AN OSCILLATOR DRIVEN BY ALGEBRAICALLY DECORRELATING NOISE

CHISTOPHE GOMEZ, GAUTAM IYER, HAI LE, AND ALEXEI NOVIKOV

Abstract. We consider a stochastically forced nonlinear oscillator driven by a stationary Gaussian noise that has an algebraically decaying covariance function. It is well known that such noise processes can be renormalized to converge to *fractional* Brownian motion, a process that has memory. In contrast, we show that the renormalized limit of the nonlinear oscillator driven by this noise converges to diffusion driven by standard (not fractional) Brownian motion, and thus retains no memory in the scaling limit. The proof is based on the study of a fast-slow system using the perturbed test function method.

1. Introduction

Consider a stochastically forced nonlinear oscillator with 1 degree of freedom:

$$\ddot{x}_t + f(x_t) = \varepsilon v(t), \quad x_0 \in \mathbb{R}, \quad \dot{x}_0 = y_0 \in \mathbb{R}\,. \tag{1.1}$$

Here $f\colon \mathbb{R} \to \mathbb{R}$ is a given smooth function, and v is a stochastic process representing the noise. Our interest is to study the asymptotic long-time behavior of x when the noise v has an algebraically decaying covariance function. In this case, the rescaled noise converges to fractional Brownian motion (fBm), a process that has a memory. We aim to study how the nonlinear dynamics affects the limiting behavior, and study whether or not the rescaled oscillator also converges to a process with memory.

A similar question was studied by Komorowski et al. [KNR14] in the case of a passive tracer particle advected by a periodic shear flow. In this case, it turns out that there is a certain parameter regime where the time rescaled dynamics is Markovian and the memory effect of the noise is forgotten. However, there is also a regime (namely the "very long-time" behavior when the Hurst index of the driving noise is larger than 1/2) where the memory effect persists. In contrast, for the oscillator (1.1), we will show that the memory effect *never* persists, and the effective long-time behavior is always Markovian. The main reason for this is that the oscillatory nature of the deterministic dynamics cancels slowly decaying correlations, and destroys the memory effect. To study (1.1) we cannot use the limit theorem for additive functionals fBm's used by [KNR14]. Instead, we recast (1.1) as a fast-slow system and use the perturbed test function method. Even though this method was introduced in the '76 [PSV76, Kus84], it has applications to many problems arising

2010 *Mathematics Subject Classification.* Primary 37H05, 60H10; Secondary 60F05, 60G15.

Key words and phrases. Hamiltonian systems, Random perturbations, Algebraic decaying correlations, Fractional processes, Diffusion-approximation.

This work has been partially supported by the National Science Foundation under grants NSF DMS-1813943, DMS-1515187 to AN and HL, NSF DMS-1814147 to GI, and the Center for Nonlinear Analysis.

today [FGPSl07, DRV21, CV21], and is the only method we presently know that can be used to prove our main result.

To describe our setup, we first introduce the Hamiltonian

$$H(x,y) \stackrel{\text{def}}{=} \frac{1}{2}y^2 + \int_0^x f(s)\,ds\,,$$

and recast (1.1) as a stochastically perturbed Hamiltonian system

$$\dot{x}_t = y_t = \partial_y H(x_t, y_t)\,, \tag{1.2a}$$

$$\dot{y}_t = \ddot{x}_t = -f(x_t) + \varepsilon v(t) = -\partial_x H(x_t, y_t) + \varepsilon v(t)\,. \tag{1.2b}$$

Setting $X_t \stackrel{\text{def}}{=} (x_t, y_t)^T$, we rewrite (1.2a)–(1.2b) compactly as

$$\dot{X}_t = \nabla^\perp H(X_t) + \varepsilon v(t) e_2, \quad X_0 = (x_0, y_0) \in \mathbb{R}^2\,. \tag{1.3}$$

where $\nabla^\perp \stackrel{\text{def}}{=} (\partial_y, -\partial_x)^T$, and $e_2 = (0,1)^T$.

To study the long-time behavior, we consider the time rescaled process $X^\varepsilon(t) \stackrel{\text{def}}{=} X(t/\varepsilon^2)$, and observe

$$\dot{X}^\varepsilon_t = \frac{1}{\varepsilon^2}\nabla^\perp H(X^\varepsilon_t) + \frac{1}{\varepsilon} v\left(\frac{t}{\varepsilon^2}\right) e_2\,, \quad X^\varepsilon_0 = X_0 = (x_0, y_0) \in \mathbb{R}^2. \tag{1.4}$$

When the noise v is white, the behavior of X^ε as $\varepsilon \to 0$ is completely described by the averaging principle of Freidlin and Wentzell [FW93, FW94, FW98]. To briefly explain this, we note that in the absence of noise (i.e. when $v \equiv 0$), the process X^ε travels very fast along level sets of the Hamiltonian. In the presence of noise, the process X^ε will also diffuse slowly across these level sets. To capture the limiting behavior we factor out the fast motion by projecting to the *Reeb graph* of the Hamiltonian, which has the effect of identifying all closed trajectories of the Hamiltonian system to points. Now, as $\varepsilon \to 0$, this projection converges to a diffusion driven by a standard Brownian motion and we refer the reader to [FW93, FW94, FW98] for details. (We also remark that when the noise v is white, one may also study the behavior of X on time scales shorter than $1/\varepsilon^2$. In some scenarios, a robust and stable limiting behavior is also observed on these time scales and has been studied by several authors [You88, YJ91, Bak11, HV14a, HV14b, HKPG16, TV16, HIK$^+$18].)

In this paper, our interest is to study the case when the driving noise in (1.4) is colored in time and has algebraically decaying correlations. It is well known that this noise can be renormalized to converge to a fractional Brownian motion (fBm), with Hurst parameter that is determined from the decay rate of the correlation function. When the Hurst parameter is not half (which happens when the correlation function decays like $1/t^\gamma$ with $\gamma \neq 1$), the renormalized limit of the noise has memory and is a non-Markovian process. We aim to study whether the memory effect is also present in the $\varepsilon \to 0$ limit of the nonlinear oscillator (1.4), where the noise is combined with the Hamiltonian dynamics.

At present, we are only able to analyze the scenario where the Hamiltonian is smooth with exactly one non-degenerate critical point. In this case, it is convenient to think about the dynamics of X^ε in terms of action-angle coordinates. The angular coordinate of X^ε changes very fast, and has no meaningful limit as $\varepsilon \to 0$. The action coordinate of X^ε, on the other hand, changes slowly as a result of the interaction between the noise and the averaged angular coordinate and has a non-trivial limit $\varepsilon \to 0$. To study this we use the Hamiltonian itself as a proxy for the action coordinate and state our results in terms of convergence of the process

$H(X^\varepsilon)$. (To relate this to the Freidlin–Wentzell framework, we note that when H has exactly one non-degenerate critical point, the Reeb graph has exactly one vertex and one edge, and $H(X^\varepsilon)$ is precisely the projection of X^ε onto the Reeb graph.)

In this paper, we prove two results. They show that even though the driving noise has memory, the time correlations are destroyed by the oscillatory dynamics of (1.4), and $H(X^\varepsilon)$ converges to a diffusion driven by standard Brownian motion (a process without memory). Roughly speaking our main results are as follows:

(1) If the Hamiltonian H is quadratic, and the noise v is any stationary Gaussian process with an algebraically decaying covariance function, then $H(X^\varepsilon)$ converges to a diffusion process driven by a Brownian motion. In this particular case, the limiting diffusion is a rescaling of the square of the 2-dimensional Bessel process.

(2) If the Hamiltonian is not quadratic (but still smooth with exactly one non-degenerate critical point), then we can show $H(X^\varepsilon)$ converges to a diffusion driven by Brownian motions, provided the noise v is chosen suitably. (The noise v is still a stationary Gaussian process with an algebraically decaying covariance function. However, it must be expressible as a superposition of Ornstein–Uhlenbeck processes as we make use of certain exponential mixing estimates in the proof.)

It is well known [Taq75, Taq77, Mar05] that stationary Gaussian noise with an algebraically decaying covariance function can be renormalized to converge to an fBm. Thus, in the above results, the renormalized noise converges to an fBm (a process with memory), whereas the processes $H(X^\varepsilon)$ themselves converge to a memory-less diffusion. Another striking point, which we will elaborate on shortly, is the fact that when the noise has long-range correlations the noise increments are *diverging* even though the process $H(X^\varepsilon)$ converges. So not only do the oscillatory dynamics of (1.4) destroy the memory effect but they also coerce $H(X^\varepsilon)$ into converging on a time scale where the driving noise diverges.

1.1. **Main results for a quadratic Hamiltonian.** We now state our results precisely when H is quadratic. Hereafter we assume that the noise v is a stationary Gaussian process with covariance function

$$R(t) \stackrel{\text{def}}{=} \boldsymbol{E} v(t)v(0) = \boldsymbol{E} v(t+s)v(s)\,. \tag{1.5}$$

Our main result when the Hamiltonian is quadratic is as follows:

Proposition 1.1. *Let H be the quadratic Hamiltonian*

$$H(x) \stackrel{\text{def}}{=} \frac{|x|^2}{2}\,, \qquad x = (x_1, x_2) \in \mathbb{R}^2\,, \tag{1.6}$$

and suppose the noise v is a stationary Gaussian process whose covariance function, R, is of the form

$$R(t) = \frac{L(t)}{t^\gamma} \tag{1.7}$$

for some $\gamma \in (0,2)$ and a function L that is slowly varying at infinity. If $\gamma \in (0,1]$, we further assume that L has the slow increase property

$$\lim_{t\to\infty} \frac{L'(t)}{L(t)/t} = 0\,. \tag{1.8}$$

Then the family of processes $H(X^\varepsilon)_{\varepsilon>0}$ converges in distribution to $H(W^D)$, where W^D is a 2D Brownian motion with $W_0 = X_0$ and covariance matrix D given by

$$D_{11} = D_{22} \stackrel{\text{def}}{=} \lim_{x\to\infty} \int_0^x R(z)\cos(z)\,dz \qquad \text{and} \qquad D_{12} = D_{21} = 0\,. \tag{1.9}$$

Remark 1.2. For $\gamma \in (0,1]$ we note that (1.8) implies that $R' < 0$ near infinity. Since $\cos(z)$ oscillates periodically, the limit in (1.9) exists and is finite. For $\gamma \in (1,2)$, we note $R \in L^1(\mathbb{R})$ and so the limit in (1.9) also exists and is finite.

As mentioned earlier, it is well known [Taq75, Taq77, Mar05] that the renormalized noise converges to a fractional Brownian motion. For the reader's convenience, we state a result to that effect next.

Proposition 1.3. *Let v be a stationary Gaussian process with covariance function R given by (1.7) for some $\gamma \in (0,2)$ and a slowly varying function L. If $\gamma \in (1,2)$, we additionally suppose*

$$\int_0^\infty R(t)\,dt = 0\,. \tag{1.10}$$

Let

$$\sigma(\varepsilon) = \begin{cases} L(\varepsilon^{-2})^{1/2}\varepsilon^\gamma & \gamma \neq 1\,, \\ L(\varepsilon^{-2})^{1/2}\varepsilon|\ln\varepsilon|^{1/2} & \gamma = 1\,, \end{cases} \quad \text{and} \quad u^\varepsilon(t) \stackrel{\text{def}}{=} \frac{1}{\sigma(\varepsilon)}\int_0^t v\left(\frac{s}{\varepsilon^2}\right)ds\,. \tag{1.11}$$

Then, as $\varepsilon \to 0$, the family of processes $(u^\varepsilon)_{\varepsilon>0}$ converges in distribution to $\sigma_{\mathcal{H}} B^{\mathcal{H}}$, where $B^{\mathcal{H}}$ is a standard fractional Brownian motion with Hurst index $\mathcal{H} = 1 - \gamma/2$, and

$$\sigma_{\mathcal{H}}^2 \stackrel{\text{def}}{=} \begin{cases} \dfrac{1}{\mathcal{H}|2\mathcal{H}-1|} & \mathcal{H} \neq \dfrac{1}{2}\,, \\ 1 & \mathcal{H} = \dfrac{1}{2}\,. \end{cases} \tag{1.12}$$

To reiterate our main point, we note that Proposition 1.1 implies

$$H\Big(X_0 + \frac{1}{\varepsilon^2}\int_0^t \nabla^\perp H(X_s^\varepsilon)\,ds + \frac{1}{\varepsilon}\int_0^t v\Big(\frac{s}{\varepsilon^2}\Big)e_2\,ds\Big) \xrightarrow{\varepsilon\to 0} H(W_t^D)\,, \tag{1.13}$$

for a Brownian motion W^D with covariance matrix D. However,

$$\frac{1}{\sigma(\varepsilon)}\int_0^t v\Big(\frac{s}{\varepsilon^2}\Big)\,ds \xrightarrow{\varepsilon\to 0} \sigma_{\mathcal{H}} B_t^H\,,$$

for a fBm $B^{\mathcal{H}}$ with Hurst index $\mathcal{H}$. Of course, $H(W^D)$ is a Markov process with no memory, but $B^{\mathcal{H}}$ is a non-Markovian process with memory.

Note further that when $\gamma \in (0,1)$, we see $\varepsilon \ll \sigma(\varepsilon)$ and so the term $\frac{1}{\varepsilon}\int_0^t v(s/\varepsilon^2)\,ds$ appearing in (1.13) diverges as $\varepsilon \to 0$. On the other hand when $\gamma \in (1,2)$ (and equation (1.10) holds), we see $\sigma(\varepsilon) \ll \varepsilon$, and so this term vanishes. In both cases, the oscillatory Hamiltonian term contributes non-trivially and as a result, the time correlations are forgotten and $H(X^\varepsilon)$ converges to a memoryless Markov process.

We now explain the reason for the assumption (1.10), which is required in Proposition 1.3 when $\gamma \in (1,2)$, but not in Proposition 1.1. If $\gamma \in (1,2)$ by the central limit theorem one can show that $\frac{1}{\varepsilon}\int_0^t v(s/\varepsilon^2)\,ds$ converges to a Brownian motion. The covariance of this Brownian motion is proportional to $\int_0^\infty R(t)\,dt$. If

this is 0 (as required by assumption (1.10)), then one can divide $\int_0^t v(s/\varepsilon^2)\,ds$ by the smaller factor $\varepsilon^\gamma = \sigma(\varepsilon)$, and look for a non-trivial limit. This is precisely what is given by Proposition 1.3. The reason the assumption (1.10) is not needed for Proposition 1.1 is because in (1.13) the noise is only scaled by a factor of $1/\varepsilon$ and not $1/\sigma(\varepsilon)$.

Finally, we briefly summarize the main idea of the proofs of Propositions 1.1 and 1.3. First, Proposition 1.3 only requires convergence of Gaussian processes and can be proved by directly computing covariances. Similar results are well known [Taq75, Taq77, Mar05]. Since the exact result we need isn't readily available, we prove it in Appendix B.

The proof of Proposition 1.1 requires a little more work. The main simplification obtained from the assumption that H is quadratic is that the evolution of the angle coordinate is independent of the action coordinate. As a result one can perform a (time dependent) rotation in space and prove convergence of $H(X^\varepsilon)$ by studying integrals of the form

$$\frac{1}{\varepsilon}\int_0^t v\Big(\frac{s}{\varepsilon^2}\Big)\cos\Big(\frac{s}{\varepsilon^2}\Big)\,ds \qquad \text{and} \qquad \frac{1}{\varepsilon}\int_0^t v\Big(\frac{s}{\varepsilon^2}\Big)\sin\Big(\frac{s}{\varepsilon^2}\Big)\,ds\,.$$

These are now Gaussian and one can study convergence by computing covariances directly. We note that the study of similar integrals arises in the study of random media where several authors [Mar05, FGPSl07, Gar97] have used oscillatory dynamics to decorrelate the medium.

1.2. **Main results for Hamiltonians with only one critical point.** We now study the case where the Hamiltonian has exactly one non-degenerate critical point but is not necessarily quadratic. In this case we will need to work with a stationary Gaussian noise process v that can be obtained by super-imposing Ornstein–Uhlenbeck processes and written in the form

$$v(t) \stackrel{\text{def}}{=} \int_S e^{-\mu|p|^{2\beta}(t-u)}\mathcal{B}(du,dp)\,.$$

Here $S \subseteq \mathbb{R}$ is a bounded symmetric open interval, μ, β are constants, and $\mathcal{B}$ is a Gaussian random measure that is white in time and colored in space and will be constructed explicitly in Section 3 below to ensure that the covariance of v decays algebraically. Our main theorem can now be stated as follows.

Theorem 1.4. *Let $H\colon \mathbb{R}^2 \to \mathbb{R}$ be a smooth Hamiltonian with exactly one non-degenerate minimum at $(0,0)$, and v be the stationary Gaussian noise described above. Then*

$$H(X^\varepsilon) \xrightarrow[\varepsilon\to 0]{\mathcal{D}} \mathcal{X}\,, \tag{1.14}$$

where $\mathcal{X}$ is a diffusion driven by standard Brownian motion.

Theorem 1.4 follows immediately from the more general Theorem 4.3, below, concerning fast-slow systems, and the proof of Theorem 1.4 is presented shortly after the statement of Theorem 4.3. Moreover, Theorem 4.3 also explicitly characterizes the limiting system and can be used to explicitly characterize the diffusion $\mathcal{X}$ in Theorem 1.4.

Note that when H is not quadratic the convergence of $H(X^\varepsilon)$ can't be reduced to Gaussian integrals as in Proposition 1.1, and so the proof of Theorem 1.4 is more

involved. We prove Theorem 1.4 by first switching to action-angle coordinates and converting (1.4) to a fast-slow system, where the slow variables are mean-zero in the fast variable. We then use the perturbed test function (PTF) method [Kus84] to prove convergence of this system. This approach provides more information than just the convergence of $H(X^\varepsilon)$, as stated in Theorem 1.4, and we refer the reader to Theorem 4.3, below, for the precise statement. Of course, the convergence of action coordinate is equivalent to convergence of the processes $H(X^\varepsilon)$. However, our proof also identifies an asymptotic diffusive behavior of the angle coordinate on the slow scale.

We remark that our fast-slow system is different from that recently considered by Hairer and Li [HL20]. Indeed, in [HL20], the authors proved an averaging principle for a coupled fast-slow system driven by fBm, with Hurst index $\mathcal{H} > \frac{1}{2}$, and they obtained convergence in probability to the naively averaged system. However, they require the fast variable to be driven by an independent *standard* Brownian motion. In the scenario that arises in the present paper, the fast variable is driven by noise that converges to a *fractional* Brownian motion when renormalized, and the same noise is also used to drive the slow variable. Moreover the limiting system we obtain is not the naively averaged one, but a diffusion with an averaged generator akin to the limiting behavior of fast-slow systems driven by standard Brownian noise (see for instance [PS08]).

Presently we are only able to handle the case when H has exactly one non-degenerate critical point. In the general case, the Reeb graph of H will have multiple vertices, and the limiting process will have gluing conditions at the vertices. The Freidlin–Wentzell theory obtains these gluing conditions by considering regions corresponding to small neighborhoods of these vertices, and restarting the process every time it exits these neighborhoods. This relies heavily on the Markov property, which is not available to us. The perturbed test function method was introduced to study convergence of non-Markov processes, and so it may still be used in our scenario. There are, however, several technical obstructions and we are presently unable to prove Theorem 1.4 when H has more than one critical point.

1.3. **Plan of this paper.** In Section 2 we prove Proposition 1.1, modulo two computational lemmas. The proofs of these lemmas are relegated to Appendix A. The proof of Proposition 1.3 is presented in appendix B. In Section 3 we construct the noise process v that will be used in Theorem 1.4. We also establish basic properties of the noise in this section and put two computationally involved proofs in appendices C and D. In Section 4 we state a precise, more general, version of Theorem 1.4 (Theorems 4.3 and 4.4), and prove Theorem 1.4. We prove regularity of the coefficients of the limiting equation in Theorem 4.3 (equation (4.13)) in Appendix E, and study the limiting equation itself in Section 5. Finally, we prove Theorem 4.3 in Sections 6, 7 and 8.

Acknowledgments. The authors thank Lenya Ryzhik for suggesting this problem to us, and for many helpful discussions.

2. Proof of convergence for a quadratic Hamiltonian.

This section is devoted to proving Proposition 1.1. The main idea behind the proof is that when H is quadratic, the deterministic dynamics rotates with constant angular speed in all trajectories. Performing a spatial rotation will now reduce the

problem to studying convergence of Gaussian processes, which can be resolved by computing covariances.

Proof of Proposition 1.1. When H is given by (1.6), we use Duhamel's formula to write the solution of (1.4) as

$$X_t^\varepsilon = \begin{pmatrix} \cos(\frac{t}{\varepsilon^2}) & \sin(\frac{t}{\varepsilon^2}) \\ -\sin(\frac{t}{\varepsilon^2}) & \cos(\frac{t}{\varepsilon^2}) \end{pmatrix} X_0 + \frac{1}{\varepsilon}\int_0^t v\Big(\frac{\tau}{\varepsilon^2}\Big) \begin{pmatrix} \sin(\frac{t-\tau)}{\varepsilon^2}) \\ \cos(\frac{t-\tau}{\varepsilon^2}) \end{pmatrix} d\tau.$$

Let $M(t)$ be the rotation matrix

$$M(t) \stackrel{\text{def}}{=} \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix},$$

and define the rotated process Y^ε by $Y_t^\varepsilon = M(-t/\varepsilon^2)X_t^\varepsilon$. Clearly $H(X^\varepsilon) = H(Y^\varepsilon)$, and so convergence of the processes $H(X^\varepsilon)$ reduces to convergence of the processes $H(Y^\varepsilon)$.

We claim that the processes Y^ε itself converge to a Brownian motion as $\varepsilon \to 0$. To see this, observe

$$Y_t^\varepsilon = X_0 + \frac{1}{\varepsilon}\int_0^t v\Big(\frac{\tau}{\varepsilon^2}\Big) \begin{pmatrix} -\sin(\frac{\tau}{\varepsilon^2}) \\ \cos(\frac{\tau}{\varepsilon^2}) \end{pmatrix} d\tau\,.$$

We will now show that the second term above converges to a Brownian motion. For this, define

$$w_1^\varepsilon(t) \stackrel{\text{def}}{=} \frac{1}{\varepsilon}\int_0^t v\left(\frac{\tau}{\varepsilon^2}\right)\sin\left(\frac{\tau}{\varepsilon^2}\right) d\tau\,, \tag{2.1}$$

$$w_2^\varepsilon(t) \stackrel{\text{def}}{=} \frac{1}{\varepsilon}\int_0^t v\left(\frac{\tau}{\varepsilon^2}\right)\cos\left(\frac{\tau}{\varepsilon^2}\right) d\tau\,. \tag{2.2}$$

Note, the noise v (when normalized by $\sigma(\varepsilon)$) converges to fBm. However, the expressions above use a normalization factor of ε instead of $\sigma(\varepsilon)$, and have an oscillatory factor. For this reason we claim $(w_1^\varepsilon, w_2^\varepsilon)$ converges to a Brownian motion. To prove convergence of w_1^ε, w_2^ε, we first state two lemmas:

Lemma 2.1. *For any $T > 0$, there exists a constant $C > 0$ such that for every $s, t \in [0, T]$, $i \in \{1, 2\}$, we have*

$$\boldsymbol{E}(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2 \leqslant \begin{cases} C|t-s|^{1-\gamma} & \gamma \in (0,1)\,, \\ C|t-s|^{2-\gamma} & \gamma \in [1,2)\,. \end{cases} \tag{2.3}$$

Lemma 2.2. *For every $s, t \geqslant 0$, and $i \in \{1, 2\}$ we have*

$$\lim_{\varepsilon\to 0} \boldsymbol{E}(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2 = D_{11}|t-s|\,, \tag{2.4}$$

$$\lim_{\varepsilon\to 0} \boldsymbol{E}(w_1^\varepsilon(t) - w_1^\varepsilon(s))(w_2^\varepsilon(t) - w_2^\varepsilon(s)) = 0\,. \tag{2.5}$$

The proofs of both Lemmas 2.1 and 2.2 are lengthy, but direct computations. Thus, for clarity of presentation, we postpone the proofs to Appendix A. Once Lemmas 2.1 and 2.2 are established, the proof of Proposition 1.1 follows quickly. Indeed, Lemma 2.1 implies that the family $(Y^\varepsilon)_{\varepsilon>0}$ is tight on $\mathcal{C}([0,T],\mathbb{R}^2)$

To see this, it suffices to show that the processes w_i^ε are tight. Without loss of generality suppose $T = 1$, and suppose $\gamma \in (0, 1)$ (the case when $\gamma \in [1, 2)$ is

similar). Choose an integer M such that $M_\gamma = M(1-\gamma) > 1$. Since w_i^ε is Gaussian process, we know

$$\boldsymbol{E}(w_i^\varepsilon(t) - w_i^\varepsilon(s))^{2M} = C_M[\boldsymbol{E}(w_i^\varepsilon(t) - w_i^\varepsilon(s))^2]^M$$

for some constant C_M. We will allow C_M to change from line to line as long as it only depends on M, and remains independent of ε. The above implies

$$\boldsymbol{E}(w_i^\varepsilon(t) - w_i^\varepsilon(s))^{2M} \leqslant C_M|t-s|^{M_\gamma}\,. \tag{2.6}$$

Since $M_\gamma > 1$ by choice, Kolmogorov's criterion (see for instance [EK86] Proposition 10.3) implies that the family w_i^ε is tight. This implies that the processes Y^ε converge in distribution along a subsequence to a continuous process Y.

Since each process Y^ε is a Gaussian process, the limiting process Y must also be a Gaussian process. Now Lemma 2.2 implies Y is a Brownian motion in $\mathbb{R}^2$ with $Y_0 = X_0$ and covariance matrix D. Since the law of the limiting process is uniquely determined, the family Y^ε itself must converge in distribution, without having to select a subsequence. Finally, since $H(X_t^\varepsilon) = H(Y_t^\varepsilon)$, we obtain convergence of $(H(X^\varepsilon))_{\varepsilon>0}$ as claimed. □

3. **Construction of the Noise.**

In this section, we will construct noise v that will be used in Theorem 1.4, and establish a few properties that will be used in the course of the proof of Theorem 1.4. We require the covariance function R to be of the form

$$R(t) \stackrel{\text{def}}{=} \int_S r(p)e^{-\mu|p|^{2\beta}|t|}\,dp\,, \tag{3.1}$$

where $S = (-r_s, r_s)$ is a symmetric, bounded, open interval, $\mu, \beta > 0$ are constants, and $r\colon S - \{0\} \to [0,\infty)$ is defined by

$$r(p) \stackrel{\text{def}}{=} \frac{\lambda(p)}{|p|^{2\alpha}}\,.$$

Here $\lambda\colon S \to \mathbb{R}$ is a smooth bounded even function such that $\lambda(0) \neq 0$ and

$$\int_S r(p)\,dp > 0\,,$$

and

$$\alpha < \frac{1}{2}\,. \tag{3.2}$$

We will now construct a stationary Gaussian process v with covariance function R. The chosen form of R allows us to construct v by superimposing Ornstein–Uhlenbeck processes as follows. Let ξ be 2D white noise and define a Gaussian random measure $\mathcal{B}$ by

$$\mathcal{B}(du,dp) = \sqrt{2\mu r(p)}\,|p|^\beta \mathbf{1}_S(p)\,\xi(du,dp)\,,$$

Clearly, the covariance of $\mathcal{B}$ is given by

$$\boldsymbol{E}\mathcal{B}(du,dp)\,\mathcal{B}(du',dp') = 2\mu\, r(p)\,|p|^{2\beta}\delta(u-u')\delta(p-p')\,du\,du'\,dp\,dp'\,.$$

Now define the measure-valued Gaussian random process V by

$$V(t,dp) = \int_{-\infty}^t e^{-\mu|p|^{2\beta}(t-u)}\,\mathcal{B}(du,dp)\,,$$

and finally define the noise v by

$$v(t) = \int_S V(t, dp)\,.$$

We now establish a few properties of the process v that will be used in the proof of Theorem 1.4.

Lemma 3.1. *The process v defined above is a stationary Gaussian process with covariance function R.*

Proof. Clearly v is a Gaussian process. To see that v is stationary with covariance function R we choose any $s \leqslant t$ and compute

$$\begin{aligned}\boldsymbol{E}v(s)v(t) &= \int_{-\infty}^{s}\int_S \exp\Big(-\mu\big(|p|^{2\beta}(s-u) + |p|^{2\beta}(t-u)\big)\Big)\cdot 2\mu|p|^{2\beta} r(p)\,du\,dp \\ &= \int_S r(p) e^{-\mu|p|^{2\beta}(t-s)}\,dp = R(t-s)\,.\end{aligned}$$

□

Lemma 3.2. *If R is given by (3.1), then*

$$\lim_{t\to\infty} t^\gamma R(t) = c_0$$

where c_0 and γ are defined by

$$\gamma \stackrel{\text{def}}{=} \frac{1-2\alpha}{2\beta} \qquad \text{and} \qquad c_0 \stackrel{\text{def}}{=} \lambda(0)\int_{p\in\mathbb{R}} \frac{e^{-\mu|p|^{2\beta}}}{|p|^{2\alpha}}\,dp\,.$$

Proof. By a change of variable, note that

$$t^\gamma R(t) = t^\gamma \int_S \frac{\lambda(p)}{|p|^{2\alpha}} e^{-\mu|p|^{2\beta}t}\,dp = \int_{(t^{1/(2\beta)}S)} \frac{1}{|p'|^{2\alpha}}\lambda\Big(\frac{p'}{t^{1/(2\beta)}}\Big) e^{-\mu|p'|^{2\beta}}\,dp'\,,$$

which converges to c_0 as $t\to\infty$. □

Remark 3.3. Note that (3.2) already implies $\gamma > 0$. If additionally we assume

$$2\alpha + 4\beta > 1 \tag{3.3}$$

then we will also have $\gamma < 2$. In this case the process v can be renormalized as in (1.11) to converge to a fBm. For Theorem 1.4, however, the assumption (3.3) is unnecessary.

Let $\{\mathcal{G}_t\}$ be the augmented filtration generated by V. That is,

$$\mathcal{G}_t \stackrel{\text{def}}{=} \sigma\big(\mathcal{N} \cup \{\sigma(V(s,\cdot)) \mid 0 \leqslant s \leqslant t\}\big)\,, \tag{3.4}$$

where $\mathcal{N}$ is the set of all null sets known at time infinity. Our next lemma computes conditional expectations of $V(\cdot, dp)$ with respect to the filtration $\mathcal{G}$.

Lemma 3.4. *We have for any $t, h \geqslant 0$*

$$\boldsymbol{E}\big[V(t+h, dp)|\mathcal{G}_t\big] = e^{-\mu|p|^{2\beta}h}\,V(t, dp) \tag{3.5}$$

and

$$\begin{aligned}\boldsymbol{E}\Big[V(t+h,dp)V(t+h,dq)\Big|\mathcal{G}_t\Big] - \boldsymbol{E}\big[V(t+h,dp)|\mathcal{G}_t\big]\boldsymbol{E}\big[V(t+h,dq)|\mathcal{G}_t\big] \\ = (1-e^{-2\mu|p|^{2\beta}h})r(p)\delta(p-q)\,dp\,dq\,.\end{aligned} \tag{3.6}$$

Our last result concerns the boundedness of our noise process.

Lemma 3.5. *Let $T > 0$, $M > 0$,*

$$D_{k,M} \stackrel{\text{def}}{=} [0,T] \times L^\infty([0,T], W_{k,M})$$

with

$$W_{k,M} \stackrel{\text{def}}{=} \{\varphi \in W^{1,k}(S) : \quad \|\varphi\|_{W^{1,k}} \leqslant M\}$$

and where $W^{1,k}(S)$ stands for the Sobolev space with $k \in (1,\infty]$. We have

$$\boldsymbol{E}\Big[\sup_{(t,\varphi)\in D_{k,M}} \Big|V\Big(\frac{t}{\varepsilon^2}, \varphi(t,\cdot)\Big)\Big|\Big] \leqslant C + \frac{C(\varepsilon)}{\varepsilon}\,, \tag{3.7}$$

and for any $n \in \mathbb{N}^*$

$$\sup_\varepsilon \sup_{t\in[0,T]} \boldsymbol{E}\Big[\sup_{\varphi\in W_{k,M}} \Big|V\Big(\frac{t}{\varepsilon^2}, \varphi\Big)\Big|^n\Big] \leqslant C_n\,, \tag{3.8}$$

where C, C_n and $C(\varepsilon)$ are three positive constants where the latter satisfies

$$\lim_{\varepsilon\to 0} C(\varepsilon) = 0.$$

Here the notation $V(t/\varepsilon^2, \varphi)$ denotes integral of a function $\varphi(p)$ with respect to the measure $V(t/\varepsilon^2, dp)$. That is,

$$V\Big(\frac{t}{\varepsilon^2}, \varphi\Big) \stackrel{\text{def}}{=} \int_S \varphi(p)\, V\Big(\frac{t}{\varepsilon^2}, dp\Big)\,,$$

and we will drop the S in the above notation for simplicity unless specified otherwise. The proofs of these two Lemmas are computationally involved, and we relegate them to Appendices C and D respectively.

4. The Main Theorem

One canonical way to analyze integrable Hamiltonian systems is to use a set of *action-angle coordinates*. These coordinates separate the slow and fast motion, and preserves the Hamiltonian structure.

4.1. **Action-angle coordinates.** The Liouville-Arnold theorem [Arn89] asserts that there exists a symplectic canonical transformation $\varphi\colon X = (x,y) \mapsto (I,\theta) \in \mathbb{R}\times\mathbb{T}$, where the action variable I and the angle variable θ satisfy

$$K(I) = H(x,y) \qquad \text{and} \qquad \{I,\theta\} = 1\,, \tag{4.1}$$

where K is a given one variable smooth enough increasing function such that $K(0) = 0$, $\{\cdot,\cdot\}$ stands for the standard Poisson bracket and is defined by

$$\{g,h\} = \partial_x g \partial_y h - \partial_y g \partial_x h\,,$$

for our Hamiltonian system with one degree of freedom. The relation on the right-hand side of (4.1) can be also stated as

$$\nabla I \cdot \nabla^\perp \theta = 1\,.$$

Note that in the action-angle coordinate the Hamiltonian is a function of the action coordinate alone.

The existence of such transformation φ is guaranteed by the Liouville-Arnold theorem, and can be constructed through a generating function [Arn89, Section 50 pp. 281], and implicitly defined by the relations

$$y = \partial_x \mathcal{S}(I,x)\,, \qquad \theta = \partial_I \mathcal{S}(I,x)\,, \qquad \text{and} \qquad H(x, \partial_x \mathcal{S}(I,x)) = K(I)\,. \tag{4.2}$$

Such a construction is not unique. For convenience in the forthcoming analysis we will choose $\mathcal{S}(I,0)=0$ so that

$$\mathcal{S}(I,x)=\int_0^x \partial_{x'}\mathcal{S}(I,x')\,dx'\,,$$
$$\partial_I\mathcal{S}(I,x)=\int_0^x \partial_{Ix'}\mathcal{S}(I,x')\,dx'\,,$$

and hence

$$\partial_I\mathcal{S}(I,0)=0\,.$$

The meaning of the second relation is that the zero angle corresponds to either the positive part of the y-axis or the negative part. Consequently, we have for any $I>0$

$$\varphi_1^{-1}(I,\theta=0)=0 \qquad \text{and} \qquad \varphi_2^{-1}(I,\theta=0)\neq 0\,, \tag{4.3}$$

and in particular,

$$\partial_I\varphi_1^{-1}(I,\theta=0)=0\,. \tag{4.4}$$

Let us remark that (4.3) follows from the fact that K cancels only for $I=0$, and then for any $I>0$, the component x and y cannot cancel at the same time.

Now let us write $\varphi=(\varphi_1,\varphi_2)$ and define

$$I_t \overset{\text{def}}{=} I(X_t)=\varphi_1(X_t)\,, \qquad \text{and} \qquad \theta_t \overset{\text{def}}{=} \theta(X_t)=\varphi_2(X_t)\,.$$

In the absence of random fluctuations (i.e. when $v=0$), the Hamiltonian system $\dot X_t=\nabla^\perp H(X_t)$ becomes

$$\dot I_t=0\,, \qquad \text{and} \qquad \dot\theta_t=\omega(I_t)\,, \qquad \text{with} \qquad \omega(I)=K'(I)\,.$$

In the presence of random fluctuations, the Hamiltonian system (1.3) becomes

$$\dot I_t=\varepsilon v(t)a(I_t,\theta_t)\,, \qquad \text{and} \qquad \dot\theta_t=\omega(I_t)+\varepsilon v(t)b(I_t,\theta_t)\,, \tag{4.5}$$

where

$$a=e_2\cdot\nabla\varphi_1\circ\varphi^{-1}\,, \quad \text{and} \quad b=e_2\cdot\nabla\varphi_2\circ\varphi^{-1}\,.$$

Example 4.1. For illustration purposes, we now explicitly compute the action-angle variables when the Hamiltonian is quadratic. From (4.2), we have

$$x=\sqrt{\frac{I}{\pi}}\cos\Big(2\pi\theta-\frac{\pi}{2}\Big) \qquad \text{and} \qquad y=\sqrt{\frac{I}{\pi}}\sin\Big(2\pi\theta-\frac{\pi}{2}\Big)$$

where θ gives the angle of $X=(x,y)^T$ from the negative part of the vertical axis. One can then see that the action variable is a multiple of the Hamiltonian,

$$K(I)=\frac{I}{2\pi} \qquad \text{and} \qquad \omega(I)=\frac{1}{2\pi},$$

while the angle variable corresponds to the angle on a trajectory, which is a circle in the case of $H(X)=|X|^2/2$, with period 1. In this case, the functions a and b are as follows

$$a(I,\theta)=2\sqrt{\pi I}\sin\Big(2\pi\theta-\frac{\pi}{2}\Big) \quad \text{and} \quad b(I,\theta)=\frac{1}{2\sqrt{\pi I}}\cos\Big(2\pi\theta-\frac{\pi}{2}\Big)\,. \tag{4.6}$$

4.2. **The main theorem.** To study the long-time behavior of (4.5), we let

$$I_t^\varepsilon = I_{t/\varepsilon^2} \qquad \text{and} \qquad \theta_t^\varepsilon = \theta_{t/\varepsilon^2}\,.$$

Now (4.5) becomes

$$\dot I_t^\varepsilon = \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big) a(I_t^\varepsilon, \theta_t^\varepsilon)\,, \qquad \dot\theta_t^\varepsilon = \frac{\omega(I_t^\varepsilon)}{\varepsilon^2} + \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big) b(I_t^\varepsilon, \theta_t^\varepsilon)\,, \tag{4.7}$$

with $I_0^\varepsilon = I_0$ and $\theta_0^\varepsilon = \theta_0$. Using the approach in [Gar97] to separate fast and slow motions, we study this system by splitting the angle variable θ_t^ε into two parts

$$\theta_t^\varepsilon = \psi_t^\varepsilon + \tau_t^\varepsilon\,.$$

The evolution of τ_t^ε (fast motion) is obtained by averaging (4.7) over the angular coordinate, and ψ_t^ε (slow motion) is the remainder. That is we require

$$\dot\tau_t^\varepsilon = \frac{\omega(I_t^\varepsilon)}{\varepsilon^2} + \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big) \langle b(I_t^\varepsilon, \cdot)\rangle \tag{4.8}$$

with initial condition $\tau_0^\varepsilon = 0$, and where we have the notation

$$\langle g\rangle \overset{\text{def}}{=} \int_{\theta=0}^1 g(\theta)\, d\theta\,.$$

Considering only the slow motion variables $(I_t^\varepsilon, \psi_t^\varepsilon)$, the system (4.7) becomes

$$\begin{aligned} \dot I_t^\varepsilon &= \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big) A(I_t^\varepsilon, \psi_t^\varepsilon, \tau_t^\varepsilon) \\ \dot\psi_t^\varepsilon &= \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big) B(I_t^\varepsilon, \psi_t^\varepsilon, \tau_t^\varepsilon) \end{aligned} \tag{4.9}$$

with

$$A(I,\psi,\tau) \overset{\text{def}}{=} a(I, \psi+\tau) \qquad \text{and} \qquad B(I,\psi,\tau) \overset{\text{def}}{=} b(I,\psi+\tau) - \langle b(I,\cdot)\rangle\,.$$

The above equations are coupled with the initial conditions $I_0^\varepsilon = I_0$, and $\psi_0^\varepsilon = \theta_0$. Note that A and B are both 1-periodic and mean-zero with respect to τ. This latter property is mandatory to deal with the long-range correlation case as illustrated in Proposition 1.1. In fact, if the system possesses a component with zero frequency, a noise with long-range correlations will charge this component and cause the system to blow up as $\varepsilon \to 0$.

It is straightforward to see this mean-zero property in τ for B, but for A we use that

$$\operatorname{Jac} \varphi^{-1}(I,\theta) = [\operatorname{Jac} \varphi(\varphi^{-1}(I,\theta))]^{-1}$$

and that $\det \operatorname{Jac} \varphi(X) = 1$ since φ is a symplectic transformation, to obtain the following relations

$$\begin{aligned} \partial_I \varphi_1^{-1}(I,\theta) &= \partial_y \varphi_2(X), \quad \partial_I \varphi_2^{-1}(I,\theta) = -\partial_x \varphi_2(X)\,, \\ \partial_\theta \varphi_1^{-1}(I,\theta) &= -\partial_y \varphi_1(X), \quad \partial_\theta \varphi_2^{-1}(I,\theta) = \partial_x \varphi_1(X)\,. \end{aligned} \tag{4.10}$$

Therefore, we have

$$a(I,\theta) = -\partial_\theta \varphi_1^{-1}(I,\theta)$$

which is clearly mean-zero with respect to θ since we have the derivative of a periodic function. Note also that from these relations the mean-zero property in θ for b is not clear. This is the reason why we introduce the compensation $\langle b\rangle$ in the definition of B.

Our main result obtains the limiting behavior of (4.9) as $\varepsilon \to 0$ under the following assumptions.

- The function K is smooth, and

$$\inf_{I\geqslant 0} K'(I) = \inf_{I\geqslant 0} \omega(I) > \omega_0 > 0 \tag{4.11}$$

  for some strictly positive number ω_0.
- There exist $r > 0$, and positive constants $c_{1,r}$, $c_{2,r} > 0$ such that for any $I \in (0, r)$

$$c_{1,r} I \leqslant K(I) \leqslant c_{2,r} I \qquad \text{and} \qquad |\omega'(I)| \leqslant \frac{c_{2,r}}{I}\,. \tag{4.12}$$

Note that these conditions imply that $K(0) = 0$ and that K is an increasing function in I. These assumptions are not too restrictive since any Hamiltonian satisfies (4.11) and (4.12) near non-degenerate critical points. The following proposition is a consequence of the above assumptions, which concerns bounds on the function a. The proof is given in Appendix E.

Proposition 4.2. *Under (4.11) and (4.12) there exist $r > 0$ and a constant $C_r > 0$ such that for any $I \in (0, r)$ we have*

$$\sup_{\theta\in\mathbb{T}} \Big(|a(I,\theta)| + |\partial_\theta a(I,\theta)| \Big) \leqslant C_r \sqrt{I}\,,$$

and

$$\sup_{\theta\in\mathbb{T}} |\partial_I a(I,\theta)| \leqslant \frac{C_r}{\sqrt{I}}\,.$$

The proposition mainly describes the behavior of $a, \partial_I a$ and $\partial_\theta a$ around $I = 0$. These bounds will allow us to prove the main result of the paper that characterizes the limiting process of the sequence of processes $(I^\varepsilon, \psi^\varepsilon)_{\varepsilon>0}$.

Theorem 4.3. *Assume (4.11) and (4.12) hold, the family $(I^\varepsilon, \psi^\varepsilon)_{\varepsilon>0}$ (defined in (4.8)–(4.9)) converges in distribution in $\mathcal{C}([0,\infty), \mathbb{R}^2)$ to a process $(I_t, \psi_t)_{t\geqslant 0}$, where $(I_t)_{t\geqslant 0}$ is the unique weak solution of the SDE*

$$\begin{aligned} dI_t = &\int_{\tau=0}^{1} a(I_t,\tau)\, dW_t(I_t,\tau)\, d\tau \\ &+ \Big[\int_{u=0}^{\infty} R(u) \int_{\tau=0}^{1} \partial_I \big(a(I, \tau+\omega(I)u) \big)_{|I=I_t} a(I_t,\tau) \\ &\qquad + a(I_t, \tau+\omega(I_t)u) \partial_I a(I_t,\tau)\, d\tau\, du \Big]\, dt\,, \end{aligned} \tag{4.13}$$

with initial condition $I_{t=0} = I_0$. Here, W is a real valued Brownian field with covariance function

$$\begin{aligned} \boldsymbol{E}[W_t(y,\phi_1) W_s(y,\phi_2)] = t\wedge s \int_{u=0}^{\infty} R(u) \\ \times \int_{\tau=0}^{1} \phi_1(\tau+\omega(y)u)\phi_2(\tau) + \phi_1(\tau)\phi_2(\tau+\omega(y)u)\, d\tau\, du\,, \end{aligned} \tag{4.14}$$

for any $\phi_1, \phi_2 \in L^2_0(\mathbb{T})$ where

$$L^2_0(\mathbb{T}) \stackrel{\text{def}}{=} \big\{ \phi \in L^2(\mathbb{T}) : \langle \phi \rangle = 0 \big\}\,.$$

Explicitly, we can write

$$W_t(y,\tau) = \frac{1}{\sqrt{2}} \sum_{n\in\mathbb{Z}^*} e^{2i\pi n\tau} R_n^{1/2}(y) (B^1_{t,n} - iB^2_{t,n})$$

where $B^1_{t,n}$ and $B^2_{t,n}$ are two independent real valued cylindrical Brownian motions satisfying $B^1_{t,-n} = B^1_{t,n}$, $B^2_{t,-n} = -B^2_{t,n}$, and $R_n = 2\int_0^\infty R(u)\cos(2\pi n\omega(I)u)\,du$.

Also, we have

$$
\begin{aligned}
d\psi_t = &\int_{\tau=0}^{1} b(I_t,\tau)\,dW_t(I_t,\tau)\,d\tau \\
&+\Big[\int_{u=0}^{\infty} R(u)\int_{\tau=0}^{1} \partial_I\big(b(I,\tau+\omega(I)u)\big)_{|I=I_t}\,a(I_t,\tau) \\
&\qquad + b(I_t,\tau+\omega(I_t)u)\partial_I a(I_t,\tau)\,d\tau\,du\Big]dt\,,
\end{aligned}
\tag{4.15}
$$

with $\psi_{t=0} = \theta_0$.

Note that the action variable I does not depend on the slow angular motion ψ. However, the distribution of the slow angular motion ψ is completely determined by the motion of the action variable I and the Brownian field.

We also note that the functions $\partial_I a$ and $\partial_I b$ appearing in (4.13) and (4.15) may be singular at $I = 0$. Nevertheless, since the minimum of H is non-degenerate, the function $a\partial_I a$ has no singularity at $I = 0$ and there is no singularity on the right-hand side of equation (4.13). The term $a\partial_I b$ appearing in (4.15) may still have an $O(I^{-1/2})$ singularity at $I = 0$. However, since the point $I = 0$ is inaccessible (Proposition 5.2), the right-hand side of (4.15) is well defined.

We now use Theorem 4.3 to prove Theorem 1.4.

Proof of Theorem 1.4. Using action angle coordinates, we convert the system (1.4) to (4.9). Since the Hamiltonian H has exactly one non-degenerate critical point, the assumptions (4.11)–(4.12) are satisfied. Now by Theorem 4.3 we see that the family of processes $(I^\varepsilon, \psi^\varepsilon)_{\varepsilon>0}$ converges in distribution to the pair (I,ψ) which solves (4.13), (4.15). Since $H(X^\varepsilon) = K(I^\varepsilon)$, we now obtain Theorem 1.4 by applying the Itô formula to $\mathcal{X} = K(I)$. □

For the quadratic Hamiltonian we can explicitly derive the stochastic differential equations for $(I_t)_{t\geqslant 0}$ and $(\psi_t)_{t\geqslant 0}$. Using equation (4.6) we obtain

$$
dI_t = 2\sqrt{mI_t}\,dB^1_t + 2m\,dt \qquad \text{and} \qquad d\psi_t = \frac{m}{2\pi\sqrt{I_t}}\,dB^2_t \tag{4.16}
$$

with

$$
m = \pi\int_0^\infty R(u)\cos(u)\,du\,,
$$

and where B^1 and B^2 are two independent standard Brownian motions. One can easily remark that $(I_t/m)_{t\geqslant 0}$ is a 2-dimensional squared Bessel process, so that $(I_t/2\pi)_{t\geqslant 0}$ has the same law as $(|W_t|^2/2)_{t\geqslant 0}$, for W being a 2-dimensional Brownian motion with covariance matrix given by (1.9). Then, we recover the result of Proposition 1.1.

The proof of Theorem 4.3 will be given in Section 6.

4.3. **Generalization for Hamiltonian systems with one degree of freedom and 2-dimensional noises.** Theorem 4.3 can be readily extended to general Hamiltonian systems with one degree of freedom and 2-dimensional noises. In fact, our proof does not depend on the particular shapes of Hamiltonians for a one-degree of freedom oscillator problems. Also, considering 2-dimensional noises brings no additional difficulties except more tedious notations in the proof of Theorem 4.3.

Here we state this extension of Theorem 4.3 for the sake of completeness, but the detailed proof will be omitted.

We consider the following one-degree of freedom Hamiltonian system

$$X_t^\varepsilon = \frac{1}{\varepsilon}\nabla^\perp H(X_t^\varepsilon) + \frac{1}{\varepsilon}\mathbf{v}\Big(\frac{t}{\varepsilon^2}\Big)\,, \qquad X_0^\varepsilon = X_0\,,$$

where the Hamiltonian H is as in Theorem 4.3, and $\mathbf{v}$ is a 2-dimensional noise with covariance function

$$\mathbf{R}_{jl}(t-s) \overset{\text{def}}{=} \boldsymbol{E}[\mathbf{v}_j(t)\mathbf{v}_l(s)] = \int_S \frac{\lambda_{jl}(p)}{|p|^{2\alpha}} e^{-\mu|p|^{2\beta}|t-s|}\,dp \qquad j,l\in\{1,2\}\,,$$

where $(\lambda_{jl})_{j,l\in\{1,2\}}$ is a 2×2 symmetric matrix valued function with assumptions similar to the ones considered in Section 3.

Introducing the same action-angle coordinates as in Section 4.1, they now satisfy

$$\dot I_t^\varepsilon = \frac{1}{\varepsilon}\mathbf{v}\Big(\frac{t}{\varepsilon^2}\Big)\cdot\mathbf{A}(I_t^\varepsilon,\psi_t^\varepsilon,\tau_t^\varepsilon) \qquad \text{and} \qquad \dot\psi_t^\varepsilon = \frac{1}{\varepsilon}\mathbf{v}\Big(\frac{t}{\varepsilon^2}\Big)\cdot\mathbf{B}(I_t^\varepsilon,\psi_t^\varepsilon,\tau_t^\varepsilon),$$

with

$$\dot\tau_t^\varepsilon = \frac{\omega(I_t^\varepsilon)}{\varepsilon^2} + \frac{1}{\varepsilon}\mathbf{v}\Big(\frac{t}{\varepsilon^2}\Big)\cdot\langle b(I_t^\varepsilon,\cdot)\rangle\,,$$

and where

$$\mathbf{A}(I,\psi,\tau) \overset{\text{def}}{=} \begin{pmatrix} a_1(I,\psi+\tau)\\ a_2(I,\psi+\tau)\end{pmatrix} \qquad \text{and} \qquad \mathbf{B}(I,\psi,\tau) \overset{\text{def}}{=} \begin{pmatrix} b_1(I,\psi+\tau)-\langle b_1(I,\cdot)\rangle\\ b_2(I,\psi+\tau)-\langle b_2(I,\cdot)\rangle\end{pmatrix},$$

with

$$a_j(I,\theta) = e_j\cdot\nabla\varphi_1(\varphi^{-1}(I,\theta)) \qquad \text{and} \qquad b_j(I,\theta) = e_j\cdot\nabla\varphi_2(\varphi^{-1}(I,\theta))\,,$$

for $j\in\{1,2\}$.

The extension of Theorem 4.3 is as follows.

Theorem 4.4. *The family $(I^\varepsilon,\psi^\varepsilon)_{\varepsilon>0}$ converges in distribution in $\mathcal{C}([0,\infty),\mathbb{R}^2)$ to a process $(I_t,\psi_t)_{t\geqslant 0}$, where $(I_t)_{t\geqslant 0}$ is the unique weak solution of the SDE*

$$\begin{aligned} dI_t = &\int_{\tau=0}^1 a(I_t,\tau)\cdot dW_t(I_t,\tau)\,d\tau\\ &+\int_{u=0}^\infty\int_{\tau=0}^1 \partial_I\big(a^T(I,\tau+\omega(I)u)\big)_{|I=I_t}\mathbf{R}(u)a(I_t,\tau)\\ &\qquad + a^T(I_t,\tau+\omega(I_t)u)\mathbf{R}(u)\partial_I a(I_t,\tau)\,d\tau\,du\,dt, \end{aligned} \tag{4.17}$$

with initial condition $I_{t=0}=I_0$. Here, W is a 2-dimensional real valued Brownian field with covariance function

$$\begin{aligned}\boldsymbol{E}[W_t^j(y,\phi_1)W_s^l(y,\phi_2)] = (t\wedge s)&\int_{u=0}^\infty R_{jl}(u)\\ &\times\int_{\tau=0}^1 \phi_1(\tau+\omega(y)u)\phi_2(\tau)+\phi_1(\tau)\phi_2(\tau+\omega(y)u)\,d\tau\,du\,,\end{aligned}$$

for any $j,l\in\{1,2\}$ and $\phi_1,\phi_2\in L_0^2(\mathbb{T})$ where

$$L_0^2(\mathbb{T}) \overset{\text{def}}{=} \big\{\phi\in L^2(\mathbb{T}) : \langle\phi\rangle = 0\big\}\,.$$

Also, we have

$$
\begin{aligned}
d\psi_t = &\int_{\tau=0}^{1} b(I_t,\tau)\cdot dW_t(I_t,\tau)\,d\tau \\
&+\int_{u=0}^{\infty}\int_{\tau=0}^{1} \partial_I\big(b^T(I,\tau+\omega(I)u)\big)_{|I=I_t}\mathbf{R}(u)a(I_t,\tau) \\
&\qquad + b^T(I_t,\tau+\omega(I_t)u)\mathbf{R}(u)\partial_I a(I_t,\tau)\,d\tau\,du\,dt\,,
\end{aligned}
\tag{4.18}
$$

with $\psi_{t=0}=\theta_0$.

5. **Properties of the Limiting Equation for the Action Variable**

In this section, we study properties of the equation (4.13). The main issue concerns the behavior of a at $I=0$, but we will see how to define a unique global solution for this equation that does not reach 0 with probability one. In the case of a quadratic Hamiltonian, we have seen in (4.16) that $(I_t)_{t\geqslant 0}$ is related to a 2-dimensional squared Bessel process (see (4.16)) which is known to reach 0 with probability zero [KS91, Proposition 3.22 p. 161].

5.1. **Existence of a solution.** First, we remark that the function a is not Lipschitz in I as it typically has a square-root singularity near 0. So it is not immediately apparent that equation (4.13) has solutions. To construct solutions to equation (4.13), define

$$
a_n(I) \overset{\text{def}}{=} \int_{\tau=0}^{1} a(I,\tau)e^{-2i\pi n\tau}\,d\tau \tag{5.1}
$$

$$
R_n(I) \overset{\text{def}}{=} 2\int_{u=0}^{\infty} R(u)\cos(2\pi n\omega(I)u)\,du\,. \tag{5.2}
$$

and let

$$
\begin{aligned}
\mathbf{a}(I) &\overset{\text{def}}{=} \sum_{n\in\mathbb{Z}^*}|a_n(I)|^2R_n(I) \\
&= 2\int_{u=0}^{\infty} R(u)\int_{\tau=0}^{1} a(I,\tau)a(I,\tau+\omega(I)u)\,d\tau\,du\,,
\end{aligned}
\tag{5.3}
$$

$$
\begin{aligned}
\mathbf{b}(I) \overset{\text{def}}{=} \int_{u=0}^{\infty} R(u)\int_{\tau=0}^{1}\Big[&\partial_I(a(I,\tau+\omega(I)u))a(I,\tau) \\
&+\partial_I a(I,\tau)a(I,\tau+\omega(I)u)\Big]\,d\tau\,du\,.
\end{aligned}
\tag{5.4}
$$

Note that equality in (5.3) is justified by using the Fourier series expansion for $a(I,\tau)$, and using the fact that $a(I,\tau)$ is mean-zero in τ. That is,

$$
a(I,\tau)=\sum_{n\in\mathbb{Z}^*}e^{2i\pi n\tau}a_n(I)\,.
$$

Definition 5.1. A weak solution in the interval $(0,\infty)$ to equation (4.13) is a filtered probability space $(\Omega,\mathcal{F},\mathcal{F}_t,\boldsymbol{P})$ satisfying the usual conditions and a continuous pair of processes (I_t,W_t) such that the followings hold.

(1) The process I takes values in $[0,\infty]$ with $I_0\in(0,\infty)$,
(2) The process W is a standard $\mathcal{F}_t$-adapted Brownian motion on the Hilbert space $L^2_0(\mathbb{T})$ with covariance function given by (4.14).

(3) For any $t, M > 0$ we have

$$\int_{s=0}^{t\wedge\zeta_M} (\mathbf{a}(I_s) + |\mathbf{b}(I_s)|)\, ds < \infty\,, \tag{5.5}$$

$$I_{t\wedge\zeta_M} = I_0 + \int_{s=0}^{t\wedge\zeta_M}\int_{\tau=0} a(I_s,\tau)\, dW_s(I_s,\tau) + \int_{s=0}^{t\wedge\zeta_M} \mathbf{b}(I_s)\, ds \quad \forall t \geqslant 0\,, \tag{5.6}$$

almost surely. Here $\zeta_M \stackrel{\text{def}}{=} \zeta_{1,M} \wedge \zeta_{2,M}$ where $\zeta_{1,M}$ and $\zeta_{2,M}$ are defined by

$$\zeta_{1,M} \stackrel{\text{def}}{=} \inf\{t \geqslant 0 \,|\, |I_t| \geqslant M\}$$
$$\zeta_{2,M} \stackrel{\text{def}}{=} \inf\{t \geqslant 0 \,|\, |I_t| \leqslant 1/M\}\,.$$

For notational convenience we denote

$$\zeta_\infty \stackrel{\text{def}}{=} \lim_{M\to\infty} \zeta_M = \inf\{t \geqslant 0 \,|\, I_t \notin (0,\infty)\} \tag{5.7}$$

to be the first exit time of I from $(0,\infty)$.

We will now work with the canonical probability space $\mathcal{C}([0,\infty),\mathbb{R})$ and canonical filtration

$$\mathcal{M}_t = \sigma(h_s,\, 0 \leqslant s \leqslant t)\,.$$

To construct a solution to (4.13), we note first that when restricted to the interval $[1/M, M]$, the function a is Lipschitz. Thus we know (4.13) has a strong solution when truncated to this interval (see Proposition 6.1, below). If we denote this truncated solution by (I^M, ζ_M) we obtain an increasing sequence of stopping times ζ_M, an increasing family of σ-algebras $\mathcal{M}_{\zeta_M}$, and a sequence of probability measures $\boldsymbol{P}_I^M$ such that

$$\boldsymbol{P}_I^{M+1} = \boldsymbol{P}_I^M \qquad \text{on} \quad \mathcal{M}_{\zeta_M} \qquad \forall M > 0\,.$$

As a result, we can define $\boldsymbol{P}^0$ on $\bigcup_M \mathcal{M}_{\zeta_M}$ by

$$\boldsymbol{P}_I^0(O) = \boldsymbol{P}_I^M(O) \qquad \text{for} \quad O \in \mathcal{M}_{\zeta_M}\,,$$

and extend $\boldsymbol{P}_I^0$ to a probability measure on the σ-algebra $\mathcal{M}_{\zeta_\infty}$ such that for any $M \in \mathbb{N}$ we have

$$\boldsymbol{P}_I^0 = \boldsymbol{P}_I^M \qquad \text{on} \quad \mathcal{M}_{\zeta_M}\,. \tag{5.8}$$

Note that

$$\mathcal{M}_{\zeta_\infty} = \sigma\Big(\bigcup_M \mathcal{M}_{\zeta_M}\Big)\,, \tag{5.9}$$

which can be easily verified from the fact that

$$\mathcal{M}_{\zeta_\infty} = \sigma(h_{t\wedge\zeta_\infty}, t \geqslant 0)\,,$$

(see for instance [SV06, Lemma 1.3.3 p. 33]). As a result $\boldsymbol{P}_I^0$ provides a solution to (4.13), that we can denote by $I = (I_t)_{t\geqslant 0}$, in the sense of Definition 5.1. Note that the process $(I_t)_{t\geqslant 0}$ is nonnegative with probability one thanks to the Portmanteau theorem [Bil99, Theorem 2.1 p. 16].

In the next two sections, we will first show (using the bounds in Proposition 4.2) that $\boldsymbol{P}_I^0(\zeta_\infty = \infty) = 1$. That is, the process I does not reach 0 or blow up in finite time. This will show that $\boldsymbol{P}_I^0$ is the unique extension on $\mathcal{M} \stackrel{\text{def}}{=} \sigma(\bigcup_{t\geqslant 0} \mathcal{M}_t)$ satisfying (5.8). Additionally, we will show that $I = (I_t)_{t\geqslant 0}$ is the unique global solution to (4.13).

5.2. **Non-explosion and inaccessibility to 0.** Our aim in this section is to show that the process I can not explode in finite time nor reach 0.

Proposition 5.2. *Consider a weak solution to (4.13) (as in definition 5.1), and let ζ_∞ be as in (5.7). Then*

$$\boldsymbol{P}_I^0(\zeta_\infty = \infty) = 1\,.$$

This result together with (5.8) implies that for any $t \geqslant 0$

$$\lim_{M\to\infty} \boldsymbol{P}_I^M(\zeta_M \leqslant t) = \lim_{M\to\infty} \boldsymbol{P}_I^0(\zeta_M \leqslant t) = 0\,.$$

Therefore, [SV06, Theorem 1.3.5 p. 34] guarantees that $\boldsymbol{P}_I^0$ is the unique extension satisfying (5.8).

Before proving Proposition 5.2 we need to introduce the *scale function*

$$p(x) \overset{\text{def}}{=} \int_1^x \exp\left(-2\int_1^\xi \frac{\mathbf{b}(\nu)}{\mathbf{a}(\nu)}\,d\nu\right) d\xi, \quad x \in (0,\infty)$$

where $\mathbf{a}$ and $\mathbf{b}$ are defined by (5.3) and (5.4) respectively. Observe

$$p''(x) = -2\frac{\mathbf{b}(x)}{\mathbf{a}(x)}p'(x). \tag{5.10}$$

Standard results express the exit probability of I in terms of the scale function, and we prove them in our context for completeness.

Lemma 5.3. *For any $0 < x_- < x_+ < \infty$ we have*

$$\boldsymbol{P}(I_{\zeta_{x_-,x_+}} = x_-) = \frac{p(x_+) - p(I_0)}{p(x_+) - p(x_-)} \quad \textit{and} \quad \boldsymbol{P}(I_{\zeta_{x_-,x_+}} = x_+) = \frac{p(I_0) - p(x_-)}{p(x_+) - p(x_-)}\,,$$

where $\zeta_{x_-,x_+} = \zeta_{1,x_+} \wedge \zeta_{2,x_-}$.

Proof of Lemma 5.3. The proof consists of applying the Itô formula for $p(I_t)$ with $t \leqslant \zeta_{x_-,x_+} \leqslant \zeta_\infty$. To do this we rewrite the stochastic integral of (4.13) in the Fourier domain, that is

$$\int_{s=0}^t \int_{\tau=0}^1 a(I_s,\tau)\,dW_s(I_s,\tau)\,d\tau = \int_{s=0}^t \sum_{n\in\mathbb{Z}^*} a_n(I_s)R_n^{1/2}(I_s)\frac{1}{\sqrt{2}}(dB_{s,n}^1 - i dB_{s,n}^2)\,,$$

providing an explicit semi-martingale representation for $(I_t)_{t\geqslant 0}$. Therefore, applying the Itô formula to $p(I_t)$ we have

$$\begin{aligned}
p(I_{t\wedge\zeta_{x_-,x_+}}) &= p(I_0) + \int_0^{t\wedge\zeta_{x_-,x_+}} p'(I_s)\,dI_s + \frac{1}{2}\int_0^{t\wedge\zeta_{x_-,x_+}} p''(I_s)d\langle I\rangle_s \\
&= p(I_0) + \int_0^{t\wedge\zeta_{x_-,x_+}} p'(I_s)\sum_{n\in\mathbb{Z}^*} a_n(I_s)R_n^{1/2}(I_s)\frac{1}{\sqrt{2}}(dB_{s,n}^1 - i dB_{s,n}^2) \\
&\quad + \int_0^{t\wedge\zeta_{x_-,x_+}} p'(I_s)\mathbf{b}(I_s)\,ds \\
&\quad + \frac{1}{2}\int_0^{t\wedge\zeta_{x_-,x_+}} -2\frac{\mathbf{b}(I_s)}{\mathbf{a}(I_s)}p'(I_s)\sum_{n\in\mathbb{Z}^*}|a_n(I_s)|^2 R_n(I_s)\,ds \\
&= p(I_0) + \int_0^{t\wedge\zeta_{x_-,x_+}} p'(I_s)\sum_{n\in\mathbb{Z}^*} a_n(I_s)R_n^{1/2}(I_s)\frac{1}{\sqrt{2}}(dB_{s,n}^1 - i dB_{s,n}^2)
\end{aligned}$$

thanks to (5.10) and (5.3). Note that the stochastic integral on the right-hand side of the last line is a martingale starting from 0, so by taking the expectation, and passing to the limit $t \to \infty$ we obtain

$$p(I_0) = \boldsymbol{E}[p(I_{\zeta_{x_-,x_+}})] = p(x_-)\boldsymbol{P}(I_{\zeta_{x_-,x_+}} = x_-) + p(x_+)\boldsymbol{P}(I_{\zeta_{x_-,x_+}} = x_+)\,.$$

Solving for $\boldsymbol{P}(I_{\zeta_{x_-,x_+}} = x_-)$ and $\boldsymbol{P}(I_{\zeta_{x_-,x_+}} = x_+)$ we obtain the desired result. □

With this lemma, we can now turn to the proof of Proposition 5.2.

Proof of Proposition 5.2. Let us start with the following remark. The term $\mathbf{b}$ can be written as

$$\mathbf{b}(I) = \frac{1}{2}\partial_I \mathbf{a}(I)\,,$$

so that

$$\begin{aligned} p(0_+) &= \lim_{x\to 0_+} \int_1^x \exp\left(-2\int_1^\xi \frac{\mathbf{b}(\nu)}{\mathbf{a}(\nu)}d\nu\right) d\xi \\ &= \lim_{x\to 0_+} \mathbf{a}(1)\int_1^x \exp\left(-\ln(\mathbf{a}(\xi))\right) d\xi \\ &= \lim_{x\to 0_+} \mathbf{a}(1)\int_1^x \frac{d\xi}{\mathbf{a}(\xi)}\,. \end{aligned}$$

Now, according to (5.3) together with Proposition 4.2, we have

$$0 \leqslant \mathbf{a}(\xi) \leqslant C\,\xi$$

since for any $n \in \mathbb{Z}^*$

$$\begin{aligned} 0 \leqslant R_n(\xi) &= 2\int_0^\infty R(u)\cos(2\pi n\omega(\xi)u)\,du \\ &= 2\int_0^\infty \Big(\int_S r(p)e^{-\mu|p|^{2\beta}u}\,dp\Big)\cos(2\pi n\omega(\xi)u)\,du \\ &= 2\int_S dp\, r(p)\Big(\int_0^\infty \cos(2\pi n\omega(\xi)u)e^{-\mu|p|^{2\beta}u}\,du\Big) \\ &= \int_S dp\, r(p)\frac{2\mu|p|^{2\beta}}{\mu^2|p|^{4\beta} + 4\pi^2 n^2\omega^2(\xi)} \leqslant \frac{C_{\omega_0}}{n^2}\int_S |p|^{2\beta} r(p)\,dp\,. \end{aligned} \tag{5.11}$$

As a result, we have

$$p(0_+) \leqslant \lim_{x\to 0_+} \mathbf{a}(1)\int_1^x \frac{d\xi}{C\xi} = -\infty\,,$$

which implies $p(0_+) = -\infty$. To conclude the proof we consider two cases, that is $p(\infty) = \infty$ and $p(\infty) < \infty$. Following the same lines as [KS91, Proposition 5.22 p. 345], if $p(\infty) = \infty$ we directly have the desired result, and if $p(\infty) < \infty$ we have

$$\lim_{x_-\to 0_+}\lim_{x_+\to\infty} \boldsymbol{P}(I_{\zeta_{x_-,x_+}} = x_+) = \lim_{x_-\to 0_+}\lim_{x_+\to\infty} \frac{p(I_0) - p(x_-)}{p(x_+) - p(x_-)} = 1\,.$$

This implies $\boldsymbol{P}(\zeta_\infty = \zeta_{1,\infty}) = 1$ meaning that if I_t exits the interval $(0,\infty)$, it almost surely does so at ∞. Here

$$\zeta_{1,\infty} = \lim_{M\to\infty} \zeta_{1,M}\,.$$

To prove that $\boldsymbol{P}(\zeta_{1,\infty} = \infty) = 1$, and then conclude the proof of the Proposition 5.2, we need the following result.

Lemma 5.4. *For any $T > 0$ we have*

$$\lim_{M\to\infty} \boldsymbol{P}\Big(\sup_{t\in[0,T]} I_{t\wedge\zeta_M} \geqslant M\Big) = 0\,.$$

Indeed, writing

$$\begin{aligned}\boldsymbol{P}(\zeta_{1,\infty} < \infty) &= \lim_{T\to\infty} \boldsymbol{P}(\zeta_{1,\infty} \leqslant T)\\ &= \lim_{T\to\infty} \lim_{M_0\to\infty} \boldsymbol{P}\Big(\zeta_{1,\infty} \leqslant T, \inf_{t\in[0,\zeta_\infty)} I_t > 1/M_0\Big),\\ &= \lim_{T\to\infty} \lim_{M_0\to\infty} \lim_{M\to\infty} \boldsymbol{P}\Big(\zeta_{1,M} \leqslant T, \inf_{t\in[0,\zeta_M)} I_t > 1/M_0\Big)\,.\end{aligned}$$

To justify the second equality, we note that if $\zeta_{1,\infty}$ is finite, then $\inf_{t\in[0,\zeta_\infty)} I_t$ must be positive since $\zeta_{1,\infty} = \zeta_\infty$ almost surely. We have for any $M > M_0$

$$\begin{aligned}\boldsymbol{P}\Big(\zeta_{1,M} \leqslant T, \inf_{t\in[0,\zeta_M]} I_t > \frac{1}{M_0}\Big) &\leqslant \boldsymbol{P}(\zeta_{1,M} \leqslant T,\, \zeta_M = \zeta_{1,M})\\ &\leqslant \boldsymbol{P}\Big(\sup_{t\in[0,T]} I_{t\wedge\zeta_M} \geqslant M\Big).\end{aligned}$$

Therefore, Lemma 5.4 will imply $\boldsymbol{P}(\zeta_{1,\infty} < \infty) = 0$.

Proof of Lemma 5.4. First we remark that using Chebychev's inequality

$$\begin{aligned}\boldsymbol{P}\Big(\sup_{t\in[0,T]} I_{t\wedge\zeta_M} \geqslant M\Big) &= \boldsymbol{P}\Big(\sup_{t\in[0,T]} I^2_{t\wedge\zeta_M} \geqslant M^2\Big)\\ &\leqslant \frac{1}{M^2}\boldsymbol{E} \sup_{t\in[0,T]} I^2_{t\wedge\zeta_{1,M}}\,.\end{aligned} \tag{5.12}$$

Let $T' \leqslant T$, we also have

$$\begin{aligned}\boldsymbol{E} \sup_{t\in[0,T']} I^2_{t\wedge\zeta_M} &\leqslant 3I_0 + 3\boldsymbol{E} \sup_{t\in[0,T']} N_t^2 + 3\boldsymbol{E} \sup_{t\in[0,T']} \Big(\int_0^{t\wedge\zeta_M} \mathbf{b}(I_s)ds\Big)^2\\ &= 3I_0 + J_{T'} + K_{T'}\,,\end{aligned}$$

and where, using (5.6)

$$N_{t\wedge\zeta_M} \overset{\text{def}}{=} I_{t\wedge\zeta_M} - I_0 - \int_0^{t\wedge\zeta_M} \mathbf{b}(I_s)\,ds = \int_{s=0}^{t\wedge\zeta_M} \int_{\tau=0}^1 a(I_s,\tau)\,dW_s(I_s,\tau)\,d\tau$$

is a martingale with quadratic variation

$$\langle N\rangle_t = \int_0^{t\wedge\zeta_M} \mathbf{a}(I_s)\,ds\,.$$

For the term $J_{T'}$ we apply Burkholder-Davis-Gundy inequality [KS91, Theorem 3.28 p. 166] together with (5.11) from which we obtain

$$\begin{aligned}J_{T'} \leqslant C\boldsymbol{E}\int_0^{T'\wedge\zeta_M} \mathbf{a}(I_s)\,ds &\leqslant C\int_0^{T'} \boldsymbol{E}[\mathbf{a}(I_{s\wedge\zeta_M})]\,ds\\ &\leqslant C_{\omega_0}\int_S |p|^{2\beta}|r(p)|\,dp \int_{s=0}^{T'}\int_{\tau=0}^1 \boldsymbol{E}[|a(I_{s\wedge\zeta_M},\tau)|^2]\,d\tau\,ds\,.\end{aligned}$$

Moreover, using Proposition 4.2, we have

$$\begin{aligned}\int_0^1 \boldsymbol{E}[|a(I_{s\wedge\zeta_M},\tau)|^2]\,d\tau &= \int_0^1 \Big(\boldsymbol{E}[|a(I_{s\wedge\zeta_M},\tau)|^2\mathbf{1}_{(I_{s\wedge\zeta_M}\leqslant r)}] \\ &\qquad + \boldsymbol{E}[|a(I_{s\wedge\zeta_M},\tau)|^2\mathbf{1}_{(I_{s\wedge\zeta_M}>r)}]\Big)\,d\tau \\ &\leqslant r\,C + C_r\boldsymbol{E}I^2_{s\wedge\zeta_M} \\ &\leqslant r\,C + C_r\boldsymbol{E}\sup_{t\in[0,s]} I^2_{t\wedge\zeta_M}\,,\end{aligned}$$

and hence

$$J_{T'} \leqslant C_1\,T + C_2\int_0^{T'} \boldsymbol{E}\Big(\sup_{t\in[0,s]} I^2_{t\wedge\zeta_M}\Big)ds\,.$$

Now, for $K_{T'}$, we use Cauchy-Schwarz inequality

$$\begin{aligned}\boldsymbol{E}\sup_{t\in[0,T']}\Big(\int_0^{t\wedge\zeta_M}\mathbf{b}(I_s)\,ds\Big)^2 &\leqslant T\boldsymbol{E}\sup_{t\in[0,T']}\int_0^{t\wedge\zeta_M}\mathbf{b}^2(I_s)\,ds \\ &\leqslant T\boldsymbol{E}\int_0^{T'}\mathbf{b}^2(I_{s\wedge\zeta_M})\,ds\,.\end{aligned}$$

From Proposition 4.2, for $I\leqslant r$, we again use the Cauchy-Schwarz inequality

$$\begin{aligned}|\mathbf{b}(I)| &\leqslant C\,|\omega'(I)|\sqrt{\int_0^1|a(I,\tau)|^2\,d\tau\int_0^1|\partial_\tau a(I,\tau)|^2\,d\tau} \\ &\quad + C_{\omega_0}\sqrt{\int_0^1|a(I,\tau)|^2\,d\tau\int_0^1|\partial_I a(I,\tau)|^2\,d\tau} \\ &\leqslant C.\end{aligned}$$

As a result,

$$\begin{aligned}\boldsymbol{E}\int_0^{T'}\mathbf{b}^2(I_{s\wedge\zeta_M})\,ds &= \boldsymbol{E}\int_0^{T'}\mathbf{b}^2(I_{s\wedge\zeta_M})\mathbf{1}_{(I_{s\wedge\zeta_M}\leqslant r)}\,ds \\ &\qquad + \boldsymbol{E}\int_0^{T'}\mathbf{b}^2(I_{s\wedge\zeta_M})\mathbf{1}_{(I_{s\wedge\zeta_M}>r)}\,ds \\ &\leqslant C_1\,T + C_{2,r}\boldsymbol{E}\int_0^{T'} I^2_{s\wedge\zeta_M}\mathbf{1}_{(I_{s\wedge\zeta_M}>r)}\,ds \\ &\leqslant C_1\,T + C_{2,r}\int_0^{T'}\boldsymbol{E}\Big(\sup_{t\in[0,s]} I^2_{t\wedge\zeta_M}\Big)\,ds\,,\end{aligned}$$

and

$$K_{T'} \leqslant C_1\,T^2 + C_2\,T\int_0^{T'}\boldsymbol{E}\Big(\sup_{t\in[0,s]} I^2_{t\wedge\zeta_M}\Big)\,ds\,.$$

Finally, we have

$$\boldsymbol{E}\sup_{t\in[0,T']} I^2_{t\wedge\zeta_M} \leqslant C_1(1+T+T^2) + C_2(1+T)\int_0^{T'}\boldsymbol{E}\Big(\sup_{t\in[0,s]} I^2_{t\wedge\zeta_M}\Big)\,ds$$

and then by Gronwall's inequality with $T'=T$

$$\boldsymbol{E}\sup_{t\in[0,T]} I^2_{t\wedge\zeta_M} \leqslant C_1(1+T+T^2)e^{C_2(1+T)T}\,.$$

This concludes the proof of the Lemma by letting $M \to \infty$ in (5.12). Consequently, the proof of Proposition 5.2 is also complete. □

5.3. **Uniqueness of solutions.** The weak uniqueness property for equation (4.13) comes from the fact that this equation has strong uniqueness on $\mathcal{M}_{\zeta_{2,M}}$ for any M, and then weak uniqueness on $\mathcal{M}_{\zeta_{2,M}}$ for any M [KS91, Proposition 3.20 p. 309], since we avoid the lack of regularity at 0 for the action variable I. However, according to Proposition 5.2, we have $\lim_{M\to\infty} \boldsymbol{P}_I^0(\zeta_{2,M} \leqslant T) = 0$ for any $T > 0$ which guarantees the weak uniqueness thanks to [SV06, Theorem 1.3.5 p. 34].

6. Proof of the main theorem (Theorem 4.3)

6.1. **Truncated Process.** We first describe the strategy used to prove Theorem 4.3. In addition to the difficulties concerning the behavior of a around $I = 0$ that we mentioned in Theorem 4.3, we have no *apriori* estimates for the process $(I_t^\varepsilon, \psi_t^\varepsilon)_{t\geqslant 0}$ that are uniform in ε, in probability. This is a problem when trying to apply [Kus84, Theorem 4 p. 48] directly to prove the tightness property, and when trying to identify the law of the subsequential limits. To bypass this we follow the strategy developed in [SV06, Chapter 11] by introducing a *truncated process* that does not suffer from the above problems. To relate the truncated process with the original one we introduce a family of stopping times. As we will see, these times go to ∞ when we remove the truncation and then finally obtain the weak convergence of the original processes.

Let $M > 0$ and consider the process $(I_t^{\varepsilon,M}, \psi_t^{\varepsilon,M})_{t\geqslant 0}$ which is the unique solution to

$$\begin{aligned} \dot{I}_t^{\varepsilon,M} &= \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big)\phi_M(I_t^{\varepsilon,M}, \psi_t^{\varepsilon,M}) a(I_t^{\varepsilon,M}, \psi_t^{\varepsilon,M} + \tau_t^{\varepsilon,M}) \\ \dot{\psi}_t^{\varepsilon,M} &= \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big)\phi_M(I_t^{\varepsilon,M}, \psi_t^{\varepsilon,M})(b(I_t^{\varepsilon,M}, \psi_t^{\varepsilon,M} + \tau_t^{\varepsilon,M}) - \langle b(I_t^{\varepsilon,M}, \cdot)\rangle) \end{aligned} \tag{6.1}$$

coupled with the initial conditions $I_0^{\varepsilon,M} = I_0$, and $\psi_0^{\varepsilon,M} = \theta_0$, with

$$\dot{\tau}_t^{\varepsilon,M} = \frac{\omega(I_t^{\varepsilon,M})}{\varepsilon^2} + \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big)\phi_M(I_t^{\varepsilon,M}, \psi_t^{\varepsilon,M})\langle b(I_t^{\varepsilon,M}, \cdot)\rangle\,.$$

Here ϕ_M is a smooth function on $\mathbb{R}^2$ such that

$$0 \leqslant \phi_M \leqslant 1, \quad \phi_M(I, \psi) = 1 \quad \text{if} \quad 1/M \leqslant I \leqslant M \quad \text{and} \quad |\psi| \leqslant M\,,$$

and

$$\phi_M(I, \psi) = 0 \quad \text{if} \quad I > 2M \quad \text{or} \quad I < 1/(2M) \quad \text{or} \quad |\psi| > 2M\,.$$

Thanks to the truncation provided by the cutoff function ϕ_M, for both action and angle variables, the process $(I_t^{\varepsilon,M}, \psi_t^{\varepsilon,M})_{t\geqslant 0}$ does not suffer from the regularity problem for the action variable around 0 as mentioned above. Also, its convergence in distribution in $\mathcal{C}([0,\infty), \mathbb{R}^2)$ can be proved using the perturbed-test function method and martingale properties [Kus84, PSV76] thanks to the boundedness of the process. All these points are developed more precisely in Section 7.

To simplify the notation, we now omit the superscript M and denote

$$Y_t^\varepsilon \overset{\text{def}}{=} \begin{pmatrix} I_t^{\varepsilon,M} \\ \psi_t^{\varepsilon,M} \end{pmatrix}\,.$$

Now the system (6.1) can be rewritten as

$$\dot{Y}_t^\varepsilon = \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big) F(Y_t^\varepsilon, \tau_t^\varepsilon) \tag{6.2}$$

with

$$\dot{\tau}_t^\varepsilon = \frac{\omega(Y_t^\varepsilon)}{\varepsilon^2} + \frac{1}{\varepsilon} v\Big(\frac{t}{\varepsilon^2}\Big) G(Y_t^\varepsilon)\,. \tag{6.3}$$

Here

$$\begin{aligned} F(Y,\tau) &\stackrel{\text{def}}{=} \begin{pmatrix} \phi_M(I,\psi) a(I,\psi+\tau) \\ \phi_M(I,\psi)(b(I,\psi+\tau) - \langle b(I,\cdot)\rangle) \end{pmatrix}, \\ G(Y) &\stackrel{\text{def}}{=} \phi_M(I,\psi)\langle b(I,\cdot)\rangle\,, \end{aligned} \tag{6.4}$$

and a, b are defined in equation (4.5). Note that F is a smooth bounded function (with bounded derivatives) 1-periodic and mean-zero with respect to τ (the truncation only affects ψ, not τ), and $\omega(Y_t^\varepsilon) = \omega(I_t^{\varepsilon,M})$.

To relate the truncated and original processes we introduce first some notations. In the remaining of the paper, the space of all possible outcomes is $\mathcal{C}([0,\infty),\mathbb{R}^2)$, and we denote the corresponding canonical filtration by $\mathcal{M}_t = \sigma(y_s, 0 \leqslant s \leqslant t)$. The laws of the truncated and original processes on $\mathcal{C}([0,\infty),\mathbb{R}^2)$ will be denoted respectively by $\boldsymbol{P}^{\varepsilon,M}$ and $\boldsymbol{P}^\varepsilon$. Now, we consider the stopping times

$$\eta_{1,M}(y) \stackrel{\text{def}}{=} \inf\{t \geqslant 0 : \|y_t\| \geqslant M\} \quad \text{and} \quad \eta_{2,M}(y) \stackrel{\text{def}}{=} \inf\{t \geqslant 0 : |y_t^1| \leqslant 1/M\}\,,$$

for any $y = (y^1, y^2) \in \mathcal{C}([0,\infty),\mathbb{R}^2)$ and

$$\eta_M \stackrel{\text{def}}{=} \eta_{1,M} \wedge \eta_{2,M}\,. \tag{6.5}$$

It is not hard to see that $\eta_{1,M}$ and $\eta_{2,M}$ are lower semi-continuous so that η_M is also lower semi-continuous, that is their lower level sets are closed subsets of $\mathcal{C}([0,\infty),\mathbb{R}^2)$. The latter property will be used to obtain the convergence of the original process from the truncated one.

From the definition of the above stopping times and the cutoff function ϕ_M it is clear that

$$\boldsymbol{P}^{\varepsilon,M} = \boldsymbol{P}^\varepsilon \qquad \text{on} \quad \mathcal{M}_{\eta_M}\,. \tag{6.6}$$

This is the relation that links the truncated and original processes and that will be also used below in Section 6.2 to show the convergence of the original process.

For now, we have the following convergence results for the truncated process. We return the index M for a moment to emphasize the truncation.

Proposition 6.1. *The family $(Y^\varepsilon)_{\varepsilon>0} \stackrel{\text{def}}{=} (I^{\varepsilon,M}, \psi^{\varepsilon,M})_{\varepsilon>0}$ converges in distribution in $\mathcal{C}([0,\infty),\mathbb{R}^2)$ to the unique solution $(Y_t^M)_{t\geqslant 0} = (I_t^M, \psi_t^M)_{t\geqslant 0}$ to the martingale problem with generator defined by*

$$\mathcal{L}h(y) \stackrel{\text{def}}{=} \int_{\tau=0}^1 \int_{u=0}^\infty R(u)\tilde{F}(0,y,\tau)\cdot\nabla(\tilde{F}(u,y,\tau)\cdot\nabla h(y))\,du\,d\tau\,, \tag{6.7}$$

where

$$\tilde{F}(u,y,\tau) \stackrel{\text{def}}{=} F(y,\tau+\omega(y)u)\,,$$

with F defined in (6.4) and starting point (I_0,ψ_0).

From this characterization, we can deduce the SDEs satisfied by the limiting process $Y_t^M = (I_t^M, \psi_t^M)$ up to the stopping time η_M.

Proposition 6.2. *Denoting* $\tilde{I}_t^M \overset{\text{def}}{=} I^M_{t\wedge\eta_M}$, *we have*

$$
\begin{aligned}
d\tilde{I}_t^M &= \int_{\tau=0}^1 a(\tilde{I}_t^M,\tau)\,dW_{t\wedge\eta_M}(\tilde{I}_t^M,\tau)\,d\tau \\
&\quad + \int_{u=0}^\infty R(u)\Big(\int_{\tau=0}^1 \partial_I\big(a(I,\tau+\omega(I)u)\big)_{|I=\tilde{I}_t^M} a(\tilde{I}_t^M,\tau) \\
&\qquad + a(\tilde{I}_t^M,\tau+\omega(\tilde{I}_t^M)u)\partial_I a(\tilde{I}_t^M,\tau)\Big)\,d\tau\,du\,dt,
\end{aligned}
\tag{6.8}
$$

with $\tilde{I}^M_{t=0} = I_0$, *and where* W *is defined in* (4.14). *Also, we have*

$$
\begin{aligned}
d\psi^M_{t\wedge\eta_M} &= \int_{\tau=0}^1 b(\tilde{I}_t^M,\tau)\,dW_{t\wedge\eta_M}(\tilde{I}_t^M,\tau)\,d\tau \\
&\quad + \int_{u=0}^\infty R(u)\Big(\int_{\tau=0}^1 \partial_I\big(b(I,\tau+\omega(I)u)\big)_{|I=\tilde{I}_t^M} a(\tilde{I}_t^M,\tau) \\
&\qquad + b(\tilde{I}_t^M,\tau+\omega(\tilde{I}_t^M)u)\partial_I a(\tilde{I}_t^M,\tau)\Big)\,d\tau\,du\,dt,
\end{aligned}
\tag{6.9}
$$

with $\psi^M_{t=0} = \theta_0$.

As we can see, this result is similar to the one of Theorem 4.3, with coefficients a and b themselves, but involving the cutoff M through the stopping time η_M. The proofs of Proposition 6.1 and 6.2 are presented in Sections 7 and 8, respectively. In the next section, we will show how these two propositions can be used to prove Theorem 4.3.

6.2. **Convergence of** $(I^\varepsilon,\psi^\varepsilon)_{\varepsilon>0}$**.** In order to prove Theorem 4.3, we need to remove the cutoff M to show the convergence in distribution of $(I^\varepsilon,\psi^\varepsilon)_{\varepsilon>0}$ to $(I_t,\psi_t)_{t\geqslant 0}$, where I, ψ are defined by (4.13)-(4.15). This is similar to Lemma 11.1.1 (p. 262) in [SV06], and we adapt the proof to our situation.

In what follows, we denote by $\boldsymbol{P}^M$ and $\boldsymbol{P}^0$ the distribution of $Y^M = (I_t^M,\psi_t^M)_{t\geqslant 0}$ (defined in Proposition 6.1) and $Y = (I_t,\psi_t)_{t\geqslant 0}$ (defined in Proposition 4.3), respectively. Note that according to (6.9) and (4.15) the distributions $\boldsymbol{P}^M$ and $\boldsymbol{P}^0$ are completely determined by their first marginal $\boldsymbol{P}_I^M$ and $\boldsymbol{P}_I^0$. Also, in view of (6.8) and (6.9), it is straightforward to see that (with η_M defined in (6.5))

$$
\boldsymbol{P}^0 = \boldsymbol{P}^M \qquad \text{on} \quad \mathcal{M}_{\eta_M}\,.
\tag{6.10}
$$

To prove this convergence, let Φ be a continuous bounded function on $\mathcal{C}([0,\infty),\mathbb{R}^2)$, and write

$$
\begin{aligned}
&\boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t,t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}}[\Phi(y_t,t\in[0,T])] \\
&\quad = \boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t,t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}^{\varepsilon,M}}[\Phi(y_t,t\in[0,T])] \\
&\qquad + \boldsymbol{E}^{\boldsymbol{P}^{\varepsilon,M}}[\Phi(y_t,t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}^M}[\Phi(y_t,t\in[0,T])] \\
&\qquad + \boldsymbol{E}^{\boldsymbol{P}^M}[\Phi(y_t,t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}^0}[\Phi(y_t,t\in[0,T])]\,.
\end{aligned}
$$

Considering the following decomposition

$$
\begin{aligned}
\boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t,t\in[0,T])] &= \boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t,t\in[0,T])\mathbf{1}_{\{\eta_M\leqslant T\}}] \\
&\quad + \boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t,t\in[0,T])\mathbf{1}_{\{\eta_M> T\}}]
\end{aligned}
$$

with

$$\begin{aligned}\boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t, t\in[0,T])\mathbf{1}_{\{\eta_M>T\}}] &= \boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_{t\wedge\eta_M}, t\in[0,T])\mathbf{1}_{\{\eta_M>T\}}]\\ &= \boldsymbol{E}^{\boldsymbol{P}^{\varepsilon,M}}[\Phi(y_{t\wedge\eta_M}, t\in[0,T])\mathbf{1}_{\{\eta_M>T\}}]\\ &= \boldsymbol{E}^{\boldsymbol{P}^{\varepsilon,M}}[\Phi(y_t, t\in[0,T])\mathbf{1}_{\{\eta_M>T\}}]\\ &= \boldsymbol{E}^{\boldsymbol{P}^{\varepsilon,M}}[\Phi(y_t, t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}^{\varepsilon,M}}[\Phi(y_t, t\in[0,T])\mathbf{1}_{\{\eta_M\leqslant T\}}]\,.\end{aligned}$$

Here we used (6.6) in the second line. Using (6.6) again we note

$$\begin{aligned}\boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t, t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}^{\varepsilon,M}}[\Phi(y_t, t\in[0,T])]&\\ \leqslant \|\Phi\|_\infty(\boldsymbol{P}^\varepsilon(\eta_M\leqslant T) + \boldsymbol{P}^{\varepsilon,M}(\eta_M\leqslant T))&\\ \leqslant 2\|\Phi\|_\infty \boldsymbol{P}^{\varepsilon,M}(\eta_M\leqslant T)\,.&\end{aligned}$$

The same results hold for $\boldsymbol{P}^0$ and $\boldsymbol{P}^M$ instead of $\boldsymbol{P}^\varepsilon$ and $\boldsymbol{P}^{\varepsilon,M}$ respectively, but using (6.10) instead of (6.6). As a result, we obtain

$$\begin{aligned}\limsup_{\varepsilon\to0}|\boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t, t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}^0}[\Phi(y_t, t\in[0,T])]|&\\ \leqslant 2\|\Phi\|_\infty \limsup_{\varepsilon\to0}\boldsymbol{P}^{\varepsilon,M}(\eta_M\leqslant T)&\\ + 2\|\Phi\|_\infty \boldsymbol{P}^0(\eta_M\leqslant T)\,.&\end{aligned}$$

Notice that the stopping time η_M is lower semi-continuous on $\mathcal{C}([0,\infty),\mathbb{R}^2)$, so that $(\eta_M\leqslant M)\in\mathcal{M}_{\eta_M}$ is a closed subset of $\mathcal{C}([0,\infty),\mathbb{R}^2)$. According to the Portmanteau theorem [Bil99, Theorem 2.1 p. 16] we have

$$\limsup_{\varepsilon\to0}|\boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t, t\in[0,T])] - \boldsymbol{E}^{\boldsymbol{P}^0}[\Phi(y_t, t\in[0,T])]|\leqslant 4\|\Phi\|_\infty \boldsymbol{P}^0(\eta_M\leqslant T)\,,$$

for any M. Moreover, we have

$$\boldsymbol{P}^0(\eta_M\leqslant T)\leqslant \boldsymbol{P}^0(\zeta_M\leqslant T) + \boldsymbol{P}^0(\zeta'_M\leqslant T)\,.$$

where for any any $y=(y^1,y^2)\in\mathcal{C}([0,\infty),\mathbb{R}^2)$ we define

$$\zeta'_M(y)\stackrel{\text{def}}{=}\inf\{t\geqslant0 : |y^2_t|\geqslant M\}\,.$$

By Proposition 5.2 we know

$$\boldsymbol{P}^0(\zeta_M\leqslant T) = \boldsymbol{P}^0_I(\zeta_M\leqslant T)\underset{M\to\infty}{\longrightarrow}0\,.$$

For the other term, let $0<\eta<1$ and write

$$\begin{aligned}\boldsymbol{P}^0(\zeta'_M\leqslant T)\leqslant \boldsymbol{P}\Big(\sup_{t\in[0,T]}|\psi_t|\geqslant M, \inf_{t\in[0,T]}I_t>\eta, \sup_{t\in[0,T]}I_t<\frac{1}{\eta}\Big)&\\ + \boldsymbol{P}\Big(\sup_{t\in[0,T]}I_t\geqslant\frac{1}{\eta}\Big) + \boldsymbol{P}\Big(\inf_{t\in[0,T]}I_t\leqslant\eta\Big)\,,&\end{aligned}$$

then using Markov inequality and (4.15), we obtain the following for some $C_\eta > 0$

$$\begin{aligned}
&\boldsymbol{P}\Big(\sup_{t\in[0,T]} |\psi_t| \geqslant M, \inf_{t\in[0,T]} I_t > \eta, \sup_{t\in[0,T]} I_t < \frac{1}{\eta}\Big) \\
&\quad \leqslant \frac{1}{M^2} \boldsymbol{E}\Big[\sup_{t\in[0,T]} |\psi_t|^2 \mathbf{1}_{\{\inf_{t\in[0,T]} I_t > \eta, \sup_{t\in[0,T]} I_t < \frac{1}{\eta}\}}\Big] \\
&\quad \leqslant \frac{2}{M^2} \boldsymbol{E}\Big[\sup_{t\in[0,T]} \Big| \int_{\tau=0}^1 \int_{s=0}^t b(I_s,\tau) dW_s(I_s,\tau) d\tau \Big|^2 \mathbf{1}_{\{\inf_{t\in[0,T]} I_t > \eta, \sup_{t\in[0,T]} I_t < \frac{1}{\eta}\}}\Big] \\
&\qquad + \frac{C_\eta}{M^2}\,.
\end{aligned}$$

Indeed, since we are on the event $\{\inf_{t\in[0,T]} I_t > \eta, \sup_{t\in[0,T]} I_t < \frac{1}{\eta}\}$ we avoid the singularity of the function b at $I = 0$ and keep I bounded. Therefore, we can directly bound the drift term in (4.15) by bounding continuous functions $a, b, \partial_I a, \partial_I b$ by some constant C_η. We will subsequently let C_η be a constant that only depends on η whose value may change from line to line. The details are as follows.

Write the Fourier expansions of a and b as

$$a(I,\tau) = \sum_{n\in\mathbb{Z}^*} e^{2i\pi n\tau} a_n(I), \qquad b(I,\tau) = \sum_{n\in\mathbb{Z}} e^{2i\pi n\tau} b_n(I)\,.$$

Considering only the first integrand in the drift term in (4.15) and substituting the above expansions, we obtain

$$\begin{aligned}
&\int_{u=0}^\infty R(u) \int_{\tau=0}^1 \partial_I \big(b(I, \tau + \omega(I)u)\big)_{|I=I_t} a(I_t,\tau)\, d\tau\, du \\
&\quad = \int_{u=0}^\infty R(u) \int_{\tau=0}^1 \Big[\sum_{n\in\mathbb{Z}} \Big((\partial_I b)_n(I_t) + 2i\pi n\omega'(I_t) u b_n(I_t)\Big) e^{2i\pi n(\tau+\omega(I_t)u)} \\
&\qquad\qquad \cdot \Big(\sum_{n'\in\mathbb{Z}^*} a_{n'}(I_t) e^{2i\pi n'\tau}\Big)\Big]\, d\tau\, du \\
&\quad = \int_{u=0}^\infty R(u) \sum_{n\in\mathbb{Z}^*} \Big((\partial_I b)_n(I_t) + 2i\pi n\omega'(I_t) u b_n(I_t)\Big) a_{-n}(I_t) e^{2i\pi n\omega(I_t)u}\, du \qquad (6.11) \\
&\quad = R_1 + R_2\,, \qquad (6.12)
\end{aligned}$$

where

$$\begin{aligned}
R_1 &\stackrel{\text{def}}{=} \int_{u=0}^\infty R(u) \sum_{n\in\mathbb{Z}^*} (\partial_I b)_n(I_t) a_{-n}(I_t) \cos(2\pi n\omega(I_t)u)\, du \\
R_2 &\stackrel{\text{def}}{=} -\int_{u=0}^\infty R(u) \sum_{n\in\mathbb{Z}^*} (\partial_I b)_n(I_t) a_{-n}(I_t) \cdot 2\pi n\omega'(I_t) u \sin(2\pi n\omega(I_t)u)\, du\,.
\end{aligned}$$

Here (6.11) followed by expanding the double sum in n and n', integrating in τ, and observing that the only terms that do not vanish are those with $n + n' = 0$. The last equality (6.12) followed since the expression is real-valued and so must equal its real part.

Using (5.11) to bound the integral in u and bounding $(\partial_I b)_n(I_t)$, $a_{-n}(I_t)$ by some constant C_η, we have that

$$|R_1| \leqslant C_\eta \sum_{n\in\mathbb{Z}^*} \frac{1}{n^2} \leqslant C_\eta\,.$$

Similarly, we can estimate that

$$\begin{aligned}
&\int_{u=0}^{\infty} R(u)\cdot 2\pi n u \sin(2\pi n\omega(I_t)u)\,du \\
&\quad= \int_{u=0}^{\infty}\Big(\int_S r(p)e^{-\mu|p|^{2\beta}u}\,dp\Big)\cdot 2\pi n u\sin(2\pi n\omega(I_t)u)\,du \\
&\quad= \int_S dp\, r(p)\Big(\int_{u=0}^{\infty} 2\pi n u\sin(2\pi n\omega(I_t)u)e^{-\mu|p|^{2\beta}u}\,du\Big) \\
&\quad= 2\int_S dp\, r(p)\frac{(2\pi n)^2\omega(I_t)\mu|p|^{2\beta}}{(\mu^2|p|^{4\beta}+4\pi^2n^2\omega^2(I_t))^2} \le \frac{C_{\omega_0}}{n^2}\int_S |p|^{2\beta}r(p)\,dp\,.
\end{aligned}$$

Therefore, we also have

$$|R_2| \le C_\eta \sum_{n\in\mathbb{Z}^*}\frac{1}{n^2} \le C_\eta\,.$$

Combining these estimates we have

$$\begin{aligned}
\Big|\int_{u=0}^{\infty} R(u)\int_{\tau=0}^{1}&\partial_I\big(b(I,\tau+\omega(I)u)\big)_{|I=I_t}a(I_t,\tau) \\
&+ b(I_t,\tau+\omega(I_t)u)\partial_I a(I_t,\tau)\,d\tau\,du\Big| \leqslant C_\eta\sum_{n\in\mathbb{Z}^*}\frac{1}{n^2}\leqslant C_\eta\,.
\end{aligned}$$

For the stochastic integral term, we introduce a smooth function b^η such that for any $\tau\in(0,1)$ we have

$$b^\eta(I,\tau) = b(I,\tau) \quad \text{if} \quad I>\eta \qquad \text{and} \qquad b^\eta(I,\tau)=0 \quad \text{if} \quad I<\frac{\eta}{2}\,.$$

Then

$$\begin{aligned}
&\boldsymbol{E}\Big[\sup_{t\in[0,T]}\Big|\int_{\tau=0}^{1}\int_{s=0}^{t} b(I_s,\tau)dW_s(I_s,\tau)d\tau\Big|^2\mathbf{1}_{\{\inf_{t\in[0,T]}I_t>\eta,\,\sup_{t\in[0,T]}I_t<\frac{1}{\eta}\}}\Big] \\
&\quad= \boldsymbol{E}\Big[\sup_{t\in[0,T]}\Big|\int_{\tau=0}^{1}\int_{s=0}^{t} b^\eta(I_s,\tau)dW_s(I_s,\tau)d\tau\Big|^2\mathbf{1}_{\{\inf_{t\in[0,T]}I_t>\eta,\,\sup_{t\in[0,T]}I_t<\frac{1}{\eta}\}}\Big] \\
&\quad\leqslant C_\eta\,,
\end{aligned}$$

where the last inequality follows from the Burkholder-Davis-Gundy inequality to bound the supremum of a continuous local martingale. Finally, for any $0<\eta<1$ we have

$$\limsup_{M\to\infty}\boldsymbol{P}^0(\zeta'_M\leqslant T)\leqslant \boldsymbol{P}\Big(\sup_{t\in[0,T]}I_t\geqslant\frac{1}{\eta}\Big)+\boldsymbol{P}\Big(\inf_{t\in[0,T]}I_t\leqslant\eta\Big),$$

and

$$\begin{aligned}
&\lim_{\eta\to0}\boldsymbol{P}\Big(\inf_{t\in[0,T]}I_t\leqslant\eta\Big)=\boldsymbol{P}\Big(\inf_{t\in[0,T]}I_t=0\Big)=0\,, \\
&\lim_{\eta\to0}\boldsymbol{P}\Big(\sup_{t\in[0,T]}I_t\geqslant\frac{1}{\eta}\Big)=0 \qquad \text{(by Lemma 5.4)}\,.
\end{aligned}$$

As a result, we obtain

$$\lim_{\varepsilon\to0}\boldsymbol{E}^{\boldsymbol{P}^\varepsilon}[\Phi(y_t,t\in[0,T])]=\boldsymbol{E}^{\boldsymbol{P}^0}[\Phi(y_t,t\in[0,T])]\,,$$

which concludes the proof of Theorem 4.3.

It remains to prove Proposition 6.1, and we do this in the next section.

7. Proof of Proposition 6.1

We will prove convergence of the family $(Y^\varepsilon)_{\varepsilon>0}$ by first showing tightness, and then showing all subsequential limits are solutions of a well-posed martingale problem with generator (6.7). This follows the classical approach to prove weak convergence of a sequence of diffusions using martingale problem that is well developed in [SV06]. However, because the processes involved are not diffusions (and often not even Markovian), we use pseudo-generators and the perturbed test function method (see for instance [Kus84]).

7.1. **Pseudo-generator and Perturbed test function method.** In this section, we follow [Kus84, Chapter 3]. Even though the pair of processes $(Y_t^\varepsilon, \tau_t^\varepsilon)_{t\geqslant 0}$ is Markov for each $\varepsilon > 0$, the process $(Y_t^\varepsilon)_{t\geqslant 0}$ by itself is *not* Markov. Since analyzing the pair $(Y_t^\varepsilon, \tau_t^\varepsilon)_{t\geqslant 0}$ is difficult, we study the process $(Y_t^\varepsilon)_{t\geqslant 0}$ on its own by using the pseudo-generator. This allows us to apply martingale techniques to the non-Markovian process $(Y_t^\varepsilon)_{t\geqslant 0}$.

7.1.1. *Pseudo-generator.* Let

$$\mathcal{G}_t^\varepsilon \stackrel{\text{def}}{=} \mathcal{G}_{t/\varepsilon^2}\,,$$

where $\mathcal{G}_t$ is defined by (3.4), and $\mathcal{S}^\varepsilon$ be the set of all measurable functions $h(t)$, adapted to the filtration $(\mathcal{G}_t^\varepsilon)$, for which $\sup_{t\leqslant T} \boldsymbol{E}[|h(t)|] < +\infty$, and where $T > 0$ is fixed. The p-lim and the pseudo-generator are defined as follows. Let h and h^δ in $\mathcal{S}^\varepsilon$ for all $\delta > 0$. We say that $h = \text{p-lim}_\delta\, h^\delta$ if

$$\sup_{t,\delta} \boldsymbol{E}|h^\delta(t)| < +\infty \quad \text{and} \quad \lim_{\delta\to 0} \boldsymbol{E}|h^\delta(t) - h(t)| = 0 \quad \forall t \geqslant 0\,.$$

We say that $h \in \mathcal{D}(\mathcal{A}^\varepsilon)$ the domain of $\mathcal{A}^\varepsilon$ and $\mathcal{A}^\varepsilon h = g$ if h and g are in $\mathcal{S}^\varepsilon$ and

$$\text{p-lim}_{\delta\to 0}\Big(\frac{\boldsymbol{E}_t^\varepsilon h(t+\delta) - h(t)}{\delta} - g(t)\Big) = 0\,,$$

where $\boldsymbol{E}_t^\varepsilon$ is the conditional expectation given $\mathcal{G}_t^\varepsilon$. The difference with respect to the usual infinitesimal generator is that we have *average* limit (via p-lim) rather than pointwise limit. This greatly helps us average out the noise in the limit $\varepsilon \to 0$. A useful result [Kus84, Theorem 1, p. 39] about the pseudo-generator $\mathcal{A}^\varepsilon$ that will be central in our proof is the following.

Proposition 7.1 (Theorem 1, p. 30 in [Kus84])**.** *Let $h \in \mathcal{D}(\mathcal{A}^\varepsilon)$. Then*

$$M_h^\varepsilon(t) = h(t) - h(0) - \int_0^t \mathcal{A}^\varepsilon h(u)\, du$$

is a $(\mathcal{G}_t^\varepsilon)$-martingale.

In other words, we introduce an operator $\mathcal{A}^\varepsilon$ on functions of Y^ε having similar properties to those of an infinitesimal operator of a Markov process; more specifically, it satisfies a martingale problem.

7.1.2. *Perturbed test function method.* The main idea behind the perturbed test function method is the following. To characterize the limiting process through a martingale problem with generator $\mathcal{A}$, we can then try to compare $\mathcal{A}^\varepsilon h$ with $\mathcal{A}h$ for some test function h. Unfortunately, $\mathcal{A}^\varepsilon h$ may have singular terms in ε, and we cannot proceed that way directly. Instead, try and extract the effective statistical behavior of the system under consideration from $\mathcal{A}^\varepsilon$ by introducing h^ε an appropriate

perturbation of h. In other words, if we can show $\lim_{\varepsilon\to 0} \boldsymbol{E}|h^\varepsilon(t) - h(t)| = 0$ and $\lim_{\varepsilon\to 0} \boldsymbol{E}|\mathcal{A}^\varepsilon h^\varepsilon(t) - \mathcal{A}h(Y_t^\varepsilon)| = 0$ for each t, then we can use Proposition 7.1 to conclude that the limiting process $(Y_t)_{t\geqslant 0}$ satisfies the martingale problem with generator $\mathcal{A}$. We begin by showing tightness of $(Y^\varepsilon)_{\varepsilon>0}$, which implies imply weak subsequential convergence of stochastic processes.

Before going into the proofs, we remind the reader that $(Y^\varepsilon)_{\varepsilon>0}$ is defined via a truncated process in equation (6.2) and there is an omitted index M in the functions F and G to indicate the truncation. This allows us to bound several expressions in the proofs in terms of M.

7.2. **Tightness.** In this section, we prove the tightness of the family $(Y^\varepsilon)_{\varepsilon>0}$, seen as a family of continuous-time processes. Then, according to [Bil99, Theorem 13.4] it suffices to prove the tightness of $(Y^\varepsilon)_{\varepsilon>0}$ in $\mathcal{D}([0,T],\mathbb{R}^2)$ (which is the set of càdlàg functions with values in $\mathbb{R}^2$ equipped with the Skorokhod topology).

Proposition 7.2. *The family $(Y^\varepsilon)_{\varepsilon>0}$ is tight in $\mathcal{D}([0,T],\mathbb{R}^2)$.*

The proof of Proposition 7.2 consists of applying [Kus84, Theorem 4, p. 48]. In what follows, let h be a bounded smooth function on $\mathbb{R}^2$ with bounded derivatives, and set $h_0^\varepsilon(t) = h(Y_t^\varepsilon)$. Our goal is to construct perturbations to h_0^ε as described above. The pseudo-generator at h_0^ε is then given by

$$\mathcal{A}^\varepsilon h_0^\varepsilon(t) = \frac{1}{\varepsilon}\int V\Big(\frac{t}{\varepsilon^2}, dp\Big) F\big(Y_t^\varepsilon, \tau_t^\varepsilon\big)\cdot\nabla h(Y_t^\varepsilon) \tag{7.1}$$

which is simply differentiating $h_0^\varepsilon(t)$ with respect to t (the derivative $\dot{Y}_t^\varepsilon$ is stated in equation (6.2)). The goal now is to modify the test function h_0^ε using a small perturbation h_1^ε so that the pseudo-generator $\mathcal{A}^\varepsilon(h_0^\varepsilon + h_1^\varepsilon)$ does not blow up in ε anymore. The first perturbation of h_0^ε is defined as follows.

$$\begin{aligned}
h_1^\varepsilon(t) &= \frac{1}{\varepsilon}\int_t^\infty du \int \boldsymbol{E}_t^\varepsilon\Big[V\Big(\frac{u}{\varepsilon^2}, dp\Big) F\Big(Y_t^\varepsilon, \tau_t^\varepsilon + \frac{u-t}{\varepsilon^2}\,\omega(Y_t^\varepsilon)\Big)\Big]\cdot\nabla h(Y_t^\varepsilon)\\
&= \varepsilon\int_0^\infty du \int \boldsymbol{E}_t^\varepsilon\Big[V\Big(u + \frac{t}{\varepsilon^2}, dp\Big) F\Big(Y_t^\varepsilon, \tau_t^\varepsilon + u\,\omega(Y_t^\varepsilon)\Big)\Big]\cdot\nabla h(Y_t^\varepsilon)\\
&= \varepsilon\int_0^\infty du \int \sum_{n\in\mathbb{Z}^*} \boldsymbol{E}_t^\varepsilon\Big[V\Big(u + \frac{t}{\varepsilon^2}, dp\Big)\Big] e^{2i\pi n(\tau_t^\varepsilon + u\,\omega(Y_t^\varepsilon))} F_n(Y_t^\varepsilon)\cdot\nabla h(Y_t^\varepsilon),
\end{aligned} \tag{7.2}$$

where

$$F_n(y) \stackrel{\text{def}}{=} \int_0^1 F(y,\tau) e^{-2i\pi n\tau}\, d\tau\,.$$

According to Lemma 3.4, we have

$$\boldsymbol{E}_t^\varepsilon V\Big(u + \frac{t}{\varepsilon^2}, dp\Big) = e^{-\mu|p|^{2\beta}u} V\Big(\frac{t}{\varepsilon^2}, dp\Big),$$

so that

$$\begin{aligned}
h_1^\varepsilon(t) &= \varepsilon\sum_{n\in\mathbb{Z}^*} e^{2i\pi n\tau_t^\varepsilon}\int V\Big(\frac{t}{\varepsilon^2}, dp\Big) F_n(Y_t^\varepsilon)\cdot\nabla h(Y_t^\varepsilon)\int_0^\infty e^{-\mu|p|^{2\beta}u} e^{2i\pi n\omega(Y_t^\varepsilon)u}\, du\\
&= \varepsilon\sum_{n\in\mathbb{Z}^*} e^{2i\pi n\tau_t^\varepsilon}\int \frac{V(t/\varepsilon^2, dp)}{\mu|p|^{2\beta} - 2i\pi n\,\omega(Y_t^\varepsilon)} F_n(Y_t^\varepsilon)\cdot\nabla h(Y_t^\varepsilon)
\end{aligned}$$

To apply [Kus84, Theorem 4 p. 48] and then prove Proposition 7.2, we only have to prove the two following lemmas.

Lemma 7.3. *For any $T > 0$, and $\eta > 0$*

$$\lim_{\varepsilon\to 0} \boldsymbol{P}\Big(\sup_{t\in[0,T]} |h_1^\varepsilon(t)| > \eta\Big) = 0 \qquad \textit{and} \qquad \lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \boldsymbol{E}|h_1^\varepsilon(t)| = 0\,.$$

Lemma 7.4. *For any $T > 0$, $\big\{\mathcal{A}^\varepsilon\big(h_0^\varepsilon + h_1^\varepsilon\big)(t), \varepsilon \in (0,1), 0 \leqslant t \leqslant T\big\}$ is uniformly integrable.*

Before going through the proofs of these two lemmas, we make one remark that the test function h_1^ε, above, is well defined since we sum over $n \neq 0$. This is because F has mean zero respect to the τ, and explains why this mean-zero condition is crucial to the analysis. For $n = 0$, the above function h_1^ε would not be defined in the case of long-range correlations.

Proof of Lemma 7.3. We rewrite h_1^ε as

$$h_1^\varepsilon(t) = \varepsilon V\Big(\frac{t}{\varepsilon^2}, \varphi_{t,\varepsilon}\Big),$$

with

$$\varphi_{t,\varepsilon}(p) \stackrel{\text{def}}{=} \sum_{n\in\mathbb{Z}^*} e^{2i\pi n\tau_t^\varepsilon} \frac{1}{\mu|p|^{2\beta} - 2i\pi n\,\omega(Y_t^\varepsilon)} F_n(Y_t^\varepsilon)\cdot\nabla h(Y_t^\varepsilon)\,. \tag{7.3}$$

By the Cauchy-Schwarz inequality

$$\begin{aligned}
|\varphi_{t,\varepsilon}(p)| &\leqslant C_{h,\omega_0} \sum_{n\in\mathbb{Z}^*} \frac{1}{|n|} \|F_n(Y_t^\varepsilon)\| \\
&\leqslant C_{h,\omega_0} \Big(\sum_{n\in\mathbb{Z}^*} \frac{1}{|n|^2}\Big)^{1/2} \Big(\sum_{n\in\mathbb{Z}^*} \|F_n(Y_t^\varepsilon)\|^2\Big)^{1/2} \\
&\leqslant C_{h,\omega_0} \Big(\sum_{n\in\mathbb{Z}^*} \frac{1}{|n|^2}\Big)^{1/2} \Big(\int_0^1 \|F(Y_t^\varepsilon,\tau)\|^2\,d\tau\Big)^{1/2} \\
&\leqslant C_{h,\omega_0} \sup_{|y|\leqslant 2M} \Big(\int_0^1 \|F(y,\tau)\|^2\,d\tau\Big)^{1/2}.
\end{aligned}$$

If $\beta \geqslant 1/2$, we have the Lipschitz bound

$$\begin{aligned}
|\varphi_{t,\varepsilon}(p) - \varphi_{t,\varepsilon}(q)| &\leqslant (|p|^{2\beta} - |q|^{2\beta}) C_{h,\omega_0,F} \sum_{n\in\mathbb{Z}^*} \frac{1}{|n|^2} \\
&\leqslant ||p| - |q|| \sup_{r\in S} |r|^{2\beta-1} C_{h,\omega_0,F} \sum_{n\in\mathbb{Z}^*} \frac{1}{|n|^2} \leqslant L|p-q|
\end{aligned}$$

with

$$L = \sup_{r\in S} |r|^{2\beta-1} C_{h,\omega_0,F} \sum_{n\in\mathbb{Z}^*} \frac{1}{|n|^2} < \infty\,.$$

Using inequality (3.7) of Lemma 3.5 with $k = \infty$, along with the Chebychev's inequality, we have

$$\boldsymbol{P}\Big(\sup_{t\in[0,T]} |h_1^\varepsilon(t)| > \eta\Big) \leqslant \frac{1}{\eta} \boldsymbol{E} \sup_{t\in[0,T]} |h_1^\varepsilon(t)| \leqslant C'_{h,\omega_0,F}\, C(\varepsilon)\,,$$

with $\lim_\varepsilon C(\varepsilon) = 0$.

If $\beta < 1/2$, the function $\varphi_{t,\varepsilon}$ belongs to $W^{1,k}(S)$ with $k \in (1, 1/(1-2\beta))$ since

$$\int_S |\partial_p \varphi_{t,\varepsilon}(p)|^k \, dp \leqslant C''_{h,\omega_0,F} \int_S |p|^{k(2\beta-1)} \, dr < \infty \,,$$

so that using inequality (3.7) of Lemma 3.5 but with $k \in (1, 1/(1-2\beta))$ leads again to

$$\boldsymbol{P}\Big(\sup_{t\in[0,T]} |h_1^\varepsilon(t)| > \eta \Big) \leqslant C'_{h,\omega_0,F}\, C(\varepsilon) \,.$$

To prove the second limit in the lemma, one can just remark that

$$\boldsymbol{E}|h_1^\varepsilon(t)| \leqslant \boldsymbol{E} \sup_{t\in[0,T]} |h_1^\varepsilon(t)|$$

for any $t \in [0,T]$, and then concludes the proof of Lemma 7.3. □

Proof of Lemma 7.4. For proving uniform integrability, it is sufficient to show that

$$\sup_{\varepsilon,t} \boldsymbol{E}|\mathcal{A}^\varepsilon(h_0^\varepsilon + h_1^\varepsilon)(t)|^2 \leqslant C \,.$$

Following the definition of pseudo-generator in Section 7.1.1, we have that

$$\begin{aligned}
\mathcal{A}^\varepsilon(h_1^\varepsilon)(t) &= -\mathcal{A}^\varepsilon(h_0^\varepsilon)(t) \\
&\quad + \frac{1}{\varepsilon} \int_t^\infty du \int \boldsymbol{E}_t^\varepsilon \Big[V\Big(\frac{u}{\varepsilon^2}, dp\Big) \Big] \frac{d}{ds} \Big[F\Big(Y_s^\varepsilon, \tau_s^\varepsilon + \frac{u-s}{\varepsilon^2}\, \omega(Y_s^\varepsilon) \Big) \cdot \nabla h(Y_s^\varepsilon) \Big] \Big|_{s=t}
\end{aligned}$$

which is essentially differentiating $h_1^\varepsilon(t)$ with respect to t (we however remark that we do not differentiate the t in $\boldsymbol{E}_t^\varepsilon$). After some lengthy but straightforward computations (similarly to those used to obtain (7.2)), we obtain

$$\begin{aligned}
\mathcal{A}^\varepsilon(h_1^\varepsilon)(t) &= -\mathcal{A}^\varepsilon(h_0^\varepsilon)(t) + \sum_{n,m\in\mathbb{Z}^*} e^{2i\pi(n+m)\tau_t^\varepsilon} \iint \frac{V(t/\varepsilon^2, dp) V(t/\varepsilon^2, dp')}{\mu|p|^{2\beta} - 2i\pi n\omega(Y_t^\varepsilon)} \\
&\qquad \cdot \Big(F_n(Y_t^\varepsilon)^T \nabla^2 h(Y_t^\varepsilon) F_m(Y_t^\varepsilon) + \nabla h(Y_t^\varepsilon)^T \mathrm{Jac} F_n(Y_t^\varepsilon) F_m(Y_t^\varepsilon) \\
&\qquad\quad - \frac{2i\pi n}{\mu|p|^{2\beta} - 2i\pi n\omega(Y_t^\varepsilon)} F_n(Y_t^\varepsilon) \cdot \nabla h(Y_t^\varepsilon) F_m(Y_t^\varepsilon) \cdot \nabla\omega(Y_t^\varepsilon) \Big) \\
&\quad + \sum_{n\in\mathbb{Z}^*} 2i\pi n\, e^{2i\pi n\tau_t^\varepsilon} \iint \frac{V(t/\varepsilon^2, dp) V(t/\varepsilon^2, dp')}{\mu|p|^{2\beta} - 2i\pi n\omega(Y_t^\varepsilon)} \\
&\qquad \cdot F_n(Y_t^\varepsilon) \cdot \nabla h(Y_t^\varepsilon) G(Y_t^\varepsilon) \\
&= -\mathcal{A}^\varepsilon(h_0^\varepsilon)(t) + V\Big(\frac{t}{\varepsilon^2}, \varphi_{t,\varepsilon}^1\Big) V\Big(\frac{t}{\varepsilon^2}, \mathbf{1}_S\Big) + V\Big(\frac{t}{\varepsilon^2}, \varphi_{t,\varepsilon}^2\Big) V\Big(\frac{t}{\varepsilon^2}, \mathbf{1}_S\Big) \\
&\stackrel{\text{def}}{=} -\mathcal{A}^\varepsilon(h_0^\varepsilon)(t) + \mathcal{P}_1^\varepsilon(t) + \mathcal{P}_2^\varepsilon(t) \qquad (7.4)
\end{aligned}$$

where $\varphi_{t,\varepsilon}^j$, $j = 1,2$ are defined similarly as in (7.3). As in the proof of Lemma 7.3, it is not hard to see that the $\varphi_{t,\varepsilon}^j$, $j = 1,2$ are Lispschitz in p on S, and that

$$\begin{aligned}
|\varphi_{t,\varepsilon}^1(p)| &\leqslant C_{h,\omega_0,\nabla\omega} \sum_{m\in\mathbb{Z}^*} |F_m(Y_t^\varepsilon)| \sum_{n\in\mathbb{Z}^*} \frac{1}{|n|} \Big((1+|n|)|F_n(Y_t^\varepsilon)| + \|\mathrm{Jac} F_n(Y_t^\varepsilon)\| \Big) \\
&\leqslant \tilde{C}_{h,\omega_0,\nabla\omega} \sup_{|y|\leqslant 2M} \int_0^1 \Big(|F(y,\tau)|^2 + |\partial_\tau F(y,\tau)|^2 + \|\mathrm{Jac} F(y,\tau)\|^2 \Big) \, d\tau \,.
\end{aligned}$$

The same lines give also

$$|\varphi^2_{t,\varepsilon}(p)| \leqslant C_{h,\omega_0} \sup_{|y|\leqslant 2M} \big|G(y)\big| \Big(\int_0^1 |\partial_\tau F(y,\tau)|^2 \, d\tau \Big)^{1/2}.$$

Then, we can use inequality (3.8) of Lemma 3.5, and obtain

$$\sup_{\varepsilon,t} \boldsymbol{E}[|\mathcal{P}^\varepsilon_j(t)|^2] \leqslant C, \quad j = 1, 2 \ , .$$

which concludes the proof of Lemma 7.4, and then Proposition 7.2 as well. □

7.3. **Identification of the limit.** In this section, we identify all the limit points of $(Y^\varepsilon)_{\varepsilon>0}$ via a well-posed martingale problem as stated in the following proposition. By abuse of notations, we still denote by $(Y^\varepsilon)_{\varepsilon>0}$ a converging subsequence and by $(Y_t)_{t\geqslant 0}$ a limit point.

Proposition 7.5. *All the limit points $(Y_t)_{t\geqslant 0}$ of $(Y^\varepsilon)_{\varepsilon>0}$ are solutions of a well-posed martingale problem with generator*

$$\mathcal{L}h(y) = \int_0^1 d\tau \int_0^\infty du\, R(u) \tilde{F}(0,y,\tau) \cdot \nabla_y (\tilde{F}(u,y,\tau) \cdot \nabla_y h(y)) \tag{7.5}$$

with

$$\tilde{F}(u,y,\tau) = F(y, \tau + \omega(y)u) \, .$$

To prove this proposition we use the notion of pseudo-generator introduced in Section 7.1.1 and the perturbed-test-function technique that we have already used in Section 7.2 for the proof of tightness. Thanks to the pseudo-generator we can characterize the subsequential limits of $(Y^\varepsilon)_{\varepsilon>0}$ as solutions of a well-posed martingale problem.

The outline of the proof is as follows. Recall in Section 7.2 we saw that $\mathcal{A}^\varepsilon(h^\varepsilon_0)$ has a singular $1/\varepsilon$ term, and modified the test function h^ε_0 with a small perturbation h^ε_1 to remove this singular term. As a result, $\mathcal{A}^\varepsilon(h^\varepsilon_0 + h^\varepsilon_1)$ is not singular any more. However, it is not yet the generator of a martingale problem as $\mathcal{A}^\varepsilon(h^\varepsilon_0 + h^\varepsilon_1)$ still has oscillatory terms in the form of $e^{2i\pi n \tau^\varepsilon_t}$. We will show that these terms essentially vanish as $\varepsilon \to 0$ by introducing a small perturbation $h^\varepsilon_2(t)$ to cancel these oscillations, and then construct another perturbation $h^\varepsilon_3(t)$ to cancel oscillations that result from computing $\mathcal{A}^\varepsilon(h^\varepsilon_2)(t)$. After we have constructed the perturbed test function

$$h^\varepsilon(t) = h^\varepsilon_0(t) + h^\varepsilon_1(t) + h^\varepsilon_2(t) + h^\varepsilon_3(t) \, ,$$

we show that the pseudo-generator

$$\mathcal{A}^\varepsilon(h^\varepsilon_0 + h^\varepsilon_1 + h^\varepsilon_2 + h^\varepsilon_3)(t)$$

has the desired form of a generator related to a martingale problem, plus a negligible term. Combining with tightness, we can show that the limiting process $(Y_t)_{t\geqslant 0}$ satisfies a martingale problem.

We now carry out this approach. For convenience, the proofs of needed auxiliary lemmas will be deferred to Sections 7.4-7.8.

Proof of Proposition 7.5. We split $\mathcal{A}^\varepsilon(h^\varepsilon_0 + h^\varepsilon_1)$ into three parts

$$\mathcal{A}^\varepsilon(h^\varepsilon_0 + h^\varepsilon_1)(t) = \mathcal{A}^\varepsilon_1(t) + \mathcal{A}^\varepsilon_2(t) + \mathcal{P}^\varepsilon_2(t) \, ,$$

where $\mathcal{P}_2^\varepsilon$ is defined in (7.4), and $\mathcal{A}_1^\varepsilon(t)$, $\mathcal{A}_2^\varepsilon(t)$ are given by

$$\mathcal{A}_1^\varepsilon(t) = \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} e^{2i\pi(n+m)\tau_t^\varepsilon} \iint V\Big(\frac{t}{\varepsilon^2}, dp\Big) V\Big(\frac{t}{\varepsilon^2}, dp'\Big) W_{n,m}(p, Y_t^\varepsilon)\,,$$

$$\mathcal{A}_2^\varepsilon(t) = \sum_{n\in\mathbb{Z}^*} \iint V\Big(\frac{t}{\varepsilon^2}, dp\Big) V\Big(\frac{t}{\varepsilon^2}, dp'\Big) W_n(p, Y_t^\varepsilon)\,. \tag{7.6}$$

Here

$$\begin{aligned} W_{n,m}(p, Y_t^\varepsilon) \overset{\text{def}}{=} \frac{1}{\mu|p|^{2\beta} - 2i\pi n\omega(Y_t^\varepsilon)} \Big(& F_n(Y_t^\varepsilon)^T \nabla^2 h(Y_t^\varepsilon) F_m(Y_t^\varepsilon) \\ & + \nabla h(Y_t^\varepsilon)^T \mathrm{Jac} F_n(Y_t^\varepsilon) F_m(Y_t^\varepsilon) \Big) \\ & - \frac{2i\pi n}{(\mu|p|^{2\beta} - 2i\pi n\omega(Y_t^\varepsilon))^2} F_n(Y_t^\varepsilon)\cdot\nabla h(Y_t^\varepsilon) F_m(Y_t^\varepsilon)\cdot\nabla\omega(Y_t^\varepsilon)\,, \end{aligned} \tag{7.7}$$

and

$$W_n(p, Y_t^\varepsilon) = W_{n,-n}(p, Y_t^\varepsilon)\,.$$

The term $\mathcal{A}_2^\varepsilon$ no longer has oscillatory components as $\mathcal{A}_2^\varepsilon$ only picks up the terms with $n + m = 0$. We now show that this term is related to the generator of a martingale problem.

Lemma 7.6. *For all $t \geqslant 0$, we have*

$$\lim_{\varepsilon\to 0} \boldsymbol{E}\Big| \int_0^t (\mathcal{A}_2^\varepsilon(u) - \mathcal{L}h(Y_u^\varepsilon))\, du\Big| = 0\,.$$

with

$$\mathcal{L}h(y) = \sum_{n\in\mathbb{Z}^*} \int dp\, r(p) W_n(p, y)\,,$$

where W_n is defined through (7.7).

We present the proof of Lemma 7.6 in Section 7.7. To show the well-posedness of the limiting martingale problem, we need to rewrite $\mathcal{L}$ as a second-order differential operator, which eventually gives us the form in (7.5). We do this in Section 7.8, below.

For the oscillatory term $\mathcal{A}_1^\varepsilon$, we introduce two more small perturbations h_2^ε and h_3^ε to prove that the oscillations will vanish as $\varepsilon \to 0$. Note that $\mathcal{P}_2^\varepsilon$ also has oscillatory terms. These can be treated in a manner similar to the oscillatory terms in $\mathcal{A}_1^\varepsilon$, so that these terms vanish in the limit $\varepsilon \to 0$. Since the treatment of this will be similar to the treatment of $\mathcal{A}_1^\varepsilon$, we omit the details.

Similarly to the construction of h_1^ε in equation (7.2), we introduce the second perturbation

$$\begin{aligned} h_2^\varepsilon(t) = \int_t^\infty du \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} & e^{2i\pi(n+m)(\tau_t^\varepsilon + (u-t)\omega(Y_t^\varepsilon)/\varepsilon^2)} \\ \cdot \iint \Big(\boldsymbol{E}_t^\varepsilon \Big[V\Big(\frac{u}{\varepsilon^2}, dp\Big) V\Big(\frac{u}{\varepsilon^2}, dp'\Big)\Big] & - \boldsymbol{E}[V(0, dp) V(0, dp')] \Big) W_{n,m}(p, Y_t^\varepsilon)\,. \end{aligned}$$

Making the change of variable $u \to t + \varepsilon^2 u$, we have

$$h_2^\varepsilon(t) = \varepsilon^2 \int_0^\infty du \sum_{\substack{n,m\in\mathbb{Z}^* \\ n+m\neq 0}} e^{2i\pi(n+m)(\tau_t^\varepsilon + u\omega(Y_t^\varepsilon))} W_{n,m}(p, Y_t^\varepsilon)$$
$$\cdot \iint \left(\boldsymbol{E}_t^\varepsilon \left[V\Big(u + \frac{t}{\varepsilon^2}, dp\Big) V\Big(u + \frac{t}{\varepsilon^2}, dp'\Big) \right] - \boldsymbol{E}[V(0,dp)V(0,dp')] \right).$$

According to formula (3.6), one has

$$\boldsymbol{E}_t^\varepsilon \left[V\Big(u + \frac{t}{\varepsilon^2}, dp\Big) V\Big(u + \frac{t}{\varepsilon^2}, dp'\Big) \right] - \boldsymbol{E}[V(0,dp)V(0,dp')]$$
$$= e^{-\mu(|p|^{2\beta} + |p'|^{2\beta})u} \Big(V\Big(\frac{t}{\varepsilon^2}, dp\Big) V\Big(\frac{t}{\varepsilon^2}, dp'\Big) - r(p)\delta(p-p')\, dp\, dp' \Big),$$

so that

$$h_2^\varepsilon(t) = \varepsilon^2 \sum_{\substack{n,m\in\mathbb{Z}^* \\ n+m\neq 0}} e^{2i\pi(n+m)\tau_t^\varepsilon}$$
$$\times \iint \frac{V(t/\varepsilon^2, dp)V(t/\varepsilon^2, dp') - r(p)\delta(p-p')\, dp\, dp'}{\mu(|p|^{2\beta} + |p'|^{2\beta}) - 2i\pi(n+m)\omega(Y_t^\varepsilon)} W_{n,m}(p, Y_t^\varepsilon).$$

We claim that h_2^ε is uniformly small.

Lemma 7.7. *We have*

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \boldsymbol{E}|h_2^\varepsilon(t)| = 0.$$

For clarity of presentation we postpone the proof lf Lemma 7.7 to Section 7.4. As before, following the definition of pseudo-generator in Section 7.1.1, we have

$$\mathcal{A}^\varepsilon(h_2^\varepsilon)(t) = -\mathcal{A}_1^\varepsilon(t) + \mathcal{D}_\varepsilon(t) + \mathcal{R}_\varepsilon^1(t),$$

with

$$\mathcal{D}_\varepsilon(t) \overset{\text{def}}{=} \sum_{\substack{n,m\in\mathbb{Z}^* \\ n+m\neq 0}} e^{2i\pi(n+m)\tau_t^\varepsilon} \int dp\, r(p) W_{n,m}(p, Y_t^\varepsilon),$$

and $\mathcal{R}_\varepsilon^1(t)$ to be the remainder (an explicit formula for $\mathcal{R}_\varepsilon^1(t)$ is presented in to Section 7.5, below). Even though oscillatory terms are present in $\mathcal{R}_\varepsilon^1$, we make the following claim.

Lemma 7.8. *We have*

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \boldsymbol{E}|\mathcal{R}_\varepsilon^1(t)| = 0.$$

The proof of this lemma is presented in Section 7.5, below. To treat the oscillations in $\mathcal{D}_\varepsilon(t)$ we introduce the third perturbation

$$h_3^\varepsilon(t) = \int_t^\infty du \sum_{\substack{n,m\in\mathbb{Z}^* \\ n+m\neq 0}} e^{2i\pi(n+m)(\tau_t^\varepsilon + (u-t)\omega(Y_t^\varepsilon)/\varepsilon^2)} e^{-\varepsilon(u-t)} \int dp\, r(p) W_{n,m}(p, Y_t^\varepsilon)$$
$$= \varepsilon^2 \sum_{\substack{n,m\in\mathbb{Z}^* \\ n+m\neq 0}} \frac{e^{2i\pi(n+m)\tau_t^\varepsilon}}{\varepsilon^3 - 2i\pi(n+m)\omega(Y_t^\varepsilon)} \int dp\, r(p) W_{n,m}(p, Y_t^\varepsilon)$$

We will see that

$$\mathcal{A}^\varepsilon(h_3^\varepsilon)(t) = -\mathcal{D}_\varepsilon(t) + \mathcal{R}_\varepsilon^2(t)$$

with explicit formula for $\mathcal{R}_\varepsilon^2(t)$ presented in Section 7.6, below. We claim h_3^ε and R_ε^2 both vanish as $\varepsilon \to 0$.

Lemma 7.9. *We have*

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \boldsymbol{E}|h_3^\varepsilon(t)| = 0 \qquad \textit{and} \qquad \lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \boldsymbol{E}|\mathcal{R}_\varepsilon^2(t)| = 0\,.$$

We will prove Lemma 7.9 in Section 7.6, below. As a result, we now have a perturbed test function

$$h(t) = h_0^\varepsilon(t) + h_1^\varepsilon(t) + h_2^\varepsilon(t) + h_3^\varepsilon(t)\,,$$

with $h_0^\varepsilon(t) = h(Y_t^\varepsilon)$, for which

$$\mathcal{A}^\varepsilon(h_0^\varepsilon + h_1^\varepsilon + h_2^\varepsilon + h_3^\varepsilon)(t) = \mathcal{A}_2^\varepsilon(t) + o(\varepsilon)\,. \tag{7.8}$$

Using Theorem 7.1 we see that for all $N \geqslant 1$, any bounded continuous function Ψ, and every sequence $0 < s_1 < \cdots < s_N \leqslant s < t$ we must have

$$\boldsymbol{E}\Big[\Psi\big(Y_{s_1}^\varepsilon, \ldots, Y_{s_N}^\varepsilon\big)\Big(M_h^\varepsilon(t) - M_h^\varepsilon(s)\Big)\Big] = 0\,. \tag{7.9}$$

Together with Lemmas 7.3, 7.6, 7.7, 7.8, 7.9, and estimate (7.8), we have

$$\lim_{\varepsilon\to 0} \boldsymbol{E}\Big[\Psi\big(Y_{s_1}^\varepsilon, \ldots, Y_{s_N}^\varepsilon\big)\Big(h(Y_t^\varepsilon) - h(Y_s^\varepsilon) - \int_s^t \mathcal{L}h(u)\,du\Big)\Big] = 0\,, \tag{7.10}$$

where $\mathcal{L}$ is defined in (7.5). Combining with $Y_{t=0}^\varepsilon = Y_0$ for every $\varepsilon > 0$, the equation (7.10) implies that any subsequential limit $(Y_t)_{t\geqslant 0}$ of $(Y_\varepsilon)_{\varepsilon>0}$ is a solution of the martingale problem for the generator $\mathcal{L}$ with initial condition $Y_{t=0} = Y_0$. We claim that the solution to this martingale problem is unique (we prove this in Section 7.8, below). Given this, along with the tightness proven in Proposition 7.2, Prokhorov's theorem implies the weak convergence of the sequence of processes $(Y^\varepsilon)_{\varepsilon>0}$ to the process $(Y_t)_{t\geqslant 0}$ as desired. This completes the proof of Proposition 6.1. □

7.4. **Proof of Lemma 7.7.** We now prove Lemma 7.7. By using Lemma 3.5, the proof is similar that of Lemma 7.4.

Proof of Lemma 7.7. We split the integrand in h_2^ε into two parts as follows

$$\begin{aligned} h_2^\varepsilon(t) &= \varepsilon^2 \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} e^{2i\pi(n+m)\tau_t^\varepsilon} \\ &\qquad \cdot \iint \frac{V(t/\varepsilon^2, dp)V(t/\varepsilon^2, dp') - r(p)\delta(p-p')\,dp\,dp'}{\mu(|p|^{2\beta} + |p'|^{2\beta}) - 2i\pi(n+m)\omega(Y_t^\varepsilon)} W_{n,m}(p, Y_t^\varepsilon) \\ &\stackrel{\text{def}}{=} h_{21}^\varepsilon(t) - h_{22}^\varepsilon(t) \end{aligned}$$

where h_{21}^ε contains the term $V(t/\varepsilon^2, dp)V(t/\varepsilon^2, dp')$ and h_{22}^ε contains the term $r(p)\delta(p-p')$.

For h_{22}^ε, it is straightforward to see that

$$|h_{22}^\varepsilon(t)| \leqslant \varepsilon^2 C_{h,\omega_0,\nabla\omega} \int dp\, r(p)$$

$$\cdot \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} \frac{1}{|n(n+m)|}|F_m(Y_t^\varepsilon)|\Big(\big(1+|n|\big)|F_n(Y_t^\varepsilon)| + \|\mathrm{Jac}F_n(Y_t^\varepsilon)\|\Big)$$

$$\leqslant \varepsilon^2 C_{h,\omega_0,\nabla\omega} \int dp\, r(p)$$

$$\cdot \sum_{n\in\mathbb{Z}^*} \frac{1}{|n|^2} \sup_{|y|\leqslant 2M} \int_0^1 \Big(|F(y,\tau)|^2 + |\partial_\tau F(y,\tau)|^2 + \|\mathrm{Jac}F(y,\tau)\|^2\Big) d\tau\,. \tag{7.11}$$

We rewrite

$$h_{21}^\varepsilon(t) = \varepsilon^2 V\Big(\frac{t}{\varepsilon^2}, \varphi_{1,t,\varepsilon}\Big),$$

with

$$\varphi_{1,t,\varepsilon}(p) = V\Big(\frac{t}{\varepsilon^2}, \varphi_{2,t,p,\varepsilon}\Big)$$

and

$$\varphi_{2,t,p,\varepsilon}(p') = \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} e^{2i\pi(n+m)\tau_t^\varepsilon} W_{n,m}(p, Y_t^\varepsilon)$$

$$\cdot \frac{1}{\mu(|p|^{2\beta}+|p'|^{2\beta}) - 2i\pi(n+m)\omega(Y_t^\varepsilon)}\,.$$

For readers' convenience, we pause to recall the meaning of the above notation. The function $\varphi_{1,t,\varepsilon}$ of the variable p is obtained by applying the measure $V(t/\varepsilon^2, dp')$ to the function $\varphi_{2,t,p,\varepsilon}(p')$, where $\varphi_{2,t,p,\varepsilon}$ (a function of the variable p'), is defined as above. Finally, $h_{21}^\varepsilon(t)$ is obtained by applying the measure $V(t/\varepsilon^2, dp)$ to the function $\varphi_{1,t,\varepsilon}(p)$.

For h_{22}^ε, one has

$$|\varphi_{2,t,p,\varepsilon}(p')| \leqslant C\,,$$

uniformly in p and p'. Following the same lines as for the first perturbation of the test function h_1^ε but using inequality (3.8) of Lemma 3.5 (with $k=\infty$ if $\beta \geqslant 1/2$ and $k \in (1, 1/(1-2\beta))$ if $\beta < 1/2$), we have for some $C'>0$

$$\sup_{p\in S} |\varphi_{1,t,\varepsilon}(p)| \leqslant \sup_{\varphi\in W_{k,C'}} \Big|V\Big(\frac{t}{\varepsilon^2}, \varphi\Big)\Big|\,.$$

In fact, the derivative of $\varphi_{2,t,p,\varepsilon}$ in p' can be uniformly bounded in p in the corresponding $L^k(S)$ space.

Now, let us look at the derivative of $\varphi_{1,t,\varepsilon}(p)$ in p. One can write

$$\partial_p \varphi_{1,t,\varepsilon}(p) = |p|^{2\beta-1} V\Big(\frac{t}{\varepsilon^2}, \varphi'_{2,t,p,\varepsilon}\Big) + V\Big(\frac{t}{\varepsilon^2}, \varphi''_{2,t,p,\varepsilon}\Big),$$

where $\varphi'_{2,t,p,\varepsilon}$ and $\varphi''_{2,t,p,\varepsilon}$ belong to some $W_{k,C''}$, and where C'' does not depend on p. Then,

$$|\partial_p \varphi_{1,t,\varepsilon}(p)| \leqslant (|p|^{2\beta-1}+1) \sup_{\varphi\in W_{k,C''}} \Big|V\Big(\frac{t}{\varepsilon^2}, \varphi\Big)\Big|\,.$$

As a result, considering $\tilde{C} = \max(C', C'')$, and

$$\tilde{h}_{21}^\varepsilon(t) \overset{\text{def}}{=} \frac{h_{21}^\varepsilon(t)}{\sup_{\varphi\in W_{k,\tilde{C}}} |V(t/\varepsilon^2, \varphi)|} = \varepsilon^2 V\Big(\frac{t}{\varepsilon^2}, \tilde{\varphi}_{1,t,\varepsilon}\Big)\,,$$

with

$$\tilde{\varphi}_{1,t,\varepsilon} = \frac{\varphi_{1,t,\varepsilon}}{\sup_{\varphi\in W_{k,\tilde{C}}} |V(t/\varepsilon^2,\varphi)|},$$

we have

$$\|\tilde{\varphi}_{1,t,\varepsilon}\|_{W^{1,k}(S)} \leqslant 1.$$

Therefore,

$$|\tilde{h}^{\varepsilon}_{21}(t)| \leqslant \varepsilon^2 \sup_{\varphi\in W_{k,\tilde{C}}} \Big|V\Big(\frac{t}{\varepsilon^2},\varphi\Big)\Big|,$$

and hence

$$\boldsymbol{E}|h^{\varepsilon}_{21}(t)| \leqslant \varepsilon^2 \boldsymbol{E} \sup_{\varphi\in W_{k,\max(1,\tilde{C})}} \Big|V\Big(\frac{t}{\varepsilon^2},\varphi\Big)\Big|^2 = \varepsilon^2 \boldsymbol{E} \sup_{\varphi\in W_{k,\max(1,\tilde{C})}} |V(0,\varphi)|^2 .$$

Using Lemma 3.5, this concludes the proof. □

7.5. **Proof of Lemma 7.8.** We devote this section to proving Lemma 7.8. By following the computations in (7.4), we obtain the following explicit formula for $\mathcal{R}^{\varepsilon}_1(t)$:

$$\begin{aligned}
\mathcal{R}^{\varepsilon}_1(t) = \varepsilon v\Big(\frac{t}{\varepsilon^2}\Big) \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} \iint \frac{2i\pi(n+m)e^{2i\pi(n+m)\tau^{\varepsilon}_t}}{\mu(|p|^{2\beta}+|p'|^{2\beta}) - 2i\pi(n+m)\omega(Y^{\varepsilon}_t)} \\
\cdot \Big(G(Y^{\varepsilon}_t) - \frac{\nabla\omega(Y^{\varepsilon}_t)\cdot F(Y^{\varepsilon}_t,\tau^{\varepsilon}_t)}{\mu(|p|^{2\beta}+|p'|^{2\beta}) - 2i\pi(n+m)\omega(Y^{\varepsilon}_t)}\Big) \\
\cdot \Big(V\Big(\frac{t}{\varepsilon^2},dp\Big)V\Big(\frac{t}{\varepsilon^2},dp'\Big) - r(p)\delta(p-p')\,dp\,dp'\Big)W_{n,m}(p,Y^{\varepsilon}_t) \\
+ \varepsilon v\Big(\frac{t}{\varepsilon^2}\Big) \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} \iint \frac{e^{2i\pi(n+m)\tau^{\varepsilon}_t}}{\mu(|p|^{2\beta}+|p'|^{2\beta}) - 2i\pi(n+m)\omega(Y^{\varepsilon}_t)} \\
\cdot \Big(V\Big(\frac{t}{\varepsilon^2},dp\Big)V\Big(\frac{t}{\varepsilon^2},dp'\Big) - r(p)\delta(p-p')\,dp\,dp'\Big) \\
\cdot \nabla_y W_{n,m}(p,y)_{|y=Y^{\varepsilon}_t}\cdot F(Y^{\varepsilon}_t,\tau^{\varepsilon}_t) .
\end{aligned}$$

From here, we can see that $\mathcal{R}^{\varepsilon}_1(t)$ can be treated in the same manner as we treated $h^{\varepsilon}_2(t)$ in Section 7.4. The completes the proof.

7.6. **Proof of Lemma 7.9.** We devote this section to proving Lemma 7.9. The bounds for $h^{\varepsilon}_3(t)$ can be obtained in the same manner as the bounds for h^{ε}_{22} in equation (7.11). As a result, we obtain

$$\lim_{\varepsilon\to 0} \sup_{t\in[0,T]} \boldsymbol{E}|h^{\varepsilon}_3(t)| = 0 .$$

Treating $R^{\varepsilon}_2(t)$ in the same manner as $R^{\varepsilon}_1(t)$ we obtain

$$\begin{aligned}
\mathcal{R}^{\varepsilon}_2(t) = \varepsilon v\Big(\frac{t}{\varepsilon^2}\Big) \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} \int dp\, \frac{2i\pi(n+m)r(p)e^{2i\pi(n+m)\tau^{\varepsilon}_t}}{\varepsilon^3 - 2i\pi(n+m)\omega(Y^{\varepsilon}_t)} \\
\cdot \Big(G(Y^{\varepsilon}_t) - \frac{\nabla\omega(Y^{\varepsilon}_t)\cdot F(Y^{\varepsilon}_t,\tau^{\varepsilon}_t)}{\varepsilon^3 - 2i\pi(n+m)\omega(Y^{\varepsilon}_t)}\Big)W_{n,m}(p,Y^{\varepsilon}_t)
\end{aligned}$$

$$+\varepsilon^3 \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} \int dp\, \frac{r(p)e^{2i\pi(n+m)\tau_t^\varepsilon}}{\varepsilon^3 - 2i\pi(n+m)\omega(Y_t^\varepsilon)} W_{n,m}(p, Y_t^\varepsilon)$$

$$+\varepsilon v\Big(\frac{t}{\varepsilon^2}\Big) \sum_{\substack{n,m\in\mathbb{Z}^*\\ n+m\neq 0}} \int dp\, \frac{r(p)e^{2i\pi(n+m)\tau_t^\varepsilon}}{\varepsilon^3 - 2i\pi(n+m)\omega(Y_t^\varepsilon)} \nabla_y W_{n,m}(p, y)_{|y=Y_t^\varepsilon} \cdot F(Y_t^\varepsilon, \tau_t^\varepsilon)\,.$$

From here, we can see that each of the terms in $\mathcal{R}_2^\varepsilon(t)$ can be treated in the same manner as $h_2^\varepsilon(t)$ in Section 7.4. This concludes the proof.

7.7. **Proof of Lemma 7.6.** We devote this section to proving Lemma 7.6. The proof analyzes how $V(t/\varepsilon^2, dp)V(t/\varepsilon^2, dp')$ in $\mathcal{A}_2^\varepsilon$ (defined through Formula (7.6)) averages in the limit $\varepsilon \to 0$. This is essentially similar to the analysis of the covariance in the proof of Proposition 1.1 in Section 2 in the case when the Hamiltonian was quadratic.

Let $\eta > 0$ and write

$$\begin{aligned}
\int_0^t (\mathcal{A}_2^\varepsilon(u) - \mathcal{L}^M h(Y_u^\varepsilon))\, du &= \Big(\int_0^{[t/(\varepsilon\eta)]\varepsilon\eta} + \int_{[t/(\varepsilon\eta)]\varepsilon\eta}^t \Big)(\mathcal{A}_2^\varepsilon(u) - \mathcal{L}h(Y_u^\varepsilon))\, du \\
&= \sum_{q=0}^{[t/(\varepsilon\eta)]-1} \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} (\mathcal{A}_2^\varepsilon(u) - \mathcal{L}h(Y_u^\varepsilon))\, du \\
&\quad + \int_{[t/(\varepsilon\eta)]\varepsilon\eta}^t (\mathcal{A}_2^\varepsilon(u) - \mathcal{L}h(Y_u^\varepsilon))\, du \\
&\stackrel{\text{def}}{=} R_1^\varepsilon(t) + R_2^\varepsilon(t)\,.
\end{aligned}$$

Following the computations of Lemma 7.7, we have

$$\boldsymbol{E}|R_2^\varepsilon(t)| \leqslant \varepsilon \tilde{C}\,.$$

so that we only have to take care of $R_1^\varepsilon(t)$. For this term we consider the decomposition

$$
\begin{aligned}
R_1^\varepsilon(t) &= \sum_{q=0}^{[t/(\varepsilon\eta)]-1} \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} (\mathcal{A}_2^\varepsilon(u) - \mathcal{L}h(Y_u^\varepsilon))\, du \\
&= \sum_{q=0}^{[t/(\varepsilon\eta)]-1} \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du \\
&\qquad \cdot \sum_{n\in\mathbb{Z}^*} \iint \Big(V\Big(\frac{u}{\varepsilon^2}, dp\Big) V\Big(\frac{u}{\varepsilon^2}, dp'\Big) - r(p)\delta(p-p')\, dp\, dp' \Big) \\
&\qquad \cdot W_n(p, Y_{q\varepsilon\eta}^\varepsilon) \\
&\quad + \sum_{q=0}^{[t/(\varepsilon\eta)]-1} \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du \\
&\qquad \cdot \sum_{n\in\mathbb{Z}^*} \iint \Big(V\Big(\frac{u}{\varepsilon^2}, dp\Big) V\Big(\frac{u}{\varepsilon^2}, dp'\Big) - r(p)\delta(p-p')\, dp\, dp' \Big) \\
&\qquad \cdot \Big(W_n(p, Y_u^\varepsilon) - W_n(p, Y_{q\varepsilon\eta}^\varepsilon) \Big) \\
&\stackrel{\text{def}}{=} R_{11}^\varepsilon(t) + R_{12}^\varepsilon(t)\,.
\end{aligned}
$$

For $R_{12}^\varepsilon(t)$, we have

$$
\boldsymbol{E}|R_{12}^\varepsilon(t)| \leqslant \frac{C_{h,\omega_0,\nabla\omega,F}}{\varepsilon} \sum_{q=0}^{[t/(\varepsilon\eta)]-1} \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du \int_{q\varepsilon\eta}^{u} du' \int dp\, r(p) \leqslant \eta\tilde{C}\,,
$$

following the computations in the proof of Lemma 7.7. It only remains to treat $R_{11}^\varepsilon(t)$, and this term requires more care. We have

$$
\boldsymbol{E}|R_{11}^\varepsilon(t)| \leqslant C_{h,\omega_0,\nabla\omega,F^M} \sum_{q=0}^{[t/(\varepsilon\eta)]-1} \sqrt{I_{q,\eta}^\varepsilon}\,,
$$

with

$$
\begin{aligned}
I_{q,\eta}^\varepsilon = \boldsymbol{E}\Big[\Big| \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du \iint \Big[V\Big(\frac{u}{\varepsilon^2}, dp\Big) V\Big(\frac{u}{\varepsilon^2}, dp'\Big) \\
- r(p)\delta(p-p')\, dp\, dp' \Big] F_{q,\eta}(u,p) \Big|^2 \Big]
\end{aligned}
$$

where

$$
F_{q,\eta}(u,p) = \sum_{n\in\mathbb{Z}^*} \Big(W_n(p, Y_u^\varepsilon) - W_n(p, Y_{q\varepsilon\eta}^\varepsilon) \Big)\,.
$$

Using Gaussian property of V, we have

$$
\begin{aligned}
I_{q,\eta}^\varepsilon &\stackrel{\text{def}}{=} \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du_1 \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du_2 \int F_{q,\eta}(u_1,p_1) F_{q,\eta}(u_2,p_2) \\
&\quad \cdot \Big(\boldsymbol{E}\Big[V\Big(\frac{u_1}{\varepsilon^2}, dp_1\Big) V\Big(\frac{u_2}{\varepsilon^2}, dp_2\Big) \Big] \boldsymbol{E}\Big[V\Big(\frac{u_1}{\varepsilon^2}, dp_1'\Big) V\Big(\frac{u_2}{\varepsilon^2}, dp_2'\Big) \Big] \\
&\qquad + \boldsymbol{E}\Big[V\Big(\frac{u_1}{\varepsilon^2}, dp_1\Big) V\Big(\frac{u_2}{\varepsilon^2}, dp_2'\Big) \Big] \boldsymbol{E}\Big[V\Big(\frac{u_1}{\varepsilon^2}, dp_1'\Big) V\Big(\frac{u_2}{\varepsilon^2}, dp_2\Big) \Big] \Big),
\end{aligned}
$$

and by the symmetry between u_1 and u_2, we can write

$$I^{\varepsilon}_{q,\eta} \leqslant C_{h,\omega_0,\nabla\omega,F} \iint dp\, dp'\, r(p) r(p') \cdot \int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du_1 \int_{q\varepsilon\eta}^{u_1} du_2\, e^{-\mu(|p|^{2\beta}+|p'|^{2\beta})(u_1-u_2)/\varepsilon^2}.$$

Now, we split the integration domain of p into the region where $|p| \leqslant \nu$, and the region where $|p| > \nu$. This leads to the decomposition

$$I^{\varepsilon}_{q,\eta} \leqslant J^{\varepsilon}_{1,q,\nu} + J^{\varepsilon}_{2,q,\nu}.$$

For the second term (when $|p| > \nu$), we have

$$\begin{aligned}\int_{q\varepsilon\eta}^{(q+1)\varepsilon\eta} du_1 \int_{q\varepsilon\eta}^{u_1} du_2\, e^{-\mu(|p|^{2\beta}+|p'|^{2\beta})(u_1-u_2)/\varepsilon^2} &\leqslant \frac{\varepsilon^2}{\mu|p|^{2\beta}} \int_0^{\varepsilon\eta} (1 - e^{-\mu|p|^{2\beta}u_1/\varepsilon^2})\, du \\ &\leqslant C_{\eta,\nu}\varepsilon^3\,,\end{aligned}$$

so that

$$\sum_{q=0}^{[t/(\varepsilon\eta)]-1} \sqrt{J^{\varepsilon}_{2,q,\eta}} \leqslant \varepsilon^{1/2} C_{\eta,\nu,T}\,.$$

For the first term (when $|p| \leqslant \nu$), we have

$$\begin{aligned}\sum_{q=0}^{[t/(\varepsilon\eta)]-1} \sqrt{J^{\varepsilon}_{1,q,\nu}} &\leqslant C_{T,\eta}\Big(\int_{\{|p|\leqslant\nu\}\times S} dp\, dp'\, r(p) r(p') \Big)^{1/2} \\ &\leqslant C_{T,\eta}\Big(\int_{\{|p|\leqslant 1/\nu\}} \frac{dp}{|p|^{2\alpha}} \Big)^{1/2}.\end{aligned}$$

Since $\alpha < 1/2$, the dominated convergence theorem implies that this converges to 0. This concludes the proof of Lemma 7.6.

7.8. **Well-posedness of the martingale problem.** We devote this section to showing that the martingale problem with generator $\mathcal{L}$ (see equation (7.5)) is well-posed. First, we rewrite $\mathcal{L}h(y)$ as

$$\mathcal{L}h(y) = \frac{1}{2}\sum_{j,l=1}^{2} \mathfrak{a}_{jl}(y)\partial^2_{x_l x_j} h(y) + \sum_{j=1}^{2} \partial_{x_j} h(y) \mathfrak{b}_j(y),$$

where

$$\mathfrak{a}_{jl}(y) \overset{\text{def}}{=} \sum_{n\in\mathbb{Z}^*} R_n(y) F_{j,n}(y) F_{-l,n}(y) = \sum_{n\in\mathbb{Z}^*} R_n(y) F_{j,n}(y) \overline{F_{l,n}(y)} \tag{7.12}$$

with (equivalently as in formula (5.2))

$$\begin{aligned} R_n(y) &\overset{\text{def}}{=} \int dp\, r(p) \Big(\frac{1}{\mu|p|^{2\beta} - 2i\pi n\omega(y)} + \frac{1}{\mu|p|^{2\beta} + 2i\pi n\omega(y)} \Big) \\ &= \int dp\, r(p) \frac{2\mu|p|^{2\beta}}{\mu|p|^{4\beta} + 4\pi^2 n^2 \omega^2(y)} \\ &= 2\int_0^{\infty} du\, \boldsymbol{E}[v(u)v(0)] \cos(2\pi n\omega(y)u)\,, \end{aligned} \tag{7.13}$$

and

$$\mathfrak{b}_j(y) \stackrel{\text{def}}{=} \sum_{n\in\mathbb{Z}^*} \int dp\, \frac{r(p)}{\mu|p|^{2\beta} - 2i\pi n\omega(y)}$$

$$\cdot \Big(\sum_{l=1}^{2} \Big(\partial_{x_l} F_{j,n}(y) F_{l,-n}(y) - \frac{2i\pi n}{\mu|p|^{2\beta} - 2i\pi n\omega(y)} F_{j,n}(y) F_{l,-n}(y) \partial_{x_l}\omega(y)\Big)\Big). \tag{7.14}$$

The uniqueness of the martingale problem is guaranteed by [KS91, Corollary 4.9 p. 317, Proposition 3.20 p. 309, and Theorem 2.5 p. 287] together with [Fri06, Theorem 1.2 p. 129], where the latter provides the locally Lipschitz behavior of a nonnegative square root. It is not hard to see that $\mathfrak{b}$ is locally Lipschitz with respect to y, as well as $\mathfrak{a}(y)$ is a symmetric nonnegative matrix, which is $\mathcal{C}^2(\mathbb{R}^2)$ with respect to y. The smoothness property is direct from (7.12) and (7.13). Regarding the nonnegativity, we have for any $x \in \mathbb{R}^2$

$$x^T \mathfrak{a}(y) x = \big\langle \sigma^*(y)(x), \sigma^*(y)(x) \big\rangle_{\mathscr{H}} \geqslant 0 \tag{7.15}$$

with the Hilbert space

$$\mathscr{H} = \Big\{ u : \mathbb{Z}^* \to \mathbb{C} : \overline{u_{-n}} = u_n \quad \text{and} \quad \sum_{n\in\mathbb{Z}^*} |u_n|^2 < \infty \Big\},$$

the adjoint σ_M^* is an operator from $\mathbb{R}^2$ to $\mathscr{H}$ defined by

$$\sigma_M^*(y)(n,x) \stackrel{\text{def}}{=} R_n^{1/2}(y) F_{1,n}(y) x_1 + R_n^{1/2}(y) F_{2,n}(y) x_2\,,$$

and σ, from $\mathscr{H}$ to $\mathbb{R}^2$, is defined by

$$\sigma(y)(u,j) \stackrel{\text{def}}{=} \sum_{n\in\mathbb{Z}^*} R_n^{1/2}(y) F_{j,n}(y) u_{-n} \qquad j = 1,2\,. \tag{7.16}$$

Hence, we have an explicit self-adjoint Hilbert-Schmidt square root for $\mathfrak{a}$ that we use in Section 8 to derive a SDE for the process $(Y_t)_{t\geqslant 0}$. It is also clear that σ is locally Lipschitz with respect to y, so that we can conclude the uniqueness of the martingale problem.

Let us finish by rewriting the martingale problem in a more convenient way and let us start with $\mathfrak{b}$. Using the fact that

$$\begin{aligned}\frac{-2i\pi n \partial_{x_l}\omega(y)}{(\mu|p|^{2\beta} - 2i\pi n\omega(y))^2} &= \partial_{y_l} \int_0^\infty e^{-(\mu|p|^{2\beta} - 2i\pi n\omega(y))u}\, du \\ &= 2i\pi n \partial_{x_l}\omega(y) \int_0^\infty du\, u\, e^{-(\mu|p|^{2\beta} - 2i\pi n\omega(y))u}\,,\end{aligned}$$

we sum Fourier series in (7.14) to obtain

$$\mathfrak{b}(y) = \int_0^1 d\tau \int_0^\infty du\, R(u) \mathrm{Jac}_y[\tilde{F}(u,y,\tau)] \tilde{F}(0,y,\tau)\,,$$

where

$$\tilde{F}(u,y,\tau) = F(y, \tau + \omega(y)u)\,.$$

For $\mathfrak{a}$ we have the following formulation

$$\begin{aligned}\mathfrak{a}_{jl}(y) &= \int_0^1 d\tau \int_0^\infty du\, R(u) \Big(\tilde{F}_j(u,y,\tau)\tilde{F}_l(0,y,\tau) + \tilde{F}_j(0,y,\tau)\tilde{F}_l(u,y,\tau) \Big) \\ &= 2\int_0^1 d\tau \int_0^\infty du\, R(u) \tilde{F}_j(u,y,\tau)\tilde{F}_l(0,y,\tau), \quad j,l \in \{1,2\}\,.\end{aligned}$$

Finally, to conclude, we just have to remark that according to (6.4), as well as the definition of $\mathfrak{a}$ involving the integration of the periodic variable over a whole period, we obtain the formulation (7.5).

8. **Proof of Proposition 6.2**

The proof of this proposition follows the treatment in [KS91, Proposition 4.6 p. 315] and consists of deriving the SDE corresponding to the martingale problem with generator (6.7). To prove this correspondence, we consider first the function $h_j(y) = y_j$ for any $j \in \{1, 2\}$, so that

$$M_j^0(t) = Y_{j,t} - Y_{j,0} - \int_0^t \mathfrak{b}_j(Y_s)\, ds \tag{8.1}$$

is a martingale, where $\mathfrak{b}$ is defined as $\mathfrak{b}^M$ in (7.14). Proceeding the same way choosing $h_{jl}(y) = y_j y_l$ one can show that for any $j, l \in \{1, 2\}$

$$M_j^0(t) M_l^0(t) - \int_0^t \mathfrak{a}_{jl}(Y_s)\, ds$$

is a martingale. (Recall, $\mathfrak{a}$ is defined by (7.12).) Therefore M^0 is a martingale with quadratic variation given by

$$t \mapsto \int_0^t \mathfrak{a}_{jl}(Y_s)\, ds\,.$$

Using (7.15), we note $\mathfrak{a} = \sigma\sigma^*$, where σ is defined by (7.16). We can now apply a standard martingale representation theorem involving Hilbert spaces (see [DPZ96, Theorem 8.2] for instance). This guarantees the existence of a complex valued cylindrical Brownian motion B ($B = (B^1 - iB^2)/\sqrt{2}$, with B^1 and B^2 being two independent real valued cylindrical Brownian motions) on the Hilbert space $\mathcal{T}$, possibly defined on an extension of the probability space under consideration, and such that with probability one

$$M^0(t) = \int_0^t \sigma(Y_s)\, dB_s$$

for all $t \geqslant 0$. As a result, from (8.1) we obtain the following SDE for $(Y_t)_{t \geqslant 0}$,

$$dY_t = \sigma(Y_t) dB_t + \mathfrak{b}(Y_t) dt\,.$$

Note that, because of symmetries, and that for the cylindrical Brownian motion on $\mathcal{T}$ we necessarily have $B^1_{-n} = B^1_n$ and $B^2_{-n} = -B^2_n$, this defines a real valued equation.

Regarding σ, one can write

$$\begin{aligned} \sigma_j(Y_t) dB_t &= \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}^*} \tilde{F}_{j,n}(Y_t) R_n^{1/2}(Y_t)(dB^1_{t,n} - i dB^2_{t,n}) \\ &= \int_0^1 d\tau \tilde{F}_j(Y_t, \tau)\, dW_t(Y_t, \tau) \end{aligned} \tag{8.2}$$

where

$$dW_t(y, \tau) = \frac{1}{\sqrt{2}} \sum_{n \in \mathbb{Z}^*} e^{2i\pi n\tau} R_n^{1/2}(y)(dB^1_{t,n} - i dB^2_{t,n})$$

is a real valued Brownian field with covariance function

$$\begin{aligned}\boldsymbol{E}W_t(y,\tau)W_s(y,\tau') &= t\wedge s\sum_{n\in\mathbb{Z}^*} e^{2i\pi n(\tau-\tau')}R_n(y)\\ &= t\wedge s\sum_{n\in\mathbb{Z}^*} e^{2i\pi n(\tau-\tau')}\int dp\, r(p)\\ &\qquad\cdot\int_0^\infty du\left[e^{-(\mu|p|^{2\beta}+2i\pi n\omega(y))u}+e^{-(\mu|p'|^{2\beta}-2i\pi n\omega(y))u}\right]\\ &= t\wedge s\int_0^\infty du\, R(u)\\ &\qquad\cdot\sum_{n\in\mathbb{Z}^*}\left[e^{2i\pi n(\tau-\tau'+\omega(y)u)}+e^{2i\pi n(\tau-\tau'-\omega(y)u)}\right].\end{aligned}$$

so that for any test functions ϕ_1, ϕ_2 in the Hilbert space

$$L^2_0(\mathbb{T})=\left\{\phi\in L^2(\mathbb{T}):\int_0^1\phi(\tau)d\tau=0\right\},$$

the covariance function reads

$$\begin{aligned}\boldsymbol{E}W_t(y,\phi_1)W_s(y,\phi_2) &= t\wedge s\int_0^\infty du\, R(u)\\ &\qquad\cdot\int_0^1 d\tau\Big[\phi_1(\tau+\omega(y)u)\phi_2(\tau)+\phi_1(\tau)\phi_2(\tau+\omega(y)u)\Big].\end{aligned}$$

To determine the SDE satisfied by $(I^M_t)_{t\geqslant 0}$ we just have to consider the first coordinate of $(Y_t)_{t\geqslant 0}$. Then, we have

$$dI^M_t=\int_0^1 d\tau\, a^M(I^M_t,\tau)\,dW_t(I^M_t,\tau)+\int_0^1 d\tau\int_0^\infty du\, R(u)\tilde{A}(u,I^M_t,\tau)\,dt$$

with

$$\begin{aligned}\tilde{A}(u,I,\tau) &= \partial_I\big(a^M(I,\tau+\omega(I)u)\big)a^M(I,\tau)\\ &\quad+\partial_\tau a^M(I,\tau+\omega(I)u)\big(b^M(I,\tau)-\langle b^M(I)\rangle\big).\end{aligned}$$

Using the fact that a^M is mean-zero in τ, and integrating by parts, we have

$$\begin{aligned}&\int_0^1 d\tau\,\partial_\tau a^M(I,\tau+\omega(I)u)\big(b^M(I,\tau)-\langle b^M(I)\rangle\big)\\ &\qquad=\int_0^1 d\tau\partial_\tau a^M(I,\tau+\omega(I)u)b^M(I,\tau)-\int_0^1 d\tau a^M(I,\tau+\omega(I)u)\partial_\tau b^M(I,\tau).\end{aligned}$$

Now, we restrict our study for $t\leqslant\eta_M$ so that, according to (4.10), we have for $I\in(1/M,M)$

$$\partial_\tau b^M(I,\tau)=\partial_\tau b(I,\tau)=\partial_\tau\partial_I\varphi_1^{-1}(I,\tau)=\partial_I\partial_\tau\varphi_1^{-1}(I,\tau)=-\partial_I a(I,\tau),$$

so that

$$\int_0^1 d\tau\tilde{A}(u,I,\tau)=\int_0^1 d\tau\,\partial_I\big(a(I,\tau+\omega(I)u)\big)a(I,\tau)+a(I,\tau+\omega(I)u)\partial_I a(I,\tau).$$

We then obtain the desired result for the stopped process $\tilde{I}_t^M = I_{t\wedge\eta_M}^M$. Now for the process $\psi_{t\wedge\eta_M}^M$, we look at the second coordinate of $(Y_t)_{t\geqslant 0}$, and we have

$$d\psi_t^M = \int_0^1 d\tau\, b^M(I_t^M,\tau)\, dW_t(I_t^M,\tau) + \int_0^1 d\tau \int_0^\infty du\, R(u)\tilde{B}(u,I_t^M,\tau)\, dt$$

with

$$\begin{aligned}\tilde{B}(u,I,\tau) &= \partial_I\big(b^M(I,\tau+\omega(I)u) - \langle b^M(I)\rangle\big)a^M(I,\tau)\\ &\quad + \partial_\tau(b^M(I,\tau+\omega(I)u) - \langle b^M(I)\rangle)\big(b^M(I,\tau) - \langle b^M(I)\rangle\big)\,.\end{aligned}$$

Using again that a^M and $\partial_\tau b^M$ are mean-zero in τ, together with an integration by parts, we obtain

$$\begin{aligned}\int_0^1 d\tau\, \tilde{B}(u,I,\tau) &= \int_0^1 d\tau\, \partial_I\big(b^M(I,\tau+\omega(I)u)\big)a^M(I,\tau)\\ &\quad - \int_0^1 d\tau\, b^M(I,\tau+\omega(I)u)\partial_\tau b^M(I,\tau)\,.\end{aligned}$$

Since for $I\in(1/M,M)$ we know

$$\partial_\tau b^M(I,\tau) = -\partial_I a(I,\tau)\,,$$

the proof of Proposition 6.2 is complete.

Appendix A. Proofs of Lemmas 2.1 and 2.2

In this appendix we prove Lemmas 2.1 and 2.2. While the proofs are a bit lengthy, they are a direct computation.

Proof of Lemma 2.1. From (2.1) we note

$$w_1^\varepsilon(t) = \varepsilon\int_0^{t/\varepsilon^2}\sin(\tau)v(\tau)\,d\tau\,, \tag{A.1}$$

and hence

$$\begin{aligned}\boldsymbol{E}(w_1^\varepsilon(t)-w_1^\varepsilon(s))^2 &= \varepsilon^2\int_{s/\varepsilon^2}^{t/\varepsilon^2}\int_{s/\varepsilon^2}^{t/\varepsilon^2}\sin(\theta)\sin(\tau)R(\theta-\tau)\,d\theta\,d\tau\\ &= \frac{\varepsilon^2}{2}\int_{s/\varepsilon^2}^{t/\varepsilon^2}\int_{s/\varepsilon^2}^{t/\varepsilon^2}[\cos(\theta-\tau)-\cos(\theta+\tau)]\,R(\theta-\tau)\,d\theta\,d\tau\\ &= \frac{1}{2}\big(I_1^\varepsilon(s,t)-I_2^\varepsilon(s,t)\big)\,,\end{aligned} \tag{A.2}$$

where

$$I_1^\varepsilon(s,t) \overset{\text{def}}{=} \varepsilon^2\int_{s/\varepsilon^2}^{t/\varepsilon^2}\int_{s/\varepsilon^2}^{t/\varepsilon^2}\cos(\theta-\tau)R(\theta-\tau)\,d\theta\,d\tau\,,$$

$$\text{and}\qquad I_2^\varepsilon(s,t) \overset{\text{def}}{=} \varepsilon^2\int_{s/\varepsilon^2}^{t/\varepsilon^2}\int_{s/\varepsilon^2}^{t/\varepsilon^2}\cos(\theta+\tau)R(\theta-\tau)\,d\theta\,d\tau\,.$$

We first analyze $I_1^\varepsilon(s,t)$. Making a change of variables $z=\theta-\tau$ and integrating by parts with respect to variable τ, we have

$$I_1^\varepsilon(s,t) = \varepsilon^2\int_{s/\varepsilon^2}^{t/\varepsilon^2}\int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau}\cos(z)R(z)\,dz\,d\tau$$

$$= \varepsilon^2\Big[\tau \int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau} \cos(z)R(z)\,dz\Big]_{\tau=s/\varepsilon^2}^{\tau=t/\varepsilon^2}$$
$$- \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \tau\Big(\cos\Big(\frac{s}{\varepsilon^2}-\tau\Big)R\Big(\frac{s}{\varepsilon^2}-\tau\Big) - \cos\Big(\frac{t}{\varepsilon^2}-\tau\Big)R\Big(\frac{t}{\varepsilon^2}-\tau\Big)\Big)\,d\tau$$

The first term on the right reduces to

$$\varepsilon^2\Big[\tau \int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau} \cos(z)R(z)\,dz\Big]_{\tau=s/\varepsilon^2}^{\tau=t/\varepsilon^2} = (t-s)\int_0^{(t-s)/\varepsilon^2} \cos(z)R(z)\,dz\,.$$

For the second term on the right, we note

$$\varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \tau\cos\Big(\frac{s}{\varepsilon^2}-\tau\Big)R\Big(\frac{s}{\varepsilon^2}-\tau\Big)\,d\tau$$
$$= \varepsilon^2 \int_0^{(t-s)/\varepsilon^2} \cos(z)R(z)z dz + s\int_0^{(t-s)/\varepsilon^2} \cos(z)R(z)dz\,,$$

and

$$\varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \cos\Big(\frac{t}{\varepsilon^2}-\tau\Big)R\Big(\frac{t}{\varepsilon^2}-\tau\Big)\tau d\tau$$
$$= -\varepsilon^2 \int_0^{(t-s)/\varepsilon^2} \cos(z)R(z)z dz + t\int_0^{(t-s)/\varepsilon^2} \cos(z)R(z)dz\,.$$

Hence,

$$I_1^\varepsilon(s,t) = 2(t-s)\int_0^{(t-s)/\varepsilon^2} \cos(z)R(z)\,dz - 2\varepsilon^2\int_0^{(t-s)/\varepsilon^2} z\cos(z)R(z)\,dz$$
$$= I_{11}^\varepsilon - 2I_{12}^\varepsilon\,. \tag{A.3}$$

Clearly

$$|I_{11}^\varepsilon| \leqslant 2(t-s)\sup_{x\in\mathbb{R}}\int_0^x \cos(z)R(z)\,dz\,. \tag{A.4}$$

Note that the supremum on the right-hand side is finite for all $\gamma \in (0,2)$, as by Remark 1.2 we know $\lim_{x\to\infty}\int_0^x \cos(z)R(z)\,dz$ exists and is finite.

To estimate I_{12}, we divide the analysis into cases.

Case I: $\gamma \in (0,1)$. Integrating by parts we note

$$I_{12}^\varepsilon = \varepsilon^2\Big[\sin(z)zR(z)\Big]_{z=0}^{z=(t-s)/\varepsilon^2} - \varepsilon^2\int_0^{(t-s)/\varepsilon^2} \sin(z)(R'(z)z+R(z))\,dz$$
$$= \varepsilon^{2\gamma}(t-s)^{1-\gamma}L\Big(\frac{t-s}{\varepsilon^2}\Big) - \varepsilon^2\int_0^{(t-s)/\varepsilon^2} \sin(z)(R'(z)z+R(z))\,dz\,. \tag{A.5}$$

For the first term on the right we note that

$$\varepsilon^{2\gamma}(t-s)^{1-\gamma}L\Big(\frac{t-s}{\varepsilon^2}\Big) \leqslant 2\varepsilon^{2\gamma}L\Big(\frac{1}{\varepsilon^2}\Big)(t-s)^{1-\gamma}\,,$$

for all sufficiently small ε.

For the second term, we note that (1.8) implies

$$|R'(z)z| = |L'(z)z^{1-\gamma} - \gamma L(z)z^{-\gamma}| \leqslant 2|R(z)|\,,$$

for all sufficiently large z. Thus

$$\varepsilon^2 \int_0^{(t-s)/\varepsilon^2} \sin(z)(R'(z)z + R(z))\,dz \leqslant 3\varepsilon^2 \int_0^{(t-s)/\varepsilon^2} |R(z)|\,dz \tag{A.6}$$
$$\leqslant 4\varepsilon^{2\gamma} L\Big(\frac{1}{\varepsilon^2}\Big)(t-s)^{1-\gamma}\,,$$

where the last inequality followed from Karamata's theorem [BIKS18, Th 6.2.1], and the fact that L is slowly varying. Combining these, we see

$$|I_{12}^\varepsilon| \leqslant 10\varepsilon^{2\gamma} L\Big(\frac{1}{\varepsilon^2}\Big)(t-s)^{1-\gamma}\,, \tag{A.7}$$

for all sufficiently small ε.

Case II: $\gamma = 1$. We follow the proof in the case $\gamma < 1$, with a few minor changes. For the first term on the right of (A.5), we note that $L(z) \leqslant 2R(0)z$ for all sufficiently small z. Since L is slowly varying at infinity we must certainly have $L(z) \leqslant z$ for all sufficiently large z. Hence we can find a finite constant C such that $L(z) \leqslant Cz$ for all $z \geqslant 0$. As a result, we have

$$\varepsilon^2 L\Big(\frac{t-s}{\varepsilon^2}\Big) \leqslant C(t-s)\,.$$

For the remaining terms, note that the function $g(x) \stackrel{\text{def}}{=} \int_0^x R(z)\,dz$ is slowly varying (see for instance [BGT89] Proposition 1.5.9a, p. 26). Thus using (A.6) we see that

$$\varepsilon^2 \int_0^{(t-s)/\varepsilon^2} \sin(z)(R'(z)z + R(z))\,dz \leqslant 3\varepsilon^2 \int_0^{(t-s)/\varepsilon^2} |R(z)|\,dz \tag{A.8}$$
$$= 3\varepsilon^2 g\Big(\frac{t-s}{\varepsilon^2}\Big)\,. \tag{A.9}$$

Since $g(0) = 0$, $g'(0) = R(0) < \infty$, and g is slowly varying at infinity, we must have $g(z) \leqslant Cz$ for all $z \geqslant 0$ and a finite constant C. (As before, we allow the constant C to change from line to line, provided it does not depend on ε, t and s.) Consequently,

$$|I_{21}^\varepsilon| \leqslant C(t-s)\,,$$

as desired.

Case III: $\gamma \in (1,2)$. Directly integrating I_{12}^ε and using Karamata's theorem we see

$$|I_{12}^\varepsilon| \leqslant 4\varepsilon^{2(\gamma-1)} L\Big(\frac{1}{\varepsilon^2}\Big)(t-s)^{2-\gamma}\,.$$

We now turn our attention to the term I_2^ε. Substituting $z = \theta - \tau$, we note

$$I_2^\varepsilon = \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau} \cos(z+2\tau)R(z)\,dz\,d\tau$$
$$= \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau} [\cos(z)\cos(2\tau) - \sin(z)\sin(2\tau)]R(z)\,dz\,d\tau \tag{A.10}$$
$$= I_{21}^\varepsilon - I_{22}^\varepsilon\,, \tag{A.11}$$

where

$$I_{21}^\varepsilon \stackrel{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau} \cos(z)\cos(2\tau)R(z)\,dz\,d\tau\,, \tag{A.12}$$

$$\text{and} \qquad I_{22}^{\varepsilon} \stackrel{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau} \sin(z)\sin(2\tau)R(z)\,dz\,d\tau\,.$$

We again divide the analysis into cases.

Case I: $\gamma \in (0,1)$. Integrating by parts with respect to τ, we observe

$$I_{21}^{\varepsilon} = \frac{\varepsilon^2}{2}\Big[\sin(2\tau)\int_{s/\varepsilon^2-\tau}^{t/\varepsilon^2-\tau}\cos(z)R(z)\,dz\Big]_{\tau=s/\varepsilon^2}^{\tau=t/\varepsilon^2} \tag{A.13}$$

$$-\varepsilon^2\int_{s/\varepsilon^2}^{t/\varepsilon^2}\Big(\cos\Big(\frac{s}{\varepsilon^2}-\tau\Big)R\Big(\frac{s}{\varepsilon^2}-\tau\Big) - \cos\Big(\frac{t}{\varepsilon^2}-\tau\Big)R\Big(\frac{t}{\varepsilon^2}-\tau\Big)\Big)\sin(2\tau)\,d\tau\,. \tag{A.14}$$

and hence by Karamata's theorem,

$$|I_{21}^{\varepsilon}| \leqslant C\varepsilon^2\int_0^{(t-s)/\varepsilon^2}|R(z)|\,dz \leqslant C\varepsilon^{2\gamma}L\Big(\frac{1}{\varepsilon^2}\Big)(t-s)^{1-\gamma}\,, \tag{A.15}$$

for some finite constant C.

Case II: $\gamma = 1$. As in the previous case, we know from (A.15) that

$$|I_{21}^{\varepsilon}| \leqslant C\varepsilon^2\int_0^{(t-s)/\varepsilon^2}|R(z)|\,dz\,.$$

Using the same argument as that used to bound (A.9) we obtain

$$|I_{21}^{\varepsilon}| \leqslant C(t-s)$$

as desired.

Case III: $\gamma \in (1,2)$. As before, set $g(x) \stackrel{\text{def}}{=} \int_0^x |R(z)|\,dz$, and note

$$|I_{21}^{\varepsilon}| \leqslant \varepsilon^2\int_{s/\varepsilon^2}^{t/\varepsilon^2}\Big(g\Big(\frac{t}{\varepsilon^2}-\tau\Big)-g\Big(\frac{s}{\varepsilon^2}-\tau\Big)\Big)\,d\tau = 2\varepsilon^2\int_0^{(t-s)/\varepsilon^2} g(z)\,dz\,.$$

Let $G(x) = \int_0^x g(z)\,dz$ and note that G is regularly varying with index $2-\gamma$. Thus

$$|I_{21}^{\varepsilon}| \leqslant 2\varepsilon^2 G\Big(\frac{t-s}{\varepsilon^2}\Big) = 3\varepsilon^{2\gamma-2}\tilde{L}\Big(\frac{1}{\varepsilon^2}\Big)(t-s)^{2-\gamma}$$

The estimates for I_{22}^{ε} are identical to those for I_{21}^{ε}. Combining the above estimates we obtain (2.3) for $i=1$. The proof when $i=2$ is identical, and this finishes the proof of Lemma 2.1. □

Proof of Lemma 2.2. We first prove equation (2.4) for $i=1$. To do this, we follow the same computation as in the proof of Lemma 2.1 up to (A.3). Now note

$$\lim_{\varepsilon\to 0} I_{11}^{\varepsilon} = 2\int_0^{\infty}\cos(z)R(z)\,dz\,,$$

where the above integral converges absolutely for $\gamma \in (1,2)$, and conditionally for $\gamma \in (0,1]$ (see Remark 1.2). When $\gamma \neq 1$, the proof of Lemma 2.1 already shows that I_{12}^{ε}, I_{21}^{ε} and I_{22}^{ε} all vanish as $\varepsilon \to 0$. When $\gamma = 1$, the proof of Lemma 2.1 shows

$$|I_{12}^{\varepsilon}| + |I_{21}^{\varepsilon}| \leqslant C\varepsilon^2\int_0^{(t-s)/\varepsilon^2}|R(z)|\,dz = C\varepsilon^2 g\Big(\frac{t-s}{\varepsilon^2}\Big)\,,$$

where $g(x) = \int_0^x |R(z)|\,dz$. Since g is slowly varying this vanishes as $\varepsilon \to 0$. This proves (2.4) as claimed.

To prove (2.5), we note

$$\begin{aligned}
&\boldsymbol{E}(w_1^\varepsilon(t) - w_1^\varepsilon(s))(w_2^\varepsilon(t) - w_2^\varepsilon(s)) \\
&\quad = \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} R(\theta - \tau) \sin(\theta) \cos(\tau)\, d\theta\, d\tau \\
&\quad = \frac{\varepsilon^2}{2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} R(\theta - \tau)[\sin(\theta + \tau) + \sin(\theta - \tau)]\, d\theta\, d\tau\,, \\
&\quad = \tilde{I}_1^\varepsilon + \tilde{I}_2^\varepsilon\,,
\end{aligned}$$

where

$$\begin{aligned}
\tilde{I}_1^\varepsilon &\stackrel{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(\theta - \tau) R(\theta - \tau)\, d\theta\, d\tau, \\
\tilde{I}_2^\varepsilon &\stackrel{\text{def}}{=} \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(\theta + \tau) R(\theta - \tau)\, d\theta\, d\tau.
\end{aligned}$$

We claim $\tilde{I}_1^\varepsilon = 0$. To see this, make the changes of variables

$$\theta' \stackrel{\text{def}}{=} \frac{(t+s)}{\varepsilon^2} - \theta\,, \quad \tau' \stackrel{\text{def}}{=} \frac{(t+s)}{\varepsilon^2} - \tau\,, \tag{A.16}$$

and rewrite $\tilde{I}_1^\varepsilon$ as

$$\begin{aligned}
\tilde{I}_1^\varepsilon &= \varepsilon^2 \int_{t/\varepsilon^2}^{s/\varepsilon^2} \int_{t/\varepsilon^2}^{s/\varepsilon^2} \sin(\tau' - \theta') R(\tau' - \theta')\, d\theta'\, d\tau' \\
&= -\varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2}^{t/\varepsilon^2} \sin(\theta' - \tau') R(\theta' - \tau')\, d\theta'\, d\tau' \\
&= -\tilde{I}_1^\varepsilon\,.
\end{aligned}$$

This implies $\tilde{I}_1^\varepsilon = 0$ as claimed.

We now turn our attention to $\tilde{I}_2^\varepsilon$. Making the change of variable $z = \theta - \tau$ we note

$$\begin{aligned}
\tilde{I}_2^\varepsilon &= \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2 - \tau}^{t/\varepsilon^2 - \tau} \sin(z + 2\tau) R(z)\, dz\, d\tau \\
&= \varepsilon^2 \int_{s/\varepsilon^2}^{t/\varepsilon^2} \int_{s/\varepsilon^2 - \tau}^{t/\varepsilon^2 - \tau} [\sin(z)\cos(2\tau) + \cos(z)\sin(2\tau)] R(z)\, dz\, d\tau\,.
\end{aligned}$$

This is similar to the expression for I_2^ε (equation (A.10)) and following the proof of Lemma 2.1 we see $\tilde{I}_2^\varepsilon \to 0$ as $\varepsilon \to 0$. Since $\tilde{I}_1^\varepsilon$ and $\tilde{I}_2^\varepsilon$ both vanish as $\varepsilon \to 0$ we obtain (2.5) as claimed. □

Appendix B. Proof of Proposition 1.3

In this section we show convergence of the rescaled processes u^ε (equation (1.11)) to fBm. The result is similar to well known results [Taq75, Taq77, Mar05], and the proof is presented here for completeness.

Proof of Proposition 1.3 when $\gamma \in (0,1)$. Using (1.5) and (1.11) we note

$$\boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 = \frac{2}{\sigma(\varepsilon)^2} \int_{r_1=s}^{t} \int_{r_2=s}^{r_1} R\Big(\frac{r_1 - r_2}{\varepsilon^2}\Big)\, dr_2\, dr_1\,.$$

By (1.7) and the uniform convergence theorem [BGT89, Th. 1.2.1], we know

$$\frac{1}{\sigma(\varepsilon)^2}\Big|R\Big(\frac{r_1 - r_2}{\varepsilon^2}\Big)\Big| = \frac{1}{L(\varepsilon^{-2})|r_1 - r_2|^\gamma} L\Big(\frac{r_1 - r_2}{\varepsilon^2}\Big) \leqslant \frac{2}{(r_1 - r_2)^\gamma}\,, \tag{B.1}$$

for all $r_1, r_2 \in [s,t]$ and all sufficiently small ε. Hence

$$\boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 \leqslant 4 \int_{r_1=s}^{t} \int_{r_2=s}^{r_1} \frac{dr_2\, dr_1}{(r_1 - r_2)^\gamma} = \frac{4(t-s)^{2-\gamma}}{(1-\gamma)(2-\gamma)}\,. \tag{B.2}$$

Moreover (1.7), (B.1) and the dominated convergence theorem imply

$$\lim_{\varepsilon \to 0} \boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 = 2 \int_{r_1=s}^{t} \int_{r_2=s}^{r_1} \frac{dr_2\, dr_1}{(r_1 - r_2)^\gamma} = \frac{(t-s)^{2\mathcal{H}}}{(2\mathcal{H}-1)\mathcal{H}}\,.$$

Since u^ε is Gaussian this implies that the finite-dimensional distributions of u^ε converge to that of $\sigma_{\mathcal{H}} B^{\mathcal{H}}$. For convergence in law, note (B.2) implies

$$\begin{aligned}
\boldsymbol{E}[(u^\varepsilon(s) - u^\varepsilon(r))^2 (u^\varepsilon(t) - u^\varepsilon(s))^2] &\leqslant \boldsymbol{E}[(u^\varepsilon(s) - u^\varepsilon(r))^4]^{1/2} \boldsymbol{E}[(u^\varepsilon(t) - u^\varepsilon(s))^4]^{1/2} \\
&\leqslant 3\boldsymbol{E}[(u^\varepsilon(s) - u^\varepsilon(r))^2] \boldsymbol{E}[(u^\varepsilon(t) - u^\varepsilon(s))^2] \\
&\leqslant C|t-r|^{4\mathcal{H}}\,.
\end{aligned}$$

Since $4\mathcal{H} > 1$, Theorems 13.4 and 13.5 in [Bil99] imply that u^ε converges to $\sigma_{\mathcal{H}} B^{\mathcal{H}}$ in law. □

Proof of Proposition 1.3 when $\gamma \in (1,2)$. Using (1.5), (1.11) and a change of variable we note

$$\begin{aligned}
\boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 &= \frac{2}{\sigma(\varepsilon)^2} \int_{r_1=s}^{t} \int_{r_2=s}^{r_1} R\Big(\frac{r_1 - r_2}{\varepsilon^2}\Big)\, dr_2\, dr_1 \\
&= \frac{2}{\sigma(\varepsilon)^2} \int_{r_1=s}^{t} \int_{z=0}^{r_1 - s} R\Big(\frac{z}{\varepsilon^2}\Big)\, dz\, dr_1 \\
&= \frac{-2}{\sigma(\varepsilon)^2} \int_{r_1=s}^{t} \int_{z=r_1-s}^{\infty} R\Big(\frac{z}{\varepsilon^2}\Big)\, dz\, dr_1\,,
\end{aligned} \tag{B.3}$$

where the last equality followed from (1.10). Since L is slowly varying, (1.7) implies

$$\frac{1}{\sigma(\varepsilon)^2} R\Big(\frac{z}{\varepsilon^2}\Big) \leqslant R(0) \wedge \Big(\frac{C}{z^{\gamma'}}\Big)\,, \tag{B.4}$$

where $\gamma' = (1+\gamma)/2$ and C is a finite constant that is independent of ε. We will subsequently allow C to change from line to line, as long as it doesn't depend on ε.

Note (B.3) and (B.4) immediately imply

$$\boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 \leqslant 2C \int_{r_1=s}^{t} \int_{z=r_1-s}^{\infty} \frac{dz\, dr_1}{z^{\gamma'}} = C(t-s)^{2-\gamma'}\,. \tag{B.5}$$

Moreover, by (1.7), (B.4) and the dominated convergence theorem, we see

$$\lim_{\varepsilon \to 0} \boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 = 2 \int_{r_1=s}^{t} \int_{z=r_1-s}^{\infty} \frac{dz\, dr_2}{z^\gamma} = \frac{(t-s)^{2\mathcal{H}}}{(1-2\mathcal{H})\mathcal{H}}\,.$$

Now the remainder of the proof is identical to the case when $\gamma \in (0,1)$. □

Proof of Proposition 1.3 when $\gamma = 1$. In this case we write

$$\begin{aligned}\boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 &= \frac{2}{\sigma(\varepsilon)^2}\int_{r=s}^{t}\int_{z=0}^{r-s} R\Big(\frac{z}{\varepsilon^2}\Big)\,dz\,dr \\ &= \frac{2}{L(\varepsilon^{-2})|\ln\varepsilon|}\int_{r=s}^{t}\int_{z=0}^{1} R(z)\,dz\,dr + \frac{2}{\sigma(\varepsilon)^2}\int_{r=s}^{t}\int_{z=\varepsilon^2}^{r-s} R\Big(\frac{z}{\varepsilon^2}\Big)\,dz\,dr\end{aligned}$$

The first term on the right vanishes as $\varepsilon \to 0$. Using the uniform convergence theorem [BGT89, Th 1.2.1] on the second term we see

$$\lim_{\varepsilon\to 0}\boldsymbol{E}(u^\varepsilon(t) - u^\varepsilon(s))^2 = \lim_{\varepsilon\to 0}\frac{2}{|\ln\varepsilon|}\int_{r=s}^{t}\int_{z=\varepsilon^2}^{r-s}\frac{1}{z}\,dz\,dr = t - s\,,$$

and the convergence is uniform when s, t belong to any bounded interval. Hence by the same argument as in the previous two cases we see that (u^ε) converges in law to $\sigma_{\mathcal{H}} B^{\mathcal{H}}$ as claimed. □

Appendix C. Proof of Lemma 3.4

First, to prove (3.5) it suffices to show, for all $n \geqslant 1$ and $0 \leqslant t_1 \leqslant \ldots \leqslant t_{n+1}$, that

$$\boldsymbol{E}\big[V(t_{n+1}, dp)|V(t_1,\cdot),\ldots,V(t_n,\cdot)\big] = e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}V(t_n, dp).$$

For $n = 1$ and $0 \leqslant t_1 \leqslant t_2$, we write

$$V(t_2, dp) = e^{-\mu|p|^{2\beta}(t_2-t_1)}V(t_1, dp) + Y,$$

where Y and $V(t_1, dp)$ are independent. In fact, since they are mean-zero Gaussian variables, we have

$$\begin{aligned}\boldsymbol{E}[Y(\varphi)V(t_1,\psi)] &= \boldsymbol{E}[V(t_2,\varphi)V(t_1,\psi)] - \boldsymbol{E}[V(t_1,\varphi_{t_2-t_1})V(t_1,\psi)] \\ &= \int dp\, r(p)\varphi(p)\psi(p)(e^{-\mu|p|^{2\beta}(t_2-t_1)} - e^{-\mu|p|^{2\beta}(t_2-t_1)}) \\ &= 0,\end{aligned}$$

for all φ, ψ bounded continuous functions where $\varphi_s(p) = e^{-\mu|p|^{2\beta}s}\varphi(p)$. As a result, we have

$$\boldsymbol{E}\big[V(t_2, dp)|V(t_1,\cdot)\big] = e^{-\mu|p|^{2\beta}(t_2-t_1)}V(t_1, dp).$$

Now, let us fix $n \geqslant 2$ and assume that for all family $(s_j)_{j\in\{1,\ldots,n\}}$ such that $0 \leqslant s_1 \leqslant \ldots \leqslant s_n$

$$\boldsymbol{E}\big[V(s_n, dp)|V(s_1,\cdot),\ldots,V(s_{n-1},\cdot)\big] = e^{-\mu|p|^{2\beta}(s_n-s_{n-1})}V(s_{n-1}, dp),$$

Then, we write

$$V(t_{n+1}, dp) = e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}V(t_n, dp) + Y,$$

where Y and $V(t_n, dp)$ are independent as explained above, so that

$$\begin{aligned}
&\boldsymbol{E}\big[V(t_{n+1},dp)|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\\
&\quad= e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}V(t_n,dp)+\boldsymbol{E}\big[Y|V(t_1,\cdot),\dots,V(t_{n-1},\cdot)\big]\\
&\quad= e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}V(t_n,dp)+\boldsymbol{E}\big[V(t_{n+1},dp)|V(t_1,\cdot),\dots,V(t_{n-1},\cdot)\big]\\
&\qquad- e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}\boldsymbol{E}\big[V(t_n,dp)|V(t_1,\cdot),\dots,V(t_{n-1},\cdot)\big]\\
&\quad= e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}V(t_n,dp)\\
&\qquad+(e^{-\mu|p|^{2\beta}(t_{n+1}-t_{n-1})}-e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}e^{-\mu|p|^{2\beta}(t_n-t_{n-1})})V(t_{n-1},dp)\\
&\quad= e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}V(t_n,dp),
\end{aligned}$$

which concludes the proof of (3.5) by induction.

Second, to prove (3.6) it suffices to show that for all $n\geqslant 1$, $0\leqslant t_1\leqslant\ldots\leqslant t_{n+1}\leqslant\tilde{t}_{n+1}$ and φ,ψ bounded continuous functions, that

$$\begin{aligned}
&\boldsymbol{E}\big[V(\tilde{t}_{n+1},\varphi)V(t_{n+1},\psi)|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\\
&\quad=\boldsymbol{E}\big[V(\tilde{t}_{n+1},\varphi)|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\boldsymbol{E}\big[V(t_{n+1},\psi)|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\\
&\qquad+\int dp\,r(p)\varphi(p)\psi(p)\Big(e^{-\mu|p|^{2\beta}(\tilde{t}_{n+1}-t_{n+1})}\\
&\qquad\qquad-e^{-\mu|p|^{2\beta}(\tilde{t}_{n+1}-t_n)}e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}\Big).
\end{aligned}$$

This last relation is a consequence of the following lemma, which is a consequence of [Roz87, Theorem 10.1 and Theorem 10.2].

Lemma C.1. *Let $(X,Y,Z_1,\dots,Z_n)$ be a Gaussian vector on a probability space $(\Omega,\mathcal{F},\boldsymbol{P})$, and $\mathcal{G}=\sigma(Z_1,\dots,Z_n)$ be the σ-field generated by $Z_1,\dots,Z_n$. Then, the couple $(X-\boldsymbol{E}[X|\mathcal{G}],Y-\boldsymbol{E}[Y|\mathcal{G}])$ is independent of $\mathcal{G}$, and*

$$\boldsymbol{E}[XY|\mathcal{G}]=\boldsymbol{E}[X|\mathcal{G}]\boldsymbol{E}[Y|\mathcal{G}]+\boldsymbol{E}[(X-\boldsymbol{E}[X|\mathcal{G}])(Y-\boldsymbol{E}[Y|\mathcal{G}])]. \tag{C.1}$$

Proof of Lemma C.1. The proof of the independence of $(X-\boldsymbol{E}[X|\mathcal{G}],Y-\boldsymbol{E}[Y|\mathcal{G}])$ with respect to $\mathcal{G}$ is a consequence of Theorem 10.1 and Theorem 10.2 of [Roz87]. Consequently,

$$\boldsymbol{E}\Big[\Big(X-\boldsymbol{E}[X|\mathcal{G}]\Big)\Big(Y-\boldsymbol{E}[Y|\mathcal{G}]\Big)\Big|\mathcal{G}\Big]=\boldsymbol{E}\Big[\Big(X-\boldsymbol{E}[X|\mathcal{G}]\Big)\Big(Y-\boldsymbol{E}[Y|\mathcal{G}]\Big)\Big],$$

and then

$$\begin{aligned}
\boldsymbol{E}[XY|\mathcal{G}]&=\boldsymbol{E}\Big[\Big(X-\boldsymbol{E}[X|\mathcal{G}]+\boldsymbol{E}[X|\mathcal{G}]\Big)\Big(Y-\boldsymbol{E}[Y|\mathcal{G}]+\boldsymbol{E}[Y|\mathcal{G}]\Big)\Big|\mathcal{G}\Big]\\
&=\boldsymbol{E}[X|\mathcal{G}]\boldsymbol{E}[Y|\mathcal{G}]+\boldsymbol{E}\Big[\Big(X-\boldsymbol{E}[X|\mathcal{G}]\Big)\Big(Y-\boldsymbol{E}[Y|\mathcal{G}]\Big)\Big]\\
&\quad+\boldsymbol{E}[Y|\mathcal{G}]\underbrace{\boldsymbol{E}\big[\big(X-\boldsymbol{E}[X|\mathcal{G}]\big)\big|\mathcal{G}\big]}_{=0}+\boldsymbol{E}[X|\mathcal{G}]\underbrace{\boldsymbol{E}\big[\big(Y-\boldsymbol{E}[Y|\mathcal{G}]\big)\big|\mathcal{G}\big]}_{=0}.
\end{aligned}$$

□

As a result, for all φ,ψ bounded continuous functions, with

$$X=V(\tilde{t}_{n+1},\varphi)\qquad\text{and}\qquad Y=V(t_{n+1},\psi),$$

we have

$$\begin{aligned}\boldsymbol{E}\big[V(\tilde{t}_{n+1},\varphi)V(t_{n+1},\psi)|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]&=\boldsymbol{E}\big[V(\tilde{t}_{n+1},\varphi)|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\\&\quad\cdot\boldsymbol{E}\big[V(t_{n+1},\psi)|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\\&\quad+P,\end{aligned}$$

where, using (3.5),

$$\begin{aligned}P&=\boldsymbol{E}\Big[\Big(X-\boldsymbol{E}\big[X|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\Big)\Big(Y-\boldsymbol{E}\big[Y|V(t_1,\cdot),\dots,V(t_n,\cdot)\big]\Big)\Big]\\&=\boldsymbol{E}\Big[\Big(V(\tilde{t}_{n+1},\varphi)-V(t_n,\varphi_{\tilde{t}_{n+1}-t_n})\Big)\Big(V(t_{n+1},\psi)-V(t_n,\psi_{t_{n+1}-t_n})\Big)\Big]\\&=\int dp\,r(p)\varphi(p)\psi(p)(e^{-\mu|p|^{2\beta}(\tilde{t}_{n+1}-t_{n+1})}-e^{-\mu|p|^{2\beta}(\tilde{t}_{n+1}-t_n)}e^{-\mu|p|^{2\beta}(t_{n+1}-t_n)}),\end{aligned}$$

which concludes the proof of (3.6).

Appendix D. Proof of Lemma 3.5

Let us start with the following remark. For any $(t,\varphi)\in D_{k,M}$, it is straightforward that $\varphi(t,\cdot)\in W_{k,M}$, so that

$$\boldsymbol{E}\sup_{(t,\varphi)\in D_{k,M}}\Big|V\Big(\frac{t}{\varepsilon^2},\varphi(t,\cdot)\Big)\Big|\leqslant\boldsymbol{E}\sup_{(t,\varphi)\in\tilde{D}_{k,M}}\Big|V\Big(\frac{t}{\varepsilon^2},\varphi\Big)\Big|,$$

with

$$\tilde{D}_{k,M}\overset{\text{def}}{=}[0,T]\times W_{k,M}.$$

As a result, to prove inequality (3.7), we only need to prove that

$$\boldsymbol{E}\sup_{(t,\varphi)\in\tilde{D}_{k,M}}\Big|V\Big(\frac{t}{\varepsilon^2},\varphi\Big)\Big|\leqslant C+\frac{C(\varepsilon)}{\varepsilon},$$

where $C>0$ and $C(\varepsilon)>0$ goes to 0 as $\varepsilon\to0$. To this end we first remark that there is a compact embedding of $\tilde{D}_{k,M}$ into $[0,T]\times C^0(S)$ equipped with the metric

$$\|(t,\varphi)\|_\varepsilon:=\sqrt{|t|}/\varepsilon+\|\varphi\|_\infty. \tag{D.1}$$

Let us also recall [Bre11, Theorem 8.8] that there exists $C_k>0$ such that for any $\varphi\in W^{1,k}(S)$, we have

$$\|\varphi\|_\infty\leqslant C_k\|\varphi\|_{W^{1,k}}.$$

We also remark that for the pseudometric

$$d_{V,\varepsilon}((t,\varphi),(s,\psi))=\boldsymbol{E}[|V(t/\varepsilon^2,\varphi)-V(s/\varepsilon^2,\psi)|^2]^{1/2},$$

for which

$$\begin{aligned}d^2_{V,\varepsilon}((t,\varphi),(s,\psi))&=\int_S dp\,r(p)(\varphi^2(p)+\psi^2(p)-2\varphi(p)\psi(p)e^{-\mu|p|^{2\beta}|t-s|/\varepsilon^2})\\&=\int_S dp\,r(p)(\varphi^2(p)+\psi^2(p))(1-e^{-\mu|p|^{2\beta}|t-s|/\varepsilon^2})\\&\quad+\int_S dp\,r(p)(\varphi(p)-\psi(p))^2e^{-\mu|p|^{2\beta}|t-s|/\varepsilon^2}\\&\leqslant C_{k,M,r}\|(t,\varphi)-(s,\psi)\|^2_\varepsilon,\end{aligned} \tag{D.2}$$

we have

$$diam_{d_{V,\varepsilon}}(\tilde{D}_{k,M}) \leqslant \sqrt{2C_{k,M,r}(T+M)} \stackrel{\text{def}}{=} c_{k,M,r}\,.$$

Note now that it is not easy to see if $\tilde{D}_{k,M}$ is relatively compact equipped with $d_{V,\varepsilon}$, while it is required to apply [AT07, Theorem 1.5.1 p. 41]. The problem is that $d_{V,\varepsilon}$ is only a pseudometric. For instance, $d_{V,\varepsilon}$ does not separate well points, since we have for any s and t

$$d_{V,\varepsilon}((t,0),(s,0)) = 0\,.$$

However, according to (D.2), the application $d_{V,\varepsilon}$ provides a true metric on $\tilde{D}_{k,M} - [0,T]\times\{0\}$. Then, instead of working directly with $\tilde{D}_{k,M}$ let us consider now the increasing sequence of $\|\cdot\|_\varepsilon$-relatively compact subset of $\tilde{D}_{k,M}$

$$D^n_{k,M} \stackrel{\text{def}}{=} \tilde{D}_{k,M} \cap [0,T]\times\left\{\varphi \in W^{1,k}(S) :\ \|\varphi\|_\infty \geqslant 1/n\right\}.$$

Now, since $d_{V,\varepsilon}$ is a metric on every $D^n_{k,M}$, it is straightforward to see from (D.2) that all the $D^n_{k,M}$ are also $d_{V,\varepsilon}$-relatively compact from the sequential characterization of compactness.

We can now apply [AT07, Theorem 1.5.1 p. 41 and Lemma 1.5.2 p. 44] over all the $\overline{D^n_{k,M}}^{d_{V,\varepsilon}}$, we have

$$\boldsymbol{E}\sup_{(t,\varphi)\in D^n_{k,M}} V\Big(\frac{t}{\varepsilon^2},\varphi\Big) \leqslant K\Big(c_{k,M,r}\ln(c_{k,M,r}) + \int_0^{c_{k,M,r}} \sqrt{\ln(N_\varepsilon(r))}dr\Big)\,,$$

where $N_\varepsilon(r)$ is the smallest number of balls covering $D^n_{k,M}$ with radius r, which are defined by

$$B(X,r) = \{Y \in D^n_{k,M} :\ d_{V,\varepsilon}(X,Y) < r\}\,.$$

Because of (D.2), it is clear that

$$N_\varepsilon(r) \leqslant \mathcal{N}_\varepsilon\big(r/C^{1/2}_{k,M,r}\big)\,,$$

where $\mathcal{N}_\varepsilon(u)$ stands for the smallest number of balls with radius u associated to the metric defined by the norm $\|\cdot\|_\varepsilon$. Since $D^n_{k,M}$ is defined as a product space, one can determine the smallest number of balls with radius r that cover each of its components. For each parts, the metric is defined by the corresponding parts of the r.h.s of (D.1). Therefore, the smallest number of balls with radius r that cover $[0,T]$ is of order $1/(r^2\varepsilon^2)$, and for $\{\varphi \in W^{1,k}(S) :\ \ \|\varphi\|_{W^{1,k}} \leqslant M\}$ it is of order $\exp(1/r)$ [BS67, Theorem 5.2 p. 311]. As a result, for any $n \in \mathbb{N}^*$

$$\begin{aligned}\boldsymbol{E}\sup_{(t,\varphi)\in D^n_{k,M}} V\Big(\frac{t}{\varepsilon^2},\varphi\Big) &\leqslant C_1 + C_2\int_0^{c_{k,M,r}} \sqrt{\ln\Big(\frac{C\exp(1/r)}{r^2\varepsilon^2}\Big)}\,dr\,,\\ &\leqslant C_1' + C_2'\int_0^{c_{k,M,r}} \frac{dr}{\sqrt{r}} + \frac{C_3'}{\varepsilon}\int_0^{\varepsilon c_{k,M,r}} \sqrt{\ln\Big(\frac{1}{u}\Big)}\,du\,.\end{aligned}$$

Finally, setting

$$C(\varepsilon) \stackrel{\text{def}}{=} C_3'\int_0^{\varepsilon c_{k,M,r}} \sqrt{\ln\Big(\frac{1}{u}\Big)}\,du\,,$$

and using the monotone convergence theorem, we have

$$E_1 \stackrel{\text{def}}{=} \boldsymbol{E}\sup_{(t,\varphi)\in \tilde{D}_{k,M}} V\Big(\frac{t}{\varepsilon^2},\varphi\Big) = \lim_{n\to\infty} \boldsymbol{E}\sup_{(t,\varphi)\in D^n_{k,M}} V\Big(\frac{t}{\varepsilon^2},\varphi\Big) \leqslant C_1'' + \frac{C(\varepsilon)}{\varepsilon}\,.$$

We conclude the proof of the bound (3.7) using that

$$\boldsymbol{E} \sup_{(t,\varphi)\in\tilde{D}_{k,M}} \left|V\Big(\frac{t}{\varepsilon^2},\varphi\Big)\right| \leqslant 2E_1$$

by symmetry.

For proving (3.8), we first remark that by stationarity in t we have

$$\boldsymbol{E} \sup_{\varphi\in W_{k,M}} \left|V\Big(\frac{t}{\varepsilon^2},\varphi\Big)\right|^n = \boldsymbol{E} \sup_{\varphi\in W_{k,M}} |V(0,\varphi)|^n\,,$$

so that following the same lines as above by removing the t-dependence we have for the first order moment,

$$\boldsymbol{E} \sup_{\varphi\in W_{k,M}} |V(0,\varphi)| \leqslant 2\boldsymbol{E} \sup_{\varphi\in W_{k,M}} V(0,\varphi) \leqslant C_1\,.$$

Finally, for the arbitrary order moments we write

$$\begin{aligned} E_n' \overset{\text{def}}{=} \boldsymbol{E} \sup_{\varphi\in W_{k,M}} |V(0,\varphi)|^n &= \int_0^\infty du\, \boldsymbol{P}\Big(\sup_{\varphi\in W_{k,M}} |V(0,\varphi)|^n > u\Big) \\ &= \int_0^\infty du\, \boldsymbol{P}\Big(\sup_{\varphi\in W_{k,M}} |V(0,\varphi)| > u^{1/n}\Big) \\ &\leqslant 2\int_0^\infty du\, \boldsymbol{P}\Big(\sup_{\varphi\in W_{k,M}} V(0,\varphi) > u^{1/n}\Big)\,. \end{aligned}$$

Denoting

$$E_1'' = \boldsymbol{E} \sup_{\varphi\in W_{k,M}} V(0,\varphi)\,,$$

we have

$$\begin{aligned} E_n' &\leqslant 2E_1''^n + 2\int_{E_1''^n}^\infty du\, \boldsymbol{P}\Big(\sup_{\varphi\in W_{k,M}} V(0,\varphi) - E_1'' > u^{1/n} - E_1''\Big) \\ &\leqslant 2E_1''^n + 2\int_{E_1''^n}^\infty du\, e^{-(u^{1/n}-E_1'')^2/(2\sigma_D^2)}\,, \end{aligned}$$

where the last line is given by [AT07, Theorem 2.1.1 p. 50] and

$$\sigma_D^2 \overset{\text{def}}{=} \sup_{\varphi\in W_{k,M}} \boldsymbol{E}V(0,\varphi)^2 = \sup_{\varphi\in W_{k,M}} \int dp\, r(p)\varphi(p)^2 \leqslant C_k^2 M^2 \int dp\, r(p)\,.$$

As a result,

$$\begin{aligned} E_n' &\leqslant 2E_1''^n + 2(n-1)\int_{E_1''}^\infty dv\, v^{n-1}\, e^{-(v-E_1'')^2/(2\sigma_D^2)} \\ &\leqslant 2E_1''^n + 2(n-1)\int_0^\infty dv\, v^{n-1}\, e^{-v^2/(2\sigma_D^2)}\,, \end{aligned}$$

which concludes the proof of the bound (3.8). □

Appendix E. Proof of Proposition 4.2

The estimates for a, $\partial_\theta a$, and $\partial_I a$ we prove in this section are based on the following estimates on the inverse map φ^{-1}.

Lemma E.1. *There exist $r > 0$ small enough, and a constant $C_r > 0$ such that for any $I \in (0, r)$*

$$\sup_{\theta\in\mathbb{T}} \Big(\|\varphi^{-1}(I,\theta)\| + \|\partial_\theta \varphi^{-1}(I,\theta)\| \Big) \leqslant C_r \sqrt{I}\,,$$

and

$$\sup_{\theta\in\mathbb{T}} \|\partial_I \varphi^{-1}(I,\theta)\| \leqslant \frac{C_r}{\sqrt{I}}\,.$$

The proof of this lemma is postponed to the end of this appendix.

In what follows we denote by

$$B_r = \varphi^{-1}((0,r)\times\mathbb{T})$$

with $r > 0$ small enough so that B_r describes a small neighborhood of $(0,0) \in \mathbb{R}^2$.

E.1. **Bound for a.** Let us remind the reader that

$$a(I,\theta) = e_2 \cdot \nabla\varphi_1(\varphi^{-1}(I,\theta)) = \partial_y \varphi_1(\varphi^{-1}(I,\theta))\,,$$

so that differentiating w.r.t. y the relation $K(I(x,y)) = H(x,y)$ we have

$$\partial_y I(x,y) = \partial_y \varphi_1(x,y) = \partial_y H(x,y)/\omega(I)\,.$$

Now, using that $\partial_y H(0,0) = 0$, we have for any $I \in (0,r)$

$$\begin{aligned} |a(I,\theta)| &\leqslant \omega_0^{-1} |\partial_y H(\varphi^{-1}(I,\theta)) - \partial_y H(0,0)| \\ &\leqslant \omega_0^{-1} \sup_{B_r} \|\nabla\partial_y H\| \cdot \|\varphi^{-1}(I,\theta)\| \leqslant C_r\sqrt{I}. \end{aligned}$$

E.2. **Bound for $\partial_\theta a$.** Using that

$$a(I,\theta) = \partial_y H(\varphi^{-1}(I,\theta))/\omega(I)\,, \tag{E.1}$$

and now differentiating w.r.t. θ, we obtain

$$\partial_\theta a(I,\theta) = \partial_\theta \varphi^{-1}(I,\theta) \cdot \nabla\partial_y H(\varphi^{-1}(I,\theta))/\omega(I)\,.$$

Then, by the Cauchy-Schwarz inequality

$$|\partial_\theta a(I,\theta)| \leqslant \omega_0^{-1} \|\partial_\theta \varphi^{-1}(I,\theta)\| \sup_{B_r} \|\nabla\partial_y H\| \leqslant C_r\sqrt{I}\,.$$

E.3. **Bound for $\partial_I a$.** Differentiating (E.1) w.r.t. I, we obtain

$$\begin{aligned} \partial_I a(I,\theta) &= -\frac{\omega'(I)}{\omega^2(I)} \partial_y H(\varphi^{-1}(I,\theta)) + \frac{1}{\omega(I)} \partial_I \varphi^{-1}(I,\theta) \cdot \nabla\partial_y H(\varphi^{-1}(I,\theta)) \\ &\stackrel{\text{def}}{=} T_1 + T_2\,. \end{aligned}$$

For T_1, using (4.12) and proceeding as for the function a to deal with the term $\partial_y H$ we have

$$|T_1| \leqslant \frac{C_r}{\sqrt{I}}\,.$$

For T_2, applying the Cauchy-Schwarz inequality together with Lemma E.1 yields

$$|T_2| \leqslant \frac{C'}{\sqrt{I}}\,,$$

which concludes the proof of Proposition 4.2.

E.4. **Proof of Lemma E.1.** Before going into the proof of the lemma, let us remind the reader that H admits a unique minimum at $(0,0)$ so that $\nabla^2 H(0,0)$ is a positive definite matrix

$$\nabla^2 H(0,0) = \begin{pmatrix} \lambda_1 & \lambda_{12} \\ \lambda_{12} & \lambda_2 \end{pmatrix}$$

with $\lambda_1, \lambda_2 > 0$. In this case, the inner product corresponding to this matrix defines a norm on $\mathbb{R}^2$ which is equivalent to the Euclidean norm $|\cdot|$. In other words, there exist two constants $\lambda, \bar{\lambda} > 0$ such that for any $X = (x,y)^T \in \mathbb{R}^2$

$$\lambda |X|^2 \leqslant X^T \nabla^2 H(0,0) X \leqslant \bar{\lambda} |X|^2 . \tag{E.2}$$

E.4.1. *Bound for φ^{-1}.* Since $(0,0)$ is the unique minimum to H, a Taylor expansion at the second order gives us

$$K(I) = H(X) = \frac{1}{2} X^T \nabla^2 H(0,0) X + o(|X|^2) , \tag{E.3}$$

with $X = (x,y)^T$. For any $X \in B_r$ we can have using (E.2) that

$$K(I) \geqslant \tilde{\lambda} |X|^2 \geqslant \tilde{\lambda} y^2 \tag{E.4}$$

with $\tilde{\lambda} = \lambda/2 - \mu > 0$ for some small μ, yielding directly

$$|\varphi_2^{-1}(I,\theta)| = |y| \leqslant \sqrt{K(I)/\tilde{\lambda}} = C\sqrt{K(I) - K(0)} \leqslant C \sup_{J \in (0,r)} \sqrt{\omega(J)} \sqrt{I} = C' \sqrt{I}.$$

In the same way, we also have

$$|\varphi_1^{-1}(I,\theta)| \leqslant C_r' \sqrt{I} .$$

E.4.2. *Bound for $\partial_\theta \varphi^{-1}$.* Differentiating w.r.t. x and then y the relation $K(I(X)) = H(X)$ we have

$$\partial_x \varphi_1(X) = \partial_x H(X)/\omega(I) \qquad \text{and} \qquad \partial_y \varphi_1(X) = \partial_y H(X)/\omega(I) ,$$

so that using (4.10) we obtain

$$\partial_\theta \varphi_2^{-1}(I,\theta) = \frac{\partial_x H(\varphi^{-1}(I,\theta))}{\omega(I)} \quad \text{and} \quad \partial_\theta \varphi_1^{-1}(I,\theta) = \frac{-\partial_y H(\varphi^{-1}(I,\theta))}{\omega(I)} . \tag{E.5}$$

Therefore, we have

$$\|\partial_\theta \varphi^{-1}(I,\theta)\| \leqslant \|\nabla H(\varphi^{-1}(I,\theta))\|/\omega_0 ,$$

with

$$\begin{aligned} \|\nabla H(\varphi^{-1}(I,\theta))\| &= \|\nabla H(\varphi^{-1}(I,\theta)) - \nabla H(0,0)\| \\ &\leqslant \sup_{B_r} \|\nabla^2 H\| \cdot \|\varphi^{-1}(I,\theta)\| \leqslant C_r \sqrt{I} . \end{aligned} \tag{E.6}$$

E.4.3. *Bound for $\partial_I \varphi^{-1}$.* Differentiating (E.5) w.r.t. I we obtain

$$\partial^2_{I\theta}\varphi_2^{-1}(I,\theta) = -\frac{\omega'(I)}{\omega^2(I)}\partial_x H(\varphi^{-1}(I,\theta)) + \frac{1}{\omega(I)}\partial_I\varphi^{-1}(I,\theta)\cdot\nabla\partial_x H(\varphi^{-1}(I,\theta))$$
$$\partial^2_{I\theta}\varphi_1^{-1}(I,\theta) = -\frac{\omega'(I)}{\omega^2(I)}\partial_y H(\varphi^{-1}(I,\theta)) + \frac{1}{\omega(I)}\partial_I\varphi^{-1}(I,\theta)\cdot\nabla\partial_y H(\varphi^{-1}(I,\theta))\,,$$

so that using (4.4) and (E.6)

$$|\partial_I\varphi_1^{-1}(I,\theta)| \leqslant \frac{C_1}{\sqrt{I}} + \frac{1}{\omega_0}\sup_{B_r}\|\nabla^2 H\|\int_0^\theta \|\partial_I\varphi^{-1}(I,\theta')\|d\theta'\,,$$

and in the same way for φ_2^{-1}

$$|\partial_I\varphi_2^{-1}(I,\theta)| \leqslant \frac{C_2}{\sqrt{I}} + |\partial_I\varphi_2^{-1}(I,0)| + \frac{1}{\omega_0}\sup_{B_r}\|\nabla^2 H\|\int_0^\theta \|\partial_I\varphi^{-1}(I,\theta')\|\,d\theta'.$$

Letting

$$U(I,\theta) \stackrel{\text{def}}{=} |\partial_I\varphi_1^{-1}(I,\theta)| + |\partial_I\varphi_2^{-1}(I,\theta)|\,,$$

and gathering the last two inequalities, we have

$$U(I,\theta) \leqslant \frac{C}{\sqrt{I}} + |\partial_I\varphi_2^{-1}(I,0)| + C'\int_0^\theta U(I,\theta')\,d\theta'\,.$$

Then, applying the Gronwall lemma, we obtain

$$\sup_{\theta\in(0,1)} U(I,\theta) \leqslant \Big(\frac{C}{\sqrt{I}} + |\partial_I\varphi_2^{-1}(I,0)|\Big)e^{C'}\,.$$

Moreover, differentiating w.r.t. I the relation

$$H(\varphi^{-1}(I,\theta)) = K(I)$$

we have

$$\partial_I\varphi_1^{-1}(I,\theta)\partial_x H(\varphi^{-1}(I,\theta)) + \partial_I\varphi_2^{-1}(I,\theta)\partial_y H(\varphi^{-1}(I,\theta)) = \omega(I)$$

and setting $\theta = 0$ we have from (4.3) and (4.4)

$$\partial_I\varphi_2^{-1}(I,0)\partial_y H(0,\varphi_2^{-1}(I,0)) = \omega(I)\,.$$

Also, using that

$$\partial_y H(0,y) = \partial_y H(0,0) + \partial^2_{yy}H(0,0)y + o(y) = \lambda_2 y + o(y)\,,$$

so that for any $(0,y)\in B_r$ we have, for some small μ,

$$|\partial_y H(0,y)| \geqslant (\lambda_2-\mu)|y|\,.$$

As a result,

$$|\partial_I\varphi_2^{-1}(I,0)| \leqslant \frac{\omega(I)}{(\lambda_2-\mu)|\varphi_2^{-1}(I,0)|}\,,$$

and from (4.3) together with (E.3), we have

$$K(I) = H(0,\varphi_2^{-1}(I,0)) \leqslant \bar\lambda_2\big(\varphi_2^{-1}(I,0)\big)^2,$$

with $\bar\lambda_2 = \lambda_2/2 + \mu$ so that

$$|\varphi_2^{-1}(I,0)| \geqslant \sqrt{K(I)/\bar\lambda_2} \geqslant \sqrt{c_1 I/\bar\lambda_2}\,.$$

Finally, we obtain

$$\sup_{\theta\in\mathbb{T}} |\partial_I \varphi_1^{-1}(I,\theta)| + |\partial_I \varphi_2^{-1}(I,\theta)| \leqslant \frac{C_r}{\sqrt{I}},$$

which gives the desired result and concludes the proof of the lemma.

References

[Arn89] V. I. Arnold. *Mathematical Methods of Classical Mechanics*. Springer, 1989.

[AT07] R. J. Adler and J. E. Taylor. *Random fields and geometry*. Springer Monographs in Mathematics. Springer, New York, 2007.

[Bak11] Y. Bakhtin. Noisy heteroclinic networks. *Probab. Theory Related Fields*, 150(1-2):1–42, 2011. doi:10.1007/s00440-010-0264-0.

[BGT89] N. H. Bingham, C. M. Goldie, and J. L. Teugels. *Regular variation*, volume 27 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1989.

[BIKS18] V. V. Buldygin, K.-H. Indlekofer, O. I. Klesov, and J. G. Steinebach. *Pseudo-regularly varying functions and generalized renewal processes*, volume 91 of *Probability Theory and Stochastic Modelling*. Springer, Cham, 2018. doi:10.1007/978-3-319-99537-3.

[Bil99] P. Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. doi:10.1002/9780470316962. A Wiley-Interscience Publication.

[Bre11] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Universitext. Springer, New York, 2011.

[BS67] M. Š. Birman and M. Z. Solomjak. Piecewise-polynomial approximations of functions of the classes w_α^p. *Mathematics of the USSR-Sbornik*, 2(3):295–317, apr 1967. doi:10.1070/sm1967v002n03abeh002343.

[CV21] N. Caillerie and J. Vovelle. Diffusion-approximation for a kinetic equation with perturbed velocity redistribution process. *Ann. Appl. Probab.*, 31(3):1299–1335, 2021. doi:10.1214/20-aap1619.

[DPZ96] G. Da Prato and J. Zabczyk. *Ergodicity for infinite-dimensional systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1996. doi:10.1017/CBO9780511662829.

[DRV21] A. Debussche, A. Rosello, and J. Vovelle. Diffusion-approximation for a kinetic spray-like system with random forcing. *Discrete & Continuous Dynamical Systems - S*, 14(8):2751–2803, 2021. doi:10.3934/dcdss.2021039.

[EK86] S. N. Ethier and T. G. Kurtz. *Markov processes: characterization and convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, Inc., New York, 1986. doi:10.1002/9780470316658.

[FGPSl07] J.-P. Fouque, J. Garnier, G. Papanicolaou, and K. Sø lna. *Wave propagation and time reversal in randomly layered media*, volume 56 of *Stochastic Modelling and Applied Probability*. Springer, New York, 2007.

[Fri06] A. Friedman. *Stochastic differential equations and applications*. Dover Publications, Inc., Mineola, NY, 2006. Two volumes bound as one, Reprint of the 1975 and 1976 original published in two volumes.

[FW93] M. I. Freidlin and A. D. Wentzell. Diffusion processes on graphs and the averaging principle. *Ann. Probab.*, 21(4):2215–2245, 1993. doi:10.1214/aop/1176989018.

[FW94] M. I. Freidlin and A. D. Wentzell. Random perturbations of Hamiltonian systems. *Mem. Amer. Math. Soc.*, 109(523):viii+82, 1994. doi:10.1090/memo/0523.

[FW98] M. Freidlin and M. Weber. Random perturbations of nonlinear oscillators. *Ann. Probab.*, 26(3):925–967, 1998. doi:10.1214/aop/1022855739.

[Gar97] J. Garnier. Multi-scaled diffusion-approximation. Applications to wave propagation in random media. *ESAIM Probab. Statist.*, 1:183–206, 1995/97. doi:10.1051/ps:1997107.

[HIK+18] M. Hairer, G. Iyer, L. Koralov, A. Novikov, and Z. Pajor-Gyulai. A fractional kinetic process describing the intermediate time behaviour of cellular flows. *Ann. Probab.*, 46(2):897–955, 2018. doi:10.1214/17-AOP1196.

[HKPG16] M. Hairer, L. Koralov, and Z. Pajor-Gyulai. From averaging to homogenization in cellular flows—an exact description of the transition. *Ann. Inst. Henri Poincaré Probab. Stat.*, 52(4):1592–1613, 2016. doi:10.1214/15-AIHP690.

[HL20] M. Hairer and X.-M. Li. Averaging dynamics driven by fractional brownian motion. *The Annals of Probability*, 48(4):1826 – 1860, 2020. doi:10.1214/19-AOP1408.

[HV14a] P. H. Haynes and J. Vanneste. Dispersion in the large-deviation regime. Part 1: shear flows and periodic flows. *J. Fluid Mech.*, 745:321–350, 2014. doi:10.1017/jfm.2014.64.

[HV14b] P. H. Haynes and J. Vanneste. Dispersion in the large-deviation regime. Part 2. Cellular flow at large Péclet number. *J. Fluid Mech.*, 745:351–377, 2014. doi:10.1017/jfm.2014.65.

[KNR14] T. Komorowski, A. Novikov, and L. Ryzhik. Homogenization driven by a fractional Brownian motion: the shear layer case. *Multiscale Model. Simul.*, 12(2):440–457, 2014. doi:10.1137/13092068X.

[KS91] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991. doi:10.1007/978-1-4612-0949-2.

[Kus84] H. J. Kushner. *Approximation and weak convergence methods for random processes, with applications to stochastic systems theory*. MIT Press Series in Signal Processing, Optimization, and Control, 6. MIT Press, Cambridge, MA, 1984.

[Mar05] R. Marty. Asymptotic behavior of differential equations driven by periodic and random processes with slowly decaying correlations. *ESAIM Probab. Stat.*, 9:165–184 (electronic), 2005. doi:10.1051/ps:2005009.

[PS08] G. A. Pavliotis and A. M. Stuart. *Multiscale methods – Averaging and homogenization*, volume 53 of *Texts in Applied Mathematics*. Springer, New York, 2008.

[PSV76] G. C. Papanicolaou, D. Stroock, and S. R. S. Varadhan. Martingale approach to some limit theorem. *Duke Univ. Math. Series III*, IV:1–120, 1976.

[Roz87] Y. A. Rozanov. *Introduction to random processes*. Springer Series in Soviet Mathematics. Springer-Verlag, Berlin, 1987. doi:10.1007/978-3-642-72717-7. Translated from the Russian by Birgit Röthinger.

[SV06] D. W. Stroock and S. R. S. Varadhan. *Multidimensional diffusion processes*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1997 edition.

[Taq77] M. S. Taqqu. Law of the iterated logarithm for sums of non-linear functions of Gaussian variables that exhibit a long range dependence. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 40(3):203–238, 1977. doi:10.1007/BF00736047.

[Taq75] M. S. Taqqu. Weak convergence to fractional Brownian motion and to the Rosenblatt process. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 31:287–302, 1974/75. doi:10.1007/BF00532868.

[TV16] A. Tzella and J. Vanneste. Dispersion in rectangular networks: Effective diffusivity and large-deviation rate function. *Phys. Rev. Lett.*, 117:114501, Sep 2016. doi:10.1103/PhysRevLett.117.114501.

[YJ91] W. R. Young and S. Jones. Shear dispersion. *Physics of Fluids A: Fluid Dynamics (1989-1993)*, 3(5):1087–1101, 1991.

[You88] W. R. Young. Arrested shear dispersion and other models of anomalous diffusion. *J. Fluid Mech.*, 193:129–149, Aug 1988. doi:10.1017/S0022112088002083.

Aix Marseille Univ, CNRS, I2M, Marseille, France
Email address: christophe.gomez@univ-amu.fr

Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA
Email address: gautam@math.cmu.edu

Department of Mathematics, Pennsylvania State University, State College PA 16802
Email address: hvl2@psu.edu

Department of Mathematics, Pennsylvania State University, State College PA 16802
Email address: anovikov@math.psu.edu