A sublinear variance bound for solutions of a random Hamilton-Jacobi equation

Ivan Matic* and James Nolen†

September 28, 2012

Abstract

We estimate the variance of the value function for a random optimal control problem. The value function is the solution w^ϵ of a random Hamilton-Jacobi equation with random Hamiltonian $H(p, x, \omega) = K(p) - V(x/\epsilon, \omega)$ in dimension $d \geq 2$. It is known that homogenization occurs as $\epsilon \to 0$, but little is known about the statistical fluctuations of w^ϵ. Our main result shows that the variance of the solution w^ϵ is bounded by $O(\epsilon/|\log \epsilon|)$. The proof relies on a modified Poincaré inequality of Talagrand.

1 Introduction

In this paper we study the random optimal control problem

$$u(t, x, \omega) = \sup_{\gamma \in \mathcal{A}_{t,x}} g(\gamma(t)) - \mathcal{L}(\gamma, \omega), \quad x \in \mathbb{R}^d, \ \ t > 0 \tag{1.1}$$

in dimension $d \geq 2$, where the supremum is taken over the set of admissible paths

$$\mathcal{A}_{t,x} = \{\gamma \in W^{1,\infty}([0, t]; \mathbb{R}^d) \mid \gamma(0) = x\}.$$

The upper-semicontinuous payoff function $g : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ is given. The cost functional $\mathcal{L}$ has the form

$$\mathcal{L}(\gamma, \omega) = \int_0^t L(\gamma'(s), \gamma(s), \omega)\, ds = \int_0^t K(\gamma'(s)) + V(\gamma(s), \omega)\, ds,$$

where $K(p) : \mathbb{R}^d \to [0, \infty)$ is convex and grows super-linearly in $|p|$. The function $V(x, \omega)$ is a scalar random field that is statistically stationary and ergodic with respect to certain translations in x. The parameter $\omega \in \Omega$ denotes a sample from a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Thus, the value function $u(t, x, \omega)$ is random. Our main result shows that the variance of $u(x, t, \omega)$ grows only sublinearly in t as $t \to \infty$.

Under certain conditions on g and L, $u(t, x, \omega)$ is uniformly continuous and is a viscosity solution [5] of the random Hamilton-Jacobi equation

$$\begin{cases} u_t = H(Du, x, \omega), \quad x \in \mathbb{R}^d, \quad t > 0 \\ u(0, x) = g(x), \quad x \in \mathbb{R}^d, \end{cases} \tag{1.2}$$

*Dept. of Mathematics, Duke University, Box 90320, Durham, NC 27708, USA. (matic@math.duke.edu).
†Dept. of Mathematics, Duke University, Box 90320, Durham, NC 27708, USA. (nolen@math.duke.edu).

where $H(p,x) = K^*(p) - V(x)$, K^* being the Legendre transform of K. For simplicity, consider the case where $g(x) = \eta \cdot x$ is a linear function. Then for each $\epsilon > 0$, the function $w^\epsilon(t,x,\omega) = \epsilon u(t/\epsilon, x/\epsilon, \omega)$ solves the initial value problem

$$\begin{cases} w^\epsilon_t = H(Dw^\epsilon, \frac{x}{\epsilon}, \omega), \quad x \in \mathbb{R}^d, \quad t > 0 \\ w^\epsilon(0,x) = \eta \cdot x = g(x). \end{cases} \tag{1.3}$$

For certain Hamiltonians $H(p,x,\omega)$ which are convex in p, statistically stationary and ergodic with respect to translation in x, it is known [16, 19] that as $\epsilon \to 0$, homogenization occurs (see also [3, 14] for alternative proofs and [10, 11, 13, 18] for related results). This means that the functions $w^\epsilon(t,x,\omega)$ converge locally uniformly in $[0,\infty) \times \mathbb{R}^d$, as $\epsilon \to 0$ to the deterministic function $\bar{w}(t,x)$ which solves

$$\begin{cases} \bar{w}_t = \bar{H}(D\bar{w}), \quad x \in \mathbb{R}^d, \quad t > 0 \\ \bar{w}(0,x) = g(x). \end{cases} \tag{1.4}$$

The function $\bar{H}(p) : \mathbb{R}^d \to \mathbb{R}$ is called the effective Hamiltonian. We may think of this convergence as kind of law of large numbers for w^ϵ, although the limit $\bar{w}$ and the effective Hamiltonian $\bar{H}$ are not determined by a simple averaging. Beyond this convergence result, relatively little is known about the properties of $\bar{H}$, about the rate of convergence $w^\epsilon \to \bar{w}$, or about the statistical behavior of $w^\epsilon - \mathbb{E}[w^\epsilon]$, where $\mathbb{E}[\cdot]$ denotes expectation with respect to the probability measure $\mathbb{P}$. Our work pertains to this last issue: in terms of $w^\epsilon(t,x,\omega)$, our estimate on the variance of u implies that $\mathrm{var}(w^\epsilon(t,x,\omega)) \leq C\epsilon/|\log \epsilon|$, as $\epsilon \to 0$.

Before stating our main result, let us make some definitions and assumptions more precise. We will suppose the random field $V(x,\omega)$ has the following special structure. Let $a < b$ be two real numbers. Let $\Omega = \{a,b\}^{\mathbb{Z}^d}$ be the set of all functions $\omega : \mathbb{Z}^d \to \{a,b\}$. Let the probability measure $\mathbb{P}$ be the shift-invariant product measure on Ω determined by $\mathbb{P}(\omega_k = a) = \alpha$ and $\mathbb{P}(\omega_k = b) = \beta$, for all $k \in \mathbb{Z}^d$, where $\alpha \in (0,1)$ and $\beta = 1 - \alpha$. Thus the random variables $\{\omega_k\}_{k \in \mathbb{Z}^d}$ are independent and identically distributed. Now for $k \in \mathbb{Z}^d$, let $Q_k = k + [0,1)^d$ denote the unit cube with corner at the point k. Given $\omega \in \Omega$, define $V(x,\omega) : \mathbb{R}^d \times \Omega \to \{a,b\}$ by

$$V(x,\omega) = \sum_{k \in \mathbb{Z}^d} \omega_k \mathbb{I}_{Q_k}(x), \tag{1.5}$$

with $\mathbb{I}_{Q_k}$ is the indicator function for the set Q_k. Thus, $x \mapsto V$ is piecewise constant, taking values a or b on the unit cubes. By construction, the law of $V(x,\omega)$ is the same as that of $V(x+k,\omega)$ for any $k \in \mathbb{Z}^d$. This precise construction of the field $V(x,\omega)$ is not essential for our result to hold. In particular, the function could be mollified so that it is uniformly continuous, or $V(x,\omega)$ could depend on the values of ω_k for k in a bounded neighborhood of x. Nevertheless, the choice of $\mathbb{P}$ as the product measure on $\Omega = \{a,b\}^{\mathbb{Z}^d}$ is motivated by the main analytical tool presented below in Theorem 1.2.

We suppose that $K : \mathbb{R}^d \to [0,\infty)$ is convex, $K(0) = 0$, and that

$$\lim_{|z| \to \infty} \frac{K(z)}{|z|} = +\infty.$$

For the case of dimension $d = 2$ we will make use of an extra non-degeneracy condition: for some $\nu > 1$,

$$K(z) \geq |z|^\nu \quad \forall\, z \in B_{1/2}(0). \tag{1.6}$$

Given V and K, let $L(p,x,\omega) = K(x) + V(x,\omega)$ and let u be defined by (1.1). The following estimate of the variance of u for large t is our main result:

Theorem 1.1 *Let $d \geq 2$. Let $x \in \mathbb{R}^d$ and suppose that $g : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$ is upper-semicontinuous and satisfies*

$$g(y) < g(x) + C_1(1 + |y - x|), \quad \forall y \in \mathbb{R}^d. \tag{1.7}$$

There is a constant $C > 0$, depending only on C_1, K, α, β, and $|b - a|$, such that

$$\text{var}(u(t, x, \omega)) \leq C \frac{t}{\log t}, \quad \forall\, t \geq 2. \tag{1.8}$$

The main tool that we use to control the variance of $u(t, x, \omega)$, is the following theorem, which is a slight variation of an inequality of Talagrand (see [20], Theorem 1.5). This result holds for product spaces of the form $\Omega_J = \{a, b\}^J$, J being a finite set, and $\mathbb{P}$ being the product measure on Ω_J with marginals $\mathbb{P}(\omega_j = a) = \alpha \in (0, 1)$ and $\mathbb{P}(\omega_j = b) = \beta = 1 - \alpha$, for all $j \in J$. Let us define $\phi_j \omega$ to be the element of $\{a, b\}^J$ which is identical to ω except that the j-th component of $\phi_j \omega$ is different from ω_j. That is, $\phi_j \omega = \omega'$, where $\omega'_k = \omega_k$ for $k \neq j$, and $\omega'_j \neq \omega_j$. For each random variable $f : \Omega_J \to \mathbb{R}$ define $\sigma_j f(\omega) = f(\phi_j \omega)$ and

$$\rho_j f(\omega) = \frac{\sigma_j f(\omega) - f(\omega)}{2}.$$

Theorem 1.2 *There is a constant $C > 0$, independent of $|J|$, such that*

$$\text{var}(f) \leq C \sum_{j \in J} \frac{\|\rho_j f\|_2^2}{1 + \log \frac{\|\rho_j f\|_2}{\|\rho_j f\|_1}} \tag{1.9}$$

holds for all $f \in L^2(\Omega_J)$.

The idea of using this inequality to estimate the variance of $f(\omega) = u(t, x, \omega)$ comes from the work of Benjamini, Kalai, and Schramm [6] who used this inequality to estimate the distance variance in first passage percolation, a problem which has some features similar to the control problem (1.1). Specifically, they consider the length of minimal paths between two points in the integer lattice $\mathbb{Z}^d$ under a random metric. Each edge e in the nearest-neighbor graph is assigned an independent random weight $\omega_e \in \{a, b\}$, and the length of a path between two points $x, y \in \mathbb{Z}^d$ is defined as the sum of the edge weights along a path connecting x and y. They proved that $\text{var}(d_\omega(0, v)) \leq C|v| / \log |v|$, where $d_\omega(0, v)$ is the length of the shortest path connecting 0 and v. See [7] for some extensions of that result. The main difficulty in applying the ideas of [6] to the present setting comes from the continuum nature of the problem and from the different structure of the cost functional $\mathcal{L}(\gamma, \omega)$, which necessitates more control on the optimizing paths.

As we have mentioned, for $d \geq 2$ there are relatively few results that apply directly to the random fluctuation of $u(t, x, \omega)$ (as $t \to \infty$) or $w^\epsilon(t, x, \omega)$ (as $\epsilon \to 0$). In [15], Rezakhanlou derived conditions under which a central limit theorem holds for $w^\epsilon(t, x)$ where w^ϵ is the solution of the Hamilton-Jacobi equation (1.3), i.e. whether $\epsilon^{-1/2}(w^\epsilon - \bar{w})$ converges in law to some nontrivial stochastic process as $\epsilon \to 0$. In the case $d = 1$ those conditions can be verified for Hamiltonians having the form $H(p, x, \omega) = K(p) - V(x, \omega)$, and the limit distribution can be computed (see Corollary 2.6 in [15]). For $d \geq 2$, however, it is difficult to verify those conditions. Indeed, our result shows that we may have $\text{var}(w^\epsilon) = o(\epsilon)$, which is less than what a CLT as in [15] would suggest. As this paper was being written, we learned of another work by Armstrong, Cardaliaguet, and Souganidis [2], who study the rate of convergence $w^\epsilon \to \bar{w}$. Our Theorem 1.1 pertains to the variance of w^ϵ, i.e the statistical error $w^\epsilon - \mathbb{E}[w^\epsilon]$, but does not give an estimate of the bias $\mathbb{E}[w^\epsilon] - \bar{w}$.

By analogy with conjectures for first passage percolation and for various other models (see [6, 12, 9, 4, 1, 17] and the many references therein) one expects that $\text{var}(u)$ scales algebraically, like

$t^{2\chi}$ for some exponent $\chi < 1/2$. Hence the bound (1.8) is not expected to be optimal. For example, suppose that $K(p) = \frac{1}{2}|p|^2$ so that $H(p,x) = \frac{1}{2}|p|^2 - V(x)$, and consider (1.2) with the addition of a second-order term:

$$u_t = \frac{1}{2}\Delta u + \frac{1}{2}|\nabla u|^2 - V(x,\omega). \tag{1.10}$$

Suppose $g \equiv 0$. The function $z = e^u$ solves the linear equation $z_t = \frac{1}{2}\Delta z - V(x,\omega)z$ with $z(0,x) \equiv 1$, and it admits the representation

$$z(t,x,\omega) = \mathbb{E}^B[e^{-\int_0^t V(x+B_s,\omega)\,ds}],$$

where $\mathbb{E}^B$ denotes expectation with respect to a standard Brownian motion B_s, independent of the random potential V. This formally shows the connection between $u = \log z$ and directed polymer models, in which z is a partition function. For certain discrete directed polymer models, Seppäläinen [17] has shown that in 1+1 dimensions the variance of $\log z$ is $O(t^{2/3})$ (i.e. $\chi = 2/3$). In the continuum setting, if we formally replace the potential V by a space-time white noise $\dot{W}$ which is independent of B_s, then for $d = 1$ equation (1.10) becomes the KPZ equation (e.g. see [8]). In this case, it is also know that $\operatorname{var} u = \operatorname{var}(\log z)$ is $O(t^{2/3})$ [4]. In light of these results, extension of the estimate (1.8) to solutions of the second order equation (1.10) and an improvement of (1.8) to a sublinear bound of the form $O(t^{2\chi})$ are interesting directions for future work.

The paper is organized as follows. In Section 2 we derive some properties of the paths γ which nearly optimize (1.1). Section 3 contains the main argument for the proof of Theorem 1.1. Section 4 and Section 5 contain proofs of some technical estimates needed in Section 3.

Acknowledgement The work of JN was partially funded by grant DMS-1007572 from the US National Science Foundation.

2 Properties of optimizing paths

Without loss of generality, let us suppose $x = 0$ and simply write $u(t,\omega) = u(t,0,\omega)$ and

$$\mathcal{A}_t = \mathcal{A}_{t,0} = \{\gamma \in W^{1,\infty}([0,t];\mathbb{R}^d) \mid \gamma(0) = 0\}.$$

So, we are studying the quantity

$$u(t,\omega) = \sup_{\gamma \in \mathcal{A}_t} g(\gamma(t)) - \mathcal{L}(\gamma,\omega). \tag{2.11}$$

for some function $g : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$, with $g(0) \in \mathbb{R}$. For $\delta > 0$ and $\omega \in \Omega$, let $M_\delta(t,\omega)$ be the set of all paths $\gamma \in \mathcal{A}_t$ such that

$$g(\gamma(t)) - \mathcal{L}(\gamma,\omega) \geq u(t,\omega) - \delta.$$

For each $\delta > 0$, this set is non-empty, and we refer to these paths as δ-approximate optimizers. Observe that $M_{\delta_2}(t,\omega) \subset M_{\delta_1}(t,\omega)$ if $0 < \delta_2 < \delta_1$. If an optimal path γ exists, meaning that $u(t,\omega) = g(\gamma(t)) - \mathcal{L}(\gamma,\omega)$, then it is certainly an approximate optimizer for any $\delta > 0$. In this section we derive some useful properties of approximate optimizers which will be used in the proof of Theorem 1.1. We will often use $M_\delta(\omega)$ to denote the set $M_\delta(t,\omega)$, the dependence of t being understood.

Deterministic Bounds

First, we have a few estimates which do not involve the random structure of the control problem.

Lemma 2.1 *Let $\gamma \in M_\delta(t,\omega)$. Then for any $r_1, r_2 \in [0,t]$,*

$$\int_{r_1}^{r_2} L(\gamma'(s),\gamma(s),\omega)\,ds \le (r_2-r_1)b + (r_2-r_1)K\left(\frac{\gamma(r_2)-\gamma(r_1)}{r_2-r_1}\right) + \delta, \tag{2.12}$$

and

$$\int_{r_1}^{r_2} L(\gamma'(s),\gamma(s),\omega)\,ds \ge (r_2-r_1)a + (r_2-r_1)K\left(\frac{\gamma(r_2)-\gamma(r_1)}{r_2-r_1}\right). \tag{2.13}$$

Proof of Lemma 2.1: Given $\gamma \in M_\delta(\omega)$, define a new path $\hat{\gamma} \in \mathcal{A}_t$ according to

$$\hat{\gamma}(s) = \gamma(r_1) + (s-r_1)\frac{\gamma(r_2)-\gamma(r_1)}{r_2-r_1} \quad s \in [r_1,r_2]$$

and $\hat{\gamma}(s) = \gamma(s)$ for $s \notin [r_1,r_2]$. Thus we have replaced a section of γ with a straight-line path connecting the same points. Since $\gamma \in \mathcal{M}_\delta(\omega)$, we must have $\mathcal{L}(\gamma,\omega) \le \mathcal{L}(\hat{\gamma},\omega) + \delta$. In particular,

$$\begin{aligned}\int_{r_1}^{r_2} L(\gamma',\gamma,\omega)\,ds &\le \int_{r_1}^{r_2} L(\hat{\gamma}',\gamma,\omega)\,ds + \delta \\ &\le (r_2-r_1)b + (r_2-r_1)K\left(\frac{\gamma(r_2)-\gamma(r_1)}{r_2-r_1}\right) + \delta. \end{aligned} \tag{2.14}$$

This proves (2.12). The lower bound (2.13) follows from Jensen's inequality, the convexity of K, and the fact that $V(x,\omega) \ge a$. □

Lemma 2.2 *Suppose that $g : \mathbb{R}^d \to \mathbb{R} \cup \{-\infty\}$, $g(0) \in \mathbb{R}$, and*

$$g(y) < g(0) + C_1(1+|y|), \quad \forall\ y \in \mathbb{R}^d. \tag{2.15}$$

There is a constant R depending only on K, C_1, and $b-a$ such that

$$|\gamma(t_2)-\gamma(t_1)| \le R(1+|t_1-t_2|), \quad \forall\, t_2, t_1 \in [0,t]$$

holds for all paths $\gamma \in M_1(t,\omega)$ and all $t > 1$.

Proof of Lemma 2.2: We first show there is a constant R_0 depending only on K, C_1, and $b-a$ such that

$$|\gamma(t)-\gamma(0)| \le tR_0 \tag{2.16}$$

holds for all $\gamma \in M_1(t,\omega)$ and all $t \ge 1$. Define the path $\hat{\gamma}(s) = \gamma(0) = 0$ for all $s \in [0,t]$. We have

$$u(t,\omega) \ge g(0) - \mathcal{L}(\hat{\gamma},\omega) \ge g(0) - tb. \tag{2.17}$$

By (2.13), we also have the lower bound

$$\mathcal{L}(\gamma,\omega) \ge tK\left(\frac{\gamma(t)-\gamma(0)}{t}\right) + at.$$

Since $\gamma \in M_1(\omega)$, we may combine these two estimates with $u(t,\omega) \leq 1 + g(\gamma(t)) - \mathcal{L}(\gamma,\omega)$ to conclude

$$K\left(\frac{\gamma(t)-\gamma(0)}{t}\right) \leq \frac{1}{t} + (b-a) + \frac{g(\gamma(t)) - g(0)}{t} \leq 1 + (b-a) + C_1\left(\frac{1+|\gamma(t)|}{t}\right).$$

Since $K(p)$ grows super-linearly in $|p|$, (2.16) follows. If we assume $t \in [1, 50]$, then (2.16) implies the desired result with $R = 100R_0$, since $|\gamma(t_2)-\gamma(t_1)| \leq |\gamma(t_2)-\gamma(0)|+|\gamma(t_1)-\gamma(0)| \leq R_0(t_1+t_2)$.

Next, suppose $t \geq 50$ and consider γ at integer times $k \in [1, t-1] \cap \mathbb{Z}$. We will show that there is a constant R_1, independent of $t > 1$, such that at least five of the times $k \in [1, t-1] \cap \mathbb{Z}$ must satisfy both

$$|\gamma(k) - \gamma(k-1)| \quad \leq \quad R_1 \quad \text{and} \quad |\gamma(k+1) - \gamma(k)| \leq R_1. \tag{2.18}$$

Arguing by way of contradiction, let us suppose (2.18) does not hold for more than four times $k \in [1, t-1] \cap \mathbb{Z}$. Then $|\gamma(j+1) - \gamma(j)| > R_1$ must hold for at least $\lfloor \frac{t}{2} \rfloor - 5$ of the times $j \in [1, t-1] \cap \mathbb{Z}$. Since $t > 50$ we have $\lfloor \frac{t}{2} \rfloor - 5 > t/3$. This implies that

$$u(t,\omega) \leq 1 + g(\gamma(t)) - \mathcal{L}(\gamma,\omega) \leq 1 + g(\gamma(t)) - at - \frac{t}{3}\min_{|q|\geq R_1} K(q).$$

On the other hand, by Lemma 2.1 and (2.16), we know that

$$u(t,\omega) \geq g(\gamma(t)) - bt - tK(\frac{\gamma(t)-\gamma(0)}{t}) - 1 \geq g(\gamma(t)) - bt - t\max_{|q|\leq R_0} K(q) - 1$$

holds for all $\gamma \in M_1(\omega)$. Combining these two bounds we obtain

$$\frac{1}{3}\min_{|q|\geq R_1} K(q) \leq 1 + (b-a) + \max_{|q|\leq R_0} K(q).$$

If $R_1 > R_0$ is sufficiently large (depending only on $b-a$, R_0, and K) this forces a contradiction. So, (2.18) must hold.

Now we conclude the proof. Let $R_2 > R_1$, and suppose that for some $t_1, t_2 \in [0, t]$ with $t_1 < t_2$ and $1 \leq |t_2 - t_1| \leq 2$, we have $|\gamma(t_2) - \gamma(t_1)| \geq R_2$. Let $k \in [1, t-1] \cap \mathbb{Z}$ be such that (2.18) holds. Since there are at least five such times, we may suppose, without loss of generality, that $(k-1, k+1) \cap (t_1, t_2) = \emptyset$. Let's also suppose $k+1 \leq t_1$ (the case $k-1 \geq t_2$ is similar). Consider the path $\hat{\gamma}$ defined by

$$\hat{\gamma}(s) = \begin{cases} \gamma(s), & \text{for } s \in [0, k-1] \cup [t_2, t], \\ \gamma(k-1) + (s-k+1)(\gamma(k+1) - \gamma(k-1)), & \text{for } s \in [k-1, k], \\ \gamma(s+1), & \text{for } s \in [k, t_1 - 1], \end{cases}$$

and for $s \in [t_1 - 1, t_2]$

$$\hat{\gamma}(s) = \gamma(t_1) + (\gamma(t_2) - \gamma(t_1))\frac{s - t_1 + 1}{t_2 - t_1 + 1}.$$

Then we have

$$\begin{aligned}
\mathcal{L}(\hat{\gamma}) - \mathcal{L}(\gamma) &\leq 4(b-a) + \int_{k-1}^{k} K(\hat{\gamma}'(s))\,ds + \int_{t_1-1}^{t_2} K(\hat{\gamma}'(s))\,ds \\
&\quad - \int_{k-1}^{k+1} K(\gamma'(s))\,ds - \int_{t_1}^{t_2} K(\gamma'(s))\,ds \\
&\leq 4(b-a) + K(\gamma(k+1)-\gamma(k-1)) + (t_2-t_1+1)K\left(\frac{\gamma(t_2)-\gamma(t_1)}{t_2-t_1+1}\right) \\
&\quad -2K\left(\frac{\gamma(k+1)-\gamma(k-1)}{2}\right) - (t_2-t_1)K\left(\frac{\gamma(t_2)-\gamma(t_1)}{t_2-t_1}\right) \\
&\leq M + (t_2-t_1+1)K\left(\frac{\gamma(t_2)-\gamma(t_1)}{t_2-t_1+1}\right) - (t_2-t_1)K\left(\frac{\gamma(t_2)-\gamma(t_1)}{t_2-t_1}\right), \quad (2.19)
\end{aligned}$$

where

$$M = 4(b-a) + \max_{|z|\leq 2R_1} K(z).$$

Let $\Delta t = t_2 - t_1$ and $\sigma = (\Delta t + 1)/(\Delta t)$ and $z = (\gamma(t_2) - \gamma(t_1))/(t_2 - t_1 + 1)$. The inequality (2.19) has the form

$$\mathcal{L}(\hat{\gamma}) - \mathcal{L}(\gamma) \leq M + (\Delta t + 1)K(z) - \Delta t K(\sigma z). \quad (2.20)$$

The properties of K (convexity and super-linear growth) imply that if R_2 sufficiently large, then

$$\inf_{|z|\geq R_2/3} K(\sigma z) - \sigma K(z) > M + 1.$$

Applying this at (2.20) we conclude $\mathcal{L}(\hat{\gamma}) - \mathcal{L}(\gamma) < -1$, which contradicts the fact that $\gamma \in M_1(t,\omega)$. Therefore, we must have $|\gamma(t_2) - \gamma(t_1)| \leq R_2$ if $1 \leq |t_1 - t_2| \leq 2$. This and the triangle inequality now imply the desired result for all $t_1, t_2 \in [0,t]$. ☐

By Lemma 2.2, we have

$$\{\gamma(s) \mid s \in [0,t]\} \subset \bigcup_{k\in[0,t)\cap\mathbb{Z}} \overline{B_{2R}(\gamma(k))}, \quad (2.21)$$

for all $\gamma \in M_1(t,\omega)$, which immediately implies that approximate optimizers visit at most $O(t)$ cubes:

Corollary 2.3 *There is a constant $C_d > 0$ such that if $\gamma \in M_1(t,\omega)$, then there are at most $C_d(1+t)$ indices $j \in \mathbb{Z}^d$ for which $\gamma(s) \in Q_j$ for some $s \in [0,t]$.*

Remark 2.4 *Here, and later, the constant implied by the notation $O(t)$ is independent of ω.*

Important cubes

Our method of estimating the variance of u involves bounding the random variable $|\sigma_j u - u|$. So, we must understand when changing the value of ω_j leads to a large change in the value of $u(t,\omega)$. Given a path $\gamma \in \mathcal{A}_t$ and an index $j \in \mathbb{Z}^d$, define

$$\pi_j(\gamma) = |\{s \in [0,t] \mid \gamma(s) \in Q_j\}|,$$

which is the total time that the path γ occupies the cube Q_j. Observe that for any path $\gamma \in \mathcal{A}_t$, we have

$$\mathcal{L}(\gamma, \phi_j\omega) \le \mathcal{L}(\gamma, \omega) + (b - \omega_j)\pi_j(\gamma). \tag{2.22}$$

In particular, if $\gamma \in M_\delta(t, \omega)$, then

$$\begin{aligned} \sigma_j u(t, \omega) = u(t, \phi_j\omega) &\ge g(\gamma(t)) - \mathcal{L}(\gamma, \phi_j\omega) \\ &\ge g(\gamma(t)) - \mathcal{L}(\gamma, \omega) - (b - \omega_j)\pi_j(\gamma) \\ &\ge u(t, \omega) - (b - \omega_j)\pi_j(\gamma) - \delta. \end{aligned} \tag{2.23}$$

From this we deduce that if $\omega_j = b$ or if there is $\gamma \in M_\delta(t, \omega)$ for which $\pi_j(\gamma) = 0$, then it must be true that $u(t, \omega) - \sigma_j u(t, \omega) \le \delta$. On the other hand, this also shows that if $u(t, \omega) - \sigma_j u(t, \omega) > \delta$, then $\omega_j = a$ and $\pi_j(\gamma) > 0$ must hold for all $\gamma \in M_\delta(t, \omega)$. This motivates the following definition. We say that the cube Q_j is *important* if $\omega_j = a$ and for some $\delta > 0$ we have

$$\pi_j(\gamma) > 0, \ \ \forall\, \gamma \in M_\delta(t, \omega). \tag{2.24}$$

Observe that if (2.24) holds for some $\delta > 0$, then it also holds for all $\delta' \in (0, \delta]$. So, Q_j is important if $\omega_j = a$ and for δ sufficiently small every δ-approximate optimizer spends time in cube Q_j. Let $\mathcal{I}_j \subset \Omega$ denote the event that the cube Q_j is important:

$$\begin{aligned} \mathcal{I}_j &= \{\omega \in \Omega \mid \omega_j = a;\ \ \exists \delta > 0 \ \text{ such that } \ \pi_j(\gamma) > 0 \ \ \forall\, \gamma \in M_\delta(t, \omega)\} && (2.25) \\ &= \bigcup_{n \ge 1} \left\{\omega \in \Omega \mid \omega_j = a;\ \ \pi_j(\gamma) > 0 \ \ \forall\, \gamma \in M_{1/n}(t, \omega)\right\}. && (2.26) \end{aligned}$$

The above analysis shows that

$$\{\omega \in \Omega \mid u(t, \omega) > \sigma_j u(t, \omega)\} \subset \mathcal{I}_j, \quad \forall\, j \in \mathbb{Z}^d \tag{2.27}$$

so we have

$$\mathbb{P}(\sigma_j u < u) \le \mathbb{P}(\mathcal{I}_j). \tag{2.28}$$

Observe that $\mathbb{P}(\mathcal{I}_j)$ depends on t, in addition to j.

It will be useful to further classify some cubes as *very important*. To this end, we define a set of cubes

$$N(\delta, \omega) = \bigcup_{\gamma \in M_\delta(\omega)} \{k \in \mathbb{Z}^d \mid \pi_k(\gamma) > 0\}.$$

This is the set of all cubes visited by some path $\gamma \in M_\delta(\omega)$. Next, we define the event $\mathcal{I}_j^+ \subset \mathcal{I}_j \subset \Omega$ that cube Q_j is *very important*:

$$\mathcal{I}_j^+ = \{\omega \in \mathcal{I}_j \mid \exists\, \delta > 0 \ \text{ such that } \ \omega_\ell = b \quad \forall\, \ell \in N(\delta, \omega) \setminus \{j\}\}. \tag{2.29}$$

On this event, Q_j is an important cube, and for any other cube Q_ℓ visited by a path $\gamma \in M_\delta(\omega)$, we have $\omega_\ell = b$, if δ is sufficiently small. On the event $\mathcal{I}_j^- = \mathcal{I}_j \setminus \mathcal{I}_j^+$, cube Q_j is important but not very important: for any $\delta > 0$ we can find a path $\gamma \in M_\delta(t, \omega)$ such that γ passes through another cube $Q_\ell \neq Q_j$, on which $\omega_\ell = a$. The following lemma shows that the only way for $(u - \sigma_j u)^2 1_{\mathcal{I}_j}$ to be large is if Q_j is very important.

Lemma 2.5 *There is a constant $C_0 > 0$, depending only on K and $|b-a|$, such that*

$$\mathbb{P}\left(\{\omega \mid \ (u-\sigma_j u)^2 1_{\mathcal{I}_j^-} < C_0\}\right) = 1$$

holds for all $t \geq 1$ and $j \in \mathbb{Z}^d$.

Proof of Lemma 2.5: If $\omega_j = b$, then $\omega \notin \mathcal{I}_j^-$, so obviously $(u-\sigma_j u)^2 1_{\mathcal{I}_j^-} = 0$. Hence, we may assume $\omega_j = a$ and $\omega \in \mathcal{I}_j^-$. When $\omega_j = a$, we clearly have $\sigma_j u \leq u$, since $\mathcal{L}(\gamma, \phi_j \omega) \geq \mathcal{L}(\gamma, \omega)$ in this case. So, we must bound $u - \sigma_j u$ from above.

Consider an approximate optimizer $\gamma \in M_\delta(\omega)$ for some $\delta \leq 1$. If $\pi_j(\gamma) \leq 1$ then $u - \sigma_j u \leq (b-a) + \delta$ according to (2.23). So, we must consider the possibility that $\pi_j(\gamma) > 0$ is large. Since $\omega \in \mathcal{I}_j^-$, we may assume the path γ also passes through another cube Q_ℓ, with $\ell \neq j$, for which $\omega_\ell = a$. We will construct a new path $\hat{\gamma}$ such that $\pi_j(\hat{\gamma}) \leq 1$ and $\mathcal{L}(\hat{\gamma}, \omega) \leq \mathcal{L}(\gamma, \omega) + C$. The two paths $\hat{\gamma}$ and γ will have the same starting and ending points. This implies that the difference $u - \sigma_j u$ is bounded by a constant, since by (2.22) we have

$$\begin{aligned}
\sigma_j u = u(t, \phi_j \omega) \ &\geq \ g(\hat{\gamma}(t)) - \mathcal{L}(\hat{\gamma}, \phi_j \omega) \\
&\geq \ g(\hat{\gamma}(t)) - \mathcal{L}(\hat{\gamma}, \omega) - (b-a)\pi_j(\hat{\gamma}) \\
&\geq \ g(\hat{\gamma}(t)) - \mathcal{L}(\gamma, \omega) - C - (b-a)\pi_j(\hat{\gamma}) \geq u(t, \omega) - C - (b-a) - \delta. \quad (2.30)
\end{aligned}$$

Suppose that $[t_1, t_2]$ is the smallest interval containing all s for which $\gamma(s) \in Q_j$. We may assume $t_2 - t_1 \geq \pi_j(\gamma) > 1$. Suppose that $\gamma(t_3) \in Q_\ell$ where $\ell \neq j$ and $\omega_\ell = a$. Since $\gamma \in M_\delta(\omega)$ we may assume (by Lemma 2.1) that $\gamma(s) \in Q_j$ for all $s \in (t_1, t_2)$, so $t_3 \notin (t_1, t_2)$. Suppose that $t_3 \geq t_2$ (the case $t_3 \leq t_1$ is similar). Define the new path $\hat{\gamma}$ as follows:

(i) For $s \in [0, t_1]$, let $\hat{\gamma}(s) = \gamma(s)$.

(ii) For $s \in [t_1, t_1 + 1]$, let $\hat{\gamma}(s) = \gamma(t_1) + (s - t_1)(\gamma(t_2) - \gamma(t_1))$.

(iii) For $s \in [t_1 + 1, t_1 + 1 + (t_3 - t_2)]$, let $\hat{\gamma}(s) = \gamma(s - t_1 - 1 + t_2)$.

(iv) For $s \in [t_1 + 1 + (t_3 - t_2), t_3]$, let $\hat{\gamma}(s) = \gamma(t_3)$.

(v) For $s \in [t_3, t]$, let $\hat{\gamma}(s) = \gamma(s)$.

Much of $\hat{\gamma}$ is just a linear reparameterization of γ, and we have

$$\int_0^t L(\hat{\gamma}'(s), \hat{\gamma}(s), \omega)\, ds - \int_0^t L(\gamma'(s), \gamma(s), \omega)\, ds \leq \int_{t_1}^{t_1+1} K(\hat{\gamma}'(s))\, ds \leq K(\gamma(t_2) - \gamma(t_1)).$$

Since $|\gamma(t_2) - \gamma(t_1)|$ is bounded by the diameter of cube Q_j, we have $\mathcal{L}(\hat{\gamma}, \omega) \leq \mathcal{L}(\gamma, \omega) + C$ with $C = \max_{|p| \leq \sqrt{d}} K(p)$. $\square$

3 Proof of Theorem 1.1

In this section we prove Theorem 1.1. As we have mentioned, the main argument is similar to that of [6]. In particular, it is convenient to average $u(t, \omega)$ over a random shift of the environment.

Random shifting of the environment

We now consider an augmented probability space $\tilde{\Omega} = \Omega \times \Omega_1$ with product measure $\tilde{\mathbb{P}} = \mathbb{P} \times P_1$, and we introduce a random function $h(\omega_1) : \Omega_1 \to \mathbb{Z}^d$ to define a random shift of the environment. For $(\omega, \omega_1) \in \tilde{\Omega}$, let us define

$$\tilde{u}(t, \omega, \omega_1) = u(t, \tau_{h(\omega_1)}\omega) = \sup_{\gamma \in \mathcal{A}_t} g(\gamma(t)) - \mathcal{L}(\gamma + h(\omega_1), \omega) \tag{3.31}$$

where $\gamma + h(\omega_1)$ denotes the shifted path $t \mapsto \gamma(t) + h(\omega_1)$. We define $M_\delta(\omega, \omega_1) = M_\delta(\tau_{h(\omega_1)}\omega)$ to be the set of paths $\gamma \in \mathcal{A}_t$ for which

$$\tilde{u}(t, \omega, \omega_1) \le g(\gamma(t)) - \mathcal{L}(\gamma + h(\omega_1), \omega) + \delta.$$

We construct Ω_1, P_1, and h in such a way that $|\tilde{u}(t, \omega, \omega_1) - u(t, \omega)| = o(\sqrt{t})$, and for this reason an estimate of $\mathrm{var}(\tilde{u})$ that is sublinear in t will imply a sublinear bound for $\mathrm{var}(u)$.

The random shift $h(\omega_1)$ will lie in the set $[0, m)^d \subset \mathbb{R}^d$ where $m = \lfloor t^\zeta \rfloor$, for some positive $\zeta < 1/2$. For $d \ge 3$, it will suffice to choose $\zeta \in (\frac{1}{d}, \frac{1}{2})$. For $d = 2$, we will require that $\zeta \in \left(\frac{\nu - 1}{2\nu - 1}, 1/2\right)$, where ν was defined by the non-degeneracy condition (1.6). Denote by P_0 the product probability measure on the set $\Omega_0 = \{a, b\}^{m^2}$ and having marginal distribution $P_0'(a) = \alpha$, $P_0'(b) = \beta = 1 - \alpha \in (0, 1)$. The following statement is Lemma 3 from [6], so we omit the proof:

Lemma 3.1 *There exists a constant $C > 0$ independent of $m = \lfloor t^\zeta \rfloor$, and a function $\tilde{h} : \Omega_0 \to \{0, 1, \ldots, m-1\}$ for which the following two conditions hold:*

(i) $P_0(\tilde{h} = i) \le \frac{C}{m}$ for all $i \in [0, m-1]$, and

(ii) for every $x, y \in \Omega_0$ that differ in at most one coordinate, the difference between $\tilde{h}(x)$ and $\tilde{h}(y)$ satisfies

$$|\tilde{h}(x) - \tilde{h}(y)| \le 1.$$

Define the set $\Theta = \Theta_t = \{1, 2, \ldots, d\} \times \{1, 2, \ldots, m^2\}$. Let $\Omega_1 = \{a, b\}^\Theta$, and let P_1 be a uniform probability measure on Ω_1. Each $\omega_1 \in \Omega_1$ can be written as $\omega_1 = (\omega_1^1, \omega_1^2, \ldots, \omega_1^d)$, where each $\omega_1^i \in \Omega_0$ is a binary sequence of length m^2. Let $\vec{e}_i$ denote the i-th coordinate vector. Define $h(\omega_1) = \sum_{i=1}^d \tilde{h}(\omega_1^i)\vec{e}_i$. There exists a constant $C > 0$ independent on m such that for each $x \in \{0, 1, \ldots, m-1\}^d$ one has $P_1(h = x) \le \frac{C}{m^d}$. Moreover, if ω_1 and ω_1' differ in exactly one coordinate then we have $|h(\omega_1) - h(\omega_1')| \le 1$. Given the space $\tilde{\Omega} = \Omega \times \Omega_1$ with the product measure $\tilde{\mathbb{P}} = \mathbb{P} \times P_1$ on $\Omega \times \Omega_1$ defined in this way, we now consider the function $\tilde{u}$ defined by (3.31).

Lemma 3.2 *There is a constant $C > 0$ such that*

$$|u(t, \omega) - \tilde{u}(t, \omega, \omega_1)| \le C|h(\omega_1)| \le Ct^\zeta, \quad \forall\, t > 1 \tag{3.32}$$

holds $\tilde{\mathbb{P}}$-almost surely, and

$$\mathrm{var}\, u \le C\, \mathrm{var}\, \tilde{u} + Ct^{2\zeta}, \quad \forall\ t > 1. \tag{3.33}$$

Proof of Lemma 3.2: We will prove that $|u(t, \omega) - \tilde{u}(t, \omega, \omega_1)| \le C|h(\omega_1)|$, $\tilde{\mathbb{P}}$-almost surely, for some constant $C > 0$ independent of m and t. Given a path $\gamma \in M_\delta(\omega)$, we can modify it to construct an approximate optimizer for $\tilde{u}(t, \omega, \omega_1)$, thus estimating $\tilde{u}(t, \omega, \omega_1) - u(t, \omega)$ from below. However, we cannot simply shift γ by $-h(\omega_1)$, since we must preserve the starting and ending points.

Suppose $|h(\omega_1)| \le \sqrt{d}\kappa$, with $\kappa \in [1, m] \cap \mathbb{Z}$. Fixing a path $\gamma \in M_\delta(\omega)$, we define the new path $\hat{\gamma}$ in the following way:

(i) For $r \in [0,\kappa]$, set $\hat{\gamma}(r) = \gamma(0) + \left(\frac{r}{\kappa}\right)(\gamma(2\kappa) - h(\omega_1) - \gamma(0))$.

(ii) For $r \in [\kappa, t-\kappa]$, set $\hat{\gamma}(r) = \gamma(r+\kappa) - h(\omega_1)$.

(iii) For $r \in [t-\kappa, t]$, set $\hat{\gamma}(r) = \gamma(t) + (r-t)\frac{h(\omega_1)}{\kappa}$.

We now verify that the path $\hat{\gamma}$ yields the desired bound on $\tilde{u}(t,\omega,\omega_1) - u(t,\omega)$ and $\operatorname{var} u$. Since $\gamma \in M_\delta(\omega)$, we have

$$\begin{aligned}\tilde{u}(t,\omega,\omega_1) - u(t,\omega) &\geq -\int_0^\kappa L(\hat{\gamma}'(s),\hat{\gamma}(s)+h(\omega_1),\omega)\,ds - \int_{t-\kappa}^t L(\hat{\gamma}'(s),\hat{\gamma}(s)+h(\omega_1),\omega)\,ds \\ &\quad + \int_0^{2\kappa} L(\gamma'(s),\gamma(s),\omega)\,ds - \delta \\ &\geq -C\kappa\left(1 + \sup_{|z|\leq 2R} K(z)\right) - \delta,\end{aligned}$$

where C is a positive real number that depends only on K, $|b-a|$. In a similar way we prove that $\tilde{u}(t,\omega,\omega_1) - u(t,\omega) \leq C\kappa + \delta$. Recalling that $m = \lfloor t^\zeta \rfloor$, we obtain (3.32). Therefore

$$\begin{aligned}\operatorname{var} u &= \tilde{\mathbb{E}}[(u - \tilde{\mathbb{E}}(u))^2] \\ &= \tilde{\mathbb{E}}\left[\left(\tilde{u} - \tilde{\mathbb{E}}(\tilde{u}) + (u-\tilde{u}) - \tilde{\mathbb{E}}(u-\tilde{u})\right)^2\right] \\ &\leq 2\operatorname{var}\tilde{u} + 2\tilde{\mathbb{E}}\left[\left((u-\tilde{u}) - \tilde{\mathbb{E}}(u-\tilde{u})\right)^2\right] \\ &\leq 2\operatorname{var}\tilde{u} + 8C^2t^{2\zeta},\end{aligned}$$

which is (3.33).

□

Variance estimate for $\tilde{u}$

Given Lemma 3.2 and the choice of $\zeta < 1/2$, we now wish to establish a bound of order $t/\log t$ for the variance of $\tilde{u}(t,\omega,\omega_1)$ under $\tilde{\mathbb{P}}$. The augmented probability space was constructed in such a way that $\tilde{u}(t,\omega,\omega_1)$ is amenable to Talagrand's inequality. The function u depends on ω_j for only $O(t^d)$ of the indices $j \in \mathbb{Z}^d$:

Lemma 3.3 *There is a constant $R > 0$ such that*

$$\tilde{u}(t,\omega,\omega_1) = \tilde{u}(t,\phi_j\omega,\omega_1), \quad \forall\ j \in \mathbb{Z}^d,\ \ |j| > Rt,\ \ t > 0$$

holds $\tilde{P}$ almost surely.

Proof: Since $|h(\omega_1)| \leq m\sqrt{d}$, this is a consequence of Lemma 2.2: no approximate optimizer passes through cube j, if R is sufficiently large and $|j| > Rt$. □

In view of Lemma 3.3, we may regard $\tilde{u}$ as a function of no more than $Ct^d + dm^2$ random variables taking values in the set $\{a,b\}$. In this way, Talagrand's inequality (Theorem 1.2) implies that there is a constant $C > 0$, independent of $t > 0$, such that

$$\operatorname{var}(\tilde{u}) \leq C \sum_{j\in B_t} \frac{\|\rho_j\tilde{u}\|_2^2}{1 + \log\frac{\|\rho_j\tilde{u}\|_2}{\|\rho_j\tilde{u}\|_1}} + C\sum_{k\in\Theta_t} \frac{\|\rho_k\tilde{u}\|_2^2}{1+\log\frac{\|\rho_k\tilde{u}\|_2}{\|\rho_k\tilde{u}\|_1}}. \tag{3.34}$$

where B_t is the set $B_t = \{j \in \mathbb{Z}^d \mid \ |j| \le Rt\}$, whose cardinality is bounded by Ct^d. The norms $\|\cdot\|_2$ and $\|\cdot\|_1$ refer to the $L^2(\tilde{\Omega}, \tilde{P})$ and $L^1(\tilde{\Omega}, \tilde{P})$ norms, respectively. Observe that if $k \in \Theta_t$, then $\rho_k \tilde{f}$ corresponds to translation of the random environment:

$$\rho_k \tilde{u} = \frac{\sigma_k \tilde{u} - \tilde{u}}{2} = \frac{\tilde{u}(t,\omega,\phi_k\omega_1) - \tilde{u}(t,\omega,\omega_1)}{2}.$$

If $j \in B_t$, then $\rho_j \tilde{u}$ corresponds to a local change in the random environment over the cube Q_j:

$$\rho_j \tilde{u} = \frac{\sigma_j \tilde{u} - \tilde{u}}{2} = \frac{\tilde{u}(t,\phi_j\omega,\omega_1) - \tilde{u}(t,\omega,\omega_1)}{2}.$$

Let us first consider the second sum in (3.34). We will show that this sum is $O(t^{2\zeta})$.

Lemma 3.4 *There is a constant $C > 0$ such that*

$$\sum_{k \in \Theta_t} \frac{\|\rho_k \tilde{u}\|_2^2}{1 + \log \frac{\|\rho_k \tilde{u}\|_2}{\|\rho_k \tilde{u}\|_1}} \le Ct^{2\zeta} \tag{3.35}$$

holds for all $t > 1$.

Proof: Since there are only $|\Theta_t| = m^2 \le t^{2\zeta}$ terms in the sum and since

$$1 + \log \frac{\|\rho_k \tilde{u}\|_2}{\|\rho_k \tilde{u}\|_1} \ge 1,$$

the lemma will follow from a uniform bound on $\|\rho_k \tilde{u}\|_2$. By definition of $h(\omega_1)$, we know that $|h(\phi_k\omega_1) - h(\omega_1)| \le 1$. So, by Lemma 3.2, we have$|\tilde{u}(t,\omega,\omega_1) - \tilde{u}(t,\omega,\phi_k\omega_1)| \le C|h(\phi_k\omega_1) - h(\omega_1)| \le C$ holds $\tilde{P}$ almost surely, for all $k \in \Theta_t$, $t \ge 1$. □

Having established (3.35), we now consider the first sum in (3.34).

Proposition 3.5 *There is a constant $C > 0$ such that*

$$\sum_{j \in B_t} \frac{\|\rho_j \tilde{u}\|_2^2}{1 + \log \frac{\|\rho_j \tilde{u}\|_2}{\|\rho_j \tilde{u}\|_1}} \le C \frac{t}{\log t} \tag{3.36}$$

holds for all $t > 1$.

Since we may have $\alpha \ne \beta$, we will make use of the following fact, proved in the appendix:

Lemma 3.6 *Let $C' = \min\left\{\frac{\alpha}{\beta}, \frac{\beta}{\alpha}\right\}$ and $C'' = \max\left\{\frac{\alpha}{\beta}, \frac{\beta}{\alpha}\right\}$. For any measureable set $A \subset \Omega$,*

$$C'\mathbb{P}(A) \le \mathbb{P}(\phi_j A) \le C''\mathbb{P}(A) \tag{3.37}$$

holds for all $j \in \mathbb{Z}^d$. Also, for every nonnegative integrable ψ, we have

$$C'\mathbb{E}(\psi \circ \phi_j) \le \mathbb{E}(\psi) \le C''\mathbb{E}(\psi \circ \phi_j). \tag{3.38}$$

Proof of Proposition 3.5. Let us begin by estimating $\|\rho_j \tilde{u}\|_2^2$. By Lemma 3.6, we have

$$\|\rho_j \tilde{u}\|_2^2 = \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\sigma_j \tilde{u} > \tilde{u}}\right] + \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\sigma_j \tilde{u} < \tilde{u}}\right] \le C \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\sigma_j \tilde{u} < \tilde{u}}\right].$$

Recalling the definition (2.25), let $\tilde{\mathcal{I}}_j \subset \tilde{\Omega}$ be the event that Q_j is an important cube in the shifted environment:

$$\tilde{\mathcal{I}}_j = \{(\omega, \omega_1) \in \tilde{\Omega} \mid \tau_{h(\omega_1)} \omega \in \mathcal{I}_j\}.$$

Because of (2.27), the event $\{\sigma_j \tilde{u} < \tilde{u}\}$ is contained in the event $\tilde{\mathcal{I}}_j$. So, we have

$$\|\rho_j \tilde{u}\|_2^2 \le C \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\tilde{\mathcal{I}}_j}\right]. \tag{3.39}$$

The difference $|\sigma_j \tilde{u} - \tilde{u}|$ could be large in some cases, even on the event $\tilde{\mathcal{I}}_j$, so we will distinguish a few possible scenarios. Let $\tilde{\mathcal{I}}_j^+ \subset \tilde{\mathcal{I}}_j$ denote the event that cube Q_j is *very important* in the shifted environment:

$$\tilde{\mathcal{I}}_j^+ = \{(\omega, \omega_1) \in \tilde{\mathcal{I}}_j \mid \tau_{h(\omega_1)} \omega \in \mathcal{I}_j^+\}.$$

Similarly, let $\tilde{\mathcal{I}}_j^- = \tilde{\mathcal{I}}_j \backslash \tilde{\mathcal{I}}_j^+$ be the event that the cube Q_j is important but not very important. Since $\omega \mapsto \tau_{h(\omega_1)} \omega$ is measure preserving on Ω, we have

$$\tilde{\mathbb{P}}(\{(\omega, \omega_1) \in \tilde{\Omega} \mid (\tilde{u} - \sigma_j \tilde{u})^2 1_{\tilde{\mathcal{I}}_j^-} > C_0\}) = \mathbb{P}(\{\omega \in \Omega \mid (u - \sigma_j u)^2 1_{\mathcal{I}_j^-} > C_0\}).$$

Consequently, from Lemma 2.5 and (3.39) we have

$$\begin{aligned} \|\rho_j \tilde{u}\|_2^2 &\le C \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\tilde{\mathcal{I}}_j^+}\right] + C \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\tilde{\mathcal{I}}_j^-}\right] \\ &\le C \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\tilde{\mathcal{I}}_j^+}\right] + C C_0 \tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j). \end{aligned} \tag{3.40}$$

Whether the event $\tilde{\mathcal{I}}_j^+$ has small probability depends on the function $g(y)$, so we distinguish two cases. Let $\tilde{G} \subset \tilde{\Omega}$ denote the event that

$$|\gamma(t) - \gamma(0)| \ge t^{1/4}, \quad \forall\, \gamma \in M_1(\omega, \omega_1).$$

On this event, all approximate minimizers must travel a distance at least $O(t^{1/4})$ from their starting point $\gamma(0) = 0$. According to the following lemma, the probability that minimizers travel that far when a cube Q_j is very important must be small.

Lemma 3.7 *There are constants $\kappa_1, \kappa_2 > 0$ such that*

$$\tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j^+ \cap \tilde{G}) \le \kappa_1 e^{-\kappa_2 t^{1/4}}, \quad \forall\ t > 0, \quad j \in \mathbb{Z}^d.$$

Therefore, returning to (3.40) and using the fact that $|\sigma_j \tilde{u} - \tilde{u}| \le O(t)$ and $\tilde{\mathcal{I}}_j^+ = (\tilde{\mathcal{I}}_j^+ \cap \tilde{G}) \cup (\tilde{\mathcal{I}}_j^+ \cap \tilde{G}^C)$, we conclude

$$\|\rho_j \tilde{u}\|_2^2 \le C t^2 e^{-\kappa_2 t^{1/4}} + C \tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j) + C \tilde{\mathbb{E}}\left[(\sigma_j \tilde{u} - \tilde{u})^2 1_{\tilde{\mathcal{I}}_j^+ \cap \tilde{G}^C}\right].$$

Hence,

$$\sum_{j \in B_t} \|\rho_j \tilde{u}\|_2^2 \le C |B_t| t^2 e^{-\kappa_2 t} + C \tilde{\mathbb{E}}\left[\sum_{j \in B_t} 1_{\tilde{\mathcal{I}}_j}\right] + C \tilde{\mathbb{E}}\left[\sum_{j \in B_t} (\sigma_j \tilde{u} - \tilde{u})^2 1_{\tilde{\mathcal{I}}_j^+ \cap \tilde{G}^C}\right]. \tag{3.41}$$

With probability one, the sum $\sum_{j \in B_t} 1_{\tilde{\mathcal{I}}_j}$ is bounded by $O(t)$ because there can be at most $O(t)$ important cubes, as the total number of cubes visited is $O(t)$, by Corollary 2.3.

The last term in (3.41) is bounded as follows. First,

Lemma 3.8 *There are $\kappa_1, \kappa_2 > 0$ such that*

$$\tilde{\mathbb{P}}\left(\{(\omega,\omega_1) \mid \ (\sigma_j\tilde{u}-\tilde{u})^2 1_{\tilde{G}^C} > C_0 t^{1/2}\}\right) \le \kappa_1 e^{-\kappa_2 t^{1/4}}$$

holds for all $t > 1$ and all $j \in \mathbb{Z}^d$.

Furthermore, if $\omega \in \mathcal{I}_j^+$, then $\omega \notin \mathcal{I}_k^+$ for any $k \neq j$, since Q_j must be the only important cube. Therefore, since $|\sigma_j\tilde{u}-\tilde{u}| \le |b-a|t$ always holds (by (2.23)), we must have

$$\tilde{\mathbb{E}}\left[\sum_{j\in B_t}(\sigma_j\tilde{u}-\tilde{u})^2 1_{\tilde{\mathcal{I}}_j^+\cap\tilde{G}^C}\right] \le C_0 t^{1/2} + |b-a|^2 t^2 \kappa_1 e^{-\kappa_2 t^{1/4}}.$$

Considering (3.41), we have now shown that there is a constant $C' > 0$ for which

$$\sum_{j\in B_t}\|\rho_j\tilde{u}\|_2^2 \le C' t \tag{3.42}$$

holds for all $t > 1$.

Next we consider the denominator in (3.36). We show that there is a constant $C'' > 0$ such that

$$\log\frac{\|\rho_j\tilde{u}\|_2}{\|\rho_j\tilde{u}\|_1} \quad \ge \quad C''\log t, \tag{3.43}$$

for all $t > 1$. By the Cauchy-Schwarz inequality we see that

$$\|\rho_j\tilde{u}\|_1 = \|\rho_j\tilde{u}\cdot 1_{\sigma_j\tilde{u}\neq\tilde{u}}\|_1 \le \|\rho_j\tilde{u}\|_2\cdot\sqrt{\mathbb{P}(\sigma_j\tilde{u}\neq\tilde{u})}.$$

Since $\sigma_j\sigma_j u = u$, Lemma 3.6 implies

$$\tilde{\mathbb{P}}(\sigma_j\tilde{u}\neq\tilde{u}) = \tilde{\mathbb{P}}(\sigma_j\tilde{u}>\tilde{u}) + \tilde{\mathbb{P}}(\sigma_j\tilde{u}<\tilde{u}) \le (1+C'')\tilde{\mathbb{P}}(\sigma_j\tilde{u}<\tilde{u}).$$

Hence,

$$\frac{\|\rho_j\tilde{u}\|_2}{\|\rho_j\tilde{u}\|_1} \ge \frac{1}{\sqrt{(1+C'')\tilde{\mathbb{P}}(\sigma_j\tilde{u}<\tilde{u})}}. \tag{3.44}$$

Therefore, to bound $\log(\|\rho_j\tilde{u}\|_2/\|\rho_j\tilde{u}\|_1)$ from below, we should find an upper bound for $\tilde{\mathbb{P}}(\sigma_j\tilde{u}<\tilde{u})$. Because of (2.28), we know that

$$\tilde{\mathbb{P}}(\sigma_j\tilde{u}<\tilde{u}) \le \tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j).$$

To estimate $\tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j)$ we average in ω_1, as was done in [6]:

$$\begin{aligned}\tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j) = \tilde{\mathbb{E}}\left[1_{\tilde{\mathcal{I}}_j}\right] \quad &= \quad \tilde{\mathbb{E}}\left[\tilde{\mathbb{E}}\left[1_{\tilde{\mathcal{I}}_j}\mid\omega\right]\right] \\ &= \quad \mathbb{E}\left[\sum_{z\in[0,m-1]^d}\tilde{\mathbb{E}}\left[1_{\tilde{\mathcal{I}}_j}\mid\omega, h(\omega_1)=z\right]P_1(h(\omega_1)=z)\right].\end{aligned} \tag{3.45}$$

Observe that $(\omega,\omega_1)\in\tilde{\mathcal{I}}_j$ if and only if there is $\delta > 0$ such that

$$\pi_j(\gamma) > 0 \quad \text{for all } \ \gamma\in\tilde{M}_\delta(\omega,\omega_1),$$

which holds if and only if for some $\delta > 0$

$$\pi_{j-z}(\gamma) > 0 \quad \text{for all } \gamma \in M_\delta(t, \tau_z\omega), \quad z = h(\omega_1).$$

So, for $\mathcal{I}_j$ defined by (2.26), we have

$$\tilde{\mathbb{E}}\left[1_{\tilde{\mathcal{I}}_j} \mid \omega, h(\omega_1) = z\right] = 1_{\mathcal{I}_{j-z}}(\tau_z\omega). \tag{3.46}$$

By Lemma 3.1, we also know that $P_1(h(\omega_1) = z) \le Cm^{-d}$ for any $z \in [0, m-1]^d$. Therefore,

$$\begin{aligned}\tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j) &\le Cm^{-d}\mathbb{E}\left[\sum_z \tilde{\mathbb{E}}\left[1_{\mathcal{I}_j} \mid \omega, h(\omega_1) = z\right]\right] = Cm^{-d}\mathbb{E}\left[\sum_z 1_{\mathcal{I}_{j-z}}(\tau_z\omega)\right] \\ &= Cm^{-d}\mathbb{E}\left[\sum_z 1_{\mathcal{I}_{j-z}}(\omega)\right].\end{aligned}$$

The last equality follows from the stationarity of $\mathbb{P}$ with respect to τ_z. Now, given $\omega \in \Omega$, the sum

$$\sum_{z\in[0,m-1]^d} 1_{\mathcal{I}_{j-z}}(\omega)$$

counts the number of important cubes within the box $j - [0, m-1]^d$. These cubes are visited by all paths $\gamma \in M_\delta(\omega)$ for some $\delta > 0$ sufficiently small. Hence, $\tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j) \le Cm^{-d}\mathbb{E}[\#\Lambda_j]$ where

$$\Lambda_j = \bigcup_{n\ge 1}\left\{k \in \mathbb{Z}^d \mid \; j - k \in [0, m-1]^d, \;\; \pi_k(\gamma) > 0 \;\; \forall \gamma \in M_{1/n}(\omega)\right\}.$$

We may interpret the random variable $\#\Lambda_j$ as the number of important cubes in the box $j - [0, m-1]^d$.

Obviously we have the trivial bound $\#\Lambda_j \le O(t)$. This is because each path $\gamma \in M_\delta(\omega)$ visits at most $O(t)$ cubes total, by Corollary 2.3, so $\pi_k(\gamma) > 0$ for at most $O(t)$ indices k. Therefore,

$$\mathbb{E}[\#\Lambda_j] \le t \le (m+1)^{1/\zeta}.$$

If $d \ge 3$, we may choose $\zeta \in (1/d, 1/2)$, so that

$$\tilde{\mathbb{P}}(\sigma_j\tilde{u} < \tilde{u}) \le \tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j) \le Cm^{-d}\mathbb{E}[\#\Lambda_j] \le Cm^{-d+1/\zeta} \le Ct^{1-d\zeta}$$

with $1 - d\zeta < 0$. If $d = 2$, we need the following:

Lemma 3.9 *Let $d = 2$ and assume the non-degeneracy condition (1.6) holds. Then for each $p \in \left(\frac{\nu-1+\zeta}{\zeta\nu}, 2\right)$ there exists a constant $C > 0$ such that $\#\Lambda_j \le Cm^p$ holds with probability one, for all $j \in \mathbb{Z}^d$.*

So, for $d = 2$ we still have $\tilde{\mathbb{P}}(\sigma_j\tilde{u} < \tilde{u}) \le Cm^{-d}\mathbb{E}[\#\Lambda_j] \le Ct^{(p-2)/\zeta}$, with $(p-2)/\zeta < 0$. Therefore, returning to (3.44) we conclude that there is a constant $C > 0$ such that

$$\log\frac{\|\rho_j\tilde{u}\|_2}{\|\rho_j\tilde{u}\|_1} \ge C\log t \tag{3.47}$$

holds for all t sufficiently large. Therefore, the proof of Proposition 3.5 is reduced to a proof of Lemma 3.7, Lemma 3.8 and, in case $d = 2$, Lemma 3.9. These are proved in the next section. □

Theorem 1.1 now follows immediately from (3.34), Lemma 3.2, Lemma 3.4, and Proposition 3.5.

4 Proofs of the technical estimates

Proof of Lemma 3.7.

Observe that

$$\tilde{\mathbb{P}}(\tilde{\mathcal{I}}_j^+ \cap \tilde{G}) = \mathbb{P}(\mathcal{I}_j^+ \cap G)$$

where $G \subset \Omega$ is the event for which

$$|\gamma(t) - \gamma(0)| \geq t^{1/4}, \quad \forall\ \gamma \in M_\delta(t,\omega). \tag{4.48}$$

We will show that on the event $\mathcal{I}_j^+ \cap G$, any approximate optimizer $\gamma \in M_\delta(\omega)$ must touch a set of $O(t^{1/4})$ cubes which are almost uniformly spaced on a straight line segment of length $O(t)$ and on each of those cubes we have $V(x,\omega) = b$. Such an event can occur only with small probability.

Suppose $\omega \in \mathcal{I}_j^+ \cap G$, and let $\gamma \in M_\delta(\omega)$. Then (4.48) holds. Let $[t_1, t_2] \subset [0,t]$ be the smallest interval containing all s for which $\gamma(s) \in Q_j$. Hence, $\omega(\gamma(s)) = b$ for all $s \notin [t_1, t_2]$, since $\omega \in \mathcal{I}_j^+$. Since $\omega \in G$, we know that $|\gamma(t) - \gamma(0)| \geq t^{1/4}$, which means that either

$$|\gamma(t_1) - \gamma(0)| > \frac{t^{1/4}}{3} \qquad \text{or} \qquad |\gamma(t) - \gamma(t_2)| > \frac{t^{1/4}}{3}$$

must hold, because $\gamma(t_1), \gamma(t_2) \in \overline{Q_j}$. Let us assume that $|\gamma(t) - \gamma(t_2)| > (t^{1/4})/3$ holds; the other case is treated in a similar manner.

First, since $\omega(\gamma(s)) = b$ for all $s \in (t_2, t]$, we may assume that γ is a straight line between $\gamma(t_2)$ and $\gamma(t)$. Specifically, by redefining γ slightly, we may assume that

$$\gamma(s) = \gamma(t_2) + \frac{\gamma(t) - \gamma(t_2)}{t - t_2}(s - t_2), \quad \forall\ s \in [t_2, t],$$

for otherwise, γ would not be an optimal path. This follows from (2.13).

Next, given points $\gamma(t_2)$ and $\gamma(t)$, there is a unique pair $x_{t_2}, x_t \in \mathbb{Z}^d$ such that

$$\gamma(t_2) = x_{t_2} + y_{t_2}, \qquad \gamma(t) = x_t + y_t$$

for some $y_{t_2}, y_t \in [0,1)^d$. Therefore, if we define the linear path

$$\hat{\gamma}(s) = x_{t_2} + \frac{x_t - x_{t_2}}{t - t_2}(s - t_2), \quad s \in [t_2, t]$$

we have $|\gamma(s) - \hat{\gamma}(s)| \leq 2\sqrt{d}$ for $s \in [t_2, t]$. Therefore, for each $s \in [t_2, t]$ there must be a cube Q_ℓ such that $\text{dist}\,(\hat{\gamma}(s), Q_\ell) \leq 2\sqrt{d}$ and $\omega(Q_\ell) = b$. For $y \in \mathbb{R}^d$, let B_y denote the event that there is at least one cube Q such that $\text{dist}\,(Q, y) \leq 2\sqrt{d}$ and $\omega(Q) = b$. Then $\mathbb{P}(B_y) = 1 - \mathbb{P}(B_y^C) \leq 1 - \alpha^{C_3} < 1$, for a constant $C_3 > 0$ that depends only on the dimension d. Moreover, if $|y - z| > 5\sqrt{d}$, then B_y and B_z are independent events. Therefore, for fixed times $t_2 < t$ and a fixed pair of points $x_{t_2}, x_t \in \mathbb{Z}^d$ satisfying $|x_{t_2} - x_t| \geq (t^{1/4})/4$ we have

$$\mathbb{P}\left(\bigcap_{s \in [t_2, t]} B_{\hat{\gamma}(s)} \right) \leq (1 - \alpha^{C_3})^{C_4 t^{1/4}}$$

for some constant $C_4 > 0$ independent of ϵ.

By Lemma 2.2, we know there is a constant $R > 0$ such that $|\gamma(s) - \gamma(0)| \le tR$ for all $s \in [0,t]$. There are at most $O(t^{2d})$ possible pairs $x_{t_2}, x_t \in \mathbb{Z}^d$ satisfying $|x_{t_2} - \gamma(0)| \le Rt$ and $|x_t - \gamma(0)| \le Rt$ and $|x_{t_2} - x_t| \ge t^{1/4}$. Therefore, we conclude that

$$\mathbb{P}\left(\mathcal{I}_j^+ \cap G\right) \le O(t^{2d})(1-\alpha^{C_3})^{C_4 t^{1/4}}. \tag{4.49}$$

The last inequality immediately implies the lemma. □

Proof of Lemma 3.8

Because $\omega \mapsto \tau_{h(\omega_1)}\omega$ is measure preserving on Ω, we have

$$\tilde{\mathbb{P}}\left(\{(\omega,\omega_1) \mid \ (\sigma_j\tilde{u} - \tilde{u})^2 1_{\tilde{G}^C} > C_0 t^{1/2}\}\right) = \mathbb{P}\left(\{\omega \mid \ (\sigma_j u - u)^2 1_{G^C} > C_0 t^{1/2}\}\right)$$

where the event $G \subset \Omega$ is defined by (4.48). So, on the event G^C we know there is $\gamma \in M_\delta(\omega)$ such that

$$|\gamma(t) - \gamma(0)| < t^{1/4}. \tag{4.50}$$

Let $B_r(x)$ denote the ball of radius $r > 0$ centered at x. We may assume that there are at least two indices $j,k \in \mathbb{Z}^d \cap B_{t^{1/4}}(0)$ such that $\omega_j = a$ and $\omega_k = a$. This is because the event that $\omega_\ell = a$ for at most one of the cubes contained in $B_{t^{1/4}}(0)$ has probability less than $O(N_t\beta^{N_t})$ where $N_t \ge Ct^{1/4}$ is the number of cubes contained in $B_{t^{1/4}}(0)$.

Let $\gamma \in M_\delta(\omega)$ with $|\gamma(t) - \gamma(0)| \le t^{1/4}$. Then

$$u(t,\omega) \le g(\gamma(t)) - at + \delta.$$

Suppose $\omega_k = a$ for some $k \ne j$ and $k \in B_{t^{1/4}}(0)$. Let $x_k \in Q_k$, so that $V(x_k,\omega) = a$. Define the path $\hat\gamma$ by

$$\hat\gamma(s) = \begin{cases} \gamma(0) + s\frac{x_k - \gamma(0)}{t^{1/4}}, & s \in [0, t^{1/4}] \\ x_k, & s \in [t^{1/4}, t - t^{1/4}] \\ x_k + (s - t + t^{1/4})\frac{\gamma(t) - x_k}{t^{1/4}}, & s \in [t - t^{1/4}, t]. \end{cases} \tag{4.51}$$

Then

$$\sigma_j u(t,\omega) \ge g(\hat\gamma(t)) - \mathcal{L}(\hat\gamma,\omega) \ge g(\gamma(t)) - a(t - 2t^{1/4}) - b2t^{1/4} - 2t^{1/4}\max_{|z|\le 1} K(z).$$

Therefore,

$$u(t,\omega) - \sigma_j u(t,\omega) \le (b-a)2t^{1/4} + 2t^{1/4}\max_{|z|\le 1} K(z) + \delta.$$

Hence $(u - \sigma_j u)^2 \le C_0 t^{1/2}$ except possibly on a set of probability less than $O(N_t\beta^{N_t})$.

□

Proof of Lemma 3.9 for $d = 2$

Assuming the non-degeneracy condition (1.6), we may choose real numbers $\nu > 1$ and $\varepsilon_0 > 0$ such that $K(q) \ge |q|^\nu$ for all q that satisfy $|q| < \varepsilon_0$. Having fixed $\zeta \in (\frac{\nu-1}{2\nu-1}, \frac{1}{2})$, we see that $\frac{\nu-1+\zeta}{\zeta\nu} < 2$. So, we may choose $p \in (\frac{\nu-1+\zeta}{\zeta\nu}, 2)$.

Arguing by contradiction, we assume that $\#\Lambda_j > m^p$: there are more than m^p important cubes within the box $B_j = j - [0, m-1]^d$. Fix $\delta > 0$ small. Consider a path $\gamma \in M_\delta(t,\omega)$. Let $[t_1, t_2]$ be the smallest interval containing all s for which $\gamma(s) \in B_j$. From Lemma 2.2 it follows that γ visits only $O(|t_2 - t_1|)$ cubes between times t_1 and t_2. Thus, $|t_2 - t_1| \geq Cm^p$. Choose any one of the important cubes in B_j and let x_c denote its center point. Let us define a modified path $\hat{\gamma}$ as follows:

(i) For $s \in [0, t_1] \cup [t_2, t]$, let $\hat{\gamma}(s) = \gamma(s)$.

(ii) For $s \in [t_1, t_2]$ let

$$\hat{\gamma}(s) = \begin{cases} \gamma(t_1) + (s - t_1) \cdot \frac{x_c - \gamma(t_1)}{m}, & s \in [t_1, t_1 + m], \\ x_c, & s \in [t_1 + m, t_2 - m], \\ x_c + (s - t_2 + m) \cdot \frac{\gamma(t_2) - x_c}{m}, & s \in [t_2 - m, t_2]. \end{cases}$$

We have the bound

$$\mathcal{L}(\gamma) - \mathcal{L}(\hat{\gamma}) \geq -2m \cdot \max_{|q| \leq 1} K(q) - 2m(b - a) + \int_{t_1}^{t_2} K(\gamma'(s))\, ds. \tag{4.52}$$

We will prove that for sufficiently large m the right side of (4.52) is larger than $\delta > 0$, contradicting the fact that $\gamma \in M_\delta(t, \omega)$.

Let us denote by $J_0 \subset [t_1, t_2]$ the set of times for which $|\gamma'(s)| \geq \varepsilon_0$. Therefore, we may assume

$$\int_{J_0} K(\gamma'(s))\, ds \leq 2m \left(\max_{|q| \leq 1} K(q) + (b - a) \right) + \delta, \tag{4.53}$$

for otherwise the right side of (4.52) would be larger than δ. Let $J_1 = [t_1, t_2] \setminus J_0$; for these times $s \in J_1$ we have $|\gamma'(s)| \leq \epsilon_0$. From (4.52) we also obtain:

$$\begin{aligned} \mathcal{L}(\gamma) - \mathcal{L}(\hat{\gamma}) &\geq -2m \left(\max_{|q| \leq 1} K(q) + b - a \right) + \int_{J_1} K(\gamma'(r))\, dr \\ &\geq -2m \left(\max_{|q| \leq 1} K(q) + b - a \right) + \int_{J_1} |\gamma'(r)|^\nu\, dr \\ &\geq -2m \left(\max_{|q| \leq 1} K(q) + b - a \right) + |J_1| \cdot \left(\frac{1}{|J_1|} \int_{J_1} |\gamma'(r)|\, dr \right)^\nu. \end{aligned} \tag{4.54}$$

In the last line we applied Jensen's inequality. We will now prove that there exists a real number $\epsilon_1 > 0$ such that

$$\int_{J_1} |\gamma'(r)|\, dr \geq \epsilon_1 m^p. \tag{4.55}$$

By our assumption, the number of important cubes within B_j is more than m^p. Let us now paint all these cubes in 2^d colors so that no two adjacent cubes share the same color. By the pigeon-hole principle there are at least $m^p 2^{-d}$ important cubes having the same color. The distance between two cubes of the same color is at least 1, hence we have $\int_{t_1}^{t_2} |\gamma'(r)|\, dr \geq m^p 2^{-d}$. Therefore, since there is $C > 0$ such that $|\gamma'(s)| \leq CK(\gamma'(s))$ for all $s \in J_0$, we have

$$\begin{aligned} m^p 2^{-d} &\leq \int_{J_1} |\gamma'(r)|\, dr + \int_{J_0} |\gamma'(r)|\, dr \\ &\leq \int_{J_1} |\gamma'(r)|\, dr + C \int_{J_0} K(\gamma'(r))\, dr \\ &\leq \int_{J_1} |\gamma'(r)|\, dr + C(m + \delta). \end{aligned} \tag{4.56}$$

In the last step we have applied (4.53). This last inequality implies (4.55), since $p > 1$.

Now the inequalities (4.54) and (4.55) imply:

$$\begin{aligned}\mathcal{L}(\gamma) - \mathcal{L}(\hat{\gamma}) &\geq -2m\left(\max_{|q|\leq 1} K(q) + b - a\right) + |J_1| \cdot \left(\frac{\epsilon_1 m^p}{|J_1|}\right)^{\nu} \\ &= -2m\left(\max_{|q|\leq 1} K(q) + b - a\right) + (\epsilon_1)^{\nu} m^{p\nu} \cdot |J_1|^{1-\nu} \\ &\geq -2m\left(\max_{|q|\leq 1} K(q) + b - a\right) + (\epsilon_1)^{\nu} m^{p\nu} \cdot \left(m^{1/\zeta}\right)^{1-\nu}. \qquad (4.57)\end{aligned}$$

In the last inequality we have used $|J_1| \leq t = m^{1/\zeta}$. If we have

$$p > \frac{\nu + \zeta - 1}{\zeta \nu},$$

then $p\nu + (1-\nu)/\zeta > 1$. In this case, the right side of (4.57) is positive, and larger than δ, for t sufficiently large. Since this contradicts the approximate optimality of $\gamma \in M_\delta$, we must have $\#\Lambda_j \leq m^p$. □

5 Appendix

Proof of Lemma 3.6: The bounds in (3.37) follow from the fact that $\mathbb{P}$ is the product measure on $\Omega = \{a, b\}^{\mathbb{Z}_n^d}$, with $\mathbb{P}(\omega(j) = a) = \alpha$ and $\mathbb{P}(\omega(j) = b) = \beta$. For every nonnegative integrable ψ, (3.37) implies

$$\mathbb{E}(\psi) \;=\; \int \psi \, d\mathbb{P} \leq C'' \int \psi \, d\mathbb{P} \circ \phi_j = \int \psi(\phi_j \omega) \, d\mathbb{P} = \mathbb{E}(\psi \circ \phi_j).$$

Similarly $\mathbb{E}(\psi) \geq C' \mathbb{E}(\psi \circ \phi_j)$ for all such ψ. □

Proof of Theorem 1.2: Let us define

$$\Delta_j f(\omega) = \begin{cases} \beta(f(\phi_j \omega) - f(\omega)), & \text{if } \omega_j = a \\ \alpha(f(\phi_j \omega) - f(\omega)), & \text{if } \omega_j = b. \end{cases}$$

Then Theorem 1.2 is a slight modification of the following

Theorem 5.1 ([20], Theorem 1.5) *There is a constant $C > 0$, such that*

$$\operatorname{var}(f) \leq C \cdot \sum_{j \in J} \frac{\|\Delta_j f\|_2^2}{1 + \log \frac{\|\Delta_j f\|_2}{\|\Delta_j f\|_1}}. \qquad (5.58)$$

holds for all $f \in L^2(\Omega_J)$.

To derive Theorem 1.2 from this, we start with elementary observation

$$C' |\rho_j f(\omega)| \leq |\Delta_j f(\omega)| \leq C'' |\rho_j f(\omega)|$$

for $C' = \min\{2\alpha, 2\beta\}$ and $C'' = \max\{2\alpha, 2\beta\}$. Let $\kappa = \log(C''/C') \geq 0$. If $\log \frac{\|\rho_j f\|_2}{\|\rho_j f\|_1} \geq 2\kappa$, then

$$\log \frac{\|\Delta_j f\|_2}{\|\Delta_j f\|_1} \geq \log \frac{C'}{C''} + \log \frac{\|\rho_j f\|_2}{\|\rho_j f\|_1} \geq \frac{1}{2} \log \frac{\|\rho_j f\|_2}{\|\rho_j f\|_1}.$$

Consequently, Theorem 5.1 implies

$$\operatorname{var}(f) \le C \cdot \sum_{j \in J} \frac{\|\Delta_j f\|_2^2}{1 + \log \frac{\|\Delta_j f\|_2}{\|\Delta_j f\|_1}} \le 2(C'')^2 C \sum_{j \in J} \frac{\|\rho_j f\|_2^2}{1 + \log \frac{\|\rho_j f\|_2}{\|\rho_j f\|_1}}.$$

On the other hand, if $\log \frac{\|\rho_j f\|_2}{\|\rho_j f\|_1} \in [0, 2\kappa)$, then Theorem 5.1 implies

$$\operatorname{var}(f) \le C \cdot \sum_{j \in J} \|\Delta_j f\|_2^2 \le (1 + 2\kappa) 2(C'')^2 C \sum_{j \in J} \frac{\|\rho_j f\|_2^2}{1 + \log \frac{\|\rho_j f\|_2}{\|\rho_j f\|_1}}.$$

□

References

[1] K. Alexander and N. Zygouras, *Subgaussian concentration and rates of convergence in directed polymers*, preprint 2012. arxiv:1204.1819.

[2] S. Armstrong, P. Cardaliaguet, P.E. Souganidis, *Error estimates and convergence rates for the stochastic homogenization of Hamilton-Jacobi equations*, in preparation 2012.

[3] S. Armstrong and P.E. Souganidis, *Stochastic homogenization of level-set convex Hamilton-Jacobi equations*, preprint 2012, arXiv:1203.6303.

[4] M. Balázs, J. Quastel, T. Seppäläinen, *Fluctuation exponent of the KPZ/stochastic Burgers equation*, J. Amer. Math. Soc. **24** (2011), pp. 683-708.

[5] M. Bardi and I. Capuzzo-Dolcetta, *Optimal Control and Viscosity Solutions of Hamilton-Jacobi-Bellman Equations*, Birkhäuser, 1997.

[6] I. Benjamini, G. Kalai, O. Schramm, *First passage percolation has sublinear distance variance*, Ann. Probab. **31** (2003), pp. 1970-1978.

[7] M. Benaïm and R. Rossignol, *Exponential concentration for first passage percolation throuh modified Poincaré inequalities*, Ann. Inst. Henri Poincaré - Prob. Stat. **44** (2008), pp. 544-573.

[8] I. Corwin, *The Kardar-Parisi-Zhang equation and universality class*, Random Matrices: Theory Appl. **1** (2012), 1130001.

[9] K. Johansson, *Shape fluctuations and random matrices.* Comm. Math. Phys. **209** (2000) pp. 437-476.

[10] E. Kosygina and S.R.S. Varadhan, *Homogenization of Hamilton-Jacobi-Bellman equations with respect to time-space shifts in a stationary ergodic medium*, Comm. Pure Appl. Math, **61** (2008) pp. 816-847.

[11] E. Kosygina, F. Rezakhanlou, S.R.S. Varadhan, *Stochastic homogenization of Hamilton-Jacobi-Bellman equations*, Comm. Pure Appl. Math **59** (2006), pp. 1489-1521.

[12] J. Krug and H. Spohn. *Kinetic roughening of growing surfaces*, in Solids Far From Equilibrium, C. Godrèche, ed. Cambridge University Press, 1992.

[13] P.-L. Lions and P.E. Souganidis, *Homogenization of "viscous" Hamilton-Jacobi equations in stationary ergodic media*, Comm. PDE **30** (2005), pp. 335375.

[14] P.-L. Lions and P.E. Souganidis, *Stochastic homogenization of Hamilton-Jacobi and "viscous"-Hamilton-Jacobi equations with convex nonlinearities– revisited*, Comm. Math. Sci. **8** (2010). pp. 672-637.

[15] F. Rezakhanlou, *Central limit theorem for stochastic Hamilton-Jacobi equations*, Comm. Math. Phys. **211** (2000), pp. 413–438.

[16] F. Rezakhanlou and J. E. Tarver. *Homogenization for stochastic Hamilton-Jacobi equations*, Arch. Ration. Mech. Anal., **151** (2000), pp. 277-309.

[17] T. Seppäläinen, *Scaling for a one-dimensioal directed polymer with boundary conditions*, Ann. Probab. **40** (2012), pp. 19-73.

[18] R. Schwab, *Stochastic homogenization of Hamilton-Jacobi equations in stationary ergodic spatio-temporal media*, Indiana Univ. Math. J. **58** (2009), pp. 527-581.

[19] P.E. Souganidis, *Stochastic homogenization of Hamilton-Jacobi equations and some applications*, Asymptotic Analysis, **20** (1999), pp. 1-11.

[20] M. Talagrand, *On Russo's approximate zero-one law.*, Ann. of Probab. **22** (1994), pp. 1576-1587.