we obtain expressions where, with respect to (D.3) and (D.35), in addition to the functions $\pi_1 = p_1 \partial_\theta G_1$, $\pi_2 = p_2 \partial_\theta G_2$ and $\overline{p}\, \partial_\theta \overline{G}$ and, 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, together with the differences (D.32) to (D.34), we have to expand also the differences

$$(p_1\partial_\theta^2 G_1)^{(n)} - (p_2\partial_\theta^2 G_2)^{(n)}, \quad (\partial_\theta p_1 \partial_\theta G_1)^{(n)} - (\partial_\theta p_2 \partial_\theta G_2)^{(n)}, \quad (p_1(\partial_\theta G_1)^2)^{(n)} - (p_2(\partial_\theta G_2)^2)^{(n)},$$

when considering $\partial_\theta^2 p_1^{(n)}\, q_1 \circ \mathscr{S}_1^n - \partial_\theta^2 p_2^{(n)}\, q_2 \circ \mathscr{S}_2^n$, while inserting (D.42) 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 (D.3).

Comparing (B.1b) with (B.1a) we observe 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, and this produces a further factor ρ, which compensates the factor $1/\rho$ arising from the sum.

E Triple correlations: proof of (4.9)

We start by proving a correlation inequality which somehow generalizes (7.8) in Proposition 7.6 and allows us to estimate averages of the form

$$\Big\langle f_1 \mu^{(n_1)} \circ A_0 G \circ A_0^{n_1+1} \mu^{(n_2)} \circ A_0^{n_1+m+1} f_2 \circ A_0^{n_1+n_2+m+1} \Big\rangle, \qquad G := \prod_{i=0}^{m-1} g_i \circ A_0^i, \tag{E.1}$$

where all functions $f_1, f_2, g_0, \ldots, g_{m-1}$ are in $\mathfrak{B}_\alpha(\mathbb{T}^2, \mathbb{R})$, for some $\alpha > 0$, and $\mu^{(n)}$ is defined by (7.15), with μ as in (7.14). Since we need to apply the result only when both functions f_1 and f_2 have zero average, we confine ourselves to such a case.

Lemma E.1. *Let $f_1, f_2, \mu, G \in \mathfrak{B}_\alpha(\mathbb{T}^2, \mathbb{R})$, with $\alpha \in (0, 1]$, be such that $\langle f_1 \rangle = \langle f_2 \rangle = 0$ and μ is of the form* (7.14). *Then, for every $m \geq 0$, one has*

$$\sum_{n_1, n_2 = 1}^{\infty} \Big\langle f_1 \mu^{(n_1)} \circ A_0 G \circ A_0^{n_1+1} \mu^{(n_2)} \circ A_0^{n_1+m+1} f_2 \circ A_0^{n_1+n_2+m+1} \Big\rangle \leq C \alpha^{-1} \|f_1\|_\alpha \|f_2\|_\alpha \|G\|_\alpha, \tag{E.2}$$

Proof. If we all $N_1 = n_1 + m + 1$ and $N_2 = n_1 + n_2 + m + 2$, and proceed as in Subsection 7.2.2 by using the first line of (7.6) for $\mu^{(n_1)}$ and (7.7) for $\mu^{(n_2)}$, we write

$$\begin{aligned}
\Big\langle f_1 \mu^{(n_1)} \circ A_0 G \circ A_0^{n_1+1} \mu^{(n_2)} \circ A_0^{N_1} f_2 \circ A_0^{N_2} \Big\rangle &= \langle \mu \rangle^{n_1+n_2} \Big\langle f_1 G \circ A_0^{n_1+1} f_2 \circ A_0^{N_2} \Big\rangle \\
&+ \rho \sum_{j_1=0}^{n_1-1} \langle \mu \rangle^{j_1+n_2} \Big\langle f_1 (\tilde{v}\, \mu^{(n_1-j_1-1)} \circ A_0) \circ A_0^{j_1} G \circ A_0^{n_1+1} f_2 \circ A_0^{N_2} \Big\rangle \\
&+ \rho \sum_{j_2=0}^{n_2-1} \langle \mu \rangle^{n_1+j_2} \Big\langle f_1 G \circ A_0^{n_1+1} (\mu^{(n_2-j_2-1)} \tilde{v} \circ A_0^{n_2-j_2-1}) \circ A_0^{N_1} f_2 \circ A_0^{N_2} \Big\rangle \\
&+ \rho^2 \sum_{j_1=0}^{n_1-1} \sum_{j_2=0}^{n_2-1} \langle \mu \rangle^{j_1+j_2} \\
&\qquad \Big\langle f_1 (\tilde{v}\, \mu^{(n_1-j_1-1)} \circ A_0) \circ A_0^{j_1} G \circ A_0^{n_1+1} (\mu^{(n_2-j_2-1)} \tilde{v} \circ A_0^{n_2-j_2-1}) \circ A_0^{N_1} f_2 \circ A_0^{N_2} \Big\rangle.
\end{aligned} \tag{E.3}$$

Thus, to bound the double sum in the two last lines of (E.3), if $j_1 > j_2$, by Proposition 2.3, with

$$g_+ = f_1, \qquad g_- = (\tilde{v} \mu^{(n_1-j_1-1)} \circ A_0) G \circ A_0^{n_1+1-j_1} (\mu^{(n_2-j_2-1)} \tilde{v} \circ A_0^{n_2-j_2-1}) \circ A_0^{N_1-j_1} f_2 \circ A_0^{N_2-j_1},$$

we get

$$\Big|\Big\langle f_1(\tilde v\mu^{(n_1-j_1-1)}\circ A_0)\circ A_0^{j_1}G\circ A_0^{n_1+1}(\mu^{(n_2-j_2-1)}\tilde v\circ A_0^{n_2-j_2-1})\circ A_0^{N_1}f_2\circ A_0^{N_2}\Big\rangle\Big|$$
$$\le C(1-\rho\gamma)^{n_1+n_2}(1+\alpha j_1)\lambda^{-\alpha(j_1+j_2)/2}\|f_1\|_\alpha\|f_2\|_\alpha\|G\|_\alpha,$$

where we have bounded $j_1 > (j_1+j_2)/2$; we arrive at an identical estimate in the case where $j_1 \le j_2$, using again Proposition 2.3, now with

$$g_+ = f_1(\tilde v\mu^{(n_1-j_1-1)}\circ A_0)\circ A_0^{j_1}G\circ A_0^{n_1+1}(\mu^{(n_2-j_2-1)}\tilde v\circ A_0^{n_2-j_2-1})\circ A_0^{N_1}, \qquad g_- = f_2.$$

Summing over j_1 and j_2, we obtain immediately the bound

$$C\alpha^{-1}\rho^2(1-\rho\gamma)^{n_1+n_2}\|f_1\|_\alpha\|f_2\|_\alpha\|G\|_\alpha,$$

with C independent of ρ and α, so that summing also over n_1 and n_2 produces an estimate of the form of the r.h.s. of (E.2). Reasoning in a similar way for the term in the first line of (E.3), with n_1 and n_2 in place of j_1 and j_2, we find that

$$\sum_{n_1,n_2=0}^{\infty}\Big|\big\langle f_1 G\circ A_0^{n_1+1}f_2\circ A_0^{n_1+n_2+m+1}\big\rangle\Big| \le C\alpha^{-1}\|f_1\|_\alpha\|f_2\|_\alpha\|G\|_\alpha. \tag{E.4}$$

Finally, the sums in the second and third lines of (E.3) are dealt with in the same way by discussing separately the two cases $j_1 > n_2$ and $j_1 \le n_2$ and the two cases $j_2 > n_1$ and $j_2 \le n_1$, respectively, so as to get once more an estimate of the form of the r.h.s. of (E.2). □

Remark E.2. By looking at the proof of Lemma E.1, one sees that the result still holds if either a few functions μ appearing in $\mu^{(n_1)}$ and in $\mu^{(n_2)}$ are replaced with any functions with the same regularity (see Remark 7.18 for a similar comment) or an arbitrary number of such functions μ are replaced with different functions μ' which, besides sharing the same regularity, still admit the bounds $\|\mu'\|_\infty \le (1-\rho\gamma)$.

Observe now that

$$W_0(\psi)-W_{00}(\psi) = \rho^2\sum_{n=1}^{\infty}\sum_{i=0}^{n-1}\langle\mu\rangle^i\tilde v(A_0^{-i-1}\psi)\mu^{(-n+i+1)}(A_0^{-i-1}\psi)\,b(A_0^{-(n+1)}\psi)$$
$$= \rho^2\sum_{n=1}^{\infty}\sum_{i=0}^{n-1}\langle\mu\rangle^i\tilde v(A_0^{-i-1}\psi)\mu^{(n-i-1)}(A_0^{-n}\psi)\,b(A_0^{-(n+1)}\psi),$$

where the second relation in (7.15) has been used to obtain the second line. By rearranging the sums, we can write,

$$\big(W_0-W_{00}\big)^2 = \rho^4\sum_{i_1=0}^{\infty}\sum_{n_1=i_1+1}^{\infty}\sum_{i_2=0}^{\infty}\sum_{n_2=i_2+1}^{\infty}\langle\mu\rangle^{i_1+i_2}\tilde v\circ A_0^{-i_1-1}$$
$$\times\,\mu^{(n_1-i_1-1)}\circ A_0^{-n_1}b\circ A_0^{-n_1-1}\tilde v\circ A_0^{-i_2-1}\mu^{(n_2-i_2-1)}\circ A_0^{-n_2}b\circ A_0^{-n_2-1}. \tag{E.5}$$

To bound the average of (E.5) we order the sums distinguishing the cases $i_1 \ge i_2$ and $i_1 < i_2$; in the first case we further study separately the case $n_2 > n_1$ (which implies $i_2 \le i_1 < n_1 < n_2$) and $n_2 \le n_1$ (which implies either $i_2 < n_2 \le i_1 < n_1$ or $i_2 \le i_1 < n_2 \le n_1$), while the second case can be reduced to the first one by renaming the indices. Therefore we obtain

$$
\begin{aligned}
\big|\big\langle (W_0-W_{00})^2\big\rangle\big| \le 2\rho^4 &\sum_{i_2=0}^{\infty}\sum_{i_1=i_2}^{\infty}\sum_{n_1=i_1+1}^{\infty}\sum_{n_2=n_1+1}^{\infty}\langle\mu\rangle^{i_1+i_2} \\
&\times\big|\big\langle \tilde v\circ A_0^{n_2-i_1}\mu^{(n_1-i_1-1)}\circ A_0^{n_2-n_1+1}b\circ A_0^{n_2-n_1}\tilde v\circ A_0^{n_2-i_2}\mu^{(n_2-i_2-1)}\circ A_0\, b\big\rangle\big| \\
+\,2\rho^4 &\sum_{i_2=0}^{\infty}\sum_{n_2=i_2+1}^{\infty}\sum_{i_1=n_2}^{\infty}\sum_{n_1=i_1+1}^{\infty}\langle\mu\rangle^{i_1+i_2} \\
&\times\big|\big\langle \tilde v\circ A_0^{n_1-i_1}\mu^{(n_1-i_1-1)}\circ A_0\, b\,\tilde v\circ A_0^{n_1-i_2}\mu^{(n_2-i_2-1)}\circ A_0^{n_1-n_2+1}b\circ A_0^{n_1-n_2}\big\rangle\big| \\
+\,2\rho^4 &\sum_{i_2=0}^{\infty}\sum_{i_1=i_2}^{\infty}\sum_{n_2=i_1+1}^{\infty}\sum_{n_1=n_2}^{\infty}\langle\mu\rangle^{i_1+i_2} \\
&\times\big|\big\langle \tilde v\circ A_0^{n_1-i_1}\mu^{(n_1-i_1-1)}\circ A_0\, b\,\tilde v\circ A_0^{n_1-i_2}\mu^{(n_2-i_2-1)}\circ A_0^{n_1-n_2+1}\, b\circ A_0^{n_1-n_2}\big\rangle\big|,
\end{aligned}
$$

that we rewrite, more conveniently,

$$
\begin{aligned}
\big|\big\langle (W_0-W_{00})^2\big\rangle\big| \le 2\rho^4 &\sum_{i=0}^{\infty}\sum_{m_1,m_2,m_3=0}^{\infty}\langle\mu\rangle^{2i}\big|\big\langle b\,\mu^{(m_1)}\circ A_0\,(b\mu)\circ A_0^{m_1+1}\big(\mu^{(m_2)}\circ A_0^{m_1+2}\big)^2 \\
&\quad\times\big(\tilde v\langle\mu\rangle^{m_3}\mu^{(m_3)}\big)\circ A_0^{m_1+m_2+2}\tilde v\circ A_0^{m_1+m_2+m_3+2}\big\rangle\big| \\
+\,2\rho^4 &\sum_{i=0}^{\infty}\sum_{m_1,m_2,m_3=0}^{\infty}\langle\mu\rangle^{2i}\big|\big\langle b\,\mu^{(m_1)}\circ A_0\,(\langle\mu\rangle\tilde v)\circ A_0^{m_1+1}\langle\mu\rangle^{m_2-1} \\
&\quad\times\big(b\langle\mu\rangle\big)\circ A_0^{m_1+m_2+1}\big(\langle\mu\rangle^{m_3}\mu^{(m_3)}\big)\circ A_0^{m_1+m_2+2}\tilde v\circ A_0^{m_1+m_2+m_3+2}\big\rangle\big| \\
+\,2\rho^4 &\sum_{i=0}^{\infty}\sum_{m_1,m_2,m_3=0}^{\infty}\langle\mu\rangle^{2i}\big|\big\langle b\,\mu^{(m_1-1)}\circ A_0\,(b\mu)\circ A_0^{m_1}\big(\mu^{(m_2)}\circ A_0^{m_1+1}\big)^2 \\
&\quad\times\big(\tilde v\langle\mu\rangle^{m_3}\mu^{(m_3)}\big)\circ A_0^{m_1+m_2+1}\tilde v\circ A_0^{m_1+m_2+m_3+1}\big\rangle\big|.
\end{aligned}
\tag{E.6}
$$

Note that all the averages in (E.6) are of the form (E.1), or of a similar form as described in Remark E.2, with $f_1 = b$ and $f_2 = \tilde v$ in all cases.

The sum over i produces a factor C/ρ. Since $\langle\tilde v\rangle = \langle b\rangle = 0$, we can apply (E.2) to all three contributions in the above expression. Finally, after summing over m_1 and m_3, we get

$$
\big|\big\langle (W_0-W_{00})^2\big\rangle\big| \le C\gamma^{-1}\rho^3\sum_{m_2=0}^{\infty}\big(2\|(\mu^{(m_2)})^2\|_\alpha + \langle\mu\rangle^{m_2}\|\tilde v\|_\alpha\big),
$$

and hence, using (7.3) to deal with the first term and taking into account that $\alpha = O(1)$ in ρ in the case we are considering, we arrive at the bound (4.9).

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-Amsterdam, 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] A. Ayyer, M. Stenlund, *Exponential decay of correlations for randomly chosen hyperbolic toral automorphisms*, Chaos **17** (2007), no. 4, 043116, 7 pp.

[6] M. Blank, G. Keller, C. Liverani, *Ruelle-Perron-Frobenius spectrum for Anosov maps*, Nonlinearity **15** (2002), no. 6, 1905-1973.

[7] 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.

[8] N.N. Bogoliubov, Y.A. Mitropolsky, *Asymptotic Methods in the Theory of Non-linear Oscillations*, Gordon and Breach Science Publishers, New York, 1961.

[9] 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.

[10] 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.

[11] J. Bricmont, A. Kuipianen, *Diffusion in energy conserving coupled maps*, Comm. Math. Phys. **321** (2013), no. 2, 311-369.

[12] M. Brin, G. Stuck, *Introduction to Dynamical Systems*, Cambridge University Press, Cambridge, 2002.

[13] A. Carati, L. Galgani, F. Gangemi, R. Gangemi, *Electronic trajectories in atomic physics: the chemical bond in the* H_2^+ *ion*, Chaos **30** (2020), 063109, 8pp.

[14] R. Castorrini, C. Liverani, *Quantitative statistical properties of two-dimensional partially hyperbolic systems*, Adv. Math. **409** (2022), Paper no. 108625, 122 pp.

[15] N.I. Chernov, *Limit theorems and Markov approximations for chaotic dynamical systems.* Probab. Th. Rel. Fields **101** (1995), 321-362.

[16] 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.

[17] 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.

[18] M.F. Demers, *A gentle introduction to anisotropic Banach spaces*, Chaos, Solitons and Fractals **116** (2018), 29-42.

[19] 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.

[20] J. De Simoi, C. Liverani, *The martingale approach after Varadhan and Dolgopyat*, Hyperbolic dynamics, fluctuations and large deviations, 311-339, Proc. Sympos. Pure Math. Vol. 89, American Mathematical Society, Providence, RI, 2015.

[21] J. De Simoi, C. Liverani, *Statistical properties of mostly contracting fast-slow partially hyperbolic systems*, Invent. Math. **206** (2016), 147-227.

[22] J. De Simoi, C. Liverani, *Limit theorems for fast-slow partially hyperbolic systems*, Invent. Math. **213** (2018), no. 3, 811-1016.

[23] R.L. Devaney, *An Introduction to Chaotic Dynamical Systems*, Benjamin/Cummings, Menlo Park, 1986.

[24] H.A. Dijkstra, *Nonlinear Climate Dynamics*, Cambridge University Press, Cambridge, 2013.

[25] D. Dolgopyat, *Averaging and invariant measures*, Mosc. Math. J. **5** (2005), no. 3, 537-576.

[26] D. Dolgopyat, C. Liverani, *Energy transfer in a fast-slow Hamiltonian system*, Comm. Math. Phys. **308** (2011), no. 1, 201-225.

[27] J. Franks, *Anosov diffeomorphisms.Global Analysis*, Proc. Sympos. Pure Math. Vol. 14, American Mathematical Society, Providence, R.I., 1970.

[28] G. Gallavotti, F. Bonetto, G. Gentile, *Aspects of the Ergodic, Qualitative and Statistical Theory of Motion*, Springer, Berlin, 2004.

[29] G. Gallavotti, G. Gentile, A. Giuliani, *Resonances within chaos*, Chaos **22** (2012), no.52, 026108, 6 pp.

[30] G. Gallavotti, E.G.D. Cohen, *Dynamical ensembles in stationary states*, J. Statist. Phys. **80** (1995), no. 5-6, 931-970.

[31] M. Ghil, V. Lucarini, *The physics of climate variability and climate change*, Rev. Modern Phys. **92** (2020), no. 3, 035002, 77 pp.

[32] J.M. González-Miranda, *Synchronization and Control of Chaos. An Introduction for Scientists and Engineers*, Imperial College Press, London, 2004.

[33] J. Hale, *Ordinary Differential Equations*, John Wiley & Sons, New York, 1980.

[34] Ph.J. Holmes, *Averaging and chaotic motions in forced oscillations*, SIAM J. Appl. Math. **38** (1980), no. 1, 65-80.

[35] 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.

[36] Yu. Kifer, *Averaging and climate models*, Progr. Probab. **49** (2001), 171-188.

[37] N.M. Krylov, N.N. Bogolyubov, *Introduction to Non-Linear Mechanics*, Princeton University Press, Princeton, NJ, 1943.

[38] C.A. Kitio Kwuimy, P. Woafo, *Dynamics, chaos and synchronization of self-sustained electromechanical systems with clamped-free flexible arms*, Nonlinear Dynam. **53** (2008), 201-213.

[39] J.L. Lagrange, *Mécanique Analitique*, 1788 (edition Albert Blanchard, 1965).

[40] C. Liverani, *Transport in partially hyperbolic fast-slow systems*, Proceedings of the International Congress of Mathematicians, Rio de Janeiro 2018, Vol. III. Invited lectures, 2643-2667, World Sci. Publ., Hackensack, NJ, 2018.

[41] 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.

[42] 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.

[43] C.D. Murray, S.F. Dermott, *Solar System Dynamics*, Cambridge University Press, Cambridge, 1999.

[44] S.E. Newhouse, *On codimension one Anosov diffeomorphisms*, Amer. J. Math. **92** (1970), 761-770.

[45] L.M. Pecora, Th.L. Carroll, *Synchronization in chaotic systems*, Phys. Rev. Lett. **64** (1990), no. 8, 821-825.

[46] L.M. Pecora, Th.L. Carroll, *Synchronization of chaotic systems*, Chaos **25** (2015), no. 9, 097611, 12pp.

[47] 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.

[48] A. Pikovsky, M. Rosenblum, J. Kurths, *Synchronization. A Universal Concept in Nonlinear Sciences*, Cambridge University Press, Cambridge, 2001.

[49] J.H. Sanders, F. Verhulst, J. Murdock, *Averaging Methods in Nonlinear Dynamical Systems*, Springer, New York, 2007.

[50] M.R. Snavely, *Markov partitions for the two-dimensional torus*, Proc. Amer. Math. Soc. **113** (1991), no. 2, 517-527.

[51] S. Soldatenko, A. Bogomolov, A. Ronzhin, *Mathematical modelling of climate change and variability in the context of outdoor ergonomics*, Mathematics **9** (2021), 2920, 26pp.

[52] F. Verhulst, *Nonlinear Differential Equation and Dynamical Systems*, Springer, Berlin, 1990.