HAL
open science

Quantitative stochastic homogenization of viscous Hamilton-Jacobi equations

Scott Armstrong, Pierre Cardaliaguet

▶ **To cite this version:**

Scott Armstrong, Pierre Cardaliaguet. Quantitative stochastic homogenization of viscous Hamilton-Jacobi equations. Communications in Partial Differential Equations, 2015, 40 (3), pp.540-600. 10.1080/03605302.2014.971372 . hal-00922714

HAL Id: hal-00922714

https://hal.science/hal-00922714v1

Submitted on 30 Dec 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

QUANTITATIVE STOCHASTIC HOMOGENIZATION OF VISCOUS HAMILTON-JACOBI EQUATIONS

SCOTT N. ARMSTRONG AND PIERRE CARDALIAGUET

Abstract. We prove explicit estimates for the error in random homogenization of degenerate, second-order Hamilton-Jacobi equations, assuming the coefficients satisfy a finite range of dependence. In particular, we obtain an algebraic rate of convergence with overwhelming probability under certain structural conditions on the Hamiltonian.

1. Introduction

1.1. **Motivation and informal statement of results.** The paper is concerned with second-order (i.e., "viscous") Hamilton-Jacobi equations of the form

$$u_t^\varepsilon - \varepsilon \operatorname{tr}\left(A\left(\frac{x}{\varepsilon}\right) D^2 u^\varepsilon\right) + H\left(Du^\varepsilon, \frac{x}{\varepsilon}\right) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty). \tag{1.1}$$

Here $\varepsilon > 0$ is a small parameter which we will send to zero, the Hamiltonian $H = H(p, y)$ is convex and coercive in p and $A(y)$ is a diffusion matrix which is possibly degenerate. In addition, the coefficients H and A are *random fields*, that is, sampled by a probability measure on the space of all such coefficients.

The basic *qualitative* result concerning the stochastic homogenization of (1.1) states roughly that, under appropriate assumptions on the probability measure $\mathbb{P}$ (typically that $\mathbb{P}$ is *stationary* and *ergodic* with respect to translations on $\mathbb{R}^d$), there exists a deterministic, convex *effective Hamiltonian* $\overline{H}$ such that the solutions u^ε of (1.1) converge uniformly as $\varepsilon \to 0$, with probability one, to the solution of the effective equation

$$u_t + \overline{H}(Du) = 0 \quad \text{in } \mathbb{R}^d \times (0,\infty). \tag{1.2}$$

A formulation of this result was first proved independently by Souganidis [19] and Rezakhanlou and Tarver [18] for first order equations Hamilton-Jacobi equations and later extended to the viscous setting by Lions and Souganidis [13] and by Kosygina, Rezakhanlou and Varadhan [11]. Generalizations as well as new proofs of these results, using methods that are closer to our perspective in this paper, appeared in [3, 4, 5].

The first complete quantitative homogenization results for Hamilton-Jacobi equations in the stochastic setting were presented in [2]. Previous results in this direction were obtained by Rezakhanlou [17], who gave structural conditions on Hamilton-Jacobi equations in dimension $d = 1$ in which a central limit theorem holds, and by Matic and Nolen [15], who proved a estimate for the random part of the error for a particular class

Date: December 30, 2013.

2010 *Mathematics Subject Classification.* 35B27, 35F21, 60K35.

Key words and phrases. stochastic homogenization, error estimate, convergence rate, viscous Hamilton-Jacobi equation, first-passage percolation.

of equations. In each of these papers, the main assumption quantifying the ergodicity of $\mathbb{P}$ is a finite range dependence hypothesis in the spatial dependence of the coefficients: this roughly means that there is a length scale $\ell > 0$ such that the random elements $(A(x), H(p,x))$ and $(A(y), H(q,y))$ are independent whenever $|x-y| \geq \ell$.

Each of the quantitative homogenization results mentioned in the previous paragraph concerned first-order equations, that is, the case that $A \equiv 0$. Working with a general degenerate diffusion matrix A is much more difficult from the point of view of homogenization. On the one hand, even if A is uniformly elliptic there is no useful regularizing effect, since we send $\varepsilon \to 0$. On the other hand, the presence of A makes the dependence of the solutions on the coefficients much more complicated, as it destroys the finite speed of propagation property possessed by first-order Hamilton-Jacobi equations. The latter property is essentially what "localizes" the problem and allows one to apply the independence property, and it is this idea that played a crucial role in [2].

In this paper, we develop a quantitative theory of stochastic homogenization for (1.1) with a general degenerate diffusion matrix A. In particular, we prove an algebraic convergence rate, with overwhelming probability, as $\varepsilon \to 0$. Roughly, we show that there exists an exponent $\alpha \in (0,1)$, which is given explicitly, such that

$$\mathbb{P}\left[\sup_{(x,t)\in B_T\times[0,T]} |u^\varepsilon(x,t) - u(x,t)| \geq \varepsilon^\alpha\right] \lesssim \exp\left(-\varepsilon^{-\alpha}\right).$$

As in [2], our arguments are inspired by probability techniques developed in first passage percolation by Kesten [10], Alexander [1] and Zhang [22] and in the study of the fluctuations of the Lyapunov exponents for Brownian motion in Poissonian potentials, developed by Sznitman [20] and by Wüthrich [21]. In fact, perhaps the closest previous work to ours is that of Sznitman [20], who proved a special case of our main result, Theorem 1. See the discussion following the statement of Theorem 1, below.

Viscous Hamilton-Jacobi equations like (1.1) with random coefficients arise in particular in the study of large deviations of diffusions in random environments. Although we do not explore this topic in detail here, the main results in this paper imply quantitative bounds on the rate of the large deviations principle.

1.2. **Hypotheses.** We proceed with the precise formulation of our results, beginning with the definition of the probability space Ω, which is taken to be "the set of all coefficients" for the equation (1.1) subject to some particular structural conditions. We then define a random environment by putting a probability measure on this set. Throughout the paper, we fix the parameters $q > 1$, $n \in \mathbb{N}^*$ and $\Lambda \geq 1$.

We consider diffusion matrices $A : \mathbb{R}^d \to \mathbb{S}^d$ (here $\mathbb{S}^d$ denotes the set of d-by-d real symmetric matrices) which have Lipschitz continuous square roots, that is, we assume

$$A = \frac{1}{2}\Sigma^t\Sigma, \qquad (\Sigma^t := \text{ the transpose of } \Sigma) \tag{1.3}$$

for a matrix-valued function $\Sigma : \mathbb{R}^d \to \mathbb{R}^{n\times d}$ which satisfies, for every $y, z \in \mathbb{R}^d$,

$$\frac{1}{2}\left|\Sigma(y)\right|^2 \leq \Lambda \tag{1.4}$$

and

$$|\Sigma(y) - \Sigma(z)| \leq \Lambda |y - z|. \tag{1.5}$$

The Hamiltonian $H : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$ is required to satisfy the following: for every $y \in \mathbb{R}^d$,

$$p \mapsto H(p, y) \quad \text{is convex.} \tag{1.6}$$

For every $p, \widetilde{p} \in \mathbb{R}^d$ and $y, z \in \mathbb{R}^d$,

$$\frac{1}{\Lambda}|p|^q - \Lambda \leq H(p, y) \leq \Lambda |p|^q + \Lambda, \tag{1.7}$$

$$|H(p, y) - H(p, z)| \leq \Lambda\big(|p|^q + 1\big)|y - z|, \tag{1.8}$$

and

$$|H(p, y) - H(\widetilde{p}, y)| \leq \Lambda\big(|p| + |\widetilde{p}| + 1\big)^{q-1}|p - \widetilde{p}|. \tag{1.9}$$

Finally, we require the following additional structural condition: for every $p, y \in \mathbb{R}^d$,

$$H(p, y) \geq H(0, y) \quad \text{and} \quad \sup_{z \in \mathbb{R}^d} H(0, z) = 0. \tag{1.10}$$

We denote by Ω the set of ordered pairs (Σ, H) of functions as described above:

$$\Omega := \big\{(\Sigma, H) \,:\, \Sigma \text{ and } H \text{ satisfy (1.4), (1.5) (1.6), (1.7), (1.8), (1.9) and (1.10)}\big\}.$$

Throughout, we take A to be the $\mathbb{S}^d$–valued random field on Ω defined by (1.3). We endow Ω with the following σ–algebras. For each Borel set $U \subseteq \mathbb{R}^d$, we define $\mathcal{F}(U)$ by:

$$\mathcal{F}(U) := \sigma\text{–algebra generated by the maps } (\Sigma, H) \mapsto \Sigma(y) \text{ and } (\Sigma, H) \mapsto H(p, y), \quad \text{with } p \in \mathbb{R}^d \text{ and } y \in U.$$

The largest of these σ–algebras is denoted by $\mathcal{F} := \mathcal{F}(\mathbb{R}^d)$.

We model the random environment by a given probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, which is assumed to be *stationary* with respect to translations and have a *finite range of dependence*. The expectation with respect to $\mathbb{P}$ is denoted by $\mathbb{E}$. To state the stationary assumption, we let $\{\tau_z\}_{z \in \mathbb{R}^d}$ be the group action of translation on Ω, that is, for each $z \in \mathbb{R}^d$, we define the map $\tau_z : \Omega \to \Omega$ by

$$\tau_z(\Sigma, H) := (\tau_z \Sigma, \tau_z H)\,, \quad \text{with} \;\; (\tau_z \Sigma)(y) := \Sigma(y + z) \;\; \text{and} \;\; (\tau_z H)(p, y) := H(p, y + z).$$

We extend the action of translation to $\mathcal{F}$ by defining $\tau_z F$, for each $F \in \mathcal{F}$, by

$$\tau_z F := \big\{\tau_z \omega \,:\, \omega \in F\big\}.$$

With the above notation, the hypotheses on $\mathbb{P}$ are that

$$\text{for every } y \in \mathbb{R}^d \text{ and } F \in \mathcal{F}, \qquad \mathbb{P}\left[\tau_y F\right] = \mathbb{P}\left[F\right] \tag{1.11}$$

and

$$\text{for all Borel sets } U, V \subseteq \mathbb{R}^d \;\; \text{such that} \;\; \operatorname{dist}(U, V) \geq 1, \quad \mathcal{F}(U) \text{ and } \mathcal{F}(V) \text{ are } \mathbb{P}\text{–independent.} \tag{1.12}$$

Here $\operatorname{dist}(U, V) := \inf_{x \in U,\, y \in V} |x - y|$. The first assumption (1.11) states that the statistics of the coefficients are the same everywhere in $\mathbb{R}^d$. The second assumption (1.12)

is a quantitative ergodicity assumption, which requires independence, in well-separated regions of space, of coefficients sampled by $\mathbb{P}$.

If X is a random variable, then we may write $X(\omega)$ if we wish to display its dependence on $\omega \in \Omega$ explicitly. However, we avoid this unless it is required for clarity. For instance, we usually write $u^\varepsilon(x,t)$ for the solution of (1.1) (e.g., as in the statement of Theorem 3), but we sometimes use the notation $u^\varepsilon(x,t,\omega)$ for the same quantity.

With the important exception of (1.10), the structural hypotheses on the coefficients comprise a fairly general model of a viscous Hamilton-Jacobi equation. On the other hand, (1.10) is rather restrictive: e.g., it is not satisfied by a Hamiltonians such as $H(p,y) = |p|^2 - b(y)\cdot p$ where $b : \mathbb{R}^d \to \mathbb{R}^d$ is a random vector field. It would be very interesting if quantitative homogenization results like the ones presented here could be proved without (1.10). We expect this to be quite difficult and that the statement of any such results would likely be weaker. Indeed, and as we will see, the hypothesis plays an essential role in the analysis of the metric problem by forcing some monotonicity in the level sets of the solutions. Without such monotonicity, the problem becomes much more complicated. Note also that (1.10) is natural from the control theoretic viewpoint: see the discussion in [2] for the first-order case, where the same assumption is made.

1.3. **Quantitative homogenization of the metric problem.** As shown in [4, 5], the heart of the modern theory of stochastic homogenization of Hamilton-Jacobi equations is the so-called *metric problem*. For $\mu > 0$ and $x \in \mathbb{R}^d$, we consider the unique solutions $m_\mu(\cdot,x) \in C^{0,1}_{\mathrm{loc}}(\mathbb{R}^d)$ of

$$\begin{cases} -\operatorname{tr}\left(A(y)D^2 m_\mu(\cdot,x)\right) + H(Dm_\mu(\cdot,x),y) = \mu & \text{in } \mathbb{R}^d \setminus \overline{B}_1(x), \\ m_\mu(\cdot,x) = 0 & \text{on } \overline{B}_1(x). \end{cases} \tag{1.13}$$

The problem (1.13) inherits its name from the subadditivity of its solutions (which also explains its usefulness) and the fact that $m_\mu(y,z)$ has a probabilistic interpretation roughly as a cost of transporting a particle from y to z, under a certain cost functional corresponding to the coefficients. It can be shown in more generality than we consider here (see [5]) that the solutions m_μ of (1.13) have deterministic limits: for every $\mu > 0$, there exists a positively homogeneous, convex function $\overline{m}_\mu \in C(\mathbb{R}^d)$ such that

$$\mathbb{P}\left[\sup_{R>0}\sup_{y,z\in B_R}\lim_{t\to\infty}\frac{1}{t}\left|m_\mu(ty,tz) - \overline{m}_\mu(ty-tz)\right| = 0\right] = 1. \tag{1.14}$$

As it turns out (c.f. [5]), the limit (1.14) is strong enough to imply a full homogenization result for viscous Hamilton-Jacobi equations. The effective Hamiltonian $\overline{H}$ can be expressed in terms of $\overline{m}_\mu$ by

$$\overline{H}(p) = \inf\left\{\mu > 0 \,:\, \text{for every } y \in \mathbb{R}^d,\ \overline{m}_\mu(y) \geq p\cdot y\right\}. \tag{1.15}$$

That is, $\overline{m}_\mu$ describes the μ–sublevel set of $\overline{H}$ because, as expected, it is the solution of the homogenized metric problem. See the next section for the definition and basic properties of m_μ.

As (1.14) is of fundamental importance to the qualitative homogenization theory, quantifying it is a central task for a quantitative theory. This is our first main result,

which is actually the principal result of the paper and contains most of the difficulty in building a quantitative theory of (1.1). The rest of the main results, presented in Theorems 2 and 3 below, quantifies how the metric problem controls the solutions of the approximate cell problem and of (1.1). These are more straightforward to obtain, since they follow from deterministic comparison (i.e., pure PDE) arguments.

Theorem 1. *Fix $\mu_0 \geq 1$. There exists a constant $C > 0$, depending only on (d, q, Λ, μ_0) such that, for every $0 < \mu \leq \mu_0$, $\lambda > C\mu^{-2}$ and $y \in \mathbb{R}^d$,*

$$\mathbb{P}\Big[m_\mu(y,0) - \overline{m}_\mu(y) \leq -\lambda \Big] \leq \exp\left(-\frac{\mu^4 \lambda^2}{C|y|} \right), \tag{1.16}$$

and, if

$$\lambda \geq C \left(\frac{|y|^{\frac{2}{3}}}{\mu^2} + \frac{|y|^{\frac{1}{3}}}{\mu^4} \right) \log\left(2 + \frac{|y|}{\mu} \right), \tag{1.17}$$

then

$$\mathbb{P}\Big[m_\mu(y,0) - \overline{m}_\mu(y) \geq \lambda \Big] \leq \exp\left(-\frac{\mu^4 \lambda^2}{C|y|} \right). \tag{1.18}$$

To see how Theorem 1 quantifies (1.14), note that (at least for $\mu > 0$) the right sides of (1.16) and (1.18) are very small for $\lambda \gg |y|^{1/2}$. The first inequality therefore implies roughly that, for $t \gg 1$,

$$\frac{1}{t}\left(m_\mu(ty,0) - \overline{m}_\mu(ty) \right) \ll -t^{-1/2} \quad \text{has overwhelmingly small probability.}$$

To interpret the second statement, note that for fixed $\mu > 0$ the first term in the parentheses dominates for large $|y|$. We obtain roughly that, for every $\alpha > 0$ and $t \gg 1$,

$$\frac{1}{t}\left(m_\mu(ty,0) - \overline{m}_\mu(ty) \right) \gg t^{-1/3+\alpha} \quad \text{has overwhelmingly small probability.}$$

Note that (1.16) and (1.18) both degenerate as $\mu \to 0$. This is necessarily so: see the discussion in [2], which encounters the same phenomenon in the first-order case.

The only previous result like Theorem 1 in the case $A \not\equiv 0$ is due to Sznitman [20]. The main result of his paper is an error estimate for "metrics" associated to the long-time behavior of Brownian motion in the presence of truncated Poissonian obstacles. These distance functions, translated into our language (see [5] for the precise translation and more details on this connection), are nothing other than the functions m_μ in the special case that $A(x) = \frac{1}{2} I_d$ (here I_d is the identity matrix) and the Hamiltonian has the form $H(p,y) = \frac{1}{2}|p|^2 - V(y)$, with V the truncated Poissonian potential. Compared to [20], we are able to consider more general potentials: for example, the potential V in [20] is generated by radially symmetric bump function, here we do not need such symmetry (but we pay a price with a slightly weakened estimate). The argument in [20] achieves the localization property due to special properties of Brownian motion and a probabilistic argument which analyzes the sample paths of the diffusion.

In view of (1.15), Theorem 1 gives explicit error bounds for a numerical method for computing the effective nonlinearity $\overline{H}$. The task is reduced to the numerical computation of the maximal subsolutions m_μ, and this can be further reduced to a problem on a finite domain by the results in Section 3, see in particular Lemma 3.3.

The proof of Theorem 1 is broken into two main steps: first, an estimate on the likelihood that the random variable $m_\mu(y,0)$ is far from its mean $\mathbb{E}\left[m_\mu(y,0)\right]$, which is proved in Section 3, and second, an estimate of the difference between $\mathbb{E}\left[m_\mu(y,0)\right]$ and $\overline{m}_\mu(y,0)$, which is proved in Section 4. The proof of Theorem 1 itself is given at the beginning of Section 4.

1.4. **Quantitative results on the approximate cell problem.** After proving Theorem 1, the rest of the paper is concerned primarily with transferring the error estimates for the metric problem to error estimates for the time-dependent initial-value problem. As an intermediate step, we consider the time-independent *approximate cell problem*: for each $\delta > 0$ and $p \in \mathbb{R}^d$, we consider the unique solution $v^\delta(\cdot,p) \in C^{0,1}(\mathbb{R}^d)$ of

$$\delta v^\delta(y,p) - \operatorname{tr}\left(A(y)D^2v^\delta(y,p)\right) + H(Dv^\delta(y,p)+p,y) = 0 \qquad \text{in } \mathbb{R}^d. \tag{1.19}$$

Results concerning the well-posedness of (1.19) can be found in [6]. This problem arises naturally in the qualitative theory of homogenization of (1.1), and in fact the general homogenization theorem is equivalent to the statement that, for every $p \in \mathbb{R}^d$,

$$\mathbb{P}\left[\forall p \in \mathbb{R}^d,\ \limsup_{\delta\to 0}\left|\delta v^\delta(0,p) + \overline{H}(p)\right| = 0\right] = 1. \tag{1.20}$$

Our next result, Theorem 2 below, is a quantitative version of (1.20).

The rate of convergence of the limit in (1.20) cannot be quantified without further assumptions on the law of the Hamiltonian. Indeed, it turns out that the mixing properties of the environment control the rate at which $\delta v^\delta(0,p) + \overline{H}(p)$ converges to zero *from above* for all $p \in \mathbb{R}^d$, as well as the rate at which this quantity converges to zero from below for p's satisfying $\overline{H}(p) > 0$. However, the rate at which we can expect $\delta v^\delta(0,p) + \overline{H}(p)$ to converge to zero for $p \in \operatorname{int}\{\overline{H} = 0\}$ is determined by the behavior of H near its maximum: in view of (1.10), what is important is how large the probability $\mathbb{P}\left[H(0,0) > -\delta\right]$ is for $0 < \delta \ll 1$. This phenomenon was already encountered in [2], where an explicit example was given showing that the rate of the limit may be arbitrarily slow for p belonging to the "flat spot," defined as the interior of the level set $\{\overline{H} = 0\}$.

For simplicity, we quantify (1.20) under an assumption that rules out the existence of such a flat spot, although an inspection of the arguments in Section 5 yield more precise results (e.g., estimates for p's not belonging to the flat spot without this assumption).

Theorem 2. *Assume that the lower bound in* (1.7) *is replaced by*

$$H(p,x) \geq \frac{1}{\Lambda}|p|^q. \tag{1.21}$$

Fix $\xi \geq 1$. Then there exists $C > 0$, depending on (d,q,Λ,ξ), such that, for every $p \in \mathbb{R}^d$ with $|p| \leq \xi$ and $\delta,\lambda > 0$ satisfying

$$\lambda \geq C\delta^{\frac{1}{7+6q}}\left(1 + |\log\delta|\right),$$

we have

$$\mathbb{P}\Big[\ \big|\delta v^\delta(0,p)+\overline{H}(p)\big|\geq\lambda\Big]\leq C\lambda^{-d}\left(\lambda^{-2d}+\delta^{-2d}\right)\exp\left(-\frac{\lambda^{4(1+q)}}{C\delta}\right). \tag{1.22}$$

Notice that Theorem 2 quantifies the limit (1.20) since it implies roughly that, for every $\alpha>0$ and sufficiently small $\delta>0$,

$$\big|\delta v^\delta(0,p)+\overline{H}(p)\big|\gg\delta^{\frac{1}{7+6q}+\alpha}\quad\text{has overwhelmingly small probability.}$$

While the exponent $(7+6q)^{-1}$ is far from optimal, this quantifies the convergence with an algebraic rate.

The proof of Theorem 1.22 (given at the end of Section 5) differs substantially from the analogue in [2]. The idea, as in [2], is to link the metric problem and the approximate cell problem and then apply Theorem 1. Howeover, the connection between the two problems relies on comparison and convex geometry arguments recently introduced in [5], which account for the first proof of qualitative homogenization for (1.1) based only on the metric problem. Previous arguments were either for first–order equations [4], relied on representation formulae [11, 13], or were based on weak convergence techniques [14, 3].

1.5. **Quantitative results on the time dependent problem.** The third and final result we present in the introduction relates to the time-dependent initial-value problem: we study the convergence rate for the solution $u^\varepsilon=u^\varepsilon(x,t)$ of

$$\begin{cases} u^\varepsilon_t-\varepsilon\operatorname{tr}\left(A\left(\frac{x}{\varepsilon}\right)D^2u^\varepsilon\right)+H\left(Du^\varepsilon,\frac{x}{\varepsilon}\right)=0 & \text{in } \mathbb{R}^d\times(0,\infty),\\ u^\varepsilon(x,0)=g(x) & \text{in } \mathbb{R}^d\end{cases} \tag{1.23}$$

to the solution u of the homogenized problem

$$\begin{cases} u_t+\overline{H}(Du)=0 & \text{in } \mathbb{R}^d\times(0,\infty),\\ u(x,0)=g(x) & \text{in } \mathbb{R}^d\end{cases} \tag{1.24}$$

Here $g\in C^{0,1}(\mathbb{R}^d)$ is given. Results implying the well-posedness of (1.23) and (1.24) can be found in [6].

Theorem 3 (Error estimate for the time dependent problem)**.** *Assume the hypotheses of Theorem 1.21. Let $g\in C^{1,1}(\mathbb{R}^d)$. Then there exists a constant $C>0$, depending only on $(d,q,\Lambda,\|g\|_{C^{1,1}(\mathbb{R}^d)})$ such that, for any $\lambda,\varepsilon>0$ satisfying*

$$\lambda\geq C\varepsilon^{\frac{1}{17+12q}}\left(1+|\log\varepsilon|\right)$$

we have

$$\mathbb{P}\Big[\sup_{x\in B_T,\ t\in[0,T]}|u^\varepsilon(x,t)-u(x,t)|\geq\lambda\Big]\leq C(\lambda\varepsilon)^{-3d}\exp\left(-\frac{\lambda^{\frac{11}{2}+4q}}{C\varepsilon^{\frac{1}{2}}}\right). \tag{1.25}$$

Notice that the previous theorem implies roughly that, for every $\alpha>0$ and sufficiently small $\varepsilon>0$,

$$|u^\varepsilon(x,t)-u(x,t)|\gg\varepsilon^{\frac{1}{17+12q}+\alpha}\quad\text{has overwhelmingly small probability.}$$

To prove Theorem 3, we compare the solutions $v^\delta(\cdot,p)$ of (1.19) to the solutions of (1.23) and then apply the result of Theorem 2. That is, the proof is a deterministic, pure PDE comparison argument which, while technical, is relatively straightforward. Such arguments appeared first in the context of periodic homogenization of first-order equations in Capuzzo–Dolcetta and Ishii [7], and a similar technique is used for the proof of [2, Theorem 4], also for first order HJB equations. Here the techniques required are a bit more complicated because the comparison machinery for viscosity solutions of second-order equations is more involved. The comparison argument is summarized in Lemma 6.1 and the proof of Theorem 3 is given at the end of Section 6.

1.6. **Outline of the paper.** Section 2 is a summary of the basic (deterministic) properties of the metric problem. The proof of Theorem 1 is split between an estimate of the random error (Section 3) and an estimate of the non-random error (Section 4). Theorem 2 is proved in Section 5 and Theorem 3 in Section 6.

2. Basic properties of the metric problem

We review some basic facts, needed throughout the paper, concerning the equation

$$-\operatorname{tr}\big(A(y)D^2w\big) + H(Dw,y) = \mu. \tag{2.1}$$

Proofs of most of the results collected here can be found in [6]. As the statements are deterministic, throughout this section we fix $(\Sigma, H) \in \Omega$.

The basic regularity result for coercive Hamilton-Jacobi equations is the interior Lipschitz continuity of solutions. For a proof of the following proposition with an explicit Lipschitz constant depending on the given parameters, see [6, Theorem 3.1].

Proposition 2.1. *Suppose that $\mu \geq 0$ and $u \in C(B_1)$ is a solution of (2.1) in B_1. Then there exists $L_\mu > 0$, depending on (q,Λ) and an upper bound for μ, such that*

$$\sup_{x,y\in B_{1/2},\, x\neq y} \frac{|u(x)-u(y)|}{|x-y|} \leq L_\mu.$$

We next review some properties of the metric problem. The solutions of (1.13) can be characterized as the *maximal subsolutions* of (2.1), subject to the constraint of being nonpositive in $\overline{B}_1(x)$. In other words, we define $m_\mu(y,x)$ for each $x,y\in\mathbb{R}^d$ by

$$m_\mu(y,x) := \sup\big\{w(y) \,:\, w \in \mathrm{USC}(\mathbb{R}^d) \text{ is a subsolution of (2.1) in } \mathbb{R}^d \text{ and } w \leq 0 \text{ on } \overline{B}_1(x)\big\}. \tag{2.2}$$

Here $\mathrm{USC}(X)$ denotes the set of upper semicontinuous functions on X, which is the appropriate space for viscosity subsolutions (c.f. [6]). It is convenient to extend the definition of m_μ from $\{x\}$ to arbitrary compact $K \subseteq \mathbb{R}^d$: we define, for every $y \in \mathbb{R}^d$,

$$m_\mu(y,K) := \sup\big\{w(y) \,:\, w \in \mathrm{USC}(\mathbb{R}^d) \text{ is a subsolution of (2.1) in } \mathbb{R}^d \text{ and } w \leq 0 \text{ on } K+\overline{B}_1\big\}. \tag{2.3}$$

Note that $m_\mu(y,x) = m_\mu(y,\{x\})$. The basic properties of m_μ are summarized in the following proposition.

Proposition 2.2. *Let $\mu > 0$ and K be compact subset of $\mathbb{R}^d$.*

(i) *The function $m_\mu(\cdot, K)$ is a solution of*

$$\begin{cases} -\operatorname{tr}\big(A(y)D^2 m_\mu(\cdot,K)\big) + H(Dm_\mu(\cdot,K),y) = \mu & \text{in } \mathbb{R}^d \setminus (K+\overline{B}_1), \\ m_\mu(\cdot,K) = 0 & \text{in } K+\overline{B}_1. \end{cases} \tag{2.4}$$

(ii) *For every $y, z \in \mathbb{R}^d$ and K, K' compact subsets of $\mathbb{R}^d$,*

$$|m_\mu(y,K) - m_\mu(z,K')| \le L_\mu \left(|y-z| + \operatorname{dist}_H(K,K')\right). \tag{2.5}$$

(iii) *For $x, y, z \in \mathbb{R}^d$,*

$$m_\mu(y,x) \le m_\mu(y,z) + m_\mu(z,x) + L_\mu. \tag{2.6}$$

(iv) *There exist $0 < l_\mu \le L_\mu$ such that, for some $C, c > 0$ depending only on an upper bound for μ,*

$$c\mu \le l_\mu \le L_\mu \le C \tag{2.7}$$

and

$$l_\mu \operatorname{dist}(y, B(x,1)) \le m_\mu(y,x) \le L_\mu \operatorname{dist}(y, B(x,1)). \tag{2.8}$$

(v) *Let K be a compact subset of $\mathbb{R}^d$. Then*

$$m_\mu(y,K) \ge l_\mu(\operatorname{dist}(y,K) - 2). \tag{2.9}$$

Proof. (i) and (ii) for $K = K'$ are proved in [6]. To show (2.5), it just remains to check that

$$|m_\mu(y,K) - m_\mu(y,K')| \le L_\mu \operatorname{dist}_H(K,K'). \tag{2.10}$$

We first note that $\operatorname{dist}_H(K+\overline{B}_1, K'+\overline{B}_1) \le \operatorname{dist}_H(K,K')$. Let $w(y) = m_\mu(y,K) - L_\mu \operatorname{dist}_H(K,K')$. Then w is a subsolution to (2.1). Moreover, using the Lipschitz estimate (2.5) for $K = K'$ and the fact that $m_\mu(y,K) \le 0$ in $K + \overline{B}_1$,

$$w(y) = m_\mu(y,K) - L_\mu \operatorname{dist}_H(K,K') \le 0 \qquad \forall y \in K' + \overline{B}_1.$$

By definition of $m_\mu(\cdot, K')$, this implies that $m_\mu(\cdot,K') \ge w$. Then (2.10) follows.

(iii) Let $w(y) = m_\mu(y,x) - m_\mu(z,x) - L_\mu$. Then w is a subsolution to (2.1) in $\mathbb{R}^d$ which satisfies thanks to (2.5),

$$w(y) \le m_\mu(z,x) + L_\mu|z-y| - m_\mu(z,x) - L_\mu \le 0 \qquad \forall y \in B_1(z).$$

This implies (2.6) by definition of $m_\mu(\cdot, z)$.

For proving (iv), first note that $y \to l_\mu(|y-x|-1)_+$ is a subsolution of (2.1) for l_μ small enough: hence the left-hand side of (2.7). As $y \to L_\mu(|y-x|-1)_+$ is a supersolution (2.4) for L_μ large enough, we get the right-hand side of (2.7) by comparison (see [6]).

(v) Let $\xi : \mathbb{R}^d \to \mathbb{R}$ be a standard mollification kernel and denote, for $\varepsilon > 0$, $\xi_\varepsilon(y) := \varepsilon^{-d}\xi(y/\varepsilon)$. Set $w := \operatorname{dist}(\cdot,K) * \xi_\varepsilon$. Since $\operatorname{dist}(\cdot,K)$ is Lipschitz continuous, we have $\|Dw\|_\infty \le 1$, $\|D^2 w\|_\infty \le 1/\varepsilon$ and $\|w - \operatorname{dist}(\cdot,K)\|_\infty \le C\varepsilon$. For $\lambda = c\varepsilon\mu$ with $c > 0$ small enough, the function $\lambda(w - C\varepsilon)$ is a subsolution of (2.1) which is nonpositive on $K + \overline{B}_1$. By definition of $m_\mu(\cdot,K)$ we have therefore

$$m_\mu(y,K) \ge \lambda(w(y) - C\varepsilon) \ge \lambda(\operatorname{dist}(y,K) - 2C\varepsilon) \ge l_\mu(\operatorname{dist}(y,K) - 2),$$

for ε small enough (and changing the definition of l_μ if necessary). $\square$

We continue by introducing some notation and basic observations regarding the sublevel sets of the maximal subsolutions, which play a key role in our analysis. For each $t \geq 0$, we let $\mathcal{R}_{\mu,t}$ denote the t-sublevel set of $m_\mu(\cdot, 0)$, that is,

$$\mathcal{R}_{\mu,t} := \left\{ z \in \mathbb{R}^d \,:\, m_\mu(z,0) \leq t \right\}.$$

Note that, by (2.8), $\mathcal{R}_{\mu,0} = \overline{B}_1$. For each $\mu > 0$, we think of $\mathcal{R}_{\mu,t}$ as a "moving front" with the variable t representing time.

Proposition 2.3. *For each $\mu > 0$ and $t \geq 0$:*

(i) *$\mathcal{R}_{\mu,t}$ is a compact connected subset of $\mathbb{R}^d$.*

(ii) *For every $0 \leq s < t$,*

$$\operatorname{dist}_H\left(\mathcal{R}_{\mu,s}, \mathcal{R}_{\mu,t}\right) \leq \frac{1}{l_\mu}|s-t| + 2. \tag{2.11}$$

Proof. (i) Boundedness of $\mathcal{R}_{\mu,t}$ comes from (2.8). To prove that $\mathcal{R}_{\mu,t}$ is connected, let $\mathcal{W}$ be a connected component of $\mathcal{R}_{\mu,t}$. We claim that $\mathcal{W}$ contains the ball B_1: this will show that $\mathcal{R}_{\mu,t}$ consists in a unique connected component, i.e., that $\mathcal{R}_{\mu,t}$ is connected. For this, let us assume that, contrary to our claim, $B_1 \not\subseteq \mathcal{W}$. Since, by (2.8) again, B_1 lies in the interior of $\mathcal{R}_{\mu,t}$, the sets $\mathcal{W}$ and B_1 must have an empty intersection. Therefore $m_\mu(\cdot, 0)$ is a solution to the problem

$$-\operatorname{tr}\left(A(y)D^2 m_\mu(\cdot, K)\right) + H(Dm_\mu(\cdot, K), y) = \mu \qquad \text{in } \mathcal{W},$$

with boundary conditions $m_\mu(\cdot, 0) = t$ on $\partial\mathcal{W}$. As the constant map $w(y) = t$ is a strict subsolution of this equation, we have $m_\mu(\cdot, 0) > w = t$ in the interior of $\mathcal{W}$ by comparison (see [6]). This contradicts the definition of $\mathcal{W}$.

(ii) As $\mathcal{R}_{\mu,s} \subseteq \mathcal{R}_{\mu,t}$, we just have to prove that

$$\mathcal{R}_{\mu,t} \subseteq \mathcal{R}_{\mu,s} + (2 + \frac{1}{l_\mu}|s-t|)\overline{B}_1. \tag{2.12}$$

Set $K := \mathcal{R}_{\mu,s}$ and notice that

$$m_\mu(y, 0) \geq m_\mu(y, K) + s \quad \text{in } \mathbb{R}^d \setminus K. \tag{2.13}$$

Indeed, let $\varepsilon > 0$ and denote

$$w(y) := \begin{cases} m_\mu(y,0) & y \in K, \\ \max\left\{m_\mu(y,0), m_\mu(y,K) + s - \varepsilon\right\} & y \in \mathbb{R}^d \setminus K. \end{cases}$$

Observe that w is a subsolution of (3.11) in $\mathbb{R}^d$ which vanishes on $\overline{B}_1$. Hence $w \leq m_\mu(\cdot, 0)$ by the definition (2.2). Letting $\varepsilon \to 0$ we get (2.13). Combining estimate (2.9) with (2.13), we obtain

$$m_\mu(\cdot, 0) \geq s + l_\mu\left(\operatorname{dist}(\cdot, K) - 2\right) \quad \text{in } \mathbb{R}^d \setminus K.$$

This yields (2.12). $\square$

3. Estimating the fluctuations

The statement of Theorem 1 can be divided into two parts: (i) an estimate of the random part of the error, that is, of $m_\mu(y,0) - \mathbb{E}\left[m_\mu(y,0)\right]$; and (ii) an estimate of the nonrandom part of the error, that is, of $\mathbb{E}\left[m_\mu(y,0)\right] - \overline{m}_\mu(y)$. In this section we focus on the first part. Throughout the rest of the paper, we denote the mean of $m_\mu(y,0)$ by

$$M_\mu(y) := \mathbb{E}\left[m_\mu(y,0)\right]. \tag{3.1}$$

Proposition 3.1. *Fix $\mu_0 \geq 1$. There exists $C > 0$, depending only on (q, Λ, μ_0) such that, for every $0 < \mu \leq \mu_0$, $\lambda \geq C$ and $|y| \geq C\mu^{-2}$, we have*

$$\mathbb{P}\big[\,|m_\mu(y,0) - M_\mu(y)| > \lambda\big] \leq \exp\left(-\frac{\mu^4\lambda^2}{C|y|}\right). \tag{3.2}$$

3.1. **An heuristic proof of Proposition 3.1.** The overall strategy underlying the proof of Proposition 3.1 is similar to the argument of [2, Proposition 4.1] in the first-order case (which we recommend reading first): we localize the maximal subsolutions in their sublevel sets and apply Azuma's inequality. However, the problem is more difficult here because the presence of a nonzero diffusion renders the localization phenomenon much more subtle.

The idea to use Azuma's inequality to estimate the fluctuations of a random variable like $m_\mu(y,0)$ is due to Kesten [10], who applied it to the passage time function in first-passage percolation. Later, Zhang [22] modified the approach of [10] by introducing the idea of conditioning on the environment in the sublevel sets of the passage time function– a more geometrically natural construction which was extended to first-order equations in the proof of [2, Proposition 4.1].

In this subsection we explain the heuristic ideas underlying the argument for Proposition 3.1. Most of what is presented here is *not* rigorous, although it is made precise in the rest of the section.

We begin by fixing $y \in \mathbb{R}^d$ with $|y| \gg 1$. To estimate the distribution of the random variable $m_\mu(y,0)$, we recall that $\mathcal{R}_{\mu,t}$ denotes the t-sublevel set of $m_\mu(\cdot,0)$, that is,

$$\mathcal{R}_{\mu,t} := \left\{z \in \mathbb{R}^d \,:\, m_\mu(z,0) \leq t\right\}.$$

We think of $\mathcal{R}_{\mu,t}$ as a "moving front" which "reveals the environment" as the "time" t increases. We let $\mathcal{F}'_{\mu,t}$ denote "the σ-algebra generated by the environment in $\mathcal{R}_{\mu,t}$".

The "best guess" for the difference between the random variable $m_\mu(y,0)$ and its mean $M_\mu(y) := \mathbb{E}\left[m_\mu(y,0)\right]$, given the information about the medium in $\mathcal{F}'_{\mu,t}$, is

$$X_t := \mathbb{E}\left[m_\mu(y,0) \,\middle|\, \mathcal{F}'_{\mu,t}\right] - M_\mu(y).$$

Note that $\{X_t\}_{t\geq 0}$ is a martingale with $\mathbb{E}X_t = 0$. If we could show, for some constant $C_\mu > 0$, that

$$|X_t - X_s| \leq C_\mu(1 + |s-t|) \tag{3.3}$$

and, for $T = C_\mu|y|$,

$$X_T = m_\mu(y,0) - M_\mu(y), \tag{3.4}$$

then Azuma's inequality yields

$$\mathbb{P}\left[|m_\mu(y,0) - \mathbb{E}[m_\mu(y,0)]| > \lambda\right] \le \exp\left(-\frac{2\lambda^2}{C_\mu|y|}\right),$$

which is the estimate in Proposition 3.1, albeit with a less explicit constant C_μ.

Thus the key step in the proof of Proposition 3.1 is proving (3.3) and (3.4). The quantity $|X_t - X_s|$ for $0 < s < t$ represents "the magnitude of the improvement in the estimate for $m_\mu(y,0)$ after gaining information about the environment in $\mathcal{R}_{\mu,t} \setminus \mathcal{R}_{\mu,s}$." The reason that we can expect analogues of (3.3) and (3.4) to be true has to do with the Lipschitz estimates available for the viscous Hamilton-Jacobi equation and the fact that the maximal subsolutions "localize in their sublevel sets." The rest of this subsection is devoted to giving informal heuristic arguments for (3.3) and (3.4).

Before proceeding to the heuristic proofs, we review the ingredients in the arguments. The crucial localization property can be stated more (but still not completely) precisely as "$m_\mu(y,0)$ is (almost) measurable with respect to the σ-algebra $\mathcal{F}'_{\mu,t}$ generated by the environment inside $\mathcal{R}_{\mu,t}$, provided that $t \ge m_\mu(y,0)$." In other words, if $0 < s \le t$, then

$$m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{R}_{\mu,s}\}}(\omega) \approx \mathbb{E}\big[m_\mu(y,0) \mid \mathcal{F}'_{\mu,t}\big]\mathbb{1}_{\{y\in\mathcal{R}_{\mu,s}\}}. \tag{3.5}$$

Put yet another way: "the sublevel set $\mathcal{R}_{\mu,t}$ is (almost) measurable with respect to the σ-algebra $\mathcal{F}'_{\mu,t}$ generated by the environment inside itself." If this were true, as it is in the first-order case, then since $m_\mu(y,\mathcal{R}_{\mu,t})$ depends only on the environment in the complement of $\mathcal{R}_{\mu,t}$ we would have, by independence,

$$\mathbb{E}\big[m_\mu(y,\mathcal{R}_{\mu,t}) \,\big|\, \mathcal{F}'_{\mu,t}\big](\omega) \approx \sum_{K \text{ compact}} \mathbb{E}\left[m_\mu(y,K)\right]\mathbb{1}_{\{\mathcal{R}_{\mu,t}=K\}}. \tag{3.6}$$

Here things are *much* simpler in the first-order case because the sublevel sets strongly localize the maximal subsolutions and in particular (3.5) and (3.6) hold with equality; see [2, (4.9)] and [2, Lemma 3.1]. The sum on the right of (3.6) is interpreted by partitioning the set of compact Borel sets into a finite number of sets of small diameter in the Hausdorff distance (see below).

Unfortunately, the second-order term destroys this strong localization property. An intuition is provided by the interpretation of $m_\mu(y,0)$ in terms of optimal control theory. Under this interpretation, in the first-order case, once the environment is fixed the optimal trajectories are deterministic and stay confined within the sublevel sets. In contrast, the diffusive term requires the maximal subsolutions to "see all of the environment" because optimal trajectories are stochastic and thus may venture far outside the sublevel sets before eventually returning to B_1. What saves us is a continuous dependence-type estimate (Lemma 3.3, below) which implies that the maximal subsolutions possess a weaker form of the localization property. This allows us to replace $\mathcal{R}_{\mu,t}$ by a close approximation of it, denoted by $\mathcal{S}_{\mu,t}$, which does have the property of being measurable with respect to the σ-algebra generated by the environment inside itself. The rigorous justifications of (3.5) and (3.6) appear below in Lemmas 3.8 and 3.9, respectively.

The next ingredient we discuss is the "dynamic programming principle." Under the control theoretic interpretation that $m_\mu(y,K)$ is measuring a "cost imposed by the

environment ω to transport a particle from y to K," it is intuitively clear that

$$m_\mu(y,0)\mathbb{1}_{\{y\notin\mathcal{R}_{\mu,t}\}} \approx (t + m_\mu(y,\mathcal{R}_{\mu,t}))\,\mathbb{1}_{\{y\notin\mathcal{R}_{\mu,t}\}}. \tag{3.7}$$

In the first-order case, the quantities in (3.7) are interpreted in such a way that (3.7) holds with equality; see [2, (A.20)]. This cannot occur in the second-order case due to the fact that a diffusion "has more than one way of going between points" (here we are quoting Sznitman [20, Example 1.1]). Nevertheless, we show in Lemma 3.7 that (3.7) holds provided we allow for a suitably small error.

We now assemble the above imprecise ingredients into an imprecise (but hopefully illuminating) proof of (3.3). Below we make free use of (3.7), (3.5) and (3.6) and are not troubled by the difficulties in interpreting the sum on the right of (3.6). We also allow the constants to depend on μ.

Heuristic proof of (3.3) *and* (3.4). The reason that (3.4) should hold is straightforward: according to (2.11), if $T \geq L_\mu|y|$, then we have $\mathbb{1}_{\{y\in\mathcal{R}_{\mu,T}\}} \equiv 1$. Thus (3.5) yields (3.4).

To obtain (3.3) it suffices to prove the following two estimates for fixed $0 < s \leq t$:

$$|X_t - X_s| \leq \left|\mathbb{E}\big[m_\mu(y,\mathcal{R}\mu,t)\,\big|\,\mathcal{F}'_{\mu,t}\big] - \mathbb{E}\big[m_\mu(y,\mathcal{R}\mu,s)\,\big|\,\mathcal{F}'_{\mu,s}\big]\right| + |s-t| \tag{3.8}$$

and

$$\left|\mathbb{E}\big[m_\mu(y,\mathcal{R}_{\mu,t})\,\big|\,\mathcal{F}'_{\mu,t}\big] - \mathbb{E}\big[m_\mu(y,\mathcal{R}_{\mu,s})\,\big|\,\mathcal{F}'_{\mu,s}\big]\right| \leq C(1+|s-t|). \tag{3.9}$$

To get (3.8), we write

$$X_t = \mathbb{E}\big[m_\mu(y,\mathcal{R}_{\mu,t})\,\big|\,\mathcal{F}'_{\mu,t}\big]\mathbb{1}_{\{y\in\mathcal{R}_{\mu,t}\}} + \mathbb{E}\big[m_\mu(y,\mathcal{R}_{\mu,t})\,\big|\,\mathcal{F}'_{\mu,t}\big]\mathbb{1}_{\{y\notin\mathcal{R}_{\mu,t}\}},$$

subtract from this a similar expression for X_s and apply (3.7).

To get (3.9), we use (3.5) and (3.6) to see that

$$\begin{aligned}\mathbb{E}\big[m_\mu(y,\mathcal{R}_{\mu,t})\,\big|\,\mathcal{F}'_{\mu,t}\big] &- \mathbb{E}\big[m_\mu(y,\mathcal{R}_{\mu,s})\,\big|\,\mathcal{F}'_{\mu,s}\big] \\ &= \sum_{K,\,K'\text{ compact}} \left(\mathbb{E}\left[m_\mu(y,K)\right] - \mathbb{E}\left[m_\mu(y,K')\right]\right)\mathbb{1}_{\{\mathcal{R}_{\mu,t}=K\}}\mathbb{1}_{\{\mathcal{R}_{\mu,s}=K'\}}.\end{aligned} \tag{3.10}$$

Next we notice that the Lipschitz estimates imply that, either $\mathbb{1}_{\{\mathcal{R}_{\mu,t}=K\}}\mathbb{1}_{\{\mathcal{R}_{\mu,s}=K'\}} \equiv 0$ or else K and K' are closer than $C|s-t|$, measured in the Hausdorff distance. Using the Lipschitz estimate again, we then get

$$\left|\mathbb{E}\left[m_\mu(y,K)\right] - \mathbb{E}\left[m_\mu(y,K')\right]\right|\mathbb{1}_{\{\mathcal{R}_{\mu,t}=K\}}\mathbb{1}_{\{\mathcal{R}_{\mu,s}=K'\}} \leq C(1+|s-t|)\mathbb{1}_{\{\mathcal{R}_{\mu,t}=K\}}\mathbb{1}_{\{\mathcal{R}_{\mu,s}=K'\}}.$$

Inserting this into the right side of (3.10) and summing over all compact K and K' yields (3.3). □

The rigorous version of the heuristic proof above is given in Subsection 3.4, and it relies on the precise justification of (3.5), (3.6) and (3.7), found in Lemmas 3.8, 3.9 and 3.7, respectively, in Subsection 3.3. In order to prove the latter, we first address the problems surrounding the lack of strong localization of $\mathcal{R}_{\mu,t}$. This is the subject of the next subsection.

3.2. **Localizing the sublevel sets of the maximal subsolutions.** We consider maximal solutions of the inequality

$$\begin{cases} -\operatorname{tr}\left(A(y)D^2w\right) + H(Dw,y) \le \mu & \text{in } U, \\ w \le 0 & \text{on } \overline{B}_1, \end{cases} \tag{3.11}$$

where U is an open subset of $\mathbb{R}^d$ with $\overline{B}_1 \subseteq U$. The maximal subsolutions are defined by

$$m_\mu^U(y) := \sup\left\{w(y) \,:\, w \in \mathrm{USC}(U) \text{ satisfies (3.11)}\right\}. \tag{3.12}$$

The advantage of this quantity over m_μ is that it depends only on the environment in U. Indeed, it is immediate from (3.12) that, for each $y \in U$, the random variable $m_\mu^U(y)$ is $\mathcal{G}(U)$–measurable. In order to estimate the dependence of the original m_μ on the environment outside of one of its level sets, we compare m_μ to m_μ^U for U's only a little larger than the sublevel sets. This is the purpose of this subsection, and the key estimates are obtained below in Corollary 3.4.

We continue with the basic properties of m_μ^U needed in the arguments below: proofs of these statements can be found in [6]. It follows from the definition that

$$\overline{B}_1 \subseteq U \subseteq V \qquad \text{implies that} \qquad m_\mu^V(\cdot) \le m_\mu^U(\cdot) \quad \text{in } U.$$

In particular, since $m_\mu^{\mathbb{R}^d}(\cdot) = m_\mu(\cdot, 0)$,

$$m_\mu(\cdot,0) \le m_\mu^U(\cdot) \quad \text{in } U. \tag{3.13}$$

We note that $m_\mu^U(y) = +\infty$ if y does not belong to the connected component of U containing $\overline{B}_1$. We also define m_μ^K for any Borel set $K \in \mathcal{B}$ which contains $\overline{B}_1$ in its interior by setting $m_\mu^K := m_\mu^{\mathrm{int}K}$. As shown in [6], by Perron's method, $m_\mu^U(\cdot)$ satisfies

$$-\operatorname{tr}\left(A(y)D^2m_\mu^U\right) + H\left(Dm_\mu^U, y\right) = \mu \quad \text{in } \left\{y \in U \,:\, m_\mu^U(y) < +\infty\right\} \setminus \overline{B}_1.$$

Therefore, the interior Lipschitz estimates apply: let $\widetilde{U}$ be the connected component of U containing B_1 and let

$$U_h := \left\{x \in \widetilde{U} \,:\, \operatorname{dist}(x, \partial\widetilde{U}) > h\right\},$$

we have that $m_\mu(\cdot)$ is Lipschitz on $\overline{U}_1$ and

$$\operatorname*{ess\,sup}_{y \in U_1} \left|Dm_\mu^U(y)\right| \le L_\mu. \tag{3.14}$$

The Lipschitz constant of m_μ on $\overline{U}_1$ may depend on the geometry of U_1 and in general can be much larger than L_μ. However, (3.14) implies that the Lipschitz constant of m_μ is at most L_μ relative to each of its sublevel sets which are contained in U_1. That is, for every $t > 0$,

$$\begin{aligned} &\left\{y \in U \,:\, m_\mu^U(y) \le t\right\} \subseteq \overline{U}_1 \quad \text{implies that} \\ &\qquad \text{for every } y, z \in \left\{y \in U \,:\, m_\mu^U(y) \le t\right\}, \quad \left|m_\mu^U(y) - m^U(z)\right| \le L_\mu|y-z|. \end{aligned} \tag{3.15}$$

Here is the proof: modify and extend $m_\mu^U(\cdot)$ to the complement of $\{m_\mu^U(\cdot) \le t\}$ by setting it equal to t there. It is then clear that this modified function has gradient

bounded by L_μ in $\mathbb{R}^d$, and is therefore globally Lipschitz with constant L_μ. Since we did not alter any of the values in the t–sublevel set, we obtain the conclusion of (3.15).

Observe that (3.13) and (2.8) yield the lower growth estimate

$$l_\mu(|y|-1) \le m_\mu^U(y) \quad \text{for every } y \in U \tag{3.16}$$

and (3.15) the upper growth estimate

$$m_\mu^U(y) \le L_\mu(|y|-1) \quad \text{provided that} \quad y \in \left\{ z \in U \,:\, m_\mu^U(z) \le t \right\} \subseteq \overline{U}_1. \tag{3.17}$$

We require a slightly stronger version of (3.16), contained in the following lemma, which says that the sublevel sets of m_μ^U grow at a rate bounded uniformly from above. The proof is the same as for (2.11), so we omit it.

Lemma 3.2. *For every* $0 < s \le t$,

$$\left\{ y \in U \,:\, m_\mu^U(y) \le t \right\} \subseteq \left\{ y \in U \,:\, m_\mu^U(y) \le s \right\} + \overline{B}_{2+(t-s)/l_\mu}. \tag{3.18}$$

Following the arguments of Proposition 2.3, one can also easily check that sublevel sets of $m_\mu^U(\cdot)$ are connected: more precisely,

$$\text{if} \quad \{m_\mu^U(\cdot) \le t\} \subseteq \overline{U}_1, \quad \text{then} \quad \{m_\mu^U(\cdot) \le t\} \quad \text{is connected.}$$

We next show that m_μ^U is a very good approximation for m_μ in each of the sublevel sets of m_μ^U which are contained in $\overline{U}_1$. This result is central to the localization result and hence the analysis in this paper. The proof relies on a novel change of the dependent variable.

Lemma 3.3. *Assume* $U \subseteq \mathbb{R}^d$ *and* $t > 0$ *are such that*

$$\left\{ y \in U \,:\, m_\mu^U(y) \le t \right\} \subseteq \overline{U}_1. \tag{3.19}$$

Then

$$m_\mu^U(y) - m_\mu(y,0) \le \frac{4\Lambda L_\mu^3}{\mu l_\mu} \exp\left(\frac{4L_\mu}{l_\mu}\right) \exp\left(-\frac{\mu}{\Lambda L_\mu^2}(t - m_\mu(y,0))\right) \quad \textit{for every} \quad y \in \left\{ z \in U \,:\, m_\mu^U(z) \le t \right\}. \tag{3.20}$$

Proof. We divide the proof into two steps. In the first step we perturb m_μ in order to permit a comparison to m_μ^U. The perturbation is strong near ∂U, which forces m_μ to be larger than m_μ^U on the boundary of U. On the other hand, the perturbation is very small in the s–sublevel set of m_μ^U for $t - s \gg 1$, which allows for the comparison to be useful. In the second step of the argument, we perform the comparison and eventually deduce (3.20). Throughout we simplify the notation by writing $m_\mu(y) = m_\mu(y,0)$.

Step 1. We set $w(y) := \varphi\left(m_\mu(y)\right)$, where $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ is given by

$$\varphi(s) := s + k \exp\left(\frac{\mu}{\Lambda L_\mu^2}(s - t + k)\right) \qquad k := \max\left\{ m_\mu^U(y) - m_\mu(y,0) \,:\, m_\mu^U(y) \le t \right\}.$$

We claim that w satisfies

$$-\operatorname{tr}\left(A(y)D^2 w\right) + H(Dw, y) \ge \mu \quad \text{in } U_1 \setminus \overline{B}_1 \tag{3.21}$$

as well as

$$w \geq 0 \quad \text{on } \overline{B}_1 \qquad \text{and} \qquad w \geq m_\mu^U \quad \text{on } \partial\left\{y \in U \,:\, m_\mu^U(y) \leq t\right\}. \tag{3.22}$$

To check (3.21), we perform formal computations assuming that m_μ is smooth. While we have only that m_μ is Lipschitz, in general, it is routine to make the argument rigorous in the viscosity sense by doing the same computation on a smooth test function. Computing the derivatives of w, we find

$$\begin{aligned} Dw(y) &= \left(1 + \frac{\mu k}{\Lambda L_\mu^2} \exp\left(\frac{\mu}{\Lambda L_\mu^2}\left(m_\mu(y) - t + k\right)\right)\right) Dm_\mu(y), \\ D^2 w(y) &= \left(1 + \frac{\mu k}{\Lambda L_\mu^2} \exp\left(\frac{\mu}{\Lambda L_\mu^2}\left(m_\mu(y) - t + k\right)\right)\right) D^2 m_\mu(y) \\ &\qquad + \frac{\mu^2 k}{\Lambda^2 L_\mu^4} \exp\left(\frac{\mu}{\Lambda L_\mu^2}\left(m_\mu(y) - t + k\right)\right) Dm_\mu(y) \otimes Dm_\mu(y). \end{aligned}$$

Before we attempt to evaluate the left side of (3.21), we note that, since H is convex and $H(0,y) \leq 0$, we have, for every $\lambda \geq 0$,

$$H((1+\lambda)p, y) \geq (1+\lambda) H(p,y).$$

In order to bound the trace of $A(y)$ and the second term in the expression for $D^2 w$, we recall that $|A(y)| \leq \Lambda$ from the assumption (1.4) and that $|Dm_\mu(y)| \leq L_\mu$ for every $y \in U_1$, which follows from (3.19) and Proposition 2.1. Using these and the equation for m_μ, we obtain, for every $y \in U_1 \setminus B_1$,

$$\begin{aligned} -\operatorname{tr}\left(A(y) D^2 w\right) + H(Dw, y) &\geq \left(1 + \frac{\mu k}{\Lambda L_\mu^2} \exp\left(\frac{\mu}{\Lambda L_\mu^2}\left(m_\mu(y) - t + k\right)\right)\right) \mu \\ &\qquad - \frac{\mu^2 k}{\Lambda^2 L_\mu^4} \exp\left(\frac{\mu}{\Lambda L_\mu^2}\left(m_\mu(y) - t + k\right)\right) \Lambda L_\mu^2 \\ &= \mu. \end{aligned}$$

To check (3.22), we use the definition of k and φ and the fact that

$$\partial\left\{y \in U \,:\, m_\mu^U(y) \leq t\right\} \subseteq \left\{y \in U \,:\, m_\mu^U(y) = t\right\},$$

which follows from the continuity of m_μ^U, to deduce, for each $y \in \partial\left\{z \in U \,:\, m_\mu^U(z) \leq t\right\}$,

$$\begin{aligned} w(y) &\geq m_\mu(y) + k \exp\left(\frac{\mu}{\Lambda L_\mu^2}\left(m_\mu(y) - t + k\right)\right) \\ &\geq m_\mu(y) + k \exp\left(\frac{\mu}{\Lambda L_\mu^2}\left(m_\mu^U(y) - t\right)\right) \geq m_\mu(y) + k \geq m_\mu^U(y). \end{aligned}$$

The nonnegativity of w on $\overline{B}_1$ is obvious. This completes the proof of (3.21) and (3.22).

Step 2. By comparison principle (c.f. [6]), we obtain that

$$w \geq m_\mu^U \quad \text{in } \{y \in U \,:\, m_\mu^U \leq t\}. \tag{3.23}$$

Indeed, while the comparison principle requires a strict inequality, we may compare w to $(1-\varepsilon) m_\mu^U$ for $\varepsilon > 0$ and then send $\varepsilon \to 0$, using the observation that $(1-\varepsilon) m_\mu^U$ is

a strict subsolution of (2.1) in U by the convexity of H and the fact that 0 is a strict subsolution for $\mu > 0$. See [6, Lemma 2.4] for details.

Expressing (3.23) in terms of m_μ yields

$$m_\mu(y) - m_\mu^U(y) \geq -k \exp\left(\frac{\mu}{\Lambda L_\mu^2}(m_\mu(y) - t + k)\right) \quad \text{in } \left\{x \in U \,:\, m_\mu^U(x) \leq t\right\}. \tag{3.24}$$

We complete the argument by using the Lipschitz estimate and (3.24) to estimate k from above, and then feed the result back into (3.24). We first show the following rough bound on k:

$$k \leq t(1 - l_\mu/L_\mu). \tag{3.25}$$

Since $k = t - \min\left\{m_\mu(y,0) \,:\, m_\mu^U(y) = t\right\}$, we may select $y_0 \in U$ such that $m_\mu^U(y_0) = t$ and $k = t - m_\mu(y_0)$. Then (3.17) implies that $|y_0| \geq t/L_\mu + 1$ and (2.8) gives that $(l_\mu/L_\mu)t \leq m_\mu(y_0) = t - k$, which yields (3.25).

Fix $t_1 \in [0, t-k]$ to selected below. By (2.11), there exists $y_1 \in \mathbb{R}^d$ such that $m_\mu(y_1, 0) = t_1$ and $|y_1 - y_0| \leq (t - k - t_1)/l_\mu + 2$. Using (3.24) at y_1, and the Lipschitz estimate (3.15), we deduce

$$-k \exp\left(\frac{\mu}{\Lambda L_\mu^2}(t_1 - t + k)\right) \leq t_1 - m_\mu^U(y_1) \leq t_1 - t + \frac{L_\mu}{l_\mu}(t - k - t_1) + 2L_\mu. \tag{3.26}$$

Fix $\varepsilon := \exp(-1)$ and set $t_1 := t - k - \mu^{-1}(\Lambda L_\mu^2)$. Observe that, in view of (3.25), we have $t_1 \geq 0$ provided that $t \geq (\Lambda L_\mu^3)/(\mu l_\mu)$. Then (3.26) gives

$$k \leq \frac{1}{1-\varepsilon}\left(-\frac{\Lambda L_\mu^2}{\mu} + \frac{\Lambda L_\mu^3}{\mu l_\mu} + 2L_\mu\right) \leq 4\frac{\Lambda L_\mu^3}{\mu l_\mu}.$$

Inserting this into (3.24) yields (3.20) for $t \geq (\Lambda L_\mu^3)/(\mu l_\mu)$. We conclude by noting that (3.20) always holds for $t \leq \Lambda L_\mu^3/(\mu l_\mu)$. □

We next use the Lipschitz and growth estimates to translate the previous result that $m_\mu \approx m_\mu^U$ into a result asserting that, under the same hypotheses, the corresponding level sets are close. We define a constant a_μ, used in the rest of this section, by

$$a_\mu := 1 + \frac{\Lambda L_\mu^2}{\mu l_\mu}\left(\frac{4L_\mu}{l_\mu} + \log\left(\frac{4\Lambda L_\mu^3}{\mu l_\mu^2}\right)\right). \tag{3.27}$$

We note that, for $C > 0$ depending only on (q, Λ) and an upper bound for μ,

$$a_\mu \leq C l_\mu^{-2}\mu^{-1} \leq C\mu^{-3}. \tag{3.28}$$

Corollary 3.4. *Let $t > 0$ and assume $U \subseteq \mathbb{R}^d$ is such that*

$$\left\{x \in \mathbb{R}^d \,:\, m_\mu^U(x) \leq t\right\} \subseteq \overline{U}_{2+a_\mu}. \tag{3.29}$$

Then

$$0 \leq m_\mu^U(t) - m_\mu(t,0) \leq l_\mu \quad \textit{for every } y \in \left\{z \in U \,:\, m_\mu^U(z) \leq t\right\} \tag{3.30}$$

and

$$\left\{y \in U \,:\, m_\mu^U(y) \leq t\right\} \subseteq \mathcal{R}_{\mu,t} \subseteq \left\{y \in U \,:\, m_\mu^U(y) \leq t\right\} + \overline{B}_3. \tag{3.31}$$

Proof. Define the constant

$$h := \frac{\Lambda L_\mu^2}{\mu}\left(\frac{4L_\mu}{l_\mu} + \log\left(\frac{4\Lambda L_\mu^3}{\mu l_\mu^2}\right)\right)$$

and observe that, due to (3.18), (3.29) and the fact that $a_\mu = 1 + h/l_\mu$,

$$\left\{y \in U \,:\, m_\mu^U(y) \le t + h\right\} \subseteq \left\{y \in U \,:\, m_\mu^U(y) \le t\right\} + \overline{B}_{2+h/l_\mu} \subseteq \overline{U}_1. \tag{3.32}$$

Therefore, applying (3.20) with $t + h$ in place of t, we find that

$$\begin{aligned} l_\mu &= \frac{4\Lambda L_\mu^3}{\mu l_\mu}\exp\left(\frac{4L_\mu}{l_\mu}\right)\exp\left(-\frac{\mu}{\Lambda L_\mu^2}h\right) && \text{(definition of } h\text{)} \\ &\ge \frac{4\Lambda L_\mu^3}{\mu l_\mu}\exp\left(\frac{4L_\mu}{l_\mu}\right)\exp\left(-\frac{\mu}{\Lambda L_\mu^2}(t + h - m_\mu(\cdot))\right) && \text{(in } \left\{m_\mu^U(\cdot) \le t\right\}\text{)} \\ &\ge m_\mu^U(\cdot) - m_\mu(\cdot, 0) && \text{(by (3.20), (3.32)).} \end{aligned}$$

This completes the proof of (3.30).

Next we show that (3.31) holds. The first inclusion is immediate from (3.13). To show the second inclusion, we observe that, according to (3.30), for every $\varepsilon > 0$,

$$m_\mu(\cdot, 0) > t - l_\mu \quad \text{on } \partial\left\{y \in U \,:\, m_\mu^U(y) \le t + \varepsilon\right\}.$$

Since $\mathcal{R}_{\mu,t-l_\mu}$ is connected, after sending $\varepsilon \to 0$ we conclude that

$$\mathcal{R}_{\mu,t-l_\mu} \subseteq \left\{y \in U \,:\, m_\mu^U(y) \le t\right\}.$$

The result now follows from (2.11) which implies $\mathcal{R}_{\mu,t} \subseteq \mathcal{R}_{\mu,t-l_\mu} + \overline{B}_3$. □

3.3. **The rigorous justification of** (3.5)**,** (3.6) **and** (3.7)**.** Using the estimates in Corollary 3.4, which assert that $m_\mu \approx m_\mu^U$ in sublevel sets of m_μ^U which are a unit distance from the boundary of U, we construct a random "moving front" $\mathcal{S}_{\mu,t}$, close to $\mathcal{R}_{\mu,t}$, which is "measurable with respect to the σ-algebra generated by the environment inside itself." This is the key to obtaining the rigorous versions of (3.7), (3.5) and (3.6) we need to complete the proof of Proposition 3.1.

In order to build $\mathcal{S}_{\mu,t}$ as well as interpret sums like on the right side of (3.6), it is convenient to introduce a discretization of the set of compact subsets of $\mathbb{R}^d$ which contain $\overline{B}_1$. We denote the latter by $\mathcal{K}$ and endow it will the *Hausdorff metric* dist_H defined by

$$\begin{aligned} \mathrm{dist}_H(E, F) &:= \inf_{x\in E}\sup_{y\in F}|x - y| \vee \inf_{y\in F}\sup_{x\in E}|x - y| \\ &= \inf\left\{\varepsilon > 0 : E \subseteq F + B_\varepsilon \text{ and } F \subseteq E + B_\varepsilon\right\}. \end{aligned}$$

The metric space $(\mathcal{K}, \mathrm{dist}_H)$ is locally compact (see Munkres [16]) and thus there exists a partition $(\Gamma_i)_{i\in\mathbb{N}}$ of $\mathcal{K}$ into Borel subsets of $\mathcal{K}$ of diameter at most 1. That is, $\Gamma_i \subseteq \mathcal{K}$ and $\mathrm{diam}_H(\Gamma_i) \le 1$ for all $i \in \mathbb{N}$ and $\Gamma_i \cap \Gamma_j = \emptyset$ if $i, j \in \mathbb{N}$ are such that $i \ne j$.

For each $i \ge 1$, we take K_i to be the closure of the union of the elements of Γ_i. Observe that

$$K \in \Gamma_i \qquad \text{implies that} \qquad K \subseteq K_i \subseteq K + \overline{B}_1. \tag{3.33}$$

We introduce, for each $i \in \mathbb{N}$, compact sets

$$K_i \subseteq K_i' \subseteq K_i'' \subseteq \widetilde{K}_i$$

by setting (with the constant $a_\mu > 0$ defined in (3.27))

$$K_i' := K_i + \overline{B}_{2+a_\mu}, \qquad K_i'' := K_i' + \overline{B}_7 \qquad \text{and} \qquad \widetilde{K}_i := K_i'' + \overline{B}_4.$$

Here are the reasons we must introduce so many sets:

- K_i' provides extra room so that we can apply the localization results of the Subsection 3.2;
- K_i'' must be a little larger than K_i' so that the partition $\{E_i(t)\}_{i\in\mathbb{N}}$ of Ω that we construct below has the property that $E_i(t) \in \mathcal{G}(K_i'')$, which forces the moving front $\mathcal{S}_{\mu,t}$ we build to localize properly;
- $\widetilde{K}_i$ is larger than K_i'' so that $\operatorname{dist}(K_i'' + \overline{B}_3, \mathbb{R}^d \setminus \widetilde{K}_i) = 1$ and thus the finite range of dependence hypothesis yields

$$\mathcal{G}\left(K_i'' + \overline{B}_3\right) \qquad \text{and} \qquad \mathcal{G}\left(\mathbb{R}^d \setminus \widetilde{K}_i\right) \qquad \text{are independent.} \tag{3.34}$$

We next use the partition $\{\Gamma_i\}_{i\in\mathbb{N}}$ of $\mathcal{K}$ to construct, for each $t > 0$, a partition of Ω into disjoint events $\{E_i(t)\}_{i\in\mathbb{N}}$ which are approximations of the event $\{\omega \in \Omega : \mathcal{R}_{\mu,t}(\omega) \in \Gamma_i\}$:

$$\text{“ } E_i(t) \approx \{\omega \in \Omega : \mathcal{R}_{\mu,t}(\omega) \in \Gamma_i\} \text{ . ”} \tag{3.35}$$

The reason we do not define $E_i(t)$ with equality in (3.35) is due to the fact that, as we have explained above, the $\mathcal{R}_{\mu,t}$'s do not properly “localize.” The events $E_i(t)$ we introduce are better localized in the sense made precise by Lemma 3.6.

To construct $E_i(t)$, we first make each of the Γ_i's slightly larger by setting

$$\widetilde{\Gamma}_i := \left\{K \in \mathcal{K} : K \subseteq K_i \subseteq K + \overline{B}_4\right\}$$

and then define, for every $t > 0$ and $i \in \mathbb{N}$,

$$F_i(t) := \left\{\omega \in \Omega : \left\{y \in K_i' : m_\mu^{K_i'}(y) \le t\right\} \in \widetilde{\Gamma}_i\right\}.$$

To make a partition of Ω, we modify the $F_i(t)$'s by setting

$$E_1(t) := F_1(t) \qquad \text{and} \quad E_{i+1}(t) := F_{i+1}(t) \setminus \bigcup_{j\le i} E_j(t) \qquad \text{for every } i \in \mathbb{N}.$$

We now collect some properties of the families $\{F_i(t)\}$ and $\{E_i(t)\}$.

Lemma 3.5. *Fix $t > 0$. Then the family $(E_i(t))$ is a measurable partition of Ω. Moreover, if $\omega \in F_i(t)$ and $y \in K_i$, we have*

$$\mathcal{R}_{\mu,t} \subseteq K_i + \overline{B}_3, \qquad K_i \subseteq \mathcal{R}_{\mu,t} + \overline{B}_4, \qquad \operatorname{dist}_H(\partial\mathcal{R}_{\mu,t}, \partial K_i) \le 7 \tag{3.36}$$

and

$$\left| m_\mu^{K_i'}(y) - m_\mu(y,0) \right| \le 9L_\mu. \tag{3.37}$$

Proof. To prove that $\{E_i(t)\}_{i\in\mathbb{N}}$ is a partition of Ω, we note that, for any $\omega \in \Omega$,

$$\mathcal{R}_{\mu,t}(\omega) \in \Gamma_i \quad \text{implies that} \quad \omega \in F_i(t). \tag{3.38}$$

Indeed, notice that $\{m_\mu^{K_i'}(\cdot) \le t\} \subseteq \mathcal{R}_{\mu,t}$ by (3.13), and thus $\mathcal{R}_{\mu,t}(\omega) \in \Gamma_i$ implies that

$$\{y \in K_i' \,:\, m_\mu^{K_i'}(y) \le t\} \subseteq K_i \subseteq (K_i')_{2+a_\mu}.$$

Corollary 3.4 and the fact that $\mathcal{R}_{\mu,t}(\omega) \in \Gamma_i$ then imply that

$$K_i \subseteq \mathcal{R}_{\mu,t} + \overline{B}_1 \subseteq \left\{m_\mu^{K_i'}(\cdot) \le t\right\} + \overline{B}_4,$$

so that (3.38) holds. Since $\{\Gamma_i\}_{i\in\mathbb{N}}$ is a partition of $\mathcal{K}$, we deduce from (3.38) that

$$\Omega = \bigcup_{i\in\mathbb{N}} \{\omega \in \Omega \,:\, \mathcal{R}_{\mu,t} \in \Gamma_i\} \subseteq \bigcup_{i\in\mathbb{N}} F_i(t) \subseteq \Omega.$$

Then it is clear that $\{E_i(t)\}_{i\in\mathbb{N}}$ is a partition of Ω.

Next we show (3.36). Fix $\omega \in F_i(t)$. Owing to the definition of K_i', we have

$$\left\{y \in K_i' \,:\, m_\mu^{K_i'}(y) \le t\right\} \subseteq (K_i')_{2+a_\mu}.$$

In particular, we can apply Corollary 3.4 to $U := \mathrm{int}(K_i')$ to deduce, by the definition of $F_i(t)$ and (3.31), that

$$K_i \subseteq \left\{y \in U \,:\, m_\mu^U(y) \le t\right\} + \overline{B}_4 \subseteq \mathcal{R}_{\mu,t}(\omega) + \overline{B}_4, \tag{3.39}$$

and

$$\mathcal{R}_{\mu,t}(\omega) \subseteq \left\{y \in U \,:\, m_\mu^U(y) \le t\right\} + \overline{B}_3 \subseteq K_i + \overline{B}_3.$$

To show the last statement of (3.36), let us notice that, by definition of $\widetilde{\Gamma}_i$,

$$\left\{y \in U \,:\, m_\mu^U(y) \le t\right\} \subseteq K_i \subseteq \left\{y \in U \,:\, m_\mu^U(y) \le t\right\} + \overline{B}_4.$$

Hence

$$\mathrm{dist}_H\left(\partial\left\{y \in U \,:\, m_\mu^U(y) \le t\right\}, \partial K_i\right) \le 4.$$

On another hand, (3.31) implies that

$$\mathrm{dist}_H\left(\partial\mathcal{R}_{\mu,t}, \partial\left\{y \in U \,:\, m_\mu^U(y) \le t\right\}\right) \le 3.$$

Therefore $\mathrm{dist}_H\left(\partial\mathcal{R}_{\mu,t}(\omega), \partial K_i\right) \le 7$ and the proof of (3.36) is complete.

Let us finally check that (3.37) holds: by (3.30) we have

$$0 \le m_\mu^U(\cdot) - m_\mu(\cdot, 0) \le l_\mu \quad \text{in } \left\{y \in U \,:\, m_\mu^U(y) \le t\right\}. \tag{3.40}$$

If $x \in K_i \setminus \left\{y \in U \,:\, m_\mu^U(y) \le t\right\}$, then, by definition of $F_i(t)$, there exists $x_0 \in U$ such that $m_\mu^U(x_0) \le t$ and $|x - x_0| \le 4$. We conclude by Lipschitz estimate and (3.40) that

$$\left|m_\mu^U(x) - m_\mu(x,0)\right| \le 8L_\mu + \left|m_\mu^U(x_0) - m_\mu(x_0,0)\right| \le 9L_\mu. \qquad \square$$

We next verify that $E_i(t)$ is "localized" in the sense that it belongs to $\mathcal{G}(K_i'')$.

Lemma 3.6. *For every $0 < s \le t$ and $i, j \in \mathbb{N}$,*

$$F_i(s) \cap F_j(t) \ne \emptyset \qquad \textit{implies that} \qquad E_i(s) \in \mathcal{G}(K_j''). \tag{3.41}$$

In particular, $E_i(t) \in \mathcal{G}(K_i'')$ for every $i \in \mathbb{N}$.

Proof. We first claim that

$$0 < s \le t \quad \text{and} \quad F_i(s) \cap F_j(t) \neq \emptyset \qquad \text{implies that} \qquad K_i' \subseteq K_j''. \tag{3.42}$$

Here is the proof of (3.42): if $0 < s \le t$ and $\omega \in F_i(s) \cap F_j(t)$, then by (3.36) we have

$$K_i \subseteq \mathcal{R}_{\mu,s}(\omega) + \overline{B}_4 \subseteq \mathcal{R}_{\mu,t}(\omega) + \overline{B}_4 \subseteq K_j + \overline{B}_7.$$

Hence $K_i' \subseteq K_j' + \overline{B}_7 = K_j''$.

It is clear that $F_i(s) \in \mathcal{G}(K_i')$ for every $i \in \mathbb{N}$. If $0 < s \le t$ and $i, j \in \mathbb{N}$ are such that $F_i(s) \cap F_j(t) \neq \emptyset$, then we have $F_i(s) \in \mathcal{G}(K_j'')$ by (3.42). Observe also that (3.42) yields the expression

$$E_i(s) = F_i(s) \setminus \bigcup_{n \in \beta(i)} F_n(s),$$

where we have set $\beta(i) := \{1 \le n < i \,:\, K_n' \subseteq K_i''\}$. The lemma follows. □

We now introduce the "moving front" $\mathcal{S}_{\mu,t}$ and the filtration $\mathcal{F}_{\mu,t}$ that we use in the proof of Proposition 3.1. We define $\mathcal{S}_{\mu,t}$ by

$$\mathcal{S}_{\mu,t}(\omega) := K_i \qquad \text{if } \omega \in E_i(t). \tag{3.43}$$

Observe that (3.36) implies that $\mathcal{S}_{\mu,t}$ is a good approximation of $\mathcal{R}_{\mu,t}$:

$$\mathcal{R}_{\mu,t} \subseteq \mathcal{S}_{\mu,t} + \overline{B}_3, \qquad \mathcal{S}_{\mu,t} \subseteq \mathcal{R}_{\mu,t} + \overline{B}_4 \qquad \text{and} \qquad \operatorname{dist}_H\left(\partial\mathcal{R}_{\mu,t}, \partial\mathcal{S}_{\mu,t}\right) \le 7. \tag{3.44}$$

The filtration $\mathcal{F}_{\mu,t}$ is defined as the σ-algebra generated by events of the form

$$G \cap E_i(s), \qquad \text{where} \quad 0 < s \le t,\ i \in \mathbb{N},\ G \in \mathcal{G}(K_i''). \tag{3.45}$$

We also set $\mathcal{F}_{\mu,0} = \{\emptyset, \Omega\}$.

Observe that $\{\mathcal{F}_{\mu,t}\}_{t \ge 0}$ is indeed a filtration, since by definition it is increasing in t. It is clear from Lemma 3.6 that, for every $y \in \mathbb{R}^d$ and $t \ge 0$,

$$\{\omega \in \Omega \,:\, y \in \mathcal{S}_{\mu,t}(\omega)\} \in \mathcal{F}_{\mu,t}. \tag{3.46}$$

For convenience, below we write $\{y \in \mathcal{S}_{\mu,t}\}$ to denote the event $\{\omega \in \Omega \,:\, y \in \mathcal{S}_{\mu,t}(\omega)\}$.

The next lemma is the rigorous justification of (3.7), which follows relatively easily from (3.44) and the Lipschitz estimates.

Lemma 3.7. *For any $t > 1$,*

$$|m_\mu(y,0) - (t + m_\mu(y, \mathcal{S}_{\mu,t}))|\, \mathbb{1}_{\{y \notin \mathcal{S}_{\mu,t}\}}(\omega) \le 8L_\mu \mathbb{1}_{\{y \notin \mathcal{S}_{\mu,t}\}}(\omega). \tag{3.47}$$

Proof. By the maximality of $m_\mu(\cdot, 0)$, we have

$$\inf_{z \in \partial\mathcal{S}_{\mu,t}} m_\mu(z,0) + m_\mu(\cdot, \mathcal{S}_{\mu,t}) \le m_\mu(\cdot, 0) \quad \text{in } \mathbb{R}^d \setminus \mathcal{S}_{\mu,t}. \tag{3.48}$$

Indeed, for every $\varepsilon > 0$, the function

$$w(y) := \begin{cases} m_\mu(y,0) & y \in \mathcal{S}_{\mu,t}, \\ \max\left\{ m_\mu(y,0), \inf\limits_{z \in \partial\mathcal{S}_{\mu,t}} m_\mu(z,0) + m_\mu(y, \mathcal{S}_{\mu,t}) - \varepsilon \right\} & y \notin \mathcal{S}_{\mu,t}, \end{cases}$$

is a global subsolution which vanishes on $\overline{B}_1$, hence $w \le m_\mu(\cdot,0)$ by maximality. Sending $\varepsilon \to 0$ yields (3.48). On the other hand, the maximality of $m_\mu(\cdot,\mathcal{S}_{\mu,t})$ yields

$$m_\mu(\cdot,\mathcal{S}_{\mu,t}) \ge m_\mu(\cdot,0) - \sup_{z\in\partial\mathcal{S}_{\mu,t}} m_\mu(z,0) - L_\mu \quad \text{in } \mathbb{R}^d$$

since the right side is a subsolution which is nonpositive on $\mathcal{S}_{\mu,t} + B_1$ thanks to (2.8). We conclude that

$$\inf_{z\in\partial\mathcal{S}_{\mu,t}} m_\mu(z,0) \le m_\mu(\cdot,0) - m_\mu(\cdot,\mathcal{S}_{\mu,t}) \le \sup_{z\in\partial\mathcal{S}_{\mu,t}} m_\mu(z,0) + L_\mu \quad \text{in } \mathbb{R}^d \setminus \mathcal{S}_{\mu,t} \tag{3.49}$$

Next we observe that, according to (3.44), the Lipschitz estimate and the fact that $m_\mu(\cdot,0) \equiv t$ on $\partial\mathcal{R}_{\mu,t}$, we have

$$\sup_{z\in\partial\mathcal{S}_{\mu,t}} |m_\mu(z,0) - t| \le 7L_\mu. \tag{3.50}$$

Combining (3.49) and (3.50) gives

$$|m_\mu(\cdot,0) - (t + m_\mu(\cdot,\mathcal{S}_{\mu,t}))| \le 8L_\mu \quad \text{in } \mathbb{R}^d \setminus \mathcal{S}_{\mu,t},$$

which yields the lemma. □

We next give the rigorous justification of (3.5), which asserts that $m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,t}\}}$ is nearly $\mathcal{F}_{\mu,t}$–measurable.

Lemma 3.8. *For every $y \in \mathbb{R}^d$ and $0 < s \le t$,*

$$\big|m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} - \mathbb{E}\big[m_\mu(y,0)\,|\,\mathcal{F}_{\mu,t}\big]\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}}\big| \le 18L_\mu. \tag{3.51}$$

Proof. Fix $y \in \mathbb{R}^d$, $0 < s \le t$ and define a random variable Z by

$$Z(\omega) := \sum_{i\in\mathbb{N},\, y\in K_i} m_\mu^{K_i'}(y)\mathbb{1}_{E_i(s)}(\omega).$$

It is clear that $m_\mu^{K_i'}(y)$ is $\mathcal{G}(K_i'')$–measurable. According to Lemma 3.6, $E_i(s) \in \mathcal{F}_{\mu,s}$, and hence Z is $\mathcal{F}_{\mu,s}$–measurable by the definition (3.45) of the filtration. As

$$\{\omega\in\Omega \,:\, y\in\mathcal{S}_{\mu,s}\} = \bigcup_{i\in\mathbb{N},\, y\in K_i} E_i(s),$$

we have, from (3.37), that for every $\omega \in \Omega$,

$$\big|Z(\omega) - m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}}(\omega)\big| \le \sum_{i\in\mathbb{N},\ y\in K_i} \Big|m_\mu^{K_i'}(y) - m_\mu(y,0)\Big|\,\mathbb{1}_{E_i(s)}(\omega) \le 9L_\mu.$$

Using (3.46) and that Z is $\mathcal{F}_{\mu,t}$–measurable, we find that

$$\begin{aligned}
&\big|m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} - \mathbb{E}\big[m_\mu(y,0)\,|\,\mathcal{F}_{\mu,t}\big]\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}}\big| \\
&\qquad \le \big|m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} - Z\big| + \big|\mathbb{E}\big[Z \mid \mathcal{F}_{\mu,t}\big] - \mathbb{E}\big[m_\mu(y,0) \mid \mathcal{F}_{\mu,t}\big]\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}}\big| \\
&\qquad = \big|m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} - Z\big| + \big|\mathbb{E}\left[Z - m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} \,\big|\, \mathcal{F}_{\mu,t}\right]\big| \\
&\qquad \le 2\sup_\Omega \big|m_\mu(y,0)\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} - Z\big| \\
&\qquad \le 18L_\mu,
\end{aligned}$$

as desired. □

As a consequence of Lemma 3.8, we obtain (3.4), albeit with a small error: the statement is that for every $T \geq L_\mu(|y|-1)$,

$$\big|m_\mu(y,0) - \mathbb{E}\big[m_\mu(y,0)\,|\,\mathcal{F}_{\mu,T}\big]\big| \leq 18L_\mu. \tag{3.52}$$

Here is the proof: by the definition of $F_i(t)$, we have

$$\omega \in F_i(t) \quad \text{implies that} \quad \Big\{y \in K_i' \,:\, m_\mu^{K_i'}(y) \leq t\Big\} \subseteq K_i.$$

In light of (3.17) and the definition of $\mathcal{S}_{\mu,t}$, we find that

$$T \geq L_\mu(R-1) \quad \text{implies that} \quad \overline{B}_R \subseteq \mathcal{S}_{\mu,T}.$$

Hence $T \geq L_\mu(|y|-1)$ implies that $\mathbb{1}_{\{y\in\mathcal{S}_{\mu,T}\}} \equiv 1$, and so (3.52) follows from (3.51).

The next lemma provides the rigorous justification of (3.6).

Lemma 3.9. *For every $t>0$,*

$$0 \leq \mathbb{E}\big[m_\mu(y,\mathcal{S}_{\mu,t})\,\big|\,\mathcal{F}_{\mu,t}\big] - \sum_{i\in\mathbb{N}} \mathbb{E}\left[m_\mu(y,\widetilde{K}_i)\right]\mathbb{1}_{E_i(t)} \leq L_\mu(a_\mu+13). \tag{3.53}$$

Proof. Step 1. We first show that, for every $i\in\mathbb{N}$ and $t>0$,

$$A \in \mathcal{F}_{\mu,t} \qquad \text{implies that} \qquad A\cap E_i(t) \in \mathcal{G}(K_i''+\overline{B}_3). \tag{3.54}$$

By the definition of $\mathcal{F}_{\mu,t}$, it suffices to fix $0<s\leq t$, $j\in\mathbb{N}$, $G\in\mathcal{G}(K_j'')$ and to take $A = G\cap E_j(s)$. Then

$$A\cap E_i(t) = G\cap E_j(s)\cap E_i(t).$$

By (3.41) and (3.42), we see that either this set is empty or else $K_j'\subseteq K_i''$and $E_j(s)\in\mathcal{G}(K_i'')$. This also yields that $G\in\mathcal{G}(K_j'')\subseteq\mathcal{G}(K_i''+\overline{B}_3)$ and $E_j(s)\cap E_i(t)\in\mathcal{G}(K_i'')$, and therefore we obtain $A\cap E_i(t)\in\mathcal{G}(K_i''+\overline{B}_3)$ as desired.

Step 2. We claim that

$$\mathbb{E}\left[m_\mu\big(y,\widetilde{K}_i\big)\mathbb{1}_{E_i(t)}\,\middle|\,\mathcal{F}_{\mu,t}\right] = \mathbb{E}\left[m_\mu\big(y,\widetilde{K}_i\big)\right]\mathbb{1}_{E_i(t)}. \tag{3.55}$$

By the definition of conditional expectation, we must show that, for every $A\in\mathcal{F}_{\mu,t}$,

$$\mathbb{E}\left[m_\mu(y,\widetilde{K}_i)\mathbb{1}_{A\cap E_i(t)}\right] = \mathbb{E}\left[m_\mu(y,\widetilde{K}_i)\right]\mathbb{P}\Big[A\cap E_i(t)\Big].$$

Since $m_\mu(y,\widetilde{K}_i)$ is $\mathcal{G}(\mathbb{R}^d\backslash\widetilde{K}_i)$-measurable, this follows from (3.34) and (3.54).

Step 3. The conclusion, using (3.55) and the Lipschitz estimates. Observe that

$$\mathbb{E}\big[m_\mu(y,\mathcal{S}_{\mu,t})\,\big|\,\mathcal{F}_{\mu,t}\big] = \sum_{i\in\mathbb{N}}\mathbb{E}\big[m_\mu(y,\mathcal{S}_{\mu,t})\mathbb{1}_{E_i(t)}\,\big|\,\mathcal{F}_{\mu,t}\big] = \sum_{i\in\mathbb{N}}\mathbb{E}\big[m_\mu(y,K_i)\mathbb{1}_{E_i(t)}\,\big|\,\mathcal{F}_{\mu,t}\big]$$

and, according to the Lipschitz estimates (2.5),

$$0 \leq \Big(m_\mu(y,K_i) - m_\mu(y,\widetilde{K}_i)\Big)\mathbb{1}_{E_i(t)} \leq L_\mu(a_\mu+13)\mathbb{1}_{E_i(t)}.$$

Combining the previous two lines and applying (3.55) yields (3.53). □

3.4. **The fluctuations estimate.** We now present the proof for Proposition 3.1, which follows the heuristic argument given in Subsection 3.1.

Proof of Proposition 3.1. We break the proof into three steps. Throughout, C denotes a constant which may change from line to line and depends only on (q, Λ, μ_0).

Step 1. The application of Azuma's inequality. Fix $y \in \mathbb{R}^d$ and define a $\mathcal{F}_{\mu,t}$–adapted martingale $\{X_t\}_{t\geq 0}$ by

$$X_t := \mathbb{E}\big[m_\mu(y,0) \,|\, \mathcal{F}_{\mu,t}\big] - M_\mu(y). \tag{3.56}$$

Here $\{\mathcal{F}_{\mu,t}\}$ is the filtration defined in the previous subsection and we recall that $M_\mu(y)$ is defined by (3.1). Observe that $X_0 \equiv 0$ and, according to (3.52), for all $t \geq L_\mu|y|$,

$$|X_t(\omega) - (m_\mu(y,0) - M_\mu(y))| \leq 18L_\mu. \tag{3.57}$$

The main step in our argument is to show

$$\operatorname*{ess\,sup}_{\omega\in\Omega} |X_t(\omega) - X_s(\omega)| \leq (2a_\mu + 98)L_\mu + \frac{2L_\mu}{l_\mu}|t-s|. \tag{3.58}$$

We admit (3.58) for a moment and use it to complete the proof of the proposition. Azuma's inequality applied to the discrete martingale sequence $Y_n := X_{\alpha n}$, for $n \in \mathbb{N}$ and with $\alpha := l_\mu(a_\mu + 49)$ yields, in light of (3.58),

$$\mathbb{P}\big[\,|Y_n| > \lambda\big] \leq \exp\left(-\frac{\lambda^2}{8(2a_\mu+98)^2 L_\mu^2 n}\right).$$

Note that by (3.28) we have $l_\mu a_\mu \leq C\mu^{-2}$. Take $n := \left\lceil \frac{L_\mu|y|}{\alpha}\right\rceil + 1$ and observe that for $|y| \geq C\mu^{-2} \geq Cl_\mu(a_\mu+1)$, we have

$$n = \left\lceil \frac{2L_\mu|y|}{l_\mu(2a_\mu+98)} + 1\right\rceil \leq \frac{C|y|}{l_\mu(a_\mu+1)}.$$

The above estimate, together with (3.57), yields (3.2) for every $\lambda \geq C$.

We have left to prove (3.58), which is a consequence of the following two inequalities:

$$|X_t - X_s| \leq \big|\mathbb{E}\left[m_\mu(y,\mathcal{S}_{\mu,t}) \,|\, \mathcal{F}_{\mu,t}\right] - \mathbb{E}\left[m_\mu(y,\mathcal{S}_{\mu,s}) \,|\, \mathcal{F}_{\mu,s}\right]\big| + \frac{L_\mu}{l_\mu}|s-t| + 62L_\mu \tag{3.59}$$

and

$$\big|\mathbb{E}\left[m_\mu(y,\mathcal{S}_{\mu,t}) \,|\, \mathcal{F}_{\mu,t}\right] - \mathbb{E}\left[m_\mu(y,\mathcal{S}_{\mu,s}) \,|\, \mathcal{F}_{\mu,s}\right]\big| \leq \frac{L_\mu}{l_\mu}|s-t| + (2a_\mu+36)L_\mu. \tag{3.60}$$

These are proved in the next two steps.

Step 2. The proof of (3.59). For every $0 < s \leq t$, we have $\{y \in \mathcal{S}_{\mu,s}\} \in \mathcal{F}_{\mu,s} \subseteq \mathcal{F}_{\mu,t}$ and hence

$$X_t = \mathbb{E}\big[m_\mu(y,0) \,|\, \mathcal{F}_{\mu,t}\big]\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} + \mathbb{E}\big[m_\mu(y,0)\mathbb{1}_{\{y\notin\mathcal{S}_{\mu,s}\}} \,|\, \mathcal{F}_{\mu,t}\big] - M_\mu(y) \tag{3.61}$$

and

$$X_s = \mathbb{E}\big[m_\mu(y,0) \,|\, \mathcal{F}_{\mu,s}\big]\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} + \mathbb{E}\big[m_\mu(y,0)\mathbb{1}_{\{y\notin\mathcal{S}_{\mu,s}\}} \,|\, \mathcal{F}_{\mu,s}\big] - M_\mu(y). \tag{3.62}$$

According to (3.51), we have

$$\left|\mathbb{E}\big[m_\mu(y,0)\,|\,\mathcal{F}_{\mu,t}\big]\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}} - \mathbb{E}\big[m_\mu(y,0)\,|\,\mathcal{F}_{\mu,s}\big]\mathbb{1}_{\{y\in\mathcal{S}_{\mu,s}\}}\right| \le 36L_\mu.$$

Therefore, subtracting (3.62) from (3.61), we obtain

$$|X_t - X_s| \le \left|\mathbb{E}\big[m_\mu(y,0)\mathbb{1}_{\{y\notin\mathcal{S}_{\mu,s}\}}\,|\,\mathcal{F}_{\mu,t}\big] - \mathbb{E}\big[m_\mu(y,0)\mathbb{1}_{\{y\notin\mathcal{S}_{\mu,s}\}}\,|\,\mathcal{F}_{\mu,s}\big]\right| + 36L_\mu.$$

Using (3.47) to estimate the right side of the above inequality, we find

$$|X_t - X_s| \le \left|\mathbb{E}\big[m_\mu(y,\mathcal{S}_{\mu,s})\,|\,\mathcal{F}_{\mu,t}\big] - \mathbb{E}\big[m_\mu(y,\mathcal{S}_{\mu,s})\,|\,\mathcal{F}_{\mu,s}\big]\right|\mathbb{1}_{\{y\notin\mathcal{S}_{\mu,s}\}} + 52L_\mu. \tag{3.63}$$

According to (3.44) and (2.11),

$$\begin{aligned}\operatorname{dist}_H(\mathcal{S}_{\mu,t},\mathcal{S}_{\mu,s}) &\le \operatorname{dist}_H(\mathcal{S}_{\mu,t},\mathcal{R}_{\mu,t}) + \operatorname{dist}_H(\mathcal{R}_{\mu,t},\mathcal{R}_{\mu,s}) + \operatorname{dist}_H(\mathcal{R}_{\mu,s},\mathcal{S}_{\mu,s})\\ &\le \frac{|s-t|}{l_\mu} + 10.\end{aligned} \tag{3.64}$$

By the Lipschitz estimate (2.5), this yields

$$|m_\mu(y,\mathcal{S}_{\mu,t}) - m_\mu(y,\mathcal{S}_{\mu,s})| \le \frac{L_\mu}{l_\mu}|s-t| + 10L_\mu.$$

After inserting this into (3.63), we get (3.59).

Step 3. The proof of (3.60). We use the discrete approximation constructed in the previous subsection: by Lemma 3.9, we have

$$\begin{aligned}&\left|\mathbb{E}\big[m_\mu(y,\mathcal{S}_{\mu,t})\,\big|\,\mathcal{F}_{\mu,t}\big] - \mathbb{E}\big[m_\mu(y,\mathcal{S}_{\mu,s})\,\big|\,\mathcal{F}_{\mu,s}\big]\right|\\ &\qquad\le \left|\sum_{i\in\mathbb{N}}\mathbb{E}\left[m_\mu(y,\widetilde{K}_i)\right]\mathbb{1}_{E_i(t)} - \sum_{j\in\mathbb{N}}\mathbb{E}\left[m_\mu(y,\widetilde{K}_j)\right]\mathbb{1}_{E_j(s)}\right|\\ &\qquad\quad + 2L_\mu(1+a_\mu+12)\\ &\qquad\le \sum_{i,j\in\mathbb{N}}\mathbb{E}\left[\left|m_\mu(y,\widetilde{K}_i) - m_\mu(y,\widetilde{K}_j)\right|\right]\mathbb{1}_{E_i(t)\cap E_j(s)} + 2L_\mu(a_\mu+13).\end{aligned}$$

Next we recall from (3.64) that

$$E_i(t)\cap E_j(s)\ne\emptyset \quad\text{implies that}\quad \operatorname{dist}_H(\widetilde{K}_i,\widetilde{K}_j) \le \operatorname{dist}_H(K_i,K_j) \le \frac{|s-t|}{l_\mu} + 10.$$

Using the Lipschitz estimate (2.5), this yields

$$\begin{aligned}&\sum_{i,j\in\mathbb{N}}\mathbb{E}\left|m_\mu(y,\widetilde{K}_i) - m_\mu(y,\widetilde{K}_j)\right|\mathbb{1}_{E_i(t)\cap E_j(s)}\\ &\qquad\qquad\le \left(\frac{L_\mu}{l_\mu}|s-t| + 10L_\mu\right)\sum_{i,j\in\mathbb{N}}\mathbb{1}_{E_i(t)\cap E_j(s)} = \frac{L_\mu}{l_\mu}|s-t| + 10L_\mu.\end{aligned}$$

This completes the proof of (3.60). □

4. Estimate of the statistical bias

The main result of this section is an estimate of the nonrandom error, that is, for the difference between $\mathbb{E}\left[m_\mu(y,0)\right]$ and $\overline{m}_\mu(y)$ for $|y| \gg 1$.

Proposition 4.1. *Fix $\mu_0 \geq 1$. There exists $C > 0$, depending only on (d, q, Λ, μ_0) such that, for every $y \in \mathbb{R}^d$ and $0 < \mu \leq \mu_0$,*

$$M_\mu(y) \leq \overline{m}_\mu(y) + C\left(\frac{|y|^{\frac{2}{3}}}{\mu^2} + \frac{|y|^{\frac{1}{3}}}{\mu^4}\right)\log\left(2 + \frac{|y|}{\mu}\right). \tag{4.1}$$

The section is devoted to the proof of this result. Before we begin, we first show that Propositions 3.1 and 4.1 imply Theorem 1.

Proof of Theorem 1. The estimate (1.18) is a straightforward consequence of Proposition 3.1 and Proposition 4.1 if $|y| \geq C\mu^{-2}$. If $|y| < C\mu^{-2}$, the result also holds for $\lambda \geq 2CL_\mu\mu^{-2} = C\mu^{-2}$ because, from the Lipschitz estimates:

$$m_\mu(y,0) - \overline{m}_\mu(y) \leq 2L_\mu|y| \leq 2CL_\mu\mu^{-2} \leq \lambda.$$

Let us now check estimate (1.16). Owing to the stationary of m_μ and its subadditivity (see (2.6)), we have

$$\begin{aligned} M_\mu(y+z) =& \mathbb{E}\left[m_\mu(y+z,0)\right] \leq \mathbb{E}\left[m_\mu(y,0)\right] + \mathbb{E}\left[m_\mu(y+z,y)\right] + L_\mu \\ =& M_\mu(y) + M_\mu(z) + L_\mu. \end{aligned}$$

Hence $M_\mu(\cdot) + L_\mu$ is a subadditive quantity and we have, for every $y \in \mathbb{R}^d$,

$$M_\mu(y) + L_\mu \geq \inf_{t\geq 1} t^{-1}(M_\mu(ty) + L_\mu) = \lim_{t\to\infty} t^{-1}(M_\mu(ty) + L_\mu) = \overline{m}_\mu(y). \tag{4.2}$$

The estimate (1.16) is immediate from (4.2) and (3.2) if $|y| \geq C\mu^{-2}$. If $|y| < C\mu^{-2}$ but $\lambda \geq C\mu^{-2}$, the estimate also thanks to the same argument as above. □

4.1. Introduction of the approximating quantity.

Throughout the rest of this section, we fix $\mu_0 \geq 1$ and $0 < \mu \leq \mu_0$ and let C and c denote positive constants which may depend on (q, Λ, μ_0), but vary in each occurrence.

The proof of Proposition 4.1 is based on the introduction of an approximately *superadditive* quantity which approximates $M_\mu(y)$. Since the latter is nearly subadditive, we deduce an estimate (which depends on the quality of the approximation) for the difference of $M_\mu(y)$ and $\overline{m}_\mu(y)$. The key idea, which goes back to Alexander [1], is that $t \to \mathbb{E}\left[m_\mu(H_t,0)\right]$ is almost a superadditive quantity (here H_t is a plane defined below). However, we cannot use this quantity directly, and must settle for a further approximation. In this subsection, we introduce the relevant approximating quantity and make some preliminary estimates.

We fix a unit direction $e \in \partial B_1$. For convenience, we assume $e = e_d := (0,\dots,0,1)$. For each $t > 0$, define the plane

$$H_t := te + \{e\}^\perp = \left\{(x',t) \,:\, x' \in \mathbb{R}^{d-1}\right\} \tag{4.3}$$

and the discrete version

$$\widehat{H}_t := \left\{(n,t) \,:\, n \in \mathbb{Z}^{d-1}\right\}. \tag{4.4}$$

We also denote, for $t > 0$, the halfspaces

$$H_t^+ = \left\{(x', x_d) \in \mathbb{R}^d \,:\, x_d \geq t\right\} \qquad \text{and} \qquad H_t^- = \left\{(x', x_d) \in \mathbb{R}^d \,:\, x_d \leq t\right\}. \tag{4.5}$$

For $y \in H_t^-$ we set

$$m_\mu(H_t, y) := \min_{z \in H_t} m_\mu(z, y) = \min\left\{s \geq 0,\ \mathcal{R}_{\mu,s}(y) \cap H_t \neq \emptyset\right\}.$$

Define, for each $\sigma, t > 0$, the quantities

$$G_{\mu,\sigma}(t) := \sum_{y \in \widehat{H}_t} \mathbb{E}\left[\exp\left(-\sigma m_\mu(y, 0)\right)\right] \quad \text{and} \quad g_{\mu,\sigma}(t) := -\frac{1}{\sigma}\log G_{\mu,\sigma}(t). \tag{4.6}$$

We will see below in Lemma 4.8 that $g_{\mu,\sigma}(t)$ is a good approximation of $\mathbb{E}\left[m_\mu(H_t, 0)\right]$, at least for appropriate choice of the parameter σ.

We first show that we can restrict the sum for $G_{\mu,\sigma}(t)$ to a finite number of indices and still obtain a good approximate of $G_{\mu,\sigma}$. This follows from (2.8), which implies that far away points cannot make up a large proportion of the sum in the definition of $G_{\mu,\sigma}$. As the argument is nearly identical to that of [2, Lemmas 5.2], we omit the proof.

Lemma 4.2. *There exists $C > 0$ such that, for each $t \geq 1$, $0 < \sigma \leq 1$ and $R \geq (L_\mu/l_\mu)t$,*

$$G_{\mu,\sigma}(t) \leq C\sigma^{1-d} \sum_{y \in \widehat{H}_t \cap B_R} \mathbb{E}\left[\exp\left(-\sigma m_\mu(y, 0)\right)\right]. \tag{4.7}$$

We next show that $g_{\mu,\sigma}(t)$ well–approximates $\mathbb{E}\left[m_\mu(H_t, 0)\right]$.

Lemma 4.3. *There exists $C > 0$ such that, for every $t > C\mu^{-2}$ and $0 < \sigma \leq 1$,*

$$\mathbb{E}\left[m_\mu(H_t, 0)\right] - C\left(\frac{\sigma t}{\mu^5} + \frac{1}{\sigma}\log\left(2 + \frac{t}{\sigma\mu}\right)\right) \leq g_{\mu,\sigma}(t) \leq \mathbb{E}\left[m_\mu(H_t, 0)\right] + C. \tag{4.8}$$

Proof. The upper bound in (4.8) is relatively easy and is essentially the same as in the proof of the analogous bound in [2, Lemma 5.3].

To obtain the lower bound, we use both (3.2) and (4.7). We have

$$\begin{aligned}
\mathbb{E}\left[\exp\left(-\sigma m_\mu(y, 0)\right]\right. &= \int_0^\infty \sigma\exp(-\sigma s)\mathbb{P}\left[m_\mu(y, 0) \leq s\right]\, ds \\
&\leq \exp\left(-\sigma M_\mu(y)\right) + \int_0^{M_\mu(y)} \sigma\exp(-\sigma s)\mathbb{P}\left[m_\mu(y, 0) \leq s\right]\, ds \\
&= \left(1 + \int_0^{M_\mu(y)} \sigma\exp(\sigma\lambda)\mathbb{P}\left[m_\mu(y, 0) - M_\mu(y) \leq -\lambda\right]\, d\lambda\right)\exp\left(-\sigma M_\mu(y)\right).
\end{aligned}$$

Applying (3.2) for $|y| > C\mu^{-2}$ and using $\sigma \in (0, 1]$, we obtain

$$\mathbb{E}\left[\exp\left(-\sigma m_\mu(y, 0)\right]\right. \leq \left(1 + \int_0^{M_\mu(y)} \exp\left(\sigma\lambda - \frac{\mu^4\lambda^2}{C|y|}\right)\, d\lambda\right)\exp\left(-\sigma M_\mu(y)\right).$$

We estimate the integrand above by

$$\sigma\lambda - \frac{\mu^4\lambda^2}{C|y|} = -\frac{\mu^4}{C|y|}\left(\lambda - \frac{\sigma C|y|}{2\mu^4}\right)^2 + \frac{1}{4\mu^4}\sigma^2 C|y| \leq \frac{1}{4\mu^4}\sigma^2 C|y|$$

and so we get

$$\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,0)\right]\right. \leq \left(1 + M_\mu(y)\exp\left(\frac{1}{4\mu^4}\sigma^2 C|y|\right)\right)\exp\left(-\sigma M_\mu(y)\right). \tag{4.9}$$

Fix $t \geq C\mu^{-2}$. Summing (4.9) over $y \in \widehat{H}_t \cap B_R$, taking $R := (L_\mu/l_\mu)t$ and applying (4.7) (which holds because $|y| \geq t \geq C\mu^{-2}$), we get

$$\begin{aligned}
G_{\mu,\sigma}(t) &\leq C\sigma^{1-d}\sum_{y\in\widehat{H}_t\cap B_R}\left(1 + M_\mu(y)\exp\left(\frac{1}{4\mu^4}\sigma^2 C|y|\right)\right)\exp\left(-\sigma M_\mu(y)\right)\\
&\leq C\sigma^{1-d}\sum_{y\in\widehat{H}_t\cap B_R}\left(1 + M_\mu(y)\exp\left(\frac{1}{4\mu^4}\sigma^2 C|y|\right)\right)\exp\left(-\sigma\mathbb{E}\left[m_\mu(H_t,0)\right]\right)\\
&\leq C\sigma^{1-d}R^{d-1}\exp\left(-\sigma\mathbb{E}\left[m_\mu(H_t,0)\right]\right)\left(1 + L_\mu t\exp\left(\frac{1}{4\mu^4}\sigma^2 CR\right)\right).
\end{aligned}$$

By (2.7) we have $R \leq Ct/\mu$, and thus we deduce that

$$G_{\mu,\sigma}(t) \leq Ct^d\sigma^{1-d}\mu^{1-d}\exp\left(-\sigma\mathbb{E}\left[m_\mu(H_t,0)\right] + \frac{C\sigma^2 t}{\mu^5}\right).$$

Taking logarithms, dividing by $-\sigma$ and rearranging yields the lemma. □

In the next lemma, we compare $m_\mu(\cdot, H_t)$ with $\min_{z\in\widehat{H}_t} m_\mu(\cdot, z)$. In the first-order case, these quantities are equal and so there is nothing to prove (in fact, the first quantity is defined in terms of the second, see [2, (3.26)]). This is because of the peculiarity that, in the first-order case, the convexity of H ensures that the *minimum* of a family of solutions is a *sub*solution (and hence also a solution). Of course, this is not true in the second-order case, but nevertheless we are able to use the convexity of H to show that a *perturbation* of the second quantity is still a subsolution, and hence the desired estimate follows by the definition of the first. The idea of using convexity to perturb a supersolution into a subsolution is a natural one and was used previously for example in [13, Lemma 6.14].

Lemma 4.4. *There exists a constant $C > 0$ such that, for $t > 0$ and $y \in H^+_{t+1}$,*

$$\min_{z\in\widehat{H}_t\cap B_{L_\mu|y|/l_\mu}} m_\mu(y,z) \leq m_\mu(y,H_t) + C\left(1 + \frac{|y|}{\mu}\log\left(2 + \frac{|y|}{\mu}\right)\right)^{\frac{1}{2}}.$$

Proof. With $\theta > 0$ selected at the end of the argument, we introduce the test function

$$Z(y) := -\frac{1}{\theta}\log\left(\sum_{z\in\widehat{H}_t}\exp\left(-\theta m_\mu(y,z)\right)\right).$$

The main point of the proof is that Z is both a good approximation of $\min_{z\in\widehat{H}_t} m_\mu(\cdot, z)$ as well as (almost) a subsolution of (2.1) in H^+_{t+t}.

Step 1. We show that Z is a good approximation of $\min_{z\in\widehat{H}_t} m_\mu(\cdot, z)$. Arguing as in the proof of Lemma 4.2, we can show that there exists $C > 0$ such that, for every $t > 0$,

$y \in H_{t+1}^+$ and $R \geq (L_\mu/l_\mu)|y|$,

$$\begin{aligned}\sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right) &\leq \left(1+C\theta^{1-d}\right) \sum_{z\in\widehat{H}_t\cap B_R} \exp\left(-\theta m_\mu(y,z)\right)\\ &\leq R^{d-1}\left(1+C\theta^{1-d}\right) \max_{z\in\widehat{H}_t\cap B_R} \exp\left(-\theta m_\mu(y,z)\right).\end{aligned}$$

In particular,

$$Z(y) \geq \min_{z\in\widehat{H}_t\cap B_R} m_\mu(y,z) - \frac{1}{\theta}\log\left(R^{d-1}\left(1+C\theta^{1-d}\right)\right). \tag{4.10}$$

Step 2. We show that Z is (almost) a subsolution of (2.1) in H_t^+. The claim is that

$$-\operatorname{tr}\left(A(y)D^2Z(y)\right) + H(DZ(y),y) \leq \mu + 2\Lambda L_\mu^2\theta \quad \text{in } H_{t+1}^+. \tag{4.11}$$

As in the proof of Lemma 3.20, we check this inequality assuming that the functions are smooth. A rigorous argument in the viscosity sense follows by performing essentially identical computations on a smooth test function. It is convenient to work with

$$W(y) = \exp\{-\theta Z(y)\} = \sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right).$$

Computing the derivatives of Z in terms of W, we find

$$\begin{aligned}DZ(y) &= -\frac{1}{\theta W(y)} DW(y),\\ D^2Z(y) &= -\frac{1}{\theta W(y)} D^2W(y) + \frac{1}{\theta W^2(y)} DW(y)\otimes DW(y),\end{aligned}$$

and computing the derivatives of W in terms of $m_\mu(\cdot,z)$ yields

$$\begin{aligned}DW(y) &= -\theta \sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right) Dm_\mu(y,z),\\ D^2W(y) &= -\theta \sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right)\left(D^2m_\mu(y,z) - \theta Dm_\mu(y,z)\otimes Dm_\mu(y,z)\right),\end{aligned}$$

The Lipschitz bound $|Dm_\mu(\cdot,z)| \leq L_\mu$ thus gives the estimates

$$\left|DW(y)\right| \leq \theta L_\mu W(y), \quad \left|D^2W(y) + \theta \sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right) D^2m_\mu(y,z)\right| \leq \theta^2 L_\mu^2 W(y).$$

Assembling these together, we obtain

$$\begin{aligned}&-\operatorname{tr}\left(A(y)D^2Z(y)\right) + H(DZ(y),y)\\ &\qquad\leq \frac{1}{\theta W(y)}\operatorname{tr}\left(A(y)D^2W(y)\right) + H\left(-\frac{1}{\theta W(y)}DW(y),y\right) + \Lambda L_\mu^2\theta.\end{aligned}$$

Next, we observe that

$$\frac{1}{\theta W(y)} \operatorname{tr}\big(A(y)D^2W(y)\big) \leq -\frac{1}{W(y)} \sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right) \operatorname{tr}\big(A(y)D^2m_\mu(y,z)\big) + \Lambda L_\mu^2\theta$$

and, by the convexity of H and the definition of W,

$$\begin{aligned} H\left(-\frac{1}{\theta W(y)}DW(y), y\right) &= H\left(\frac{1}{W(y)}\sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right) Dm_\mu(y,z), y\right) \\ &\leq \frac{1}{W(y)}\sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right) H(Dm_\mu(y,z),y). \end{aligned}$$

Putting the last three inequalities together, we obtain

$$\begin{aligned} &-\operatorname{tr}\big(A(y)D^2Z(y)\big) + H(DZ(y),y) - 2\Lambda L_\mu^2\theta \\ &\quad\leq \frac{1}{W(y)}\sum_{z\in\widehat{H}_t} \exp\left(-\theta m_\mu(y,z)\right)\Big(-\operatorname{tr}\big(A(y)D^2m_\mu(y,z)\big) + H(Dm_\mu(y,z),y)\Big) \\ &\quad= \mu. \end{aligned}$$

This completes the proof of (4.11).

Step 3. The conclusion. Due to (4.11) and the convexity of H and $H(0,y)\leq 0$, we obtain that $\zeta(y) := (1+2\Lambda L_\mu^2\theta/\mu)^{-1}Z(y)$ is a subsolution of

$$-\operatorname{tr}\big(A(y)D^2\zeta\big) + H(D\zeta,y) \leq \mu \quad \text{in } H_{t+1}^+. \tag{4.12}$$

By Lipschitz estimate on m_μ, we have, for every $y\in H_{t+1}$,

$$Z(y) \leq \min_{z\in\widehat{H}_t} m_\mu(y,z) \leq L_\mu.$$

By the maximality of $m_\mu(\cdot,H_t)$, we obtain:

$$(1+2\Lambda L_\mu^2\theta/\mu)^{-1}\left(Z(y)-L_\mu)\right) \leq m_\mu(y,H_t),$$

which can be rearranged to read

$$Z(y) \leq (1+2\Lambda L_\mu^2\theta/\mu)m_\mu(y,H_t) + C \leq m_\mu(y,H_t) + \frac{2\Lambda L_\mu^3\theta|y|}{\mu} + C.$$

Using (4.10) we get

$$\min_{z\in\widehat{H}_t\cap B_R} m_\mu(y,z) \leq m_\mu(y,H_t) + \frac{1}{\theta}\log\left(R^{d-1}\left(1+C\theta^{1-d}\right)\right) + \frac{\Lambda^2L_\mu^3\theta|y|}{\mu} + C.$$

Choosing $R := (L_\mu/l_\mu)|y|$ and $\theta := |y|^{-\frac{1}{2}}\mu^{\frac{1}{2}}(\log(|y|/\mu))^{\frac{1}{2}}$ gives the lemma. □

4.2. **The (almost) superadditivity of $g_{\mu,\sigma}$.** In this subsection we show that $g_{\mu,\sigma}$ is superadditive, up to a small error, and as a corollary derive a rate of the convergence of the deterministic quantity $t^{-1}\mathbb{E}\left[m_\mu(H_t,0)\right]$ to its limit $\overline{m}_\mu(H_1)$. The statements of these assertions appear respectively in Lemmas 4.6 and 4.7, below. The arguments for these facts depend on the localization results of Section 3.3, which we first must adapt to the setting here. This is the purpose of Lemma 4.5.

It is immediate that, for any $y \in H_t^+$, the random variable $m_\mu(y,H_t)$ is $\mathcal{G}(H_{t-1}^+)$–measurable. On another hand, we cannot expect $m_\mu(H_t,0)$ to be $\mathcal{G}(H_t^-)$–measurable for the same reason that $m_\mu(y,0)$ depends on the full σ–algebra $\mathcal{F}$, due to the second–order term in our equation. Nevertheless, we argue, using the estimates of Subsection 3.3, that $m_\mu(H_t,0)$ is close to being $\mathcal{G}(H_t^-)-$measurable.

Lemma 4.5. *For every $t > 1$,*

$$\left|m_\mu(H_t,0) - \mathbb{E}\big[m_\mu(H_t,0) \mid \mathcal{G}(H_{t+7}^-)\big]\right| \le 8L_\mu.$$

Proof. We introduce the "stopping time"

$$T := \inf\left\{s \ge 0 \,:\, \mathcal{S}_{\mu,s} \cap H_t^+ \ne \emptyset\right\}.$$

Here $\mathcal{S}_{\mu,s}$ is defined by (3.43) and we recall that it is adapted to the filtration $\mathcal{F}_{\mu,s}$. We claim that

$$T \quad \text{is} \quad \mathcal{G}(H_{t+7}^-)\text{–measurable} \tag{4.13}$$

and

$$|T - m_\mu(H_t,0)| \le 4L_\mu. \tag{4.14}$$

Observe that (4.13) and (4.14) yield the lemma, since they imply

$$\begin{aligned}
&\left|m_\mu(H_t,0) - \mathbb{E}\big[m_\mu(H_t,0) \,|\, \mathcal{G}(H_{t+7}^-)\big]\right| \\
&\qquad \le |m_\mu(H_t,0) - T| + \left|\mathbb{E}\big[T - m_\mu(H_t,0) \mid \mathcal{G}(H_{t+7}^-)\big]\right| && \text{(by (4.13))} \\
&\qquad \le 8L_\mu. && \text{(by (4.14))}
\end{aligned}$$

Step 1. The proof of (4.13). For each $\omega \in \Omega$, we let $\mathbf{i}_s(\omega)$ be the unique index $i \in \mathbb{N}$ such that $\omega \in E_i(s)$. Let $\mathcal{J}_t = \{i \in \mathbb{N} \,:\, K_i' \subseteq \mathrm{int}(H_t^-)\}$. Then, by definition, $T(\omega) = \inf\{s \ge 0 \,:\, \mathbf{i}_s(\omega) \notin \mathcal{J}_t\}$. Note that

$$\{\mathbf{i}_s \in \mathcal{J}_t\} = \bigcup_{i \in \mathcal{J}_t} \{\mathbf{i}_s = i\} = \bigcup_{i \in \mathcal{J}_t} E_i(s).$$

Recall that, by construction, $E_i(s) \in \mathcal{G}(K_i'')$. If $i \in \mathcal{J}_t$, then we have $K_i'' \subseteq H_{t+7}^-$ since $K_i' \subseteq H_t^-$. Therefore $\{\mathbf{i}_s \in \mathcal{J}_t\} \in \mathcal{G}(H_{t+7}^-)$, which proves our claim.

Step 2. The proof of (4.14). Pick $y \in H_{t+4}$ be such that $\tau := m_\mu(H_{t+4},0) = m_\mu(y,0)$. According to (3.44), we have $\mathrm{dist}_H(\mathcal{R}_{\mu,\tau}, \mathcal{S}_{\mu,\tau}) \le 4$. So there exists $z \in \mathcal{S}_{\mu,\tau}$ with $|z-y| \le 4$. As $y \in H_{t+4}$, we have $z \in H_t^+$. Therefore $T \le \tau$. By (2.5), we have

$$|m_\mu(H_{t+4},0) - m_\mu(H_t,0)| \le 4L_\mu,$$

which yields that $T \le m_\mu(H_t,0) + 4L_\mu$. On the other hand, if $x \in \mathcal{S}_{\mu,T} \cap H_t$, according to (3.44) there exists $z \in \mathcal{R}_{\mu,T(\omega)}$ such that $|z-x| \le 4$. Hence

$$m_\mu(H_t,0) \le m_\mu(x,0) \le m_\mu(z,0) + 4L_\mu \le T + 4L_\mu\ .$$

This completes the proof of (4.14). □

Lemma 4.6. *There is a constant $C > 0$ such that, for every $s, t > 1$ and $0 < \sigma \le 1$,*

$$g_{\mu,\sigma}(t+s) \ge g_{\mu,\sigma}(t) + g_{\mu,\sigma}(s) - C\left(\frac{1}{\sigma} + \frac{(s+t)^{\frac{1}{2}}}{\mu}\right)\log\left(\frac{s+t}{\sigma\mu}\right). \tag{4.15}$$

Proof. Fix $s, t > 1$. We first claim that, for any $y \in H_t^+$,

$$m_\mu(y,0) \ge m_\mu(H_t,0) + m_\mu(y,H_t) - 3L_\mu. \tag{4.16}$$

Indeed, observe that the maps $y \to m_\mu(y,0)$ and $y \to m_\mu(H_t,0) + m_\mu(y,H_t)$ are both maximal solutions of (3.11) in H_{t+1}^+ with boundary conditions on H_{t+1} respectively equal to $m_\mu(y,0)$ and $m_\mu(H_t,0)$. As, for $z \in H_{t+1}$, we have by Lipschitz estimates,

$$m_\mu(z,0) \ge m_\mu(z-e_d,0) - L_\mu \ge m_\mu(H_t,0) - L_\mu,$$

we obtain by comparison (see [6])

$$m_\mu(y,0) \ge m_\mu(H_t,0) + m_\mu(y,H_t) - L_\mu \qquad \text{in } H_{t+1}^+.$$

This implies (4.16) again thanks to the Lipschitz estimates.

Using once more the Lipschitz estimate for m_μ and Lemma 4.5, (4.16) becomes

$$m_\mu(y,0) \ge \mathbb{E}\big[m_\mu(H_t,0) \mid \mathcal{G}(H_{t+7}^-)\big] + m_\mu(y,H_{t+8}) - C.$$

In light of (1.12), the random variables $\mathbb{E}\big[m_\mu(H_t,0) \mid \mathcal{G}(H_{t+7}^-)\big]$ and $m_\mu(y,H_{t+8})$ are independent and thus

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,0)\right)\right]\\ &\qquad \le \exp\left(C\sigma\right)\mathbb{E}\left[\exp\left(-\sigma\mathbb{E}\big[m_\mu(H_t,0) \mid \mathcal{G}(H_{t+7}^-)\big]\right)\right]\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,H_{t+8})\right)\right].\end{aligned}$$

Using Lemma 4.5 again, we obtain

$$\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,0)\right)\right] \quad\le\quad \exp\left(C\sigma\right)\mathbb{E}\left[\exp\left(-\sigma m_\mu(H_t,0)\right)\right]\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,H_{t+8})\right)\right].$$

Returning to the discrete setting, we have, for $R := L_\mu(s+t)/l_\mu$, and thanks to the Lipschitz estimates:

$$m_\mu(H_t,0) \ge \min_{z\in\widehat{H}_t\cap B_R} m_\mu(z,0) - L_\mu(d-1)^{\frac{1}{2}}. \tag{4.17}$$

On another hand Lemma 4.4 implies that, for $y \in \widehat{H}_{t+s} \cap B_R$,

$$\begin{aligned}m_\mu(y,H_{t+8}) &\ge \min_{z\in\widehat{H}_{t+8}} m_\mu(y,z) - C\left(1 + \frac{R}{\mu}\log\left(2+\frac{R}{\mu}\right)\right)^{\frac{1}{2}}\\ &\ge \min_{z\in\widehat{H}_t} m_\mu(y,z) - C\left(1 + \frac{R}{\mu}\log\left(2+\frac{R}{\mu}\right)\right)^{\frac{1}{2}}.\end{aligned}$$

Combining these inequalities, we obtain

$$\begin{aligned}&\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,0)\right)\right]\\ &\leq \exp\left(C\sigma\left(1+\tfrac{R}{\mu}\log\left(2+\tfrac{R}{\mu}\right)\right)^{\frac12}\right)\sum_{z,z'\in\widehat{H}_t}\mathbb{E}\left[\exp\left(-\sigma m_\mu(z,0)\right)\right]\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,z')\right)\right].\end{aligned}$$

Note that, if $y\in\widehat{H}_{t+s}$ and $z'\in\widehat{H}_t$, then $y-z'\in\widehat{H}_s$. So, in view of the definition of $G_{\mu,\sigma}$ and the stationarity of m_μ, we have

$$\sum_{z'\in\widehat{H}_t}\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,z')\right)\right]=\sum_{z'\in\widehat{H}_t}\mathbb{E}\left[\exp\left(-\sigma m_\mu(y-z',0)\right)\right]=G_{\mu,\sigma}(s).$$

Therefore

$$\mathbb{E}\left[\exp\left(-\sigma m_\mu(y,0)\right)\right]\leq\exp\left(C\sigma\left(1+\frac{R}{\mu}\log\left(2+\frac{R}{\mu}\right)\right)^{\frac12}\right)G_{\mu,\sigma}(t)G_{\mu,\sigma}(s).$$

Summing over all $y\in\widehat{H}_{t+s}\cap B_R$ and using Lemma 4.2 yields

$$G_{\mu,\sigma}(s+t)\leq CR^{d-1}\sigma^{1-d}\exp\left(C\sigma\left(1+\frac{R}{\mu}\log\left(2+\frac{R}{\mu}\right)\right)^{\frac12}\right)G_{\mu,\sigma}(t)G_{\mu,\sigma}(s).$$

Taking logarithms and dividing by σ concludes the proof. □

The rest of this section follows [2], the main differences being the values of the constants. We next use Hammersley-Fekete lemma to obtain a rate of convergence for the means $t^{-1}\mathbb{E}\left[m_\mu(H_t,0)\right]$ to their limit $\overline{m}_\mu(H_t)$.

Lemma 4.7. *There exists a constant $C>0$ such that, for every $t\geq 0$,*

$$\mathbb{E}\left[m_\mu(H_t,0)\right]\leq\overline{m}_\mu(H_t)+C\left(\frac{t}{\mu^5}\right)^{\frac12}\log\left(2+\frac{t}{\mu}\right). \tag{4.18}$$

Proof. According to Lemma 4.6, the quantity $g_{\sigma,\mu}$ is almost superadditive. More precisely, for all $s,t>0$, we have

$$g_{\sigma,\mu}(s+t)\geq g_{\sigma,\mu}(s)+g_{\sigma,\mu}(t)-\Delta_{\sigma,\mu}(s+t), \tag{4.19}$$

where

$$\Delta_{\sigma,\mu}(t):=C\left(\frac{1}{\sigma}+\frac{\sqrt{t}}{\mu}\right)\log\left(2+\frac{t}{\sigma\mu}\right).$$

Since $\Delta_{\sigma,\mu}$ is increasing on $[1,\infty)$ and

$$\int_1^\infty\frac{\Delta_{\sigma,\mu}(t)}{t^2}\,dt<\infty,$$

we may apply Hammersley-Fekete lemma to deduce that $\overline{g}_{\sigma,\mu}:=\lim_{t\to\infty}g_{\sigma,\mu}(t)/t$ exists and, for every $t>1$,

$$\frac{1}{t}g_{\sigma,\mu}(t)-4\int_{2t}^\infty\frac{\Delta_{\sigma,\mu}(s)}{s^2}\,ds\leq\overline{g}_{\sigma,\mu}.$$

An easy integration by parts yields

$$4\int_{2t}^{\infty}\frac{\Delta_{\sigma,\mu}(s)}{s^2}\,ds \le C\left(\frac{1}{\sigma t}+\frac{1}{\mu\sqrt{t}}\right)\log\left(2+\frac{t}{\sigma\mu}\right).$$

Assume now that $t \ge C\mu^{-2}$ so that, in view of the second inequality in (4.8), we have

$$\overline{g}_{\sigma,\mu} \le \lim_{t\to+\infty}\frac{1}{t}\left(\mathbb{E}\left[m_\mu(H_t,0)\right]+C\right),$$

where

$$\lim_{t\to+\infty}\frac{1}{t}\mathbb{E}\left[m_\mu(H_t,0)\right]=\overline{m}_\mu(H_1)$$

from the growth estimate (2.8) and the locally uniform convergence of $z\to m_\mu(tz,0)/t$ to $\overline{m}_\mu(z)$ as $t\to+\infty$. Combining the previous three estimates, we obtain

$$\frac{1}{t}g_{\sigma,\mu}(t) \le \overline{m}_\mu(H_1)+C\left(\frac{1}{\sigma t}+\frac{1}{\mu\sqrt{t}}\right)\log\left(2+\frac{t}{\sigma\mu}\right).$$

Multiplying by t, applying the first inequality in (4.8) and using the positive homogeneity of $\overline{m}_\mu$ yields, for $t\ge C\mu^{-2}$,

$$\mathbb{E}\left[m_\mu(H_t,0)\right] \le \overline{m}_\mu(H_t)+C\left(\frac{\sigma t}{\mu^5}+\left(\frac{1}{\sigma}+\frac{\sqrt{t}}{\mu}\right)\log\left(2+\frac{t}{\sigma\mu}\right)\right),$$

and taking $\sigma := \mu^{\frac{5}{2}}t^{-\frac{1}{2}}(\log(2+t/\mu))^{\frac{1}{2}}$ completes the proof of (4.18) when $t\ge C\mu^{-2}$. If $t<C\mu^{-2}$, then

$$\mathbb{E}\left[m_\mu(H_t,0)\right]-\overline{m}_\mu(H_t)\le Ct\le C\left(\frac{t}{\mu^2}\right)^{1/2}$$

so that (4.18) holds as well. □

4.3. **Error estimates for $M_\mu(y)-\overline{m}_\mu(y)$ and the proof of** (1.18)**.** It is the rate of convergence of $t^{-1}M_\mu(ty)$ to $\overline{m}_\mu(y)$ that we wish to estimate, not that of $t^{-1}\mathbb{E}\left[m_\mu(H_t,0)\right]$ to $\overline{m}_\mu(H_1)$. In order to reach our desired goal, we must compare the quantities $M_\mu(y)$ and $\mathbb{E}\left[m_\mu(H_t,0)\right]$. This is accomplished in two steps. The first is to show that the quantities $\mathbb{E}\left[m_\mu(H_t,0)\right]$ and

$$M_\mu(H_t):=\min_{y\in H_t}M_\mu(y)$$

are close, which then gives an estimate for the difference between $M_\mu(H_t)$ and $\overline{m}_\mu(H_t)$. The second step is to use elementary convex geometry to relate $M_\mu(y)$ to the values of $M_\mu(H)$ for all the possible planes H passing through y.

Lemma 4.8. *There exists $C>0$ such that, for each $t\ge 1$,*

$$M_\mu(H_t)\le \mathbb{E}\left[m_\mu(H_t,0)\right]+C\left(\frac{t}{\mu^5}\log\left(2+\frac{t}{\mu}\right)\right)^{\frac{1}{2}}. \tag{4.20}$$

Proof. Let $R := (L_\mu/l_\mu)t$. We may choose $\widehat{z} \in \widehat{H}_t \cap B_R$ such that

$$m_\mu(\widehat{z}, 0) \leq m_\mu(H_t, 0) + L_\mu(d-1)^{\frac{1}{2}}.$$

For every $z \in H_t$ we have $\mathbb{E}\left[m_\mu(z,0)\right] = M_\mu(z) \geq M_\mu(H_t)$ and thus, for every $\lambda > 0$,

$$\left\{\omega \in \Omega \,:\, M_\mu(H_t) - m_\mu(H_t, 0, \omega) \geq \lambda + L_\mu(d-1)^{\frac{1}{2}}\right\} \subseteq \bigcup_{z \in \widehat{H}_t \cap B_R} \left\{\omega \in \Omega \,:\, m_\mu(z, 0, \omega) \leq M_\mu(z) - \lambda\right\}. \tag{4.21}$$

Assuming $t \geq C\mu^{-2}$ and applying (3.2), we find

$$\begin{aligned} &\mathbb{P}\left[M_\mu(H_t) - m_\mu(H_t, 0) \geq \lambda + L_\mu(d-1)^{\frac{1}{2}}\right] \\ &\quad \leq CR^{d-1} \max_{z \in H_t} \mathbb{P}\left[m_\mu(z,0) - M_\mu(z) \leq -\lambda\right] \\ &\quad \leq CR^{d-1} \exp\left(-\frac{\mu^4\lambda^2}{CR}\right) \leq C\mu^{1-d}t^{d-1} \exp\left(-\frac{\mu^5\lambda^2}{Ct}\right). \end{aligned} \tag{4.22}$$

We next estimate the right side of the inequality

$$M_\mu(H_t) - \mathbb{E}\left[m_\mu(H_t, 0)\right] \leq \int_0^\infty \mathbb{P}\left[M_\mu(H_t) - m_\mu(H_t, 0) \geq \lambda\right] \, d\lambda. \tag{4.23}$$

Fix $\beta \geq 1$, to be selected below, define

$$\lambda_1 := \left(\frac{\beta t}{\mu^5} \log\left(2 + \frac{t}{\mu}\right)\right)^{\frac{1}{2}}$$

and then estimate the right side of (4.23) by

$$\begin{aligned} &\int_0^\infty \mathbb{P}\left[M_\mu(H_t) - m_\mu(H_t, 0) \geq \lambda\right] \, d\lambda \\ &\quad \leq \lambda_1 + L_\mu(d-1)^{\frac{1}{2}} + \int_{\lambda_1}^\infty \mathbb{P}\left[M_\mu(H_t) - m_\mu(H_t, 0) \geq \lambda + L_\mu(d-1)^{\frac{1}{2}}\right] \, d\lambda \\ &\quad \leq \lambda_1 + L_\mu(d-1)^{\frac{1}{2}} + C\mu^{1-d}t^{d-1} \int_{\lambda_1}^\infty \exp\left(-\frac{\mu^5\lambda^2}{Ct}\right) \, d\lambda. \end{aligned}$$

Observe that

$$\begin{aligned} \mu^{1-d}t^{d-1} \int_{\lambda_1}^\infty \exp\left(-\frac{\mu^5\lambda^2}{Ct}\right) \, d\lambda &\leq \mu^{1-d}t^{d-1} \int_{\lambda_1}^\infty \exp\left(-\frac{\mu^5\lambda_1\lambda}{Ct}\right) \, d\lambda \\ &= C\mu^{1-d}t^{d-1} \frac{t}{\mu^5\lambda_1} \exp\left(-\frac{\mu^5\lambda_1^2}{Ct}\right) \leq C\left(\frac{t}{\mu}\right)^{d+4} \left(1 + \frac{t}{\mu}\right)^{-\beta/C}. \end{aligned}$$

Taking $\beta \geq C$, the last expression on the right is bounded by C. The previous two sets of inequalities and (4.23) yield

$$M_\mu(H_t) - \mathbb{E}\left[m_\mu(H_t, 0)\right] \leq \lambda_1 + L_\mu(d-1)^{\frac{1}{2}} + C \leq \lambda_1 + C, \tag{4.24}$$

which gives, for every $t \geq C\mu^{-2}$.

$$M_\mu(H_t) \leq \mathbb{E}\left[m_\mu(H_t,0)\right] + C\left(\frac{t}{\mu^5}\log\left(2+\frac{t}{\mu}\right)\right)^{\frac{1}{2}}.$$

If $t < C\mu^{-2}$, then we obtain that

$$M_\mu(H_t) - \mathbb{E}\left[m_\mu(H_t,0)\right] \leq Ct \leq C\left(\frac{t}{\mu^2}\right)^{1/2}.$$

Hence (4.20) also holds in this case. □

Lemmas 4.7 and 4.8 give an estimate on the difference of $M_\mu(H_t)$ and $\overline{m}_\mu(H_t)$.

Corollary 4.9. *There exists $C > 0$ such that, for every $t \geq 1$,*

$$M_\mu(H_t) \leq \overline{m}_\mu(H_t) + C\left(\frac{t}{\mu^5}\right)^{\frac{1}{2}}\log\left(2+\frac{t}{\mu}\right). \tag{4.25}$$

The previous corollary yields a rate of convergence for $M_\mu(y)$ to $\overline{m}_\mu(y)$.

Proof of Proposition 4.1. The first step is to show that, for every $z \in \mathbb{R}^d$ with $|z| > 1$,

$$z \in \operatorname{conv}\left\{y \in \mathbb{R}^d \,:\, M_\mu(y) \leq \overline{m}_\mu(z) + C'\left(\frac{|z|}{\mu^5}\right)^{\frac{1}{2}}\log\left(2+\frac{|z|}{\mu}\right)\right\}, \tag{4.26}$$

where $C' > 0$ is the constant C in (4.25). Suppose on the contrary that (4.26) fails for some $z \in \mathbb{R}^d$ with $t := |z| > 1$. By elementary convex separation, there exists a plane H with $z \in H$ such that

$$M_\mu(H) > \overline{m}_\mu(z) + A \quad \text{where} \quad A := C'\left(\frac{t}{\mu^5}\right)^{\frac{1}{2}}\log\left(2+\frac{t}{\mu}\right).$$

Since H is at most a distance of $|z| = t$ from the origin, we may assume with no loss of generality that $H = H_s$ for some $s \leq t$. We deduce that

$$M_\mu(H_s) > \overline{m}_\mu(z) + A \geq \overline{m}_\mu(H_s) + C'\left(\frac{s}{\mu^5}\right)^{\frac{1}{2}}\log\left(2+\frac{s}{\mu}\right),$$

a contradiction to (4.25). Thus (4.26) holds.

Next we recall that, according to [2, Lemma 4.9], there exists $C'' > 0$ such that, for every $N \in \mathbb{N}$ and $\alpha > 0$,

$$\operatorname{conv}\left\{y \in \mathbb{R}^d \,:\, M_\mu(y) \leq \alpha\right\} \subseteq \left\{y \in \mathbb{R}^d \,:\, M_\mu(Ny) \leq (N + C''/\mu)\alpha\right\}. \tag{4.27}$$

Fix $y \in \mathbb{R}^d$ with $|y| > 1$ and apply (4.26) to $z := y/N$, with $N \in \mathbb{N}^*$ to be chosen below, to conclude that

$$y/N \in \operatorname{conv}\left\{x \in \mathbb{R}^d \,:\, M_\mu(x) \leq \overline{m}_\mu(y/N) + C'\left(\frac{|y|}{N\mu^5}\right)^{\frac{1}{2}}\log\left(2+\frac{|y|}{N\mu}\right)\right\}$$

and, after an application of (4.27),

$$\begin{aligned} M_\mu(y) &\le \left(N + \frac{C''}{\mu}\right)\left(\overline{m}_\mu(y/N) + C'\left(\frac{|y|}{N\mu^5}\right)^{\frac12}\log\left(2+\frac{|y|}{N\mu}\right)\right) \\ &\le \overline{m}_\mu(y) + C''L_\mu\frac{|y|}{N\mu} + C'\left(\frac{N|y|}{\mu^5}\right)^{\frac12}\log\left(2+\frac{|y|}{N\mu}\right) \\ &\quad + C'C''\left(\frac{|y|}{N\mu^7}\right)^{\frac12}\log\left(2+\frac{|y|}{N\mu}\right). \end{aligned}$$

We take N be the smallest integer larger than $|y|^{\frac13}\mu$ to get

$$M_\mu(y) \le \overline{m}_\mu(y) + C\frac{|y|^{\frac23}}{\mu^2} + C\frac{|y|^{\frac23}}{\mu^2}\log\left(2+\frac{|y|^{\frac23}}{\mu^2}\right) + C\frac{|y|^{\frac13}}{\mu^4}\log\left(2+\frac{|y|^{\frac23}}{\mu^2}\right).$$

This is (4.1). □

5. Error estimates for the approximate cell problem

In this section, we prove Theorem 2. As in [2], the argument consists of controlling $\delta v^\delta(\cdot;p) + \overline{H}(p)$ by the difference $m_\mu(\cdot,0) - \overline{m}_\mu(\cdot)$ and then applying Theorem 1. The idea of proving the convergence of δv^δ to $-\overline{H}(p)$ by comparing to the maximal subsolutions was introduced in [4] and extended to the second-order case in [5], although the importance of the metric problem goes back to Lions [12].

This comparison is somewhat simpler in the first-order case, because the function $-m_\mu(0,\cdot)$ can be used as a global subsolution. This is not true if $A \not\equiv 0$ and, as explained in [5], one overcomes this difficulty by looking at directions p which are *exposed points* of sublevel-sets of the form $\{\overline{H} = \mu\}$ where $\mu > 0$: indeed, for these directions, the convergence of $\delta v^\delta(\cdot;p)$to $-\overline{H}(p)$ to 0 can be proved by using the convergence of $m_\mu(\cdot,0)$ to $\overline{m}_\mu(\cdot)$, because the function $\overline{m}_\mu$ is differentiable (and hence relatively flat) along the ray that touches the plane $y \mapsto p\cdot y$. For the other directions p, there are two cases: either $\overline{H}(p) > 0$, and p can be written as a convex combination of exposed points of some level-set $\{\overline{H} = \mu\}$ where $\mu > 0$ (this is Straszewiscz's theorem): then one can translate the convergence for the exposed points to the convergence for p. In the case $\overline{H}(p) = 0$, that is, p belongs to the "flat spot," there is no relation between the approximate corrector and the metric problem, and different technique must be applied. This is not a limitation of the method, but intrinsic to the problem: as shown in [2], the limit (1.20) cannot be quantified without further information about the law of H.

Our task in this section is to quantify the argument of [5]. For this purpose we must consider a stronger version of the notion of exposed points, the so-called points of r-*strict convexity*. For these points there is a sharp control of the difference $\delta v^\delta(\cdot;p) - \big(-\overline{H}(p)\big)$ by the difference $m_\mu(\cdot,0) - \overline{m}_\mu(\cdot)$: this yields a quantitative result, which can be translated into a rate for the directions p such that $\overline{H}(p) > 0$ by a *quantitative version* of Straszewiscz's theorem.

To handle the flat spot $\{\overline{H} = 0\}$, we use the additional assumption (1.21) which essentially rules out its existence. The condition implies in particular that the flat spot is $\{0\}$, and we obtain a convergence rate for points near zero by a simple interpolation. There are other, weaker hypotheses one could use to obtain quantitative results on the flat spot. As this is not our primary focus, we leave this issue to the reader.

Throughout this section, we fix $\xi \geq 1$, which will serve as an upper bound for $|p|$. The symbols C and c denote positive constants which may change in each occurrence and depend only on (d, q, Λ, ξ).

We first recall the following standard estimates on the approximate corrector, which will be needed throughout the section (c.f. [5, 6] for the proofs): for every $p, \widetilde{p} \in B_\xi$ and $y, z \in \mathbb{R}^d$,

$$-C \leq -\Lambda|p|^q - \operatorname*{ess\,sup}_{\Omega} H(p,0) \leq \delta v^\delta(y,p) \leq -\operatorname*{ess\,inf}_{\Omega} H(p,0) \leq -\frac{1}{\Lambda}|p|^q + \Lambda \leq C, \tag{5.1}$$

$$\left|v^\delta(y,p) - v^\delta(z,p)\right| \leq C|y-z|, \tag{5.2}$$

and

$$\left|\delta v^\delta(y,p) - \delta v^\delta(y,\widetilde{p})\right| \leq C\left|p - \widetilde{p}\right|. \tag{5.3}$$

We break the proof of Theorem 2 into several steps, beginning with a bound on $-\delta v^\delta(\cdot,p) - \overline{H}(p)$ from above, for $p \in \mathbb{R}^d$ satisfying $\overline{H}(p) > 0$. This is the main focus of this section and where our arguments are different from those of [2].

5.1. **Estimate of $-\delta v^\delta - \overline{H}$ from above, off the flat spot.** The first step in the proof of Theorem 2 is an estimate from below of the quantity $\delta v^\delta(\cdot;p) + \overline{H}(p)$.

Proposition 5.1. *There exists $C > 0$ such that, for every $|p| \leq \xi$ with $\overline{H}(p) > 0$ and $\delta, \lambda > 0$ which satisfy*

$$\lambda \geq C\left(\frac{\delta^{\frac{1}{7}}}{\overline{H}(p)^{\frac{6}{7}}} + \frac{\delta^{\frac{2}{5}}}{\overline{H}(p)^{\frac{12}{5}}}\right)\log\left(2 + \frac{1}{\delta\overline{H}(p)}\right), \tag{5.4}$$

we have

$$\mathbb{P}\Big[-\delta v^\delta(0,p) \geq \overline{H}(p) + \lambda\Big] \leq C\lambda^{-3d}\exp\left(-\frac{(\overline{H}(p))^4\lambda^4}{C\delta}\right) \tag{5.5}$$

We begin by recalling some relevant definitions. For a compact subset K of $\mathbb{R}^d$, we say that $x \in K$ is an *exposed point* of K if there is a plane P such that $P \cap K = \{x\}$. If $r > 0$, we say that $x \in K$ is a *point of r-strict convexity of K* if there is a closed ball B of radius r such that $K \subseteq B$ and $x \in K \cap \partial B$. Note that

$$\text{points of } r\text{-strict convexity} \subseteq \text{exposed points} \subseteq \text{extreme points}.$$

Lemma 5.2 (Quantitative version of Straszewicz's theorem)**.** *Suppose that K is compact, convex and let $r > \operatorname{diam}(K)$. Denote by K^r the intersection of all closed balls of radius r which contain K, and let K_r be the closed convex hull of the points of r-strict convexity of K. Then*

$$\operatorname{dist}_H(K_r, K^r) \leq \frac{(\operatorname{diam}(K))^2}{r}.$$

Lemma 5.2 implies Straszewiscz's theorem since, for every $\operatorname{diam}(K) < r < \infty$,

$$K_r \subseteq \operatorname{conv}\{\text{exposed points of } K\} \subseteq K \subseteq K^r.$$

In order to put Lemma 5.2 to good use, we need the following convex analytic lemma linking the points of r-strict convexity of the sublevel sets of $\overline{H}$ to a quantitative estimate of linear approximation for the $\overline{m}_\mu$'s.

Lemma 5.3. *Let $\mu, r > 0$ and $p \in \mathbb{R}^d$ be a point of r-strict convexity of $\{q \in \mathbb{R}^d : \overline{H}(q) \le \mu\}$. Then there exists $e \in \mathbb{R}^d$ with $|e| = 1$ such that $p = D\overline{m}_\mu(e)$ and, for every $x \in B_{1/2}$,*

$$0 \le \overline{m}_\mu(e + x) - \overline{m}_\mu(e) - p \cdot x \le r|x|^2. \tag{5.6}$$

We first present the proof of Proposition 5.1. The proofs of Lemmas 5.2 and 5.3 are postponed until the end of the subsection.

Proof of Proposition 5.1. We break the proof into five steps. In the first four steps, we prove the proposition under the assumption that p is a point of r-strict convexity of a sublevel set of $\overline{H}$. In the final step, we remove the condition on p by using the convexity of $\overline{H}$ and simple convex geometry facts (Carathéodory's Theorem and Lemma 5.2).

In each of steps 1, 2, and 3, we fix $\lambda, \delta, \mu > 0$, a point $p \in \mathbb{R}^d$ of r-strict convexity of the convex set $\{q \in \mathbb{R}^d : \overline{H}(q) \le \mu\}$ with $|p| \le \xi$ and coefficients $\omega = (\Sigma, H) \in \Omega$ belonging to the event that $-\delta v^\delta(0,p) \ge \overline{H}(p) + \lambda$. We note that $\mu = \overline{H}(p)$ and, in particular, $0 < \mu \le \Lambda|p|^q \le \Lambda\xi^q \le C$.

Step 1. We modify $v^\delta(\cdot, p)$ to prepare it for comparison. With $0 < c < 1$ selected below, define

$$w(y) := v^\delta(y,p) - v^\delta(0,p) + c\lambda\left(\left(1 + |y|^2\right)^{1/2} - 1\right).$$

Observe that w satisfies

$$-\operatorname{tr}\left(A(y)D^2 w\right) + H(p + Dw, y) \ge -\delta v^\delta(y,p) - \frac{1}{4}\lambda \quad \text{in } \mathbb{R}^d,$$

provided that $c > 0$ is chosen sufficiently small, depending on (q, Λ, ξ). Select $y_0 \in \mathbb{R}^d$ such that

$$w(y_0) = \inf_{\mathbb{R}^d} w \le w(0) = 0$$

and observe that $|y_0| \le C/\lambda\delta$. Next we define

$$\widehat{w}(y) := w(y) - w(y_0) + \widehat{c}\lambda\left(\left(1 + |y - y_0|^2\right)^{1/2} - 1\right)$$

and observe, for $\widehat{c} > 0$ chosen sufficiently small, depending on (q, Λ, ξ), $\widehat{w}$ satisfies

$$-\operatorname{tr}\left(A(y)D^2\widehat{w}\right) + H(p + D\widehat{w}, y) \ge -\delta v^\delta(y,p) - \frac{1}{2}\lambda \quad \text{in } \mathbb{R}^d.$$

Consider the domain

$$V := \left\{y \in \mathbb{R}^d : \widehat{w}(y) < \lambda/4\delta\right\}.$$

It is clear from the definition of y_0 and $\widehat{w}$ that

$$V \subseteq B(y_0, R/\delta),$$

for some $R > 0$, depending only on (q, Λ, ξ). Moreover, it is simple to check that

$$\widehat{w}(y) \le \lambda/4\delta \qquad \text{implies that} \qquad -\delta v^\delta(y,p) \ge -\delta v^\delta(0,p) - \frac{\lambda}{4} \ge \mu + \frac{\lambda}{4}.$$

Hence we obtain

$$\begin{cases} -\operatorname{tr}\left(A(y)D^2\widehat{w}\right) + H(p + D\widehat{w}, y) \ge \mu + \dfrac{1}{4}\lambda & \text{in } V, \\ \widehat{w}(y_0) = 0 \quad \text{and} \quad \widehat{w} \equiv \lambda/4\delta & \text{on } \partial V. \end{cases} \tag{5.7}$$

Step 2. We modify m_μ to prepare it for comparison. According to Lemma 5.3, there exists a unit vector e such that $p = D\overline{m}_\mu(e)$ and, for every $x \in \mathbb{R}^d$ with $|x| \le 1/2$,

$$|\overline{m}_\mu(e + x) - p \cdot (e + x)| \le r|x|^2. \tag{5.8}$$

With $s > R\delta^{-1} + 1$ to be selected below, define

$$\widehat{m}(y) := m_\mu(y, y_0 - se) - m_\mu(y_0, y_0 - se) - p \cdot (y - y_0)$$

and observe that $\widehat{m}$ satisfies $\widehat{m}(y_0) = 0$ and

$$-\operatorname{tr}\left(A(y)D^2\widehat{m}\right) + H(p + D\widehat{m}, y) = \mu \quad \text{in } \mathbb{R}^d \setminus B(y_0 - se, 1) \supseteq V. \tag{5.9}$$

Step 3. We apply the comparison principle (see [6, Theorem 2.1]). Using (5.7) and (5.9), we obtain

$$0 = \widehat{m}(y_0) - \widehat{w}(y_0) \le \sup_V (\widehat{m} - \widehat{w}) = \max_{y \in \partial V} (\widehat{m} - \widehat{w}) = \max_{y \in \partial V} \widehat{m} - \lambda/4\delta.$$

A rearrangement yields

$$\frac{\lambda}{4\delta} \le \max_{y \in \partial V} \left(m_\mu(y, y_0 - se) - m_\mu(y_0, y_0 - se) - p \cdot (y - y_0)\right).$$

Using that $\partial V \subseteq B(y_0, R/\delta)$, we deduce that, for some $|y_0| \le C/\lambda\delta$, either

$$\max_{y \in B(y_0, R/\delta)} \left(m_\mu(y, y_0 - se) - p \cdot (y - (y_0 - se))\right) \ge \frac{\lambda}{8\delta} \tag{5.10}$$

or else

$$m_\mu(y_0, y_0 - se) - \overline{m}_\mu(se) = m_\mu(y_0, y_0 - se) - p \cdot (se) \le -\frac{\lambda}{8\delta}. \tag{5.11}$$

From (5.8) and positive homogeneity of $\overline{m}_\mu$, we have, for every $y \in B(y_0, R/\delta)$ and assuming $R/(s\delta) \le 1/2$, that

$$|\overline{m}_\mu(se - y_0 + y) - p \cdot (se - y_0 + y)| \le rs^{-1}|y_0 - y|^2 \le rs^{-1}R^2\delta^{-2}.$$

Hence the first alternative (5.10) implies that, for $s \ge 16rR^2/(\delta\lambda)$,

$$\max_{y \in B(y_0, R/\delta)} \left(m_\mu(y, y_0 - se) - \overline{m}_\mu(se - y_0 + y)\right)) \ge \frac{\lambda}{8\delta} - rs^{-1}R^2\delta^{-2} \ge \frac{\lambda}{16\delta}. \tag{5.12}$$

Let us choose $s := 16rR^2/(\delta\lambda)$ (so that $R/(s\delta) = \lambda/(16rR) \le 1/2$ for r large enough).

Step 4. We apply the results of Theorem 1. To summarize the previous steps, we have shown that, if $-\delta v^\delta(0,p) \ge \overline{H}(p) + \lambda$ holds, then there exists $|y_0| \le C/\lambda\delta$ such that at least one of (5.11) and (5.12) must hold. This is a purely deterministic statement

relating the convergence of (1.14) and (1.20). Using this and the Lipschitz continuity of m_μ and $\overline{m}_\mu$, we deduce that

$$\begin{aligned}
&\mathbb{P}\Big[-\delta v^\delta(0,p) \geq \overline{H}(p) + \lambda \Big] \\
&\qquad \leq \sum_{y_0 \in B(C/\lambda\delta)\cap(c\lambda/\delta)\mathbb{Z}^d} \left(\mathbb{P}\left[m_\mu(y_0, y_0 - se) - \overline{m}_\mu(se) \leq -\frac{\lambda}{16\delta} \right] \right. \\
&\qquad\qquad \left. + \mathbb{P}\left[\max_{y \in B(y_0, R/\delta)} \left(m_\mu(y, y_0 - se) - \overline{m}_\mu(se - y_0 + y) \right) \geq \frac{\lambda}{32\delta} \right] \right)
\end{aligned}$$

To estimate the first term in the sum, we apply (1.16) to find that, if $\lambda \geq C\delta\mu^{-2}$, then

$$\begin{aligned}
\mathbb{P}\left[m_\mu(y_0, y_0 - se) - \overline{m}_\mu(se) \leq -\frac{\lambda}{16\delta} \right] &= \mathbb{P}\left[m_\mu(se, 0) - \overline{m}_\mu(se) \leq -\frac{\lambda}{16\delta} \right] \\
&\leq \exp\left(-\frac{\mu^4\lambda^2}{Cs\delta^2} \right) = \exp\left(-\frac{\mu^4\lambda^3}{Cr\delta} \right).
\end{aligned}$$

To bound the second term in the sum, we first use stationarity, the Lipschitz estimate for m_μ and $\overline{m}_\mu$ and the fact that the set $B(0, R/\delta)\cap(c\lambda/\delta)\mathbb{Z}^d$ has at most $C\lambda^{-d}$ elements to obtain

$$\begin{aligned}
&\mathbb{P}\left[\max_{y \in B(y_0, R/\delta)} \left(m_\mu(y, y_0 - se) - \overline{m}_\mu(se - y_0 + y) \right) \geq \frac{\lambda}{32\delta} \right] \\
&\qquad = \mathbb{P}\left[\max_{z \in B(0, R/\delta)} \left(m_\mu(z + se, 0) - \overline{m}_\mu(se + z) \right) \geq \frac{\lambda}{32\delta} \right] \\
&\qquad \leq \sum_{z \in B(0, R/\delta)\cap(c\lambda/\delta)\mathbb{Z}^d} \mathbb{P}\left[m_\mu(se + z, 0) - \overline{m}_\mu(se + z) \geq \frac{\lambda}{64\delta} \right] \\
&\qquad \leq C\lambda^{-d} \max_{z \in B(0, R/\delta)\cap(c\lambda/\delta)\mathbb{Z}^d} \mathbb{P}\left[m_\mu(se + z, 0) - \overline{m}_\mu(se + z) \geq \frac{\lambda}{64\delta} \right].
\end{aligned}$$

As $s = 16rR^2/(\delta\lambda)$, we apply the second statement of Theorem 1 to deduce that, if λ satisfies

$$\frac{\lambda}{\delta} \geq C\left(\frac{s^{\frac{2}{3}}}{\mu^2} + \frac{s^{\frac{1}{3}}}{\mu^4} \right) \log\left(2 + \frac{s}{\mu} \right) = C\left(\frac{r^{\frac{2}{3}}}{(\delta\lambda)^{\frac{2}{3}}\mu^2} + \frac{r^{\frac{1}{3}}}{(\delta\lambda)^{\frac{1}{3}}\mu^4} \right) \log\left(2 + \frac{r}{\delta\lambda\mu} \right),$$

then we have, by (1.18),

$$\begin{aligned}
&\mathbb{P}\left[\max_{y \in B(y_0, R/\delta)} \left(m_\mu(y, y_0 - se) - \overline{m}_\mu(se - y_0 + y) \right) \geq \frac{\lambda}{32\delta} \right] \\
&\qquad \leq C\lambda^{-d} \max_{z \in B(R/\delta)\cap(c\lambda/\delta)\mathbb{Z}^d} \mathbb{P}\left[m_\mu(se + z, 0) - \overline{m}_\mu(se + z) \geq \frac{\lambda}{64\delta} \right] \\
&\qquad \leq C\lambda^{-d} \exp\left(-\frac{\mu^4\lambda^3}{Cr\delta} \right).
\end{aligned}$$

We have proved that, if p is a point of r-strict convexity of $\{q \in \mathbb{R}^d : \overline{H}(q) \le \mu\}$ with $r \ge C$ and λ satisfies

$$\frac{\lambda}{\delta} \ge C\left(\frac{r^{\frac{2}{3}}}{(\delta\lambda)^{\frac{2}{3}}\mu^2} + \frac{r^{\frac{1}{3}}}{(\delta\lambda)^{\frac{1}{3}}\mu^4}\right)\log\left(2 + \frac{r}{\delta\lambda\mu}\right), \tag{5.13}$$

then

$$\mathbb{P}\left[-\delta v^\delta(0,p) \ge \overline{H}(p) + \lambda\right] \le C\lambda^{-3d}\exp\left(-\frac{\mu^4\lambda^3}{Cr\delta}\right).$$

Note that $\lambda \ge C\delta\mu^{-2}$ is redundant in view of (5.13), and the latter condition holds provided that

$$\lambda \ge C\left(\frac{r^{\frac{2}{5}}\delta^{\frac{1}{5}}}{\mu^{\frac{6}{5}}} + \frac{r^{\frac{1}{4}}\delta^{\frac{1}{2}}}{\mu^3}\right)\left(1 + |\log r| + |\log\delta| + |\log\mu|\right).$$

Step 5. We use the convexity of H to obtain an estimate for general $|p| \le K$. Fix $p \in \mathbb{R}^d$ with $|p| \le K$ and set $\mu := \overline{H}(p)$. According to Carathéodory's Theorem (c.f. [9]) and Lemma 5.2, for each $r > 0$, there exist points $p_1, \dots, p_{d+1}$ of r-strict convexity of $\{q \in \mathbb{R}^d : \overline{H}(q) \le \mu\}$ and $\theta_1, \dots, \theta_{d+1} \in [0,1]$ such that $\sum_{i=1}^{d+1}\theta_i = 1$ and

$$\left|p - \sum_{i=1}^{d+1}\theta_i p_i\right| \le \frac{C}{r}.$$

We consider coefficients $\omega = (\Sigma, H) \in \Omega$ belonging to the event that, for every $i = 1, \dots, d+1$,

$$-\delta v^\delta(0,p_i) \le \overline{H}(p_i) + \frac{\lambda}{2},$$

where $\overline{H}(p_i) \le \mu = \overline{H}(p)$. Then, by the convexity of $\overline{H}$ and the Lipschitz continuity of the map $p \mapsto -\delta v^\delta(y,p)$, we have

$$-\delta v^\delta(0,p) \le \sum_{i=1}^{d+1}\theta_i\left(-\delta v^\delta(0,p_i)\right) + \frac{C}{r} \le \sum_{i=1}^{d+1}\theta_i\overline{H}(p_i) + \frac{\lambda}{2} + \frac{C}{r} \le \overline{H}(p) + \frac{\lambda}{2} + \frac{C}{r}.$$

Choosing r such that $\dfrac{C}{r} = \dfrac{\lambda}{4} < \dfrac{\lambda}{2}$, we obtain that

$$\left\{\omega \in \Omega : -\delta v^\delta(0,p,\omega) \ge \overline{H}(p) + \lambda\right\} \subseteq \bigcup_{i=1}^{d+1}\left\{\omega \in \Omega : -\delta v^\delta(0,p_i,\omega) \ge \overline{H}(p_i) + \frac{\lambda}{2}\right\}.$$

Then from step 4 we obtain

$$\begin{aligned}\mathbb{P}\Big[-\delta v^\delta(0,p) \ge \overline{H}(p) + \lambda\Big] &\le \sum_{i=1}^{d+1}\mathbb{P}\Big[-\delta v^\delta(0,p_i) \ge \overline{H}(p_i) + \frac{\lambda}{2}\Big] \\ &\le C\lambda^{-3d}\exp\left(-\frac{\mu^4\lambda^3}{Cr\delta}\right) = C\lambda^{-3d}\exp\left(-\frac{\mu^4\lambda^4}{C\delta}\right),\end{aligned}$$

(which is (5.5)) provided that

$$\lambda \geq C\left(\frac{\delta^{\frac{1}{5}}}{\mu^{\frac{6}{5}}\lambda^{\frac{2}{5}}} + \frac{\delta^{\frac{1}{2}}}{\mu^{3}\lambda^{\frac{1}{4}}}\right)\left(1 + |\log(\lambda)| + |\log(\delta)| + |\log(\mu)|\right).$$

This condition holds if λ and δ satisfy

$$\lambda \geq C\left(\frac{\delta^{\frac{1}{7}}}{\mu^{\frac{6}{7}}} + \frac{\delta^{\frac{2}{5}}}{\mu^{\frac{12}{5}}}\right)\log\left(2 + \frac{1}{\delta\mu}\right),$$

which is (5.4). □

Proof of Lemma 5.2. It is easy to see that $(K_r)^r = K^r$, and therefore we may assume that $K = K_r$. Let $x \in K_r$ and $y \in K$ such that $|x - y| = \operatorname{dist}_H(K, K^r)$.

The hyperplane P passing through y which is perpendicular to $x - y$ separates K from the point x. That is, K is contained in the half-space S whose boundary is P and which does not contain x. In particular, if $s = \operatorname{diam} K$, then

$$K \subseteq S \cap \overline{B}(y, s).$$

Set

$$z := y + (r^2 - s^2)^{1/2}\frac{y - x}{|y - x|}$$

and observe that $K \subseteq S \cap \overline{(}y, s) \subseteq \overline{B}(z, r)$. It follows from the definition of K^r that $K^r \subseteq \overline{B}(z, r)$. Hence $x \in B(z, r)$ and

$$\operatorname{dist}_H(K, K^r) = |x - y| = |x - z| - |y - z| \leq r - \left(r^2 - s^2\right)^{1/2}.$$

The proof is complete, as $r - (r^2 - s^2)^{1/2} \leq \frac{1}{2}s^2 r^{-1} + o(r^{-1})$ as $r \to \infty$. □

Proof of Lemma 5.3. Write $K := \{q \in \mathbb{R}^d : \overline{H}(q) \leq \mu\}$ and select a closed ball B of radius r such that $K \subseteq B$ and $p \in \partial B$. Let e be the outward-pointing unit normal vector to ∂B at p. Let $\phi(x) := \max_{y \in B}(x \cdot y)$ denote the support function of B, and observe that, for every $x \in \mathbb{R}^d$,

$$p \cdot x \leq \overline{m}_\mu(x) \leq \phi(x)$$

with equality holding at $x = e$, that is, $\phi(e) = p \cdot e$. Hence, for every $x \in \mathbb{R}^d$,

$$p \cdot x = p \cdot (x + e) - p \cdot e \leq \overline{m}_\mu(x + e) - \overline{m}_\mu(e) \leq \phi(x + e) - \phi(e).$$

Rearranging and using the fact that $|D^2\phi(e)| = r$, we obtain

$$0 \leq \overline{m}_\mu(x + e) - \overline{m}_\mu(e) - p \cdot x \leq \phi(x + e) - \phi(e) - p \cdot x \leq \frac{1}{2}r|x|^2.$$ □

5.2. **Global estimate of $-\delta v^\delta - \overline{H}$.** The estimate from below for $-\delta v^\delta - \overline{H}$ away from the flat spot is nearly verbatim from the argument of [2, Theorem 2(ii)]. We state the result without proof, refereeing to [2] for details.

Proposition 5.4. *There is a constant $C > 0$, depending only on (d, q, Λ, ξ), such that, for every $p \in \mathbb{R}^d$ with $|p| \le \xi$ and $\overline{H}(p) > 0$, $\delta, \lambda > 0$ with*

$$\lambda \ge C \left(\frac{\delta^{\frac{1}{3}}}{(\overline{H}(p))^2} + \frac{\delta^{\frac{2}{3}}}{(\overline{H}(p))^4} \right) \log \left(2 + \frac{1}{\delta \overline{H}(p)} \right), \tag{5.14}$$

we have

$$\mathbb{P}\Big[-\delta v^\delta(0,p) \le \overline{H}(p) + \lambda \Big] \le C\delta^{-2d}\lambda^{-d} \exp\left(-\frac{(\overline{H}(p))^4 \lambda^2}{C\delta} \right)$$

We next deduce Theorem 2 from Propositions 5.1 and 5.4.

Proof of Theorem 2. The condition (1.21) implies (c.f. [5, 6]) that, for every $p \in \mathbb{R}^d$,

$$-\delta v^\delta(x,p) \ge \frac{1}{\Lambda}|p|^q \quad \text{and} \quad \overline{H}(p) \ge \frac{1}{\Lambda}|p|^q. \tag{5.15}$$

Since $\delta v^\delta(y,0) = \overline{H}(0) = 0$, (5.3) yields that, for every $|p| \le \xi$,

$$\left| \delta v^\delta(0,p) + \overline{H}(p) \right| \le C|p| \qquad \forall |p| \le K.$$

Hence to obtain (1.22), we may suppose without loss of generality that $\lambda \le C|p| \le C$. If conditions (5.4) and (5.14) hold, then Propositions 5.1 and 5.4, in view of $\lambda \le C|p| \le C$ and (5.15), yield

$$\begin{aligned} \mathbb{P}\Big[\left| \delta v^\delta(0,p) + \overline{H}(p) \right| \ge \lambda \Big] &\le C\lambda^{-d} \left(\lambda^{-2d} + \delta^{-2d} \right) \exp\left(-\frac{(\overline{H}(p))^4 \lambda^4}{C\delta} \right) \\ &\le C\lambda^{-d} \left(\lambda^{-2d} + \delta^{-2d} \right) \exp\left(-\frac{\lambda^{4(1+q)}}{C\delta} \right). \end{aligned}$$

This is (1.22). Using again that $\lambda \le C|p|$ and (5.15), we find that (5.4) and (5.14) are satisfied provided that we have both

$$\lambda \ge C \left(\frac{\delta^{\frac{1}{7}}}{\lambda^{\frac{6q}{7}}} + \frac{\delta^{\frac{2}{5}}}{\lambda^{\frac{12q}{5}}} \right) \log \left(2 + \frac{1}{\delta\lambda} \right)$$

and

$$\lambda \ge C \left(\frac{\delta^{\frac{1}{3}}}{\lambda^{2q}} + \frac{\delta^{\frac{2}{3}}}{\lambda^{4q}} \right) \log \left(2 + \frac{1}{\delta\lambda} \right).$$

One easily checks that the second condition is redundant with the first, and the first term in parentheses of the first condition is the limiting one. We obtain finally that both hold if

$$\lambda \ge C\delta^{\frac{1}{7+6q}} \left(1 + |\log \delta| \right).$$

This completes the proof of the theorem. □

6. Error estimates for the time dependent problem

In this section we prove Theorem 3 on the convergence rate of the solution u^ε of (1.23) to the solution u of the homogenized problem (1.24). This part is very close to the corresponding one for first order problems (see section 7 in [2]). However the second order term induces additional technicalities, so the result is not just routine adaptation of [2]: this is the reason why we provide the details.

Before beginning the proof, we introduce some notation and recall some properties of (1.23). According to [6], there exists $L > 0$, depending on $(d, q, \Lambda, \|g\|_{C^{1,1}(\mathbb{R}^d)})$ such that, for all $\varepsilon > 0$, $x, y \in \mathbb{R}^d$ and $s, t \geq 0$,

$$|u^\varepsilon(x,t) - u^\varepsilon(y,s)| \leq L\left(|x-y| + |s-t|\right) \tag{6.1}$$

and

$$|u(x,t) - u(y,s)| \leq L\left(|x-y| + |s-t|\right). \tag{6.2}$$

It also follows easily from this that for each $\varepsilon > 0$, $x \in \mathbb{R}^d$ and $0 \leq t \leq T$,

$$|u(x,t)| + |u^\varepsilon(x,t)| \leq K + LT \leq C(1+T). \tag{6.3}$$

The key point in the proof in the proof of Theorem 3 amounts to estimate the difference $u^\varepsilon - u$ by the difference between δv^δ and $\overline{H}$. This is a purely deterministic PDE fact, which we summarize in the following lemma. Throughout, C and c denote positive constants which may vary in each occurrence and depend only on $(d, q, \Lambda, \|g\|_{C^{1,1}(\mathbb{R}^d)})$.

Lemma 6.1. *There exists a constant $C > 0$ depending only on $(d, q, \Lambda, \|g\|_{C^{1,1}(\mathbb{R}^d)})$ such that, with $L > 0$ as in (6.1) and (6.2), we have, for any $\lambda, \delta, \varepsilon > 0$, with $0 < \varepsilon < \delta < 1$ and $R := C\left(\frac{1}{\lambda\varepsilon} + \frac{1}{\delta}\right)$,*

$$\begin{aligned} \left\{\omega \in \Omega \ : \ \sup_{x \in B_T,\ t \in [0,T]} u^\varepsilon(x,t,\omega) - u(x,t) > C\left(\lambda + \frac{\varepsilon^{\frac{2}{3}}}{\delta^{\frac{1}{3}}} + \frac{\varepsilon^{\frac{1}{3}}}{\delta^{\frac{2}{3}}}\right)\right\} \\ \subseteq \left\{\omega \in \Omega \ : \ \inf_{(y,p) \in B_R \times B_L} \left(-\delta v^\delta(y,p,\omega) - \overline{H}(p)\right) < -\lambda\right\} \end{aligned} \tag{6.4}$$

and

$$\begin{aligned} \left\{\omega \in \Omega \ : \ \inf_{x \in B_T,\ t \in [0,T]} u^\varepsilon(x,t) - u(x,t) < -C\left(\lambda + \frac{\varepsilon^{\frac{2}{3}}}{\delta^{\frac{1}{3}}} + \frac{\varepsilon^{\frac{1}{3}}}{\delta^{\frac{2}{3}}}\right)\right\} \\ \subseteq \left\{\omega \in \Omega \ : \ \sup_{(y,p) \in B_R \times B_L} \left(-\delta v^\delta(y,p) - \overline{H}(p)\right) > \lambda\right\}. \end{aligned} \tag{6.5}$$

Proof. Since the arguments for (6.4) and (6.5) are nearly identical we only prove (6.4). Fix $\lambda, \delta, \varepsilon > 0$, with $\varepsilon < \delta < 1$, and a realization of the coefficients $\omega = (\Sigma, H) \in \Omega$ such that

$$\inf_{(y,p) \in B_R \times B_L} \left(-\delta v^\delta(y,p) - \overline{H}(p)\right) \geq -\lambda. \tag{6.6}$$

We are going to show that

$$\sup_{x\in B_T,\ t\in[0,T]} u^\varepsilon(x,t) - u(x,t) \le C\left(\lambda + \frac{\varepsilon^{\frac{2}{3}}}{\delta^{\frac{1}{3}}} + \frac{\varepsilon^{\frac{1}{3}}}{\delta^{\frac{2}{3}}}\right). \tag{6.7}$$

Define $\zeta : \mathbb{R}^d \to \mathbb{R}^d$ by

$$\zeta(x) := \max\left\{\frac{L}{|x|}, 1\right\} x,$$

and notice that

$$|\zeta(x)| = L \wedge |x| \quad \text{and} \quad |\zeta(x) - \zeta(y)| \le |x-y|. \tag{6.8}$$

Fix parameters $\alpha, q > 0$ and $d \in (0,1)$ to be chosen below and consider the auxiliary function $\Phi : \mathbb{R}^d \times \mathbb{R}^d \times [0,T] \times [0,T] \times \Omega \to \mathbb{R}$ given by

$$\begin{aligned}\Phi(x,y,t,s) := u^\varepsilon(x,t) - u(y,s) - \varepsilon v^\delta\left(\frac{x}{\varepsilon}; \zeta\left(\frac{x-y}{\alpha}\right)\right) - \frac{1}{2\alpha}|x-y|^2 \\ - \frac{1}{2\varepsilon}(t-s)^2 - ds - q\left(1+|x|^2\right)^{\frac{1}{2}} + q.\end{aligned} \tag{6.9}$$

Using (5.1), (6.3) and (6.8), we have, for each $(x,y,t,s) \in \mathbb{R}^d \times \mathbb{R}^d \times [0,T] \times [0,T] \times \Omega$,

$$|\Phi(x,y,t,s)| \le C(1+T) + C\varepsilon\delta^{-1} - \frac{1}{2\alpha}|x-y|^2 - \frac{1}{2\varepsilon}(t-s)^2 - ds - q\left(1+|x|^2\right)^{\frac{1}{2}} + q. \tag{6.10}$$

It follows that, for each $\omega \in \Omega$, the function Φ attains its global maximum at some point of $\mathbb{R}^d \times \mathbb{R}^d \times [0,T] \times [0,T]$. Set

$$M(\omega) := \max_{\mathbb{R}^d\times\mathbb{R}^d\times[0,T]\times[0,T]} \Phi(\cdot).$$

We now record two elementary estimates that necessarily hold for any $\omega \in \Omega$ and at any global maximum point $(x_0, y_0, t_0, s_0) \in \mathbb{R}^d \times \mathbb{R}^d \times [0,T] \times [0,T]$ of $\Phi(\cdot)$, i.e., such that

$$\Phi(x_0, y_0, t_0, s_0) = M(\omega). \tag{6.11}$$

The inequality $\Phi(x_0, y_0, t_0, s_0) \ge \Phi(0,0,0,0)$ yields, in light of (6.10), that

$$q|x_0| + \frac{1}{2\varepsilon}(t_0 - s_0)^2 \le C(1+T) + \frac{C\varepsilon}{\delta} \le CT. \tag{6.12}$$

If $s_0 \neq 0$, then by (6.2) and since $s \mapsto u(y_0, s) + (s-t_0)^2/2\varepsilon$ has a minimum at $s = s_0$, we deduce that

$$|s_0 - t_0| \le L\varepsilon. \tag{6.13}$$

Inequality (6.13) is also satisfied for a similar reason if $t_0 \neq 0$, and trivially if $s_0 = t_0 = 0$. We also claim that

$$|x_0 - y_0| \le L\alpha. \tag{6.14}$$

If not, then $y \mapsto \zeta((x_0 - y)/\alpha)$ is constant in a neighborhood of y_0 and we obtain from (6.11) that

$$y \mapsto u(y, s_0) + \frac{1}{2\alpha}|x_0 - y|^2 \quad \text{has a local minimum at } y = y_0.$$

Due to (6.2), we conclude that $|x_0 - y_0| \leq L\alpha$, which is a contradiction. In particular,

$$\zeta\left(\frac{x_0 - y_0}{\alpha}\right) = \frac{x_0 - y_0}{\alpha}. \tag{6.15}$$

We now commence with the proof of (6.7), following the classical argument for the proof of the comparison principle for viscosity solutions and ideas from [7]. Fix ω for which (6.6) holds and select a point $(x_0, y_0, t_0, s_0) \in \mathbb{R}^d \times \mathbb{R}^d \times [0,T] \times [0,T]$ as in (6.11). We assume for the moment that $s_0 > 0$ and t_0 is sufficiently large, in a sense to be explained below.

The first step is to fix $(x,t) = (x_0, t_0)$, allow (y,s) to vary, and to use the equation for u. The goal is to derive (6.20), below. From (6.11), we see that

$$(y,s) \mapsto u(y,s) + \varepsilon v^\delta\left(\frac{x_0}{\varepsilon}; \zeta\left(\frac{x_0 - y}{\alpha}\right)\right) + \frac{1}{2\alpha}|x_0 - y|^2 + \frac{1}{2\varepsilon}(t_0 - s)^2 + ds \quad \text{has a local minimum at } (y,s) = (y_0, s_0). \tag{6.16}$$

According to (5.3) and (6.8),

$$\left|\varepsilon v^\delta\left(\frac{x_0}{\varepsilon}; \zeta\left(\frac{x_0 - y}{\alpha}\right)\right) - \varepsilon v^\delta\left(\frac{x_0}{\varepsilon}; \zeta\left(\frac{x_0 - y_0}{\alpha}\right)\right)\right| \leq \frac{C\varepsilon}{\delta}\left|\zeta\left(\frac{x_0 - y}{\alpha}\right) - \zeta\left(\frac{x_0 - y_0}{\alpha}\right)\right| \leq \frac{C\varepsilon|y - y_0|}{\delta\alpha}. \tag{6.17}$$

Using (6.16), (6.17), the fact that equality holds in (6.17) at $y = y_0$ and by enlarging $C > 0$ slightly, we obtain that, for $\theta > 0$,

$$(y,s) \mapsto u(y,s) + \frac{1}{2\alpha}|x_0 - y|^2 + \frac{1}{2\varepsilon}(t_0 - s)^2 + ds + C\frac{\varepsilon}{\delta\alpha}|y - y_0| + \frac{\theta}{2}(s - s_0)^2 \quad \text{has a strict local minimum at } (y,s) = (y_0, s_0). \tag{6.18}$$

It follows that, for all sufficiently small $\beta > 0$, there exist $(y_\beta, s_\beta) \in \mathbb{R}^d \times [0,T]$ such that $(y_\beta, s_\beta) \to (y_0, s_0)$ as $\beta \to 0$ and

$$(y,s) \mapsto u(y,s) + \frac{1}{2\alpha}|x_0 - y|^2 + \frac{1}{2\varepsilon}(t_0 - s)^2 + ds + C\frac{\varepsilon}{\delta\alpha}\left(\beta + |y - y_0|^2\right)^{\frac{1}{2}} + \frac{\theta}{2}(s - s_0)^2 \quad \text{has a local minimum at } (y,s) = (y_\beta, s_\beta). \tag{6.19}$$

Using equation (1.24) satisfied by u, we obtain

$$-d + \frac{1}{\varepsilon}(t_0 - s_\beta) - \theta(s_\beta - s_0) + \overline{H}\left(\frac{x_0 - y_\beta}{\alpha} - Q_\beta\right) \geq 0,$$

where $Q_\beta := C\frac{\varepsilon}{\delta\alpha}\left(\beta + |y_\beta - y_0|^2\right)^{-\frac{1}{2}}(y_\beta - y_0)$. Since $|Q_\beta| \leq C\varepsilon/(\delta\alpha)$, the Lipschitz continuity of $\overline{H}$ yields

$$-d + \frac{1}{\varepsilon}(t_0 - s_\beta) - \theta(s_\beta - s_0) + \overline{H}\left(\frac{x_0 - y_\beta}{\alpha}\right) \geq -C\frac{\varepsilon}{\delta\alpha},$$

and, after letting $\beta \to 0$, we find

$$-d + \frac{1}{\varepsilon}(t_0 - s_0) + \overline{H}\left(\frac{x_0 - y_0}{\alpha}\right) \geq -C\frac{\varepsilon}{\delta\alpha}. \tag{6.20}$$

We next fix $(y, s) = (y_0, s_0)$ and let (x, t) vary, in order to use the equations for u^ε and v^δ. The intermediate goal is to prove (6.34), below, to complement (6.20).

From (6.11), we see that

$$\begin{aligned}(t,x) \to u^\varepsilon(x,t) - \varepsilon v^\delta\left(\frac{x}{\varepsilon}\,;\zeta\left(\frac{x-y_0}{\alpha}\right)\right) - \frac{1}{2\alpha}|x-y_0|^2 \\ - \frac{1}{2\varepsilon}(t-s_0)^2 - q\left(1+|x|^2\right)^{\frac{1}{2}} \text{ has a maximum at } (x_0,t_0).\end{aligned}$$

Using (6.15) and (6.17), the above inequality implies, as in the previous step, that

$$\begin{aligned}(t,x) \to u^\varepsilon(x,t) - \varepsilon v^\delta\left(\frac{x}{\varepsilon}\,;\frac{x_0-y_0}{\alpha}\right) - \frac{1}{2\alpha}|x-y_0|^2 - \frac{1}{2\varepsilon}(t-s_0)^2 \\ - q\left(1+|x|^2\right)^{\frac{1}{2}} - C\frac{\varepsilon}{\delta\alpha}|x-x_0| \text{ has a maximum at } (x_0,t_0).\end{aligned} \tag{6.21}$$

Let us now fix a small parameter $\theta \in (0,1)$, to be chosen later, and consider a maximum point (x_θ, t_θ) of the perturbed problem

$$\begin{aligned}(t,x) \to u^\varepsilon(x,t) - \varepsilon v^\delta\left(\frac{x}{\varepsilon}\,;\frac{x_0-y_0}{\alpha}\right) - \frac{1}{2\alpha}|x-y_0|^2 - \frac{1}{2\varepsilon}(t-s_0)^2 \\ - q\left(1+|x|^2\right)^{\frac{1}{2}} - \widetilde{C}\frac{\varepsilon}{\delta\alpha}\left(\left(\theta^2+|x-x_0|^2\right)^{\frac{1}{2}} - \theta\right) - \frac{1}{2\varepsilon}(t-t_0)^2,\end{aligned} \tag{6.22}$$

with $\widetilde{C}$ to be chosen below.

For later use, we estimate the distance from (x_0, t_0) to (x_θ, t_θ): by (6.21), we have

$$\begin{aligned}u^\varepsilon(x_\theta,t_\theta) - \varepsilon v^\delta\left(\frac{x_\theta}{\varepsilon}\,;\frac{x_0-y_0}{\alpha}\right) - \frac{1}{2\alpha}|x_\theta-y_0|^2 - \frac{1}{2\varepsilon}(t_\theta-s_0)^2 - q\left(1+|x_\theta|^2\right)^{\frac{1}{2}} - \frac{C\varepsilon}{\delta\alpha}|x_\theta-x_0| \\ \leq u^\varepsilon(x_0,t_0) - \varepsilon v^\delta\left(\frac{x_0}{\varepsilon}\,;\frac{x_0-y_0}{\alpha}\right) - \frac{1}{2\alpha}|x_0-y_0|^2 - \frac{1}{2\varepsilon}(t_0-s_0)^2 - q\left(1+|x_0|^2\right)^{\frac{1}{2}}.\end{aligned}$$

On another hand, optimality of (x_θ, t_θ) in (6.22) implies

$$\begin{aligned}u^\varepsilon(x_0,t_0) - \varepsilon v^\delta\left(\frac{x_0}{\varepsilon}\,;\frac{x_0-y_0}{\alpha}\right) - \frac{1}{2\alpha}|x_0-y_0|^2 - \frac{1}{2\varepsilon}(t_0-s_0)^2 - q\left(1+|x_0|^2\right)^{\frac{1}{2}} \\ \leq u^\varepsilon(x_\theta,t_\theta) - \varepsilon v^\delta\left(\frac{x_\theta}{\varepsilon}\,;\frac{x_0-y_0}{\alpha}\right) - \frac{1}{2\alpha}|x_\theta-y_0|^2 - \frac{1}{2\varepsilon}(t_\theta-s_0)^2 \\ - q\left(1+|x_\theta|^2\right)^{\frac{1}{2}} - \widetilde{C}\frac{\varepsilon}{\delta\alpha}\left(\left(\theta^2+|x_\theta-x_0|^2\right)^{\frac{1}{2}} - \theta\right) - \frac{1}{2\varepsilon}(t_\theta-t_0)^2.\end{aligned}$$

Putting together the above inequalities yields

$$\widetilde{C}\frac{\varepsilon}{\delta\alpha}\left(\left(\theta^2+|x_\theta-x_0|^2\right)^{\frac{1}{2}} - \theta\right) + \frac{1}{\varepsilon}(t_\theta-t_0)^2 \leq C\frac{\varepsilon}{\delta\alpha}|x_\theta-x_0|,$$

so that, if $\widetilde{C}$ is sufficiently large,

$$|x_0 - x_\theta| \le C\theta \qquad \text{and} \qquad |t_\theta - t_0| \le C\left(\frac{\varepsilon^2\theta}{\delta\alpha}\right)^{\frac{1}{2}}. \tag{6.23}$$

In particular, if $t_0 > C\left(\frac{\varepsilon^2\theta}{\delta\alpha}\right)^{\frac{1}{2}}$, then $t_\theta > 0$. We finally fix two other parameters $\sigma, \rho > 0$ (which will be sent to zero shortly) and introduce a last auxiliary function $\Psi : \mathbb{R}^d \times \mathbb{R}^d \times [0,T] \to \mathbb{R}$ defined by

$$\begin{aligned}\Psi(x,z,t) := u^\varepsilon(x,t) - \varepsilon v^\delta\left(\frac{z}{\varepsilon}\,;\frac{x_0-y_0}{\alpha}\right) - \frac{1}{2\alpha}|x-y_0|^2 - \frac{1}{2\varepsilon}(t-s_0)^2 \\ - q\left(1+|x|^2\right)^{\frac{1}{2}} - \frac{1}{2\sigma}|z-x|^2 - \widetilde{C}\frac{\varepsilon}{\delta\alpha}\left(\left(\theta^2+|x-x_0|^2\right)^{\frac{1}{2}} - \theta\right) \\ - \frac{1}{2\varepsilon}(t-t_0)^2 - \frac{\rho}{2}|(x,t)-(x_\theta,t_\theta)|^2.\end{aligned} \tag{6.24}$$

The last term in (6.24) provides some strictness and therefore, by (6.11), there exist points $(x_\sigma, z_\sigma, t_\sigma) \in \mathbb{R}^d \times \mathbb{R}^d \times [0,T]$ such that $(x_\sigma, z_\sigma, t_\sigma) \to (x_\theta, x_\theta, t_\theta)$ as $\sigma \to 0$ and

$$\Psi(x_\sigma, z_\sigma, t_\sigma) = \sup_{\mathbb{R}^d\times\mathbb{R}^d\times[0,T]} \Psi.$$

From the Lipschitz regularity of $v^\delta\left(\cdot\,;\frac{x_0-y_0}{\alpha}\right)$ recalled in (5.2), we also have

$$|x_\sigma - z_\sigma| \le K_p\sigma. \tag{6.25}$$

According to the maximum principle for semicontinuous functions [8], it follows that, for any $\eta > 0$ there exist $X_{\sigma,\eta}, Y_{\sigma,\eta} \in \mathcal{S}^d$ such that

$$\left(\frac{1}{\varepsilon}(t_\sigma - s_0) + \frac{1}{\varepsilon}(t_\sigma - t_0) + \rho(t_\sigma - t_\theta), \frac{x_\sigma - y_0}{\alpha} + \frac{x_\sigma - z_\sigma}{\sigma} + p_\sigma, X_{\sigma,\eta}\right) \in \overline{\mathcal{P}}^{2,+}u^\varepsilon(x_\theta, t_\theta)$$

$$\left(\frac{x_\sigma - z_\sigma}{\sigma}, Y_{\sigma,\eta}\right) \in \overline{\mathcal{P}}^{2,-}v^\delta\left(\frac{z_\sigma}{\varepsilon}; \frac{x_0-y_0}{\alpha}\right),$$

$$\begin{pmatrix} X_\sigma & 0 \\ 0 & \frac{1}{\varepsilon}Y_{\sigma,\eta}\end{pmatrix} \le M_\sigma + \eta M_\sigma^2, \tag{6.26}$$

where

$$p_\sigma = \frac{qx_\sigma}{(1+|x_\sigma|^2)^{\frac{1}{2}}} + \frac{\widetilde{C}\varepsilon(x_\sigma - x_0)}{\delta\alpha\left(\theta^2 + |x_\sigma - x_0|^2\right)^{\frac{1}{2}}} + \rho(x_\sigma - x_\theta),$$

$$M_\sigma = \begin{pmatrix} \frac{1}{\sigma}I_d & -\frac{1}{\sigma}I_d \\ -\frac{1}{\sigma}I_d & \frac{1}{\sigma}I_d\end{pmatrix} + \begin{pmatrix} P_\sigma & 0 \\ 0 & 0\end{pmatrix} \tag{6.27}$$

and

$$P_\sigma = \frac{1}{\alpha} I_d + q \left(\frac{I_d}{(1+|x_\sigma|^2)^{\frac{1}{2}}} - \frac{x_\sigma \otimes x_\sigma}{(1+|x_\sigma|^2)^{\frac{3}{2}}} \right) + \widetilde{C} \frac{\varepsilon}{\delta\alpha} \left(\frac{I_d}{(\theta^2+|x_\sigma - x_0|^2)^{\frac{1}{2}}} - \frac{(x_\sigma - x_0)\otimes(x_\sigma - x_0)}{(\theta^2+|x_\sigma - x_0|^2)^{\frac{3}{2}}} \right) + \rho I_d. \tag{6.28}$$

Note that

$$|p_\sigma| \le q + \frac{\widetilde{C}\varepsilon}{\delta\alpha} + \rho|x_\sigma - x_\theta| \text{ and } \|P_\sigma\| \le \frac{1}{\alpha} + 2q + C\frac{\varepsilon}{\delta\alpha\theta} + \rho. \tag{6.29}$$

From equation (1.23) satisfied by u^ε and since $t_\sigma > 0$ for σ small enough, we have

$$\frac{1}{\varepsilon}(t_\sigma - s_0) + \frac{1}{\varepsilon}(t_\sigma - t_0) + \rho(t_\sigma - t_\theta) - \varepsilon \mathrm{tr}\left(A\left(\frac{x_\sigma}{\varepsilon}\right) X_{\sigma,\eta} \right) + H\left(\frac{x_\sigma - y_0}{\alpha} + \frac{x_\sigma - z_\sigma}{\sigma} + p_\sigma, \frac{x_\sigma}{\varepsilon} \right) \le 0, \tag{6.30}$$

while, in view of equation (1.19) for v^δ, we have

$$\delta v^\delta \left(\frac{z_\sigma}{\varepsilon}; \frac{x_0 - y_0}{\alpha} \right) - \mathrm{tr}\left(A\left(\frac{z_\sigma}{\varepsilon}\right) Y_{\sigma,\eta} \right) + H\left(\frac{x_\sigma - z_\sigma}{\sigma} + \frac{x_0 - y_0}{\alpha}, \frac{z_\sigma}{\varepsilon} \right) \ge 0. \tag{6.31}$$

We multiply inequality (6.26) by the positive matrix

$$\begin{pmatrix} \Sigma\left(\frac{x_\sigma}{\varepsilon}\right) \\ \Sigma\left(\frac{z_\sigma}{\varepsilon}\right) \end{pmatrix} \begin{pmatrix} \Sigma\left(\frac{x_\sigma}{\varepsilon}\right) \\ \Sigma\left(\frac{z_\sigma}{\varepsilon}\right) \end{pmatrix}^T$$

and take the trace of the resulting expression to obtain, by (6.29),

$$\begin{aligned} &\mathrm{tr}\left(A\left(\frac{x_\sigma}{\varepsilon}\right) X_{\sigma,\eta} \right) - \frac{1}{\varepsilon}\mathrm{tr}\left(A\left(\frac{z_\sigma}{\varepsilon}\right) Y_{\sigma,\eta} \right) \\ &\le \frac{1}{\sigma}\left| \Sigma\left(\frac{x_\sigma}{\varepsilon}\right) - \Sigma\left(\frac{z_\sigma}{\varepsilon}\right) \right|^2 + \left(\frac{1}{\alpha} + 2q + C\frac{\varepsilon}{\delta\alpha\theta} + \rho \right) \left| \Sigma\left(\frac{x_\sigma}{\varepsilon}\right) \right|^2 \\ &\quad + \eta \mathrm{tr}\left(M^2_{\sigma,\eta} \begin{pmatrix} \Sigma\left(\frac{x_\sigma}{\varepsilon}\right) \\ \Sigma\left(\frac{z_\sigma}{\varepsilon}\right) \end{pmatrix} \begin{pmatrix} \Sigma\left(\frac{x_\sigma}{\varepsilon}\right) \\ \Sigma\left(\frac{z_\sigma}{\varepsilon}\right) \end{pmatrix}^t \right) \end{aligned}$$

Using (1.4) and (1.5), we obtain

$$\mathrm{tr}\left(A\left(\frac{x_\sigma}{\varepsilon}\right) X_{\sigma,\eta} \right) - \frac{1}{\varepsilon}\mathrm{tr}\left(A\left(\frac{z_\sigma}{\varepsilon}\right) Y_{\sigma,\eta} \right) \le \frac{C|x_\sigma - z_\sigma|^2}{\sigma\varepsilon^2} + C\left(\frac{1}{\alpha} + q + \frac{\varepsilon}{\delta\alpha\theta} + \rho \right) + \eta C_\sigma \tag{6.32}$$

where the constant C_σ actually depends also on all the other parameters of the problem, but is independent of η. On another hand we have by (6.25) and (6.29):

$$\left|H\left(\frac{x_\sigma - y_0}{\alpha} + \frac{x_\sigma - z_\sigma}{\sigma} + p_\sigma, \frac{x_\sigma}{\varepsilon}\right) - H\left(\frac{x_\sigma - z_\sigma}{\sigma} + \frac{x_0 - y_0}{\alpha}, \frac{z_\sigma}{\varepsilon}\right)\right| \tag{6.33}$$
$$\leq C\left(\frac{|x_\sigma - x_0|}{\alpha} + \frac{|x_\sigma - z_\sigma|}{\varepsilon} + |p_\sigma|\right) \leq C\left(\frac{|x_\sigma - x_0|}{\alpha} + \frac{\sigma}{\varepsilon} + q + \frac{\varepsilon}{\delta\alpha} + \rho|x_\sigma - x_\theta|\right).$$

Subtracting (6.31) to (6.30) and using (6.32) and (6.33) we get

$$\frac{1}{\varepsilon}(t_\sigma - s_0) + \frac{1}{\varepsilon}(t_\sigma - t_0) + \rho(t_\sigma - t_\theta) - \delta v^\delta\left(\frac{z_\sigma}{\varepsilon}; \frac{x_0 - y_0}{\alpha}\right)$$
$$\leq C\Bigg(\frac{|x_\sigma - z_\sigma|^2}{\sigma\varepsilon} + \varepsilon\left(\frac{1}{\alpha} + q + \frac{\varepsilon}{\delta\alpha\theta} + \rho\right) + \frac{|x_\sigma - x_0|}{\alpha}$$
$$+ \frac{\sigma}{\varepsilon} + q + \frac{\varepsilon}{\delta\alpha} + \rho|x_\sigma - x_\theta|\Bigg) + \varepsilon\eta C(\sigma).$$

We let η, σ and then ρ tend to 0 to obtain, thanks to (6.25),

$$\frac{1}{\varepsilon}(t_\theta - s_0) + \frac{1}{\varepsilon}(t_\theta - t_0) - \delta v^\delta\left(\frac{x_\theta}{\varepsilon}; \frac{x_0 - y_0}{\alpha}\right)$$
$$\leq C\Bigg(\varepsilon\left(\frac{1}{\alpha} + q + \frac{\varepsilon}{\delta\alpha\theta}\right) + \frac{|x_\theta - x_0|}{\alpha} + q + \frac{\varepsilon}{\delta\alpha}\Bigg).$$

Recalling (6.23) and using $0 < \delta, \varepsilon \leq 1$,

$$\frac{1}{\varepsilon}(t_0 - s_0) - \delta v^\delta\left(\frac{x_\theta}{\varepsilon}; \frac{x_0 - y_0}{\alpha}\right) \leq C\left(q + \frac{\varepsilon^2}{\delta\alpha\theta} + \frac{\theta}{\alpha} + \frac{\varepsilon}{\delta\alpha} + \left(\frac{\theta}{\delta\alpha}\right)^{\frac{1}{2}}\right). \tag{6.34}$$

We now put (6.20) and (6.34) together:

$$-\overline{H}\left(\frac{x_0 - y_0}{\alpha}\right) - \delta v^\delta\left(\frac{x_\theta}{\varepsilon}; \frac{x_0 - y_0}{\alpha}\right) \leq -d + C\left(q + \frac{\varepsilon^2}{\delta\alpha\theta} + \frac{\theta}{\alpha} + \frac{\varepsilon}{\delta\alpha} + \left(\frac{\theta}{\delta\alpha}\right)^{\frac{1}{2}}\right).$$

So if we choose

$$d = \lambda + 2C\left(q + \frac{\varepsilon^2}{\delta\alpha\theta} + \frac{\theta}{\alpha} + \frac{\varepsilon}{\delta\alpha} + \left(\frac{\theta}{\delta\alpha}\right)^{\frac{1}{2}}\right), \tag{6.35}$$

we have a contradiction with (6.6) provided $\frac{|x_\theta|}{\varepsilon} \leq R$. Thanks to (6.12) and (6.23), this latter inequality holds provided

$$R \geq C\left(\frac{1}{q\varepsilon} + \frac{\theta}{\varepsilon}\right). \tag{6.36}$$

Therefore, if (6.35) and (6.36) hold, we have either $s_0 = 0$, or $t_0 \le C\left(\frac{\varepsilon^2\theta}{\delta\alpha}\right)^{\frac{1}{2}}$. Let us first assume that $s_0 = 0$. In this case, by the Lipschitz bound (6.1), we have

$$\begin{aligned} M(\omega) &\le u^\varepsilon(x_0,t_0) - u(y_0,0) - \varepsilon v^\delta\left(\frac{x_0}{\varepsilon}\,;\zeta\left(\frac{x_0-y_0}{\alpha}\right)\right) - \frac{1}{2\alpha}|x_0-y_0|^2 - \frac{1}{2\varepsilon}t_0^2 \\ &\le u^\varepsilon(y_0,0) + L|(x_0,t_0)-(y_0,0)| - u(y_0,0) + C\frac{\varepsilon}{\delta} - \frac{1}{2\alpha}|x_0-y_0|^2 - \frac{1}{2\varepsilon}t_0^2 \\ &\le C\left(\alpha+\varepsilon+\frac{\varepsilon}{\delta}\right), \end{aligned}$$

using that $u^\varepsilon(y_0,0) = u_0(y_0) = u(y_0,0)$. On the other hand, if $t_0 \le C\left(\frac{\varepsilon^2\theta}{\delta\alpha}\right)^{\frac{1}{2}}$, then (6.13) yields

$$\begin{aligned} M(\omega) &\le u^\varepsilon(x_0,t_0) - u(y_0,s_0) - \varepsilon v^\delta\left(\frac{x_0}{\varepsilon}\,;\zeta\left(\frac{x_0-y_0}{\alpha}\right)\right) - \frac{1}{2\alpha}|x_0-y_0|^2 - \frac{1}{2\varepsilon}(t_0-s_0)^2 \\ &\le u^\varepsilon(x_0,0) - u(x_0,0) + L|(x_0,t_0)-(y_0,s_0)| + C\left(\frac{\varepsilon^2\theta}{\delta\alpha}\right)^{\frac{1}{2}} + C\frac{\varepsilon}{\delta} \\ &\qquad - \frac{1}{2\alpha}|x_0-y_0|^2 - \frac{1}{2\varepsilon}(t_0-s_0)^2 \\ &\le C\left(\alpha+\varepsilon+\frac{\varepsilon}{\delta}+\left(\frac{\varepsilon^2\theta}{\delta\alpha}\right)^{\frac{1}{2}}\right). \end{aligned}$$

As a consequence, we have, for any $(x,t) \in B_T \times [0,T]$,

$$\begin{aligned} \Phi(x,x,t,t) := u^\varepsilon(x,t) - u(x,t) - \varepsilon v^\delta\left(\frac{x}{\varepsilon}\,;0\right) - dt - q\left(1+|x|^2\right)^{\frac{1}{2}} + q \\ \le M(\omega) \le C\left(\alpha+\varepsilon+\frac{\varepsilon}{\delta}+\left(\frac{\varepsilon^2\theta}{\delta\alpha}\right)^{\frac{1}{2}}\right), \end{aligned}$$

so that, recalling the choice of d in (6.35) and using $0 \le \varepsilon \le 1$:

$$u^\varepsilon(x,t) - u(x,t) \le C\left(\lambda + q + \frac{\varepsilon^2}{\delta\alpha\theta} + \frac{\theta}{\alpha} + \frac{\varepsilon}{\delta\alpha} + \left(\frac{\theta}{\delta\alpha}\right)^{\frac{1}{2}} + \alpha\right).$$

If we choose $q = \lambda$, $\theta = \varepsilon\delta^{-1}$ and $\alpha = \varepsilon^{\frac{1}{3}}\delta^{-\frac{2}{3}}$, we get

$$u^\varepsilon(x,t) - u(x,t) \le C\left(\lambda + \frac{\varepsilon^{\frac{2}{3}}}{\delta^{\frac{1}{3}}} + \frac{\varepsilon^{\frac{1}{3}}}{\delta^{\frac{2}{3}}}\right).$$

Note also that, by definition of R, (6.36) holds. To summarize, we have proved that, if (6.7) holds, then

$$\sup_{(x,t)\in B_T\times[0,T]} \left(u^\varepsilon(x,t) - u(x,t)\right) \le C\left(\lambda + \frac{\varepsilon^{\frac{2}{3}}}{\delta^{\frac{1}{3}}} + \frac{\varepsilon^{\frac{1}{3}}}{\delta^{\frac{2}{3}}}\right).$$

This is (6.4). □

Proof of Theorem 3. From Lemma 6.1, if $\varepsilon^{\frac{2}{3}}\delta^{-\frac{1}{3}} + \varepsilon^{\frac{1}{3}}\delta^{-\frac{2}{3}} \leq c\lambda$, then

$$\left\{\omega \in \Omega \ : \ \sup_{x\in B_T,\ t\in[0,T]} |u^\varepsilon(x,t) - u(x,t)| \geq \lambda\right\} \subseteq \left\{\omega \in \Omega \ : \ \sup_{(y,p)\in B_R\times B_L} \left|\delta v^\delta(y,p) + \overline{H}(p)\right| \geq c\lambda\right\},$$

where $R := C\left(\frac{1}{\lambda\varepsilon} + \frac{1}{\delta}\right)$. Theorem 2 implies that, if inequality

$$\lambda \geq C\delta^{\frac{1}{7+6q}}\left(1 + |\log(\delta)|\right) \tag{6.37}$$

holds, then

$$\mathbb{P}\left[\sup_{(y,p)\in B_R\times B_L} \left|\delta v^\delta(y,p) + \overline{H}(p)\right| \geq c\lambda\right] \leq C\lambda^{-2d}R^d\left(\lambda^{-2d} + \delta^{-2d}\right)\exp\left(-\frac{\lambda^{4(1+q)}}{C\delta}\right).$$

Note that inequalities $\varepsilon^{\frac{2}{3}}\delta^{-\frac{1}{3}} + \varepsilon^{\frac{1}{3}}\delta^{-\frac{2}{3}} \leq c\lambda$ and (6.37) are ensured by the conditions

$$\delta := C\varepsilon^{\frac{1}{2}}\lambda^{-\frac{3}{2}} \qquad \text{and} \qquad \lambda \geq C\varepsilon^{\frac{1}{17+12q}}\left(1 + \log(\varepsilon)\right)^{\frac{1}{2}}.$$

We deduce that

$$\mathbb{P}\left[\sup_{x\in B_T,\ t\in[0,T]} |u^\varepsilon(x,t) - u(x,t)| \geq \lambda\right] \leq C\lambda^{-3d}\varepsilon^{-d}\left(\lambda^{-2d} + \varepsilon^{-d}\right)\exp\left(-\frac{\lambda^{\frac{11}{2}+4q}}{C\varepsilon^{\frac{1}{2}}}\right).$$

As $\varepsilon \leq \lambda$, this implies that (1.25) holds. □

Acknowledgements. This work was partially supported by the ANR (Agence Nationale de la Recherche) through HJnet project ANR-12-BS01-0008-01. We thank Ilya Bogdanov for indicating a nice proof of Lemma 5.2 to us on Math Overflow.

References

[1] K. S. Alexander. A note on some rates of convergence in first-passage percolation. *Ann. Appl. Probab.*, 3(1):81–90, 1993.

[2] S. N. Armstrong, P. Cardaliaguet, and P. E. Souganidis. Error estimates and convergence rates for the stochastic homogenization of Hamilton-Jacobi equations, 2012. arXiv:1206.2601v3.

[3] S. N. Armstrong and P. E. Souganidis. Stochastic homogenization of Hamilton-Jacobi and degenerate Bellman equations in unbounded environments. *J. Math. Pures Appl.*, 97:460–504, 2012.

[4] S. N. Armstrong and P. E. Souganidis. Stochastic homogenization of level-set convex Hamilton-Jacobi equations. *Int. Math. Res. Not.*, in press. Arxiv:1203.6303 [math.AP].

[5] S. N. Armstrong and H.V. Tran. In preparation. 2013. Forthcoming.

[6] S. N. Armstrong and H.V. Tran. Viscous hamilton-jacobi equations: comparison, global existence and state constraints. 2013. Forthcoming.

[7] I. Capuzzo-Dolcetta and H. Ishii. On the rate of convergence in homogenization of Hamilton-Jacobi equations. *Indiana Univ. Math. J.*, 50(3):1113–1129, 2001.

[8] M. G. Crandall, H. Ishii, and P.-L. Lions. User's guide to viscosity solutions of second order partial differential equations. *Bull. Amer. Math. Soc. (N.S.)*, 27(1):1–67, 1992.

[9] B. Grünbaum. *Convex polytopes*, volume 221 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 2003. Prepared and with a preface by Volker Kaibel, Victor Klee and Günter M. Ziegler.

[10] H. Kesten. On the speed of convergence in first-passage percolation. *Ann. Appl. Probab.*, 3(2):296–338, 1993.

[11] E. Kosygina, F. Rezakhanlou, and S. R. S. Varadhan. Stochastic homogenization of Hamilton-Jacobi-Bellman equations. *Comm. Pure Appl. Math.*, 59(10):1489–1521, 2006.
[12] P.-L. Lions. *Generalized solutions of Hamilton-Jacobi equations*, volume 69 of *Research Notes in Mathematics*. Pitman (Advanced Publishing Program), Boston, Mass., 1982.
[13] P.-L. Lions and P. E. Souganidis. Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media. *Comm. Partial Differential Equations*, 30(1-3):335–375, 2005.
[14] P.-L. Lions and P. E. Souganidis. Stochastic homogenization of Hamilton-Jacobi and "viscous"-Hamilton-Jacobi equations with convex nonlinearities—revisited. *Commun. Math. Sci.*, 8(2):627–637, 2010.
[15] Ivan Matic and James Nolen. A sublinear variance bound for solutions of a random hamilton–jacobi equation. *Journal of Statistical Physics*, 149(2):342–361, 2012.
[16] J. R. Munkres. *Topology: a first course*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1975.
[17] F. Rezakhanlou. Central limit theorem for stochastic Hamilton-Jacobi equations. *Comm. Math. Phys.*, 211(2):413–438, 2000.
[18] F. Rezakhanlou and J. E. Tarver. Homogenization for stochastic Hamilton-Jacobi equations. *Arch. Ration. Mech. Anal.*, 151(4):277–309, 2000.
[19] P. E. Souganidis. Stochastic homogenization of Hamilton-Jacobi equations and some applications. *Asymptot. Anal.*, 20(1):1–11, 1999.
[20] A.-S. Sznitman. Distance fluctuations and Lyapounov exponents. *Ann. Probab.*, 24(3):1507–1530, 1996.
[21] M. V. Wüthrich. Fluctuation results for Brownian motion in a Poissonian potential. *Ann. Inst. H. Poincaré Probab. Statist.*, 34(3):279–308, 1998.
[22] Y. Zhang. On the concentration and the convergence rate with a moment condition in first passage percolation. *Stochastic Process. Appl.*, 120(7):1317–1341, 2010.

Ceremade (UMR CNRS 7534), Université Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, 75775 Paris CEDEX 16, France
E-mail address: armstrong@ceremade.dauphine.fr

Ceremade (UMR CNRS 7534), Université Paris-Dauphine, Place du Maréchal De Lattre De Tassigny, 75775 Paris CEDEX 16, France
E-mail address: cardaliaguet@ceremade.dauphine.fr