Ann. Inst. H. Poincaré
Anal. Non Linéaire (Online first)
DOI 10.4171/AIHPC/41
© 2022 Association Publications de l'Institut Henri Poincaré
Published by EMS Press

Hölder regularity for stochastic processes with bounded and measurable increments

Ángel Arroyo, Pablo Blanc, and Mikko Parviainen

Abstract. We obtain an asymptotic Hölder estimate for expectations of a quite general class of discrete stochastic processes. Such expectations can also be described as solutions to a dynamic programming principle or as solutions to discretized PDEs. The result, which is also generalized to functions satisfying Pucci-type inequalities for discrete extremal operators, is a counterpart to the Krylov–Safonov regularity result in PDEs. However, the discrete step size ε has some crucial effects compared to the PDE setting. The proof combines analytic and probabilistic arguments.

1. Introduction

The celebrated Krylov–Safonov [15] Hölder estimate is one of the key results in the theory of nondivergence-form elliptic partial differential equations with bounded and measurable coefficients. The result, in addition to being important in its own right, also gives a flexible tool in higher regularity and existence theory due to its very general assumptions on the coefficients.

In this paper we study regularity of expectation of a quite general class of discrete stochastic processes or, equivalently, functions satisfying the dynamic programming principle (DPP)

$$u(x) = \alpha \int_{\mathbb{R}^N} u(x + \varepsilon z)\, d\nu_x(z) + \frac{\beta}{|B_\varepsilon|} \int_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 f(x), \tag{1.1}$$

where f is a Borel measurable bounded function, ν_x is a symmetric probability measure for each x with support in B_Λ, $\Lambda \geq 1$, $\varepsilon > 0$, and $\alpha + \beta = 1$, $\alpha \geq 0$, $\beta > 0$. From a stochastic point of view, our processes are generalizations of the random walk where the next step in the process is taken according to a probability measure that is a combination of ν_x and the uniform distribution on $B_\varepsilon(x)$ as described by the DPP (more details are given in Sections 2.1 and 2.3).

It is important to notice that ν_x can vary quite freely from point to point. Under continuity or other assumptions not satisfied in our case, related results have been studied for example in [7, 18, 21].

2020 Mathematics Subject Classification. Primary 35B65; Secondary 35J15, 60H30, 60J10, 91A50.
Keywords. Dynamic programming principle, local Hölder estimates, stochastic process, equations in nondivergence form, p-harmonious, p-Laplace, tug-of-war games.

The first of the two main results in this article is a Hölder estimate in the discrete setup without any further continuity assumption on the measures ν_x.

Theorem 1.1. *There exists $\varepsilon_0 > 0$ such that if u is a function satisfying (1.1) in B_{2R} with $\varepsilon < \varepsilon_0 R$, then*

$$|u(x) - u(z)| \leq \frac{C}{R^\gamma} \Big(\sup_{B_{2R}} |u| + R^2 \| f \|_\infty \Big) (|x - z|^\gamma + \varepsilon^\gamma)$$

for every $x, z \in B_R$, where $C > 0$ and $\gamma > 0$ are constants independent of ε.

The role of the discrete processes we study can be compared to the role of linear uniformly elliptic partial differential equations with bounded and measurable coefficients in the theory of PDEs.

The regularity techniques in PDEs (see [9, 15]) or in the nonlocal setting (see [7, 13]) utilize, heuristically speaking, the fact that there is information available in all scales. We also refer to [11, 31] for similar results regarding nonlocal operators with nonsymmetric kernels. For a discrete process, the step size sets a natural limit for the scale, and this limitation has some crucial effects. Indeed, the value can even be discontinuous, and our estimates are asymptotic. Such estimates suffice in many applications, for example in passing to the limit with solutions to discretized PDEs or stochastic processes. Similar results have been obtained on a grid in the context of difference equations with random coefficients in [17]. See also [19], where regularity estimates for difference equations arising from random walks are obtained using probabilistic techniques.

The proof uses a stochastic approach akin to the original proof of Krylov and Safonov in [15] with some crucial differences. The first observation, as suggested above, is that the function u in (1.1) can be presented as an expectation. The key estimate is then Theorem 5.7 stating that we can reach any set of positive measure with a positive probability before exiting a bigger cube. With this result at our disposal, the De Giorgi oscillation estimate, Lemma 5.8, follows in a straightforward manner. Indeed, we can reach a level set with a positive probability and use this in estimating the oscillation. The Hölder estimate, Theorem 1.1, then follows by the De Giorgi oscillation lemma after a finite iteration.

The proof of Theorem 5.7 is nonstandard. In the proof we would like to construct a set of cubes which is large enough and such that the set we want to reach has a high enough density in the cubes. Both conditions, however, cannot always be satisfied simultaneously in our setting. As suggested above, both the PDE and nonlocal techniques utilize the information in all scales. Concretely, a rescaling argument is used in those proofs in arbitrary small cubes. In contrast, in our case the step size ε determines the limit for the scale. If we simply drop all the cubes to smaller than of scale ε in the usual Calderón–Zygmund decomposition, we have no control on the size of the error. Therefore, the cubes of scale ε need to be taken into account separately both in the decomposition lemma, Lemma 5.4, and in the proof of the key intermediate result, Theorem 5.7.

The proof of Theorem 5.7 is based on the ε-version of the Alexandrov–Bakelman–Pucci (ABP) estimate with bounded and measurable right-hand side, Theorem 4.7.

However, the classical proof of the ABP estimate using the change of variables formula for integrals to obtain a quantity that can be estimated by the PDE does not seem directly applicable here. Instead, we adapt the nonlocal approach of Caffarelli and Silvestre [7]. A second remark is that we directly apply the ABP estimate with a discontinuous right-hand side, which is chosen to be a characteristic function of a level set. In this case, the standard ABP estimate having the L^N-norm on the right-hand side is false (Example 4.6), and therefore the statement of Theorem 4.7 is weaker, but sufficient for our purposes.

Our study is partly motivated by the aim of developing stochastic methods in connection with the p-Laplace equation and other nonlinear PDEs; see Example 2.4 and Remark 7.6. The Hölder estimate, Theorem 1.1, can be generalized to functions merely satisfying the Pucci-type inequalities

$$u(x) \le \alpha \sup_{z\in B_\Lambda} \frac{u(x+\varepsilon z)+u(x-\varepsilon z)}{2} + \frac{\beta}{|B_\varepsilon|}\int_{B_\varepsilon(x)} u(y)\,dy + \varepsilon^2 |f(x)|, \tag{1.2}$$

$$u(x) \ge \alpha \inf_{z\in B_\Lambda} \frac{u(x+\varepsilon z)+u(x-\varepsilon z)}{2} + \frac{\beta}{|B_\varepsilon|}\int_{B_\varepsilon(x)} u(y)\,dy - \varepsilon^2 |f(x)|, \tag{1.3}$$

later modified and shortened to the forms

$$\mathcal{L}_\varepsilon^+ u \ge -|f|, \quad \mathcal{L}_\varepsilon^- u \le |f|.$$

This is our second main result.

Theorem 1.2. *There exists $\varepsilon_0 > 0$ such that if u is a function satisfying (1.2) and (1.3) for every $x \in B_{2R}$ with $\varepsilon < \varepsilon_0 R$, then*

$$|u(x)-u(z)| \le \frac{C}{R^\gamma}\Big(\sup_{B_{2R}}|u| + R^2\|f\|_\infty\Big)(|x-z|^\gamma + \varepsilon^\gamma)$$

for every $x, z \in B_R$, where $C > 0$ and $\gamma > 0$ are constants independent of ε.

We refer the reader to Section 7 and in particular to Theorem 7.3 for a more detailed description.

2. Preliminaries

As above, let $\Lambda \ge 1$, $\varepsilon > 0$, $\beta \in (0,1]$ and $\alpha = 1-\beta$. Every constant may depend on Λ, α, β and the dimension N. If a constant depends on other parameters we denote it.

Throughout this paper $\Omega \subset \mathbb{R}^N$ denotes a bounded domain and, further, $B_r(x) = \{y \in \mathbb{R}^N : |x-y| < r\}$ as well as $B_r = B_r(0)$. We construct an extended domain containing all balls $B_{\Lambda\varepsilon}(x)$ with $x \in \Omega$ as follows:

$$\widetilde{\Omega}_{\Lambda\varepsilon} := \big\{x \in \mathbb{R}^n : \operatorname{dist}(x,\Omega) < \Lambda\varepsilon\big\}.$$

We follow the convention

$$\int u(x)\,dx = \int_{\mathbb{R}^N} u(x)\,dx \quad \text{and} \quad \fint_A u(x)\,dx = \frac{1}{|A|}\int_A u(x)\,dx.$$

Further,

$$\|f\|_{L^N(\Omega)} = \left(\int_\Omega |f(x)|^N\,dx\right)^{1/N}$$

and

$$\|f\|_{L^\infty(\Omega)} = \sup_\Omega |f|.$$

When no confusion arises we simply denote these as $\|\cdot\|_N$ and $\|\cdot\|_\infty$, respectively.

For each $x = (x_1, \dots, x_n) \in \mathbb{R}^N$ and $r > 0$, we define $Q_r(x)$ as the open cube of side length r and center x with faces parallel to the coordinate hyperplanes. More precisely,

$$Q_r(x) := \{y \in \mathbb{R}^N : |y_i - x_i| < r/2,\ i = 1, \dots, n\}.$$

In addition, if $Q = Q_r(x)$ and $\ell > 0$, for simplicity we denote $\ell Q = Q_{\ell r}(x)$.

2.1. Dynamic programming principle and difference operators

We consider $\mathcal{M}(B_\Lambda)$ the set of symmetric unit Radon measures with support in B_Λ and $\nu\colon \mathbb{R}^N \to \mathcal{M}(B_\Lambda)$ such that

$$x \mapsto \int u(x+z)\,d\nu_x(z) \tag{2.1}$$

defines a Borel measurable function for every $u\colon \mathbb{R}^N \to \mathbb{R}$ Borel measurable. Then for each $x \in \mathbb{R}^N$ we have a measure ν_x with support in B_Λ such that

$$\nu_x(E) = \nu_x(-E)$$

for every measurable set $E \subset \mathbb{R}^N$.

It is worth remarking that hypothesis (2.1) on Borel measurability holds, for example, when the ν_x are the pushforward of a given probability measure μ in $\mathbb{R}^N$. More precisely, if there exists a Borel measurable function $h\colon \mathbb{R}^N \times \mathbb{R}^N \to B_\Lambda$ such that

$$\nu_x = h(x,\cdot)\#\mu$$

for each x, then

$$\begin{aligned} v(x) &= \int u(x+z)\,d\nu_x(z) \\ &= \int u(x+h(x,y))\,d\mu(y) \end{aligned}$$

is measurable by Fubini's theorem.

For each $\varepsilon > 0$ we consider a generalized random walk starting at $x_0 \in \Omega$. Given the value of x_k, the next position of the process x_{k+1} is determine as follows. A biased coin is tossed. If we get heads (probability α), a vector z is chosen according to ν_{x_k} and we have $x_{k+1} = x_k + \varepsilon z$. If we get tails (probability β), x_{k+1} is distributed uniformly in the ball $B_\varepsilon(x_k)$. More details are given in Section 2.3. Denote by τ the exit time from the domain, that is,

$$\tau = \min\{n \in \mathbb{N} : x_n \notin \Omega\}.$$

Given a Borel measurable bounded function $g\colon \mathbb{R}^N \setminus \Omega \to \mathbb{R}$, we can define

$$u(x_0) := \mathbb{E}^{x_0}[g(x_\tau)],$$

where $\mathbb{E}^{x_0}$ stands for the expectation with respect to the process. We will prove in Section 3 that $u\colon \mathbb{R}^N \to \mathbb{R}$ satisfies the homogenous DPP given by

$$u(x) = \alpha \int u(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} u(y)\, dy$$

for $x \in \Omega$, and $u(x) = g(x)$ for $x \notin \Omega$. Moreover, u is the unique bounded function that satisfies the DPP.

Moreover, given the running payoff $f\colon \Omega \to \mathbb{R}$, a Borel measurable bounded function, we can define

$$u(x) := \mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i) + g(X_\tau)\right].$$

In Section 3 we prove that u is the unique bounded function that satisfies the DPP

$$u(x) = \alpha \int u(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 f(x) \tag{2.2}$$

for $x \in \Omega$ and $u(x) = g(x)$ for $x \notin \Omega$. For clarity, let us emphasize that (2.2) is what we call the DPP in this paper. This also motivates the following definitions.

Definition 2.1. We say that a bounded Borel measurable function u is a subsolution to the DPP if it satisfies

$$u(x) \le \alpha \int u(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 f(x)$$

in Ω. Analogously, we say that u is a supersolution if the reverse inequality holds. If the equality holds, we say that it is a solution to the DPP.

If we rearrange the terms in the DPP, we may alternatively use notation that is closer to difference methods.

Definition 2.2. Given a Borel measurable bounded function $u\colon \mathbb{R}^N \to \mathbb{R}$, we define $\mathcal{L}_\varepsilon u\colon \mathbb{R}^N \to \mathbb{R}$ as

$$\mathcal{L}_\varepsilon u(x) = \frac{1}{\varepsilon^2}\left(\alpha \int u(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} u(y)\, dy - u(x)\right).$$

With this notation, u is a subsolution (supersolution) if and only if $\mathcal{L}_\varepsilon u + f \ge 0\ (\le 0)$.

By defining $\delta u(x, y) := u(x + y) + u(x - y) - 2u(x)$ and recalling the symmetry condition on ν_x we can rewrite

$$\mathcal{L}_\varepsilon u(x) = \frac{1}{2\varepsilon^2}\left(\alpha \int \delta u(x, \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_1} \delta u(x, \varepsilon y)\, dy\right). \tag{2.3}$$

Our theorems actually hold for functions merely satisfying Pucci-type inequalities. For expositional clarity we leave this to Section 7.

2.2. Examples

In this section we present some recent examples and applications. The list is by no means exhaustive, and further examples could be obtained by discretizing partial differential operators with bounded and measurable coefficients.

Example 2.3 (Convergence to the solution of the PDE). Let $\phi \in C^2(\Omega)$. We can use the second-order Taylor expansion of ϕ to obtain an asymptotic expansion for $\mathcal{L}_\varepsilon \phi(x)$. Indeed, observe that

$$\delta\phi(x, \varepsilon y) = \varepsilon^2 \operatorname{Tr}\{D^2\phi(x) \cdot y \otimes y\} + o(\varepsilon^2)$$

holds as $\varepsilon \to 0$ for every $y \in B_\Lambda$, where $a \otimes b$ stands for the tensor product of vectors $a, b \in \mathbb{R}^n$, that is, the matrix with entries $(a_i b_j)_{ij}$. Hence, by the linearity of the trace,

$$\begin{aligned}\mathcal{L}_\varepsilon \phi(x) &= \frac{1}{2} \operatorname{Tr}\left\{D^2\phi(x) \cdot \left(\alpha \int z \otimes z\, d\nu_x(z) + \beta \fint_{B_1} y \otimes y\, dy\right)\right\} + o(\varepsilon^0) \\ &= \frac{1}{2} \operatorname{Tr}\left\{D^2\phi(x) \cdot \left(\alpha \int z \otimes z\, d\nu_x(z) + \frac{\beta}{N+2} I\right)\right\} + o(\varepsilon^0).\end{aligned}$$

On the other hand, since every measure $\nu_x \in \mathcal{M}(B_\Lambda)$ defines a matrix

$$A(x) := \frac{\alpha}{2} \int z \otimes z\, d\nu_x(z) + \frac{\beta}{2(N+2)} I,$$

we get

$$\lim_{\varepsilon \to 0} \mathcal{L}_\varepsilon \phi(x) = \operatorname{Tr}\{D^2\phi(x) \cdot A(x)\},$$

which is a linear second-order partial differential operator. Furthermore, for $\beta \in (0, 1]$, the operator is uniformly elliptic: given $\xi \in \mathbb{R}^N \setminus \{0\}$,

$$\langle A(x) \cdot \xi, \xi \rangle = \frac{\alpha}{2} \int |\langle z, \xi \rangle|^2\, d\nu_x(z) + \frac{\beta}{2(N+2)} |\xi|^2,$$

and estimating the integral we have

$$\frac{\beta}{2(N+2)} \le \frac{\langle A(x) \cdot \xi, \xi \rangle}{|\xi|^2} \le \frac{\alpha\Lambda^2}{2} + \frac{\beta}{2(N+2)}.$$

It also holds, using Theorem 1.1 (cf. [26, Theorem 4.9]), that under suitable regularity assumptions, the solutions $u_\varepsilon := u$ to the DPP converge to a viscosity solution $v \in C(\Omega)$ of

$$\operatorname{Tr}\{D^2 v(x) \cdot A(x)\} = f(x),$$

as $\varepsilon \to 0$.

Similar convergence results also hold in the following examples.

Example 2.4 (p-Laplacian). In [22, Section 3.2], the following process that is covered by Theorem 1.1 was considered. Let u be a p-harmonic (with $2 \le p < \infty$) function whose gradient vanishes at most at a finite number of points. When at $x \in \Omega$ such that $\nabla u(x) \neq 0$, we define a probability measure

$$\mu_{x,1} = \beta \mathcal{L}_{B_\varepsilon(x)} + \frac{1-\beta}{2}\big(\delta_{x+\varepsilon \frac{\nabla u(x)}{|\nabla u(x)|}} + \delta_{x-\varepsilon \frac{\nabla u(x)}{|\nabla u(x)|}}\big),$$

where $\mathcal{L}_{B_\varepsilon(x)}$ denotes the uniform probability distribution in $B_\varepsilon(x) \subset \mathbb{R}^N$ and δ_x the Dirac measure at x. Then we choose the next point according to the probability measure

$$\mu_x = \begin{cases} \mu_{x,1} & \text{if } |\nabla u(x)| > 0 \text{ and} \\ \mathcal{L}_{B_\varepsilon(x)} & \text{if } |\nabla u(x)| = 0. \end{cases}$$

There is a classical well-known connection between Brownian motion and the Laplace equation. This example is related to so-called tug-of-war games introduced in Peres, Schramm, Sheffield and Wilson [27] in connection with the infinity Laplace operator. Similarly, a connection exists between the p-Laplacian, $1 < p < \infty$, and different variants of tug-of-war games with noise [20, 26, 28].

There are several regularity methods devised for tug-of-war games with noise: the above papers contain a global approach, and a local approach is developed in [24] as well as [23]. However, none of these methods seem to directly apply to the present situation. Moreover, later we prove Theorem 7.3 which applies to solutions of the DPP associated to tug-of-war games with noise and the p-Laplacian; see Remark 7.6.

Example 2.5 (Ellipsoid process). A particular case of the stochastic process considered in this paper is the so-called ellipsoid process (see [2]). This process arises when ν_x is the uniform probability measure on $E_x \setminus B_1$, where E_x denotes an ellipsoid centered at the origin such that $B_1 \subset E_x \subset B_\Lambda$. Then

$$\alpha \nu_x(A) + \beta \frac{|A \cap B_1|}{|B_1|} = \frac{|A \cap E_x|}{|E_x|}$$

for every measurable set $A \subset \mathbb{R}^N$, $\alpha = \frac{|E_x \setminus B_1|}{|E_x|}$ and $\beta = \frac{|B_1|}{|E_x|}$. Hence, replacing this in (2.2) with $f = 0$ we get that the expectation related to the ellipsoid process satisfies the DPP

$$u(x) = \fint_{E_x} u(x + \varepsilon y)\, dy.$$

An asymptotic Hölder estimate was obtained in [2] under certain assumptions on the ellipticity ratio of the ellipsoids. Now Theorem 1.1 implies the Hölder estimate for u without any additional assumptions and thus improves the result in [2]. Such a mean value property over ellipsoids has been studied by Pucci and Talenti in connection with smooth solutions to PDEs in [29].

2.3. Stochastic process

Next we define the stochastic process related to the DPP (2.2). Let $x_0 \in \Omega$ be the initial position of the process. We equip $\mathbb{R}^N$ with the natural topology, and the σ-algebra $\mathcal{B}$ of the Borel measurable sets. We consider along with the positions of the process the results of the coin tosses, so our process is defined in the product space

$$H^\infty = \{x_0\} \times (\{0,1\} \times \mathbb{R}^N) \times (\{0,1\} \times \mathbb{R}^N) \times \cdots.$$

For $\omega = (x_0, (c_1, x_1), (c_2, x_2) \dots) \in H^\infty$, we define the result of the kth toss $\mathcal{C}_k(\omega) = c_k$ and the coordinate processes $X_k(\omega) = x_k$. If $\mathcal{C}_{k+1} = 0$ (probability α), the next position of the process X_{k+1} is distributed according to ν_{x_k}. And for $\mathcal{C}_{k+1} = 1$ (probability β), X_{k+1} is uniformly distributed in $B_\varepsilon(x_k)$. That is, we have the following transition probabilities:

$$\pi(x_0, (c_1, x_1), \dots, (c_k, x_k), \{c\} \times A) = \begin{cases} \alpha \nu_{x_k}\left(\dfrac{A - x_k}{\varepsilon}\right) & \text{if } c = 0, \\ \beta \dfrac{|A \cap B_\varepsilon(x_k)|}{|B_\varepsilon|} & \text{if } c = 1, \end{cases}$$

where $A \in \mathcal{B}$.

Let $\{\mathcal{F}_k\}_k$ denote the filtration of σ-algebras $\mathcal{F}_0 \subset \mathcal{F}_1 \subset \cdots$ defined as follows: $\mathcal{F}_k$ is the product σ-algebra generated by cylinder sets of the form

$$\{x_0\} \times A_1 \times \cdots \times A_k \times \mathbb{R}^N \times \mathbb{R}^N \times \cdots$$

with $A_i \in \mathcal{P}(\{0,1\}) \times \mathcal{B}$. We have that $\mathcal{C}_k$ and X_k are $\mathcal{F}_k$-measurable random variables. By the Kolmogorov extension theorem, the transition probabilities determine a unique probability measure $\mathbb{P}^{x_0}$ in H^∞ relative to the σ-algebra $\mathcal{F}_\infty = \sigma(\bigcup_k \mathcal{F}_k)$. We denote by $\mathbb{E}^{x_0}$ the corresponding expectation.

We consider τ the exit time from the domain, that is $\tau = \min\{n \in \mathbb{N} : X_n \notin \Omega\}$. We define T_A as the hitting time for A and τ_A the exit time, that is,

$$T_A = \min\{k \in \mathbb{N} : X_k \in A\} \quad \text{and} \quad \tau_A = \min\{k \in \mathbb{N} : X_k \notin A\}.$$

2.4. Stochastic estimates

In this section we establish some estimates related to τ and other stochastic results. We will prove that $\mathbb{E}^{x_0}[\tau]$ is of order $1/\varepsilon^2$. Moreover, we will prove that the second moment of $\varepsilon^2 \tau$ is bounded. We start with a rough estimate needed as a first step.

Lemma 2.6. *The process leaves the domain almost surely and moreover*

$$\mathbb{E}^{x_0}[\tau] < +\infty$$

for every $x_0 \in \Omega$.

Proof. We fix $x_0 \in \Omega$. We will prove that there exist $\lambda < 1$ and $n \in \mathbb{N}$ such that

$$\mathbb{P}^{x_0}(\tau > n + k | \tau > k) \leq \lambda \tag{2.4}$$

for every $k \in \mathbb{N}$ and $x \in \Omega$. Then inductively $\mathbb{P}^{x_0}(\tau > nk) \leq \lambda^k$ and we get

$$\mathbb{E}^{x_0}[\tau] = \sum_{i=0}^{\infty} \mathbb{P}^{x_0}(\tau > i) = \sum_{k=0}^{\infty} \sum_{j=0}^{n-1} \mathbb{P}^{x_0}(\tau > j + nk) \leq n \sum_{k=0}^{\infty} \mathbb{P}^{x_0}(\tau > nk) \leq n \sum_{k=0}^{\infty} \lambda^k,$$

which is finite.

Thus, it remains to prove (2.4). We choose n such that $n\frac{\varepsilon}{2} > \operatorname{diam} \Omega$. We consider the event E of n steps after the kth one to be uniformly distributed, and where the first coordinate increases at least $\frac{\varepsilon}{2}$. That is,

$$E = \{c_{k+1} = \cdots = c_{k+n} = 1 \text{ and } \pi_1(x_i - x_{i-1}) > \varepsilon/2 \text{ for } i = k+1, \ldots, k+n\},$$

where π_1 denotes the projection to the first coordinate. Observe that

$$\mathbb{P}^{x_0}(E) = \left(\beta \frac{|\{x : \pi_1(x) > \varepsilon/2\} \cap B_\varepsilon|}{|B_\varepsilon|}\right)^n > 0$$

is independent of k.

Assuming E we have $|x_{k+n} - x_k| \geq n\frac{\varepsilon}{2}$, hence since $n\frac{\varepsilon}{2} > \operatorname{diam} \Omega$, it must be the case that $\tau \leq n + k$. Therefore (2.4) holds for $\lambda = 1 - \mathbb{P}^{x_0}(E) < 1$. ∎

Now we construct two sequences of random variables. They will allow us to obtain bounds for the growth of the expected value of the square of the distance from the starting point x_0.

Lemma 2.7. *The sequence of random variables* $\{|X_k - x_0|^2 - Ck\varepsilon^2\}_k$ *is a supermartingale for* $C = \alpha\Lambda^2 + \beta \fint_{B_1} |x|^2\, dx$.

Proof. Observe that

$$\mathbb{E}^{x_0}[|X_{k+1} - x_0|^2 | \mathcal{F}_k](\omega) = \alpha \int |x_k + \varepsilon z - x_0|^2\, d\nu_{x_k}(z) + \beta \fint_{B_\varepsilon(x_k)} |x - x_0|^2\, dx.$$

By the symmetry of ν_{x_k} and the ball B_ε we can write

$$\begin{aligned}\mathbb{E}^{x_0}[|X_{k+1} - x_0|^2 | \mathcal{F}_k](\omega) = \alpha \int \frac{|x_k + \varepsilon z - x_0|^2 + |x_k - \varepsilon z - x_0|^2}{2}\, d\nu_{x_k}(z) \\ + \beta \fint_{B_\varepsilon} \frac{|x_k + x - x_0|^2 + |x_k - x - x_0|^2}{2}\, dx.\end{aligned}$$

Employing the parallelogram law we get

$$\begin{aligned}\mathbb{E}^{x_0}[|X_{k+1}-x_0|^2|\mathcal{F}_k](\omega) &= \alpha \int |x_k-x_0|^2+|\varepsilon z|^2\,dv_{x_k}(z)\\ &\quad+\beta \fint_{B_\varepsilon} |x_k-x_0|^2+|x|^2\,dx\\ &\leq \alpha(|x_k-x_0|^2+\varepsilon^2\Lambda^2)+\beta\left(|x_k-x_0|^2+\varepsilon^2\fint_{B_1}|x|^2\,dx\right)\\ &\leq |x_k-x_0|^2+\varepsilon^2\left(\alpha\Lambda^2+\beta\fint_{B_1}|x|^2\,dx\right),\end{aligned}$$

where we have used that v_{x_k} is supported in B_Λ.

Therefore, for $C=\alpha\Lambda^2+\beta\fint_{B_1}|x|^2\,dx$ we have

$$\begin{aligned}\mathbb{E}^{x_0}[|X_{k+1}-x_0|^2-C(k+1)\varepsilon^2|\mathcal{F}_k] &\leq |X_k-x_0|^2+C\varepsilon^2-C(k+1)\varepsilon^2\\ &= |X_k-x_0|^2-Ck\varepsilon^2,\end{aligned}$$

as we wanted to show. ∎

Lemma 2.8. *The sequence of random variables $\{|X_k-x_0|^2-Ck\varepsilon^2\}_k$ is a submartingale for $C=\beta\fint_{B_1}|x|^2\,dx$.*

Proof. As in the previous lemma we have

$$\begin{aligned}\mathbb{E}^{x_0}[|X_{k+1}-x_0|^2|\mathcal{F}_k](\omega) &= \alpha\int |x_k-x_0|^2+|\varepsilon z|^2\,dv_{x_k}(z)\\ &\quad+\beta\fint_{B_\varepsilon}|x_k-x_0|^2+|x|^2\,dx.\end{aligned}$$

By dropping the $|\varepsilon z|^2$ term we get

$$\begin{aligned}\mathbb{E}^{x_0}[|X_{k+1}-x_0|^2|\mathcal{F}_k](\omega) &\geq \alpha|x_k-x_0|^2+\beta\left(|x_k-x_0|^2+\fint_{B_\varepsilon}|x|^2\,dx\right)\\ &= |x_k-x_0|^2+\beta\varepsilon^2\fint_{B_1}|x|^2\,dx.\end{aligned}$$

Therefore, for $C=\beta\fint_{B_1}|x|^2\,dx$ we have

$$\begin{aligned}\mathbb{E}^{x_0}[|X_{k+1}-x_0|^2-C(k+1)\varepsilon^2|\mathcal{F}_k] &\geq |X_k-x_0|^2+C\varepsilon^2-C(k+1)\varepsilon^2\\ &= |X_k-x_0|^2-Ck\varepsilon^2,\end{aligned}$$

as claimed. ∎

We are ready to prove that $\mathbb{E}^{x_0}[\tau]$ is of order $1/\varepsilon^2$.

Lemma 2.9. *There exist $C_1, C_2 > 0$ such that*

$$C_1 \operatorname{dist}(x_0, \partial\Omega)^2 \leq \mathbb{E}^{x_0}[\varepsilon^2\tau] \leq C_2(\operatorname{diam}\Omega + 1)^2$$

for $\varepsilon < \frac{1}{\Lambda}$.

Proof. We consider the supermartingale $M_k = |X_k - x_0|^2 - Ck\varepsilon^2$ from Lemma 2.7. Since the increments of M_k are bounded and $\mathbb{E}^{x_0}[\tau] < +\infty$ we can apply the optional stopping theorem. We obtain $\mathbb{E}^{x_0}[M_\tau] \leq 0$ and hence

$$\operatorname{dist}(x_0, \partial\Omega)^2 \leq \mathbb{E}^{x_0}[|X_\tau - x_0|^2] \leq C\varepsilon^2\mathbb{E}^{x_0}[\tau]$$

as desired. The other inequality can be obtain by considering the submartingale from Lemma 2.8. In fact, we get

$$C\varepsilon^2\mathbb{E}^{x_0}[\tau] \leq \mathbb{E}^{x_0}[|X_\tau - x_0|^2] \leq (\operatorname{diam}\Omega + \Lambda\varepsilon)^2,$$

where we have used that x_τ is at a distance of at most $\Lambda\varepsilon$ from $x_{\tau-1} \in \Omega$ and therefore $|x_\tau - x_0| \leq \operatorname{diam}\Omega + \Lambda\varepsilon$. ∎

Now we obtain an estimate for the random variable $\varepsilon^2\tau$ necessary to bound its second moment in the subsequent corollary. We follow [4, Lemma 3.6]. The key point here is that the process is memoryless.

Lemma 2.10. *There exist $C > 0$ and $\mu = \mu(\operatorname{diam}\Omega) > 0$ such that*

$$\mathbb{P}^{x_0}(\varepsilon^2\tau \geq t) \leq Ce^{-\mu t}$$

for $\varepsilon < \frac{1}{\Lambda}$.

Proof. By Lemma 2.9 there exists $\widetilde{C} = \widetilde{C}(\operatorname{diam}\Omega) > 0$ such that $\mathbb{E}^{x_0}[\varepsilon^2\tau] \leq \widetilde{C}$. Then, by Markov's inequality, we have

$$\mathbb{P}^{x_0}(\varepsilon^2\tau \geq t) \leq \frac{\mathbb{E}^{x_0}[\varepsilon^2\tau]}{t} \leq \frac{\widetilde{C}}{t}$$

for $t > 0$ and every $x_0 \in \Omega$. Observe that

$$\mathbb{P}^{x_0}(\varepsilon^2\tau \geq \varepsilon^2 n + t | \varepsilon^2\tau \geq \varepsilon^2 n) \leq \sup_{x_n \in \Omega} \mathbb{P}^{x_n}(\varepsilon^2\tau \geq t) \leq \frac{\widetilde{C}}{t}.$$

So, for $n, k \in \mathbb{N}$, applying this bound multiple times we obtain

$$\begin{aligned}\mathbb{P}^{x_0}(\varepsilon^2\tau \geq \varepsilon^2 nk) &= \mathbb{P}^{x_0}(\varepsilon^2\tau \geq \varepsilon^2 nk | \varepsilon^2\tau \geq \varepsilon^2 n(k-1)) \\ &\quad \times \mathbb{P}^{x_0}(\varepsilon^2\tau \geq \varepsilon^2 n(k-1) | \varepsilon^2\tau \geq \varepsilon^2 n(k-2)) \\ &\quad \times \cdots \times \mathbb{P}^{x_0}(\varepsilon^2\tau \geq \varepsilon^2 n) \\ &\leq \Big(\frac{\widetilde{C}}{\varepsilon^2 n}\Big)^k. \end{aligned} \tag{2.5}$$

We define $\delta = \widetilde{C}e + 1$ and observe that

$$\mathbb{P}^{x_0}(\varepsilon^2\tau \geq t) \leq \mathbb{P}^{x_0}\Big(\varepsilon^2\tau \geq \varepsilon^2\Big\lfloor\frac{\delta}{\varepsilon^2}\Big\rfloor\Big\lfloor\frac{t}{\delta}\Big\rfloor\Big).$$

By (2.5) choosing $n = \lfloor\frac{\delta}{\varepsilon^2}\rfloor$ and $k = \lfloor\frac{t}{\delta}\rfloor$ we get

$$\mathbb{P}^{x}(\varepsilon^2\tau \geq t) \leq \Big(\frac{\widetilde{C}}{\varepsilon^2\lfloor\frac{\delta}{\varepsilon^2}\rfloor}\Big)^{\lfloor\frac{t}{\delta}\rfloor} \leq \Big(\frac{\widetilde{C}}{\delta-\varepsilon^2}\Big)^{\frac{t}{\delta}-1} \leq e^{-\frac{t}{\delta}+1},$$

where we have estimated $\varepsilon^2\lfloor\frac{\delta}{\varepsilon^2}\rfloor \geq \delta - \varepsilon^2 \geq \widetilde{C}e$. The result holds for $C = e$ and $\mu = \frac{1}{\delta}$. ∎

Corollary 2.11. *There exists $C = C(\operatorname{diam}\Omega) > 0$ such that*

$$\mathbb{E}^{x_0}[(\varepsilon^2\tau)^2] \leq C$$

for $\varepsilon < \frac{1}{\Lambda}$.

Proof. By Lemma 2.10 there exist $\widetilde{C} > 0$ and $\mu = \mu(\operatorname{diam}\Omega) > 0$ such that

$$\mathbb{P}^{x_0}((\varepsilon^2\tau)^2 \geq t) = \mathbb{P}^{x_0}(\varepsilon^2\tau \geq t^{1/2}) \leq \widetilde{C}e^{-\mu t^{1/2}}.$$

Then we can bound

$$\mathbb{E}^{x_0}[(\varepsilon^2\tau)^2] = \int_0^\infty \mathbb{P}^{x_0}((\varepsilon^2\tau)^2 \geq t)\,dt \leq \int_0^\infty \widetilde{C}e^{-\mu t^{1/2}}\,dt = \frac{2\widetilde{C}}{\mu^2}.$$ ∎

3. Dynamic programming principle: Existence and uniqueness

Recall that $\Omega \subset \mathbb{R}^N$ is a bounded domain and $g\colon \mathbb{R}^N \setminus \Omega \to \mathbb{R}$ and $f\colon \Omega \to \mathbb{R}$ are measurable bounded functions. We define

$$u(x) := \mathbb{E}^x\Big[\varepsilon^2\sum_{i=0}^{\tau-1} f(X_i) + g(X_\tau)\Big],$$

where τ is the exit time from Ω. In this section we prove that u is the unique bounded solution to the DPP (2.2) given by

$$u(x) = \alpha\int u(x+\varepsilon z)\,d\nu_x(z) + \beta\fint_{B_\varepsilon(x)} u(y)\,dy + \varepsilon^2 f(x)$$

for $x \in \Omega$, and $u(x) = g(x)$ for $x \notin \Omega$. For related arguments, see [25], and [1, 14, 30] as well as [5].

In the following lemma we prove that subsolutions are uniformly bounded. We have required subsolutions to be bounded, which is necessary as shown by Example 3.5 below, but here we prove that there is a bound that only depends on the parameters of the problem and not the solution itself.

Lemma 3.1. *There exists $C = C(\operatorname{diam}\Omega, f, g, \varepsilon) > 0$ such that $u \le C$ for every subsolution u to the DPP with boundary values g.*

Proof. We consider the space partitioned along the x_N-axis in strips of width $\frac{\varepsilon}{2}$. We define $S_k = \{y : y_N < k\varepsilon/2\}$,

$$A = \frac{|\{y \in B_\varepsilon(x) : y_N < x_N - \frac{\varepsilon}{2}\}|}{|B_\varepsilon|} = \frac{|\{y \in B_1 : y_N < -\frac{1}{2}\}|}{|B_1|},$$

and $K = \varepsilon^2 \sup_\Omega f$. We have

$$\begin{aligned} u(x) &\le \alpha \int u(x+\varepsilon z)\, dv_x(z) + \beta \fint_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 f(x) \\ &\le \alpha \sup u + \beta A \sup_{y : y_N < x_N - \frac{\varepsilon}{2}} u(y) + \beta(1-A)\sup u + K \\ &= \beta A \sup_{y : y_N < x_N - \frac{\varepsilon}{2}} u(y) + (1-\beta A)\sup u + K. \end{aligned}$$

We define $p = \beta A$ and consider $k \in \mathbb{Z}$: for $x \in S_{k+1}$ we have $\{y : y_N < x_N - \frac{\varepsilon}{2}\} \subset S_k$. Therefore we obtain

$$\sup_{S_{k+1}} u \le p \sup_{S_k} u + (1-p)\sup u + K.$$

Then, inductively, we get

$$\sup_{S_{n+k}} u \le ((1-p)\sup u + K)\sum_{i=0}^{n-1} p^i + p^n \sup_{S_k} u. \tag{3.1}$$

We assume without loss of generality that $\Omega \subset \{y : 0 < y_N < R\}$ for some $R > 0$. Then, since $u = g$ in $\mathbb{R}^N \setminus \Omega \supset S_0$, we have $\sup_{S_0} u = \sup_{S_0} g \le \sup g$. We assume that $\sup u \ge \sup g$ (if not then $\sup g$ is an upper bound for u and the proof is finished) and consider n such that $n\frac{\varepsilon}{2} > R$; then $\Omega \subset S_n$ and we have $\sup_{S_n} u = \sup u$. We apply (3.1) for such n and $k = 0$ to get

$$\begin{aligned} \sup u &\le ((1-p)\sup u + K)\sum_{i=0}^{n-1} p^i + p^n \sup g \\ &= ((1-p)\sup u + K)\frac{1-p^n}{1-p} + p^n \sup g \\ &= (1-p^n)\sup u + K\frac{1-p^n}{1-p} + p^n \sup g, \end{aligned}$$

from where we finally get the upper bound

$$\sup u \le K\frac{1-p^n}{p^n(1-p)} + \sup g. \qquad \blacksquare$$

Lemma 3.2. *There exists u_0 a subsolution to the DPP with $u_0 = g$ on $\mathbb{R}^N \setminus \Omega$.*

Proof. We consider $v(x) = K + L|x|^2$ where we will choose $K \in \mathbb{R}$ and $L > 0$ in what follows. Since it is convex we have

$$v(x) \leq \int v(x + \varepsilon z)\, d\nu_x(z).$$

Even more, since it is strictly convex,

$$v(x) < \fint_{B_\varepsilon(x)} v(y)\, dy,$$

and therefore for L large enough we get

$$v(x) \leq \alpha \int v(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} v(y)\, dy - \varepsilon^2 \| f \|_\infty$$

for every $x \in \Omega$.

We choose K small enough such that $v \leq g$ in $\widetilde{\Omega}_{\Lambda\varepsilon}$. Then u_0 given by $u_0 = v$ on Ω and $u_0 = g$ on $\mathbb{R}^N \setminus \Omega$ is a subsolution. ∎

Proposition 3.3. *There exists a solution to the DPP with boundary values g.*

Proof. We construct the solution by Perron's method. We consider $\mathcal{S}$ the set of subsolutions, that is,

$$\mathcal{S} = \{u \colon \mathbb{R}^N \to \mathbb{R} : \mathcal{L}_\varepsilon u + f \geq 0 \text{ in } \Omega \text{ and } u \leq g \text{ in } \mathbb{R}^N \setminus \Omega\}.$$

The set is not empty by Lemma 3.2. We define

$$\bar{u}(x) = \sup_{u \in \mathcal{S}} u(x)$$

for each $x \in \mathbb{R}^N$. By Lemma 3.1 we have that the function $\bar{u}$ is bounded. We will show that $\bar{u}$ is a solution to the DPP.

The function u_0 given by Lemma 3.2 belongs to $\mathcal{S}$; therefore $\bar{u} \geq u_0$. Since u_0 also takes the right boundary values, we have that $\bar{u} = g$ on $\mathbb{R}^N \setminus \Omega$.

For every $u \in \mathcal{S}$ we have that $u \leq \bar{u}$ and hence

$$u(x) \leq \alpha \int \bar{u}(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} \bar{u}(y)\, dy + \varepsilon^2 f(x).$$

If we take the supremum over all $u \in \mathcal{S}$ in the left-hand side, we obtain that $\bar{u}$ is a subsolution.

We consider $v \colon \mathbb{R}^N \to \mathbb{R}$ given by

$$v(x) = \alpha \int \bar{u}(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} \bar{u}(y)\, dy + \varepsilon^2 f(x)$$

for $x \in \Omega$ and $v = g$ in $\mathbb{R}^N \setminus \Omega$. Since $\bar{u}$ is a subsolution we have $\bar{u} \leq v$. Then

$$v(x) \leq \alpha \int v(x + \varepsilon z)\, d\nu_x(z) + \beta \fint_{B_\varepsilon(x)} v(y)\, dy + \varepsilon^2 f(x),$$

and hence v is a subsolution, which proves that $v \leq \bar{u}$. Thus $v = \bar{u}$ and we conclude that $\bar{u}$ is a solution to the DPP. ∎

Theorem 3.4. *The function*

$$u(x) = \mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i) + g(X_\tau)\right]$$

is the unique bounded solution to the DPP with boundary values g.

Proof. Given a solution v to the DPP, we have that $\{v(X_k) + \varepsilon^2 \sum_{i=0}^{k-1} f(X_i)\}_k$ is a martingale. Indeed,

$$\begin{aligned}
&\mathbb{E}^{x_0}\left[v(X_{k+1}) + \varepsilon^2 \sum_{i=0}^{k} f(X_i) \,\middle|\, \mathcal{F}_k\right](\omega) \\
&\quad = \alpha \int v(x_k + \varepsilon z)\, d\nu_{x_k}(z) + \beta \fint_{B_\varepsilon(x_k)} v(y)\, dy + \varepsilon^2 \sum_{i=0}^{k} f(x_i) \\
&\quad = v(x_k) + \varepsilon^2 \sum_{i=0}^{k-1} f(x_i).
\end{aligned}$$

Then, by Doob's stopping time theorem (recall that v and f are bounded), we have

$$v(x) = \mathbb{E}^x\left[v(X_\tau) + \varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i)\right].$$

Hence, since $v(x) = g(x)$ for $x \notin \Omega$ we obtain $v(x) = \mathbb{E}^x[g(x_\tau) + \varepsilon^2 \sum_{i=0}^{\tau-1} f(x_i)]$. Thus $v = u$, and we have proved that u is a solution to the DPP and that every solution coincides with it, so there is a unique solution. ∎

Example 3.5. The uniqueness fails if we do not assume that solutions to the DPP are bounded. For $\Omega = (-2, 2) \subset \mathbb{R}$ we consider the process given by $\varepsilon = 1$, $\alpha = \beta = 1/2$ and ν_x the uniform probability distribution on B_1 for every point $x \in \Omega$ except those in the set $\{\frac{1}{2^k}\}_{k\in\mathbb{N}}$. There we set

$$\nu_{\frac{1}{2^k}} = \frac{\delta_{\frac{1}{2^{k+1}}} + \delta_{\frac{3}{2^{k+1}}}}{2}.$$

In this case the DPP has multiple solutions: the function $u \equiv 0$ and

$$v(x) = \begin{cases} 4^k, & x = \frac{1}{2^k},\ k \in \mathbb{N}, \\ 0 & \text{otherwise}, \end{cases}$$

which is not bounded. This is why the solutions to the DPP are required to be bounded.

4. ε-ABP estimate

Regarding the classic theory of elliptic PDEs, one of the key inequalities in the Krylov–Safonov proof of Hölder regularity is the so-called Alexandrov–Bakelman–Pucci estimate (ABP estimate for short), which guarantees a pointwise bound for subsolutions of $Lu + f = 0$ by means of the L^N-norm of f. Namely, if f is a continuous bounded function in Ω and $u \in C(\bar{\Omega})$ satisfies

$$\operatorname{Tr}\{A(x) \cdot D^2 u(x)\} + f(x) \geq 0, \quad x \in \Omega,$$

then there exists a constant $C > 0$ depending only on N, $\operatorname{diam}\Omega$ and the uniform ellipticity of $A(x)$ such that

$$\sup_{\Omega} u \leq \sup_{\partial\Omega} u + C \|f\|_{L^N(\Omega)}.$$

See [9, Chapter 3] for the ABP estimate for viscosity solutions and [12, Section 9.1] for strong solutions.

Given a subsolution u, one of the key ideas in the proof of the classical ABP estimate was the use of the concavity properties of u at the set of points where the graph of u can be *touched* from above by tangent hyperplanes. This set of points (known as the *contact set* and denoted by K_u) turned out to carry all relevant information about the subsolution. To be more precise, if we denote by Γ the concave envelope of u, the ABP estimate is obtained by studying the behavior of Γ at those points in Ω where Γ and u agree. Using the concavity of Γ, the first main step in the proof consisted in obtaining an estimate of $\sup_\Omega u$ in terms of $|\nabla\Gamma(\Omega)|$. It is worth noting that the structure of the PDE does not play any role in the proof of this first estimate, which was obtained using exclusively geometric arguments.

In a second step, and in addition to the concavity of Γ, it turned out that Γ is $C^{1,1}$ in the contact set so, by virtue of Rademacher's theorem, Γ is indeed C^2 a.e. in K_u. This fact and a change of variables formula gives an inequality of the form

$$|\nabla\Gamma(\Omega)| \leq \int_\Omega |\det D^2\Gamma(x)|\, dx,$$

which allowed the equation to be used to estimate the right-hand side and, consequently, to obtain the ABP estimate.

However, in the case of the DPP, the nonlocal nature of the setting forces us to also consider noncontinuous subsolutions of the DPP, so the corresponding concave envelope Γ might not be $C^{1,1}$ as in the classical setting. In addition, there is no PDE to connect with the right-hand side of the previous inequality, and therefore we follow a different strategy in order to estimate $|\nabla\Gamma(\Omega)|$. The idea is to cover the contact set K_u by a finite collection of balls of radius $\varepsilon/4$, and then to estimate $|\nabla\Gamma(B_{\varepsilon/4}(x))|$ by means of the oscillation of Γ with respect to a supporting hyperplane touching the graph of Γ from above at $x \in K_u$. This oscillation, in turn, is estimated by using the DPP, which yields the desired ε-ABP

estimate. It is also interesting to note that one can recover the classical ABP estimate by taking limits as $\varepsilon \to 0$.

In this section we adapt the ideas from [7, Section 8], where an ABP-type estimate was obtained for continuous solutions of nonlocal integro-differential equations (see also [8,10] for similar approaches). Further references are [17] for an ABP estimate for elliptic difference equations and [13], where an ABP-type estimate is obtained using a generalization of the concept of a concave envelope as a nonlocal fractional envelope.

Let $u\colon \mathbb{R}^N \to \mathbb{R}$ be a bounded function and let Γ be the *concave envelope* of $u^+ := \max\{u, \sup_{\mathbb{R}^N\setminus\Omega} u\}$ in $\widetilde{\Omega}_{\Lambda\varepsilon}$, that is,

$$\Gamma(x) := \begin{cases} \inf\{\ell(x) : \text{for all hyperplanes } \ell \geq u^+ \text{ in } \widetilde{\Omega}_{\Lambda\varepsilon}\} & \text{if } x \in \widetilde{\Omega}_{\Lambda\varepsilon}, \\ \sup_{\mathbb{R}^N\setminus\Omega} u & \text{if } x \notin \widetilde{\Omega}_{\Lambda\varepsilon}. \end{cases}$$

Since u is not necessarily continuous, we define the "contact" set as

$$K_u := \{x \in \bar{\Omega} : \limsup_{y\to x} u^+(y) = \Gamma(x)\}. \tag{4.1}$$

Since $u^+ \leq \Gamma$, then K_u is a closed subset, and thus compact. Moreover, observe that in the particular case of u being an upper semicontinuous function in Ω, then $K_u = \bar{\Omega} \cap \{u^+ = \Gamma\}$.

As we have already pointed out, one of the key steps in the proof of our ABP-type estimate is the construction of a suitable cover of the contact set K_u by balls of radius $\varepsilon/4$. For this purpose, before stating the main result of this section we introduce the following notation. Given $\varepsilon > 0$, we denote by $\mathcal{Q}_\varepsilon(\mathbb{R}^N)$ a grid of open cubes of diameter $\varepsilon/4$ covering $\mathbb{R}^N$. Take for instance

$$\mathcal{Q}_\varepsilon(\mathbb{R}^N) := \{Q = Q_{\frac{\varepsilon}{4\sqrt{N}}}(x) : x \in \tfrac{\varepsilon}{4\sqrt{N}}\mathbb{Z}^N\}.$$

In addition, if $A \subset \mathbb{R}^N$, we write

$$\mathcal{Q}_\varepsilon(A) := \{Q \in \mathcal{Q}_\varepsilon(\mathbb{R}^N) : \bar{Q} \cap A \neq \emptyset\}, \tag{4.2}$$

so

$$A \subset \bigcup_{Q\in\mathcal{Q}_\varepsilon(A)} \bar{Q}.$$

We stress that, while not needed in the proof of the main ε-ABP estimate, the assumption of $\mathcal{Q}_\varepsilon$ being a grid is needed later in the proof of Theorem 4.7.

Now we are in condition to state the main theorem of this section. We use the notation $\mathcal{L}_\varepsilon u + f$ for convenience in some of the proofs, but this is equivalent to the DPP and stochastic notation as we recall at the end of the section.

Theorem 4.1 (ε-ABP estimate with continuous f). *Suppose that u is a bounded Borel measurable function satisfying*

$$\mathcal{L}_\varepsilon u + f \geq 0$$

in Ω for $f \in C(\overline{\Omega})$. Let Γ be the concave envelope of u^+ in $\widetilde{\Omega}_{\Lambda\varepsilon}$ and let $\mathcal{Q} = \mathcal{Q}_\varepsilon(K_u)$ be the grid of pairwise disjoint open cubes Q of diameter $\varepsilon/4$ defined in (4.2). Then

$$\sup_{\Omega} u \leq \sup_{\mathbb{R}^N \setminus \Omega} u + \frac{2^{N+3}}{\beta}(\operatorname{diam}\Omega + \Lambda\varepsilon)\left(\sum_{Q\in\mathcal{Q}} (\sup_{Q} f^+)^N\right)^{1/N} \varepsilon. \tag{4.3}$$

After proving this theorem we relate the result to the stochastic process and in Theorem 4.7 we obtain a version of the estimate where the continuity hypothesis for f is removed.

In what follows we can assume without loss of generality that $f \geq 0$ in $\overline{\Omega}$ and $\sup_{\mathbb{R}^N\setminus\Omega} u = 0$. Then $u^+ = \max\{u, 0\}$.

It turns out that in order to prove Theorem 4.1 we only need to use the information of the concave envelope in the set of contact points K_u. Indeed, since

$$\delta u(x, y) \leq \delta\Gamma(x, y) + 2[\Gamma(x) - u(x)],$$

inserting this inequality in (2.3) we get

$$\mathcal{L}_\varepsilon u(x) \leq \mathcal{L}_\varepsilon\Gamma(x) + \frac{1}{\varepsilon^2}[\Gamma(x) - u(x)].$$

Since Γ is concave, then Γ is continuous and for any fixed $x_0 \in K_u$ we have

$$\liminf_{x\to x_0} \mathcal{L}_\varepsilon u(x) \leq \mathcal{L}_\varepsilon\Gamma(x_0).$$

Hence, if $f \geq 0$ is a continuous function in $\overline{\Omega}$ and u is a bounded function satisfying

$$\begin{cases} \mathcal{L}_\varepsilon u + f \geq 0 & \text{in } \Omega, \\ u \leq 0 & \text{in } \mathbb{R}^N \setminus \Omega, \end{cases} \tag{4.4}$$

then

$$\mathcal{L}_\varepsilon\Gamma + f \geq 0 \quad \text{in } K_u. \tag{4.5}$$

Therefore, in what follows, we will use (4.5) instead of (4.4).

Before stating the first lemma of this section, we need to define the *superdifferential* of Γ at $x \in \widetilde{\Omega}_{\Lambda\varepsilon}$ as the set

$$\nabla\Gamma(x) := \left\{\xi \in \mathbb{R}^N : \Gamma(z) \leq \Gamma(x) + \langle\xi, z - x\rangle \text{ for all } z \in \widetilde{\Omega}_{\Lambda\varepsilon}\right\}. \tag{4.6}$$

Since Γ is a concave function in $\widetilde{\Omega}_{\Lambda\varepsilon}$, then $\nabla\Gamma(x) \neq \emptyset$ for every $x \in \widetilde{\Omega}_{\Lambda\varepsilon}$. Moreover, given a set $S \subset \widetilde{\Omega}_{\Lambda\varepsilon}$, we denote $\nabla\Gamma(S) = \bigcup_{x\in S} \nabla\Gamma(x)$.

In addition, if S is a compact subset of $\overline{\Omega}$, then $\nabla\Gamma(S)$ is closed. Indeed, if $\{\xi_n\}_n \subset \nabla\Gamma(S)$ is a sequence converging to $\xi_0 \in \widetilde{\Omega}_{\Lambda\varepsilon}$, by definition there exists $\{x_n\}_n \subset S$ (which by compactness we can assume converges to some $x_0 \in S$ by passing to a subsequence) such that $\Gamma(z) \leq \Gamma(x_n) + \langle\xi_n, z - x_n\rangle$ for each $z \in \widetilde{\Omega}_{\Lambda\varepsilon}$. Since Γ is concave (and thus continuous), taking limits we get that $\xi_0 \in \nabla\Gamma(x_0) \subset \nabla\Gamma(S)$. In consequence, $\nabla\Gamma(S)$ is a Lebesgue measurable set.

Lemma 4.2. *Let $u\colon \mathbb{R}^N \to \mathbb{R}$ be a bounded function such that $u \le 0$ in $\mathbb{R}^N \setminus \Omega$. Then*

$$\sup_{\Omega} u \le (\operatorname{diam}\Omega + \Lambda\varepsilon)\left(\frac{|\nabla\Gamma(K_u)|}{|B_1|}\right)^{1/N}, \tag{4.7}$$

where K_u is the contact set defined in (4.1).

Proof. Let us assume that $\sup_\Omega u > 0$ (otherwise (4.7) would follow trivially) and define

$$\rho := \frac{\sup_\Omega u}{\operatorname{diam}\Omega + \Lambda\varepsilon} > 0.$$

Since Γ is the concave envelope of u^+ and $\sup u = \sup \Gamma$, then for every $|\xi| < \rho$ there exists ℓ a supporting hyperplane of Γ in $\widetilde{\Omega}_{\Lambda\varepsilon}$ such that $\nabla\ell \equiv \xi$. Fix any $\xi \in B_\rho$. We claim that $\xi \in \nabla\Gamma(x_0)$ for some $x_0 \in K_u$. To see this, let

$$\tau := \sup_{z\in\widetilde{\Omega}_{\Lambda\varepsilon}} \{u^+(z) - \langle \xi, z\rangle\}$$

and define $\ell(z) = \tau + \langle \xi, z\rangle$ for every $z \in \widetilde{\Omega}_{\Lambda\varepsilon}$. Then $\ell \ge \Gamma \ge u^+$ in $\widetilde{\Omega}_{\Lambda\varepsilon}$. Moreover, by the definition of τ, for each $n \in \mathbb{N}$, there exists $x_n \in \widetilde{\Omega}_{\Lambda\varepsilon}$ such that

$$u^+(x_n) + \frac{1}{n} \ge \ell(x_n) \ge \Gamma(x_n).$$

On the other hand, by the definition of ℓ, for any $z \in \Omega$ we have

$$\begin{aligned} u^+(z) \le \ell(z) &= \ell(x_n) + \langle \xi, z - x_n\rangle \\ &\le \ell(x_n) + |\xi|(\operatorname{diam}\Omega + \Lambda\varepsilon) \\ &= \ell(x_n) + \frac{|\xi|}{\rho}\sup_\Omega u^+ \end{aligned}$$

for each $n \in \mathbb{N}$, where in the inequality we have used that $x_n \in \widetilde{\Omega}_{\Lambda\varepsilon}$ and $z \in \Omega$, and the definition of ρ has been recalled in the last equality. Taking the supremum for $z \in \Omega$ we obtain

$$\left(1 - \frac{|\xi|}{\rho}\right)\sup_\Omega u^+ \le \ell(x_n) \le u^+(x_n) + \frac{1}{n}$$

for each $n \in \mathbb{N}$. Hence, since $|\xi| < \rho$ and $\sup_\Omega u^+ > 0$, we can assume that $x_n \in \overline{\Omega}$ for every $n \in \mathbb{N}$. Otherwise, since $u \le 0$ in $\mathbb{R}^N \setminus \Omega$ by assumption, if $x_n \in \widetilde{\Omega}_{\Lambda\varepsilon} \setminus \overline{\Omega}$ for each sufficiently large $n \in \mathbb{N}$, letting $n \to \infty$ we would obtain a contradiction. Furthermore, by a compactness argument, we can assume without loss of generality that x_n converges to a point $x_0 \in \overline{\Omega}$. Thus, since Γ is continuous, taking limits we get

$$\limsup_{y\to x_0} u^+(y) \ge \limsup_{n\to\infty} u^+(x_n) \ge \Gamma(x_0).$$

Finally, since $u^+ \le \Gamma$, we have in particular that $x_0 \in K_u$ and $\xi \in \nabla\Gamma(x_0)$. In consequence $B_\rho \subset \nabla\Gamma(K_u)$, so $|B_1|\rho^N \le |\nabla\Gamma(K_u)|$ and (4.7) follows. ∎

The idea is to estimate the term $|\nabla\Gamma(K_u)|$ in the right-hand side of (4.7) by covering the contact set K_u with balls of radius $\varepsilon/4$ and estimating $|\nabla\Gamma(B_{\varepsilon/4}(x))|$. This is done by obtaining an upper bound for the gradients of the concave function Γ in $B_{\varepsilon/4}(x)$ which depends on the oscillation of Γ with respect to a supporting hyperplane touching the graph of Γ at x.

Lemma 4.3. *Let $\Gamma\colon \widetilde{\Omega}_{\Lambda\varepsilon} \to \mathbb{R}$ be a concave function. Then*

$$\frac{|\nabla\Gamma(B_{\varepsilon/4}(x))|}{|B_\varepsilon|} \leq \Big(\frac{2}{\varepsilon^2} \operatorname*{osc}_{y\in B_{\varepsilon/2}} \{\Gamma(x) - \Gamma(x+y) + \langle \xi, y\rangle\}\Big)^N \tag{4.8}$$

for every $x \in \Omega$ and $\xi \in \nabla\Gamma(x)$.

Proof. Fix $x \in \Omega$ and any $\xi \in \nabla\Gamma(x)$ and define the auxiliary function $\Phi\colon B_{\varepsilon/2} \to \mathbb{R}$ by

$$\Phi(y) := \Gamma(x+y) - \Gamma(x) - \langle \xi, y\rangle$$

for every $y \in B_{\varepsilon/2}$. Since Γ is a concave function, then Φ is also concave in $B_{\varepsilon/2}$, so

$$\nabla\Phi(y) = \big\{\zeta \in \mathbb{R}^N : \Phi(y+z) \leq \Phi(y) + \langle \zeta, z\rangle \text{ for all } z \text{ s.t. } y+z \in B_{\varepsilon/2}\big\} \neq \emptyset$$

for every $y \in B_{\varepsilon/2}$. Let us fix any $y \in B_{\varepsilon/4}$ and any $\zeta \in \nabla\Phi(y)$. Since Φ is concave, $\Phi \leq 0$ and $\Phi(0) = 0$ we have

$$|\zeta| \leq \frac{\operatorname{osc}_{B_{\varepsilon/2}\setminus B_{\varepsilon/4}} \Phi}{\varepsilon/2} \leq \frac{2}{\varepsilon} \operatorname*{osc}_{B_{\varepsilon/2}} \Phi =: \rho. \tag{4.9}$$

On the other hand, by the definition of Φ and $\nabla\Gamma(x+y)$ in (4.6), for any $\xi' \in \nabla\Gamma(x+y)$ we have

$$\Phi(y+z) - \Phi(y) - \langle \xi' - \xi, z\rangle = \Gamma(x+y+z) - \Gamma(x+y) - \langle \xi', z\rangle \leq 0$$

for every z such that $z + y \in B_{\varepsilon/2}$, so $\xi' - \xi \in \nabla\Phi(y)$. Then, since $|\zeta| \leq \rho$ for every $\zeta \in \nabla\Phi(y)$ by (4.9), we get that $\xi' \in \overline{B}_\rho(\xi)$ for every $\xi' \in \nabla\Gamma(x+y)$. Thus $\nabla\Gamma(x+y) \subset \overline{B}_\rho(\xi)$ for every $y \in B_{\varepsilon/4}$, so $|\nabla\Gamma(B_{\varepsilon/4}(x))| \leq |B_1|\rho^N$ and (4.8) follows. ∎

The following lemma shows that the graph of Γ stays quadratically close to a tangent hyperplane in a neighborhood of any point in which the inequality $\mathcal{L}_\varepsilon\Gamma + f \geq 0$ is satisfied. It is noteworthy to mention that this is the only result where the DPP is used.

Lemma 4.4. *Suppose that Γ is a concave function and $x_0 \in \overline{\Omega}$ satisfies $\mathcal{L}_\varepsilon\Gamma(x_0) + f(x_0) \geq 0$. Then, for any $w > 0$, the following holds:*

$$\frac{\big|\{y \in B_\varepsilon : \Gamma(x_0) - \Gamma(x_0+y) + \langle \xi, y\rangle > w\}\big|}{|B_\varepsilon|} \leq \frac{f(x_0)\varepsilon^2}{w\beta}, \tag{4.10}$$

where ξ is any vector in $\nabla\Gamma(x_0)$. Furthermore,

$$\operatorname*{osc}_{y\in B_{\varepsilon/2}} \{\Gamma(x_0) - \Gamma(x_0+y) + \langle \xi, y\rangle\} \leq \frac{2^{N+2}}{\beta} f(x_0)\varepsilon^2. \tag{4.11}$$

Proof. First observe that, since Γ is concave in $\widetilde{\Omega}_{\Lambda\varepsilon}$, then $\delta\Gamma(x_0, y) \le 0$ for every $y \in B_{\Lambda\varepsilon}$. Thus, we can estimate by zero the α-term in (2.3), so we obtain

$$\mathcal{L}_\varepsilon \Gamma(x_0) \le \frac{\beta}{2\varepsilon^2} \fint_{B_\varepsilon} \delta\Gamma(x_0, y)\, dy. \tag{4.12}$$

Since $f(x_0) \ge -\mathcal{L}_\varepsilon \Gamma(x_0)$ by assumption, using the definition of $\delta\Gamma(x_0, y)$ and the symmetry of the ball we can estimate

$$\begin{aligned} \frac{f(x_0)\varepsilon^2}{\beta} &\ge -\frac{1}{2} \fint_{B_\varepsilon} \delta\Gamma(x_0, y)\, dy \\ &= \fint_{B_\varepsilon} (\Gamma(x_0) - \Gamma(x_0 + y))\, dy \\ &= \fint_{B_\varepsilon} (\Gamma(x_0) - \Gamma(x_0 + y) + \langle \xi, y\rangle)\, dy \end{aligned}$$

for any fixed $\xi \in \nabla\Gamma(x_0)$. Let us define the auxiliary function $\Phi\colon B_\varepsilon \to \mathbb{R}$ by

$$\Phi(y) = \Gamma(x_0) - \Gamma(x_0 + y) + \langle \xi, y\rangle.$$

Observe that, for the sake of convenience, the sign of Φ has been changed with respect to the previous proof. Notice that $\Phi \ge 0$ due to the concavity of Γ. We split the ball B_ε in two sets and we study the integral of Φ over each of them. Then

$$\begin{aligned} \int_{B_\varepsilon} \Phi(y)\, dy &= \int_{B_\varepsilon \cap \{\Phi > w\}} \Phi(y)\, dy + \int_{B_\varepsilon \cap \{\Phi \le w\}} \Phi(y)\, dy \\ &\ge \int_{B_\varepsilon \cap \{\Phi > w\}} w\, dy \\ &= w|B_\varepsilon \cap \{\Phi > w\}|, \end{aligned}$$

where in the inequality we used that $\Phi \ge 0$ to estimate the second integral over $B_\varepsilon \cap \{\Phi \le w\}$. Then (4.10) follows by combination of the previous estimates.

Now we prove (4.11). If $f(x_0) = 0$, then (4.10) yields that $\Phi \le w$ a.e. for every $w > 0$. Then, since Φ is continuous and $\Phi \ge 0$, we get that $\Phi \equiv 0$, so the oscillation in (4.11) is zero as desired.

If $f(x_0) > 0$, we choose $w > 0$ so that $0 \le \Phi(y) \le w$ holds for every $y \in B_{\varepsilon/2}$. Notice that, as we already mentioned, $\Phi \ge 0$ follows directly from the concavity of Γ. To check that $\Phi \le w$, observe that the inclusion $B_{\varepsilon/2}(y) \subset B_\varepsilon$ holds for every $y \in B_{\varepsilon/2}$. Then (4.10) yields

$$\frac{\left|\{z \in B_{\varepsilon/2} : \Phi(y + z) > w\}\right|}{|B_{\varepsilon/2}|} \le 2^N \frac{|B_\varepsilon \cap \{\Phi > w\}|}{|B_\varepsilon|} \le 2^N \frac{f(x_0)\varepsilon^2}{w\beta}.$$

In particular, choosing $w = \frac{2^{N+2}}{\beta} f(x_0)\varepsilon^2$, we get that the left-hand side of the previous inequality is bounded by $1/4$, and thus there exists $z \in B_{\varepsilon/2}$ such that

$$\Phi(y \pm z) \le w = \frac{2^{N+2}}{\beta} f(x_0)\varepsilon^2.$$

Combining the inequalities for z and $-z$ we obtain

$$\frac{1}{2}\Phi(y+z)+\frac{1}{2}\Phi(y-z)\leq w,$$

so $\Phi(y)\leq w$ follows from the convexity of Φ and this completes the proof. ∎

Now we are in position to prove the main result of this section.

Proof of Theorem 4.1. Let us consider the pairwise disjoint collection of open cubes $\mathcal{Q}_\varepsilon(K_u)$ defined in (4.2). Then the following conditions are satisfied:

(1) $\operatorname{diam} Q=\varepsilon/4$;

(2) $\bar{Q}\cap K_u\neq\emptyset$ for each $Q\in\mathcal{Q}_\varepsilon(K_u)$;

(3) $K_u\subset\bigcup_{Q\in\mathcal{Q}_\varepsilon(K_u)}\bar{Q}$.

Since K_u is bounded, we can label the cubes in $\mathcal{Q}_\varepsilon(K_u)$ as $Q_1,\dots,Q_n$, where $n=n(\varepsilon)\in\mathbb{N}$. Furthermore, we select a point $x_i\in K_u\cap Q_i$ for each $i=1,\dots,n$ so that $Q_i\subset B_{\varepsilon/4}(x_i)$. From the above considerations we can estimate

$$|\nabla\Gamma(K_u)|\leq\Big|\bigcup_{i=1}^{n}\nabla\Gamma(\bar{Q}_i)\Big|\leq\sum_{i=1}^{n}|\nabla\Gamma(\bar{Q}_i)|\leq\sum_{i=1}^{n}|\nabla\Gamma(B_{\varepsilon/4}(x_i))|.$$

Combining this with the estimates from Lemmas 4.3 and 4.4 we obtain

$$|\nabla\Gamma(K_u)|\leq|B_1|\Big(\frac{2^{N+3}}{\beta}\Big)^N\Big(\sum_{i=1}^{n}f(x_i)^N\Big)\varepsilon^N.$$

Moreover, since $x_i\in Q_i$, we can estimate $f(x_i)\leq\sup_{Q_i}f$ for each $i=1,\dots,n$. Then the result follows by replacing this in the estimate from Lemma 4.2. ∎

As we saw in Section 3, solutions to the DPP can be interpreted as expected values. Thus the ε-ABP extends to this setting as well.

Corollary 4.5. *Given $f\in C(\bar{\Omega})$ such that $f\geq 0$ and the family $\mathcal{Q}=\mathcal{Q}_\varepsilon(\Omega)$ of pairwise disjoint open cubes Q of diameter $\varepsilon/4$ defined in (4.2), there exists $C>0$ such that*

$$\mathbb{E}^x\Big[\varepsilon^2\sum_{i=0}^{\tau-1}f(X_i)\Big]\leq C(\operatorname{diam}\Omega+\Lambda\varepsilon)\Big(\sum_{Q\in\mathcal{Q}}\Big(\sup_Q f\Big)^N\Big)^{1/N}\varepsilon. \tag{4.13}$$

More precisely, $C=2^{N+3}/\beta$.

Proof. We consider $u\colon\mathbb{R}^N\to\mathbb{R}$,

$$u(x)=\mathbb{E}^x\Big[\varepsilon^2\sum_{i=0}^{\tau-1}f(X_i)\Big].$$

By Theorem 3.4 we have that $\mathcal{L}_\varepsilon u=-f$ in Ω and $u=0$ in $\mathbb{R}^N\setminus\Omega$. The result follows by applying Theorem 4.1 to u. ∎

Let us observe that the ε-ABP estimate (4.3) yields the classical ABP estimate after taking limits as $\varepsilon \to 0$. Since each Q in $\mathcal{Q}_\varepsilon(K_u)$ has diameter $\varepsilon/4$, then $\varepsilon = 4\sqrt{N}|Q|^{1/N}$ and thus

$$\left(\sum_{Q \in \mathcal{Q}_\varepsilon(K_u)} \left(\sup_Q f^+ \right)^N \right)^{1/N} \varepsilon = 4\sqrt{N} \left(\sum_{Q \in \mathcal{Q}_\varepsilon(K_u)} \left(\sup_Q f^+ \right)^N |Q| \right)^{1/N}.$$

Furthermore, letting $\varepsilon \to 0$ the size of the cubes in $\mathcal{Q}_\varepsilon(K_u)$ converges to zero, and since f is continuous, we obtain the $L^N(K_u)$-norm of f^+ as the limit of Riemann sums, that is,

$$\lim_{\varepsilon \to 0} \left(\sum_{Q \in \mathcal{Q}_\varepsilon(K_u)} \left(\sup_Q f^+ \right)^N \right)^{1/N} \varepsilon = 4\sqrt{N} \| f^+ \|_{L^N(K_u)},$$

where

$$\| f^+ \|_{L^N(K_u)} := \left(\int_{K_u} f^+(x)^N \, dx \right)^{1/N}.$$

Thus, replacing this in the ε-ABP estimate (4.3) we get

$$\sup_\Omega u \leq \sup_{\mathbb{R}^N \setminus \Omega} u + \frac{2^{N+5}\sqrt{N} \operatorname{diam} \Omega}{\beta} \| f^+ \|_{L^N(K_u)} + o(\varepsilon^0),$$

which is the classical ABP estimate plus an error term vanishing when $\varepsilon \to 0$.

Observe that the error depends on f; moreover it does not vanish uniformly on f. Also observe that the ABP estimate requires f to be continuous. The standard version of the ABP estimate in the context of PDEs is with the L^N-norm of f on the right-hand side. Unfortunately, such an inequality does not hold in our setting. That is, for a general f, an inequality such as

$$\mathbb{E}^x \left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(x_i) \right] \leq C \| f \|_N$$

does not hold, as the next example shows.

Example 4.6. Let us consider $\Omega = B_2$, $\varepsilon = 1$, and $f = 1_{\mathbb{Q}^N}$, for which

$$\| f \|_N = 0.$$

Let ν be given by

$$\nu_x = \frac{\delta_{v_x} + \delta_{-v_x}}{2},$$

where v_x is such that $x + v_x \in \mathbb{Q}^N$. It follows that $\mathbb{E}^x[f(x_i)] \geq \frac{\alpha}{2}$ for any x. Then since $\mathbb{E}^0[\varepsilon^2 \tau] \geq c$ we have

$$\mathbb{E}^0 \left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(x_i) \right] \geq \frac{c\alpha}{2}.$$

We overcome the difficulty of the previous example in the following theorem, where we obtain a weaker version of the result. Fortunately it is enough for our purposes; see Lemma 5.1.

Theorem 4.7 (ε-ABP estimate with measurable f). *Given $f\colon \Omega \to \mathbb{R}$ a nonnegative bounded measurable function, there exists $C > 0$ (depending only on N) such that*

$$\mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i)\right] \leq (\varepsilon^2 + \alpha \mathbb{E}^x[\varepsilon^2 \tau])\|f\|_\infty + C(\operatorname{diam}\Omega + \Lambda\varepsilon)\|f\|_N.$$

Proof. First we extend $f\colon \Omega \to \mathbb{R}$ outside Ω defining $f(x) = 0$ for every $x \in \mathbb{R}^N \setminus \Omega$. Then we define $\tilde{f}\colon \mathbb{R}^N \to \mathbb{R}$ as the function given by

$$\tilde{f}(x) = \fint_{B_\varepsilon(x)} f(y)\,dy$$

for every $x \in \mathbb{R}^N$, so $\tilde{f}$ is continuous in $\mathbb{R}^N$ and, in particular, $\tilde{f} \in C(\bar{\Omega})$. For $i \geq 1$ we have

$$\begin{aligned}\mathbb{E}^x[f(X_i)|\mathcal{F}_{i-1}](\omega) &= \alpha \int f(x_{i-1} + \varepsilon z)\,dv_{x_{i-1}}(z) + \beta \fint_{B_\varepsilon(x_{i-1})} f(y)\,dy \\ &\leq \alpha\|f\|_\infty + \beta \tilde{f}(x_{i-1}). \end{aligned} \tag{4.14}$$

Since $\mathbb{E}^x[f(X_i)] = \mathbb{E}^x[\mathbb{E}^x[f(X_i)|\mathcal{F}_{i-1}]]$ we obtain

$$\begin{aligned}\mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i)\right] &= \mathbb{E}^x\left[\varepsilon^2 f(X_0) + \varepsilon^2 \sum_{i=1}^{\tau-1} \mathbb{E}^x[f(X_i)|\mathcal{F}_{i-1}]\right] \\ &\leq \varepsilon^2\|f\|_\infty + \mathbb{E}^x\left[\varepsilon^2 \sum_{i=1}^{\tau-1} (\alpha\|f\|_\infty + \beta \tilde{f}(X_{i-1}))\right],\end{aligned}$$

where we used (4.14). Rearranging terms we get

$$\begin{aligned}\mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i)\right] &\leq \varepsilon^2\|f\|_\infty + \alpha \mathbb{E}^x[\varepsilon^2(\tau-1)]\|f\|_\infty + \beta \mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-2} \tilde{f}(X_i)\right] \\ &\leq (\varepsilon^2 + \alpha \mathbb{E}^x[\varepsilon^2 \tau])\|f\|_\infty + \beta \mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-1} \tilde{f}(X_i)\right].\end{aligned}$$

Observe that since $\tilde{f} \in C(\bar{\Omega})$ and $\tilde{f} \geq 0$ we can apply Corollary 4.5. We obtain

$$\mathbb{E}^x\left[\varepsilon^2 \sum_{i=0}^{\tau-1} \tilde{f}(X_i)\right] \leq \frac{2^{N+3}}{\beta}(\operatorname{diam}\Omega + \Lambda\varepsilon)\left(\sum_{Q \in \mathcal{Q}_\varepsilon(\Omega)} \left(\sup_Q \tilde{f}\right)^N\right)^{1/N} \varepsilon. \tag{4.15}$$

For any fixed $Q \in \mathcal{Q}_\varepsilon(\Omega)$, let x_0 denote the center of Q, so $Q = Q_{\frac{\varepsilon}{4\sqrt{N}}}(x_0)$. Since $\operatorname{diam} Q = \varepsilon/4$, then $|x - x_0| < \varepsilon/8$ for every $x \in \overline{Q}$. Then

$$Q \subset B_\varepsilon(x) \subset B_{\varepsilon+|x-x_0|}(x_0) \subset B_{9\varepsilon/8}(x_0) \subset Q_{9\varepsilon/4}(x_0) = 9\sqrt{N}\, Q_{\frac{\varepsilon}{4\sqrt{N}}}(x_0) = 9\sqrt{N}\, Q.$$

Let $\ell = \ell(N) \in \mathbb{N}$ be the unique odd integer such that $\ell - 2 < 9\sqrt{N} \le \ell$. In consequence, $Q \subset B_\varepsilon(x) \subset \ell Q$ for every $x \in \overline{Q}$. Since $\tilde{f}$ is continuous in $\mathbb{R}^N$, there exists some $\bar{x} \in \overline{Q}$ such that $\tilde{f}(\bar{x}) = \sup_Q \tilde{f}$ and thus

$$\Big(\sup_Q \tilde{f}\Big)^N = \Big(\fint_{B_\varepsilon(\bar{x})} f(y)\, dy\Big)^N \le \fint_{B_\varepsilon(\bar{x})} f(y)^N\, dy \le \frac{1}{|B_\varepsilon|} \int_{\ell Q} f(y)^N\, dy,$$

where in the first inequality we have recalled Jensen's inequality for convex functions. Moreover, since the cubes in $\mathcal{Q}_\varepsilon(\Omega)$ form a grid, it turns out that $\overline{\ell Q}$ can be expressed as the union of the cubes $\overline{Q'}$ such that $Q' \in \mathcal{Q}_\varepsilon(\Omega)$ and $Q' \subset \ell Q$. Since any particular $Q' \in Q_\varepsilon(\Omega)$ belongs to $\operatorname{card}\{Q \in \mathcal{Q}_\varepsilon(\Omega) : Q' \subset \ell Q\} = \ell^N$ number of cubes ℓQ, we can estimate the overlap and get

$$\begin{aligned}
\sum_{Q \in \mathcal{Q}_\varepsilon(\Omega)} \Big(\sup_Q \tilde{f}\Big)^N &\le \frac{1}{|B_\varepsilon|} \sum_{Q \in \mathcal{Q}_\varepsilon(\Omega)} \int_{\ell Q} f(y)^N\, dy \\
&= \frac{1}{|B_\varepsilon|} \sum_{Q' \in \mathcal{Q}_\varepsilon(\Omega)} \operatorname{card}\big\{Q \in \mathcal{Q}_\varepsilon(\Omega) : Q' \subset \ell Q\big\} \int_{Q'} f(y)^N\, dy \\
&= \frac{\ell^N}{|B_\varepsilon|} \sum_{Q' \in \mathcal{Q}_\varepsilon(\Omega)} \int_{Q'} f(y)^N\, dy \\
&= \frac{\ell^N}{|B_\varepsilon|} \int_\Omega f(y)^N\, dy,
\end{aligned}$$

where the last equality comes from the fact that $f \equiv 0$ outside Ω. Taking the Nth root we finally obtain

$$\Big(\sum_{\mathcal{Q}_\varepsilon(\Omega)} \Big(\sup_Q \tilde{f}\Big)^N\Big)^{1/N} \le \frac{\ell}{|B_1|^{1/N}\varepsilon} \Big(\int_\Omega f(y)^N\, dy\Big)^{1/N},$$

and the result follows after inserting this in (4.15). ∎

5. De Giorgi oscillation lemma

The main goal of this section is to prove Lemma 5.8, a version of the classical De Giorgi oscillation lemma.

We follow the Krylov–Safonov argument in [15, 16]; see also [3, Chapter V, Section 7]. However, our case is partly discrete and ε sets a natural limit for the scale that can be

used in the proofs. This causes considerable changes. The key result is Theorem 5.7 where we prove that a set of positive measure is reached by the process with positive probability. This then implies the De Giorgi oscillation lemma in a straightforward manner by using a level set as the set of positive measure.

One of the key steps in this section is the use of an adapted version of the Calderón–Zygmund decomposition, Lemma 5.4. The main difference from the classical version is that we do not consider cubes of scale smaller than ε. If we simply stop the decomposition once we reach cubes of scale ε, then we would lose control between the original set and the union of cubes in the decomposition. Therefore we need a subtle additional condition in the decomposition for cubes of size ε.

The first step in our argument is to prove that sets of "large density" are reached by the process with positive probability. This is done in the following lemma. There we employ the ε-ABP estimate, Theorem 4.7, with the characteristic function of A as a right-hand side, and further estimates from Section 2.4. Recall that T_A denotes the hitting time for A and τ_A the exit time, that is,

$$T_A = \min\{k \in \mathbb{N} : X_k \in A\} \quad \text{and} \quad \tau_A = \min\{k \in \mathbb{N} : X_k \notin A\}.$$

Lemma 5.1. *There exist $\varepsilon_1 > 0$, $\theta > 0$ and $c > 0$ such that if $0 < \varepsilon < \varepsilon_1$, $x \in Q_{1/2}$, $A \subset Q_1$ and $|Q_1 \setminus A| < \theta$ then*

$$\mathbb{P}^x(T_A < \tau_{Q_1}) \geq c.$$

Proof. We denote $\tau = \tau_{Q_1}$ and $A^c = Q_1 \setminus A$. We write

$$\begin{aligned}\mathbb{E}^x[\varepsilon^2\tau] &= \mathbb{E}^x[\varepsilon^2\tau 1_{\{T_A<\tau\}}] + \mathbb{E}^x[\varepsilon^2\tau 1_{\{T_A\geq\tau\}}] \\ &\leq \mathbb{E}^x[(\varepsilon^2\tau)^2]^{\frac{1}{2}}\mathbb{P}^x(T_A < \tau)^{\frac{1}{2}} + \mathbb{E}^x\left[\varepsilon^2\sum_{i=0}^{\tau-1} 1_{A^c}(X_i)\right], \qquad (5.1)\end{aligned}$$

where we have used the Cauchy–Schwarz inequality and that $\sum_{i=0}^{\tau-1} 1_{A^c}(X_i) = \tau$ when $T_A \geq \tau$. By Theorem 4.7 we have

$$\mathbb{E}^x\left[\varepsilon^2\sum_{i=0}^{\tau-1} 1_{A^c}(X_i)\right] \leq \varepsilon^2 + \alpha\mathbb{E}^x[\varepsilon^2\tau] + C|A^c|^{1/N}.$$

Combining this inequality with (5.1) we obtain

$$\mathbb{E}^x[\varepsilon^2\tau] \leq \mathbb{E}^x[(\varepsilon^2\tau)^2]^{\frac{1}{2}}\mathbb{P}^x(T_A < \tau)^{\frac{1}{2}} + \varepsilon^2 + \alpha\mathbb{E}^x[\varepsilon^2\tau] + C\theta^{1/N}$$

and, rearranging terms,

$$\beta\mathbb{E}^x[\varepsilon^2\tau] \leq \mathbb{E}^x[(\varepsilon^2\tau)^2]^{\frac{1}{2}}\mathbb{P}^x(T_A < \tau)^{\frac{1}{2}} + \varepsilon^2 + C\theta^{1/N}.$$

By Lemma 2.9 (observe that $\operatorname{dist}(x, \mathbb{R}^N \setminus Q_1) \geq 1/4$) and Corollary 2.11 there exist $c_1, c_2 > 0$ such that $c_1 \leq \mathbb{E}^x[\varepsilon^2\tau]$ and $\mathbb{E}^x[(\varepsilon^2\tau)^2]^{\frac{1}{2}} \leq c_2$. Therefore

$$\beta c_1 \leq c_2\mathbb{P}^x(T_A < \tau)^{\frac{1}{2}} + C\theta^{1/N} + \varepsilon^2$$

and the result follows for ε and θ small enough. ∎

In the following lemma we prove that sets of positive measure are reached by the process with positive probability when $\varepsilon_0/2 \le \varepsilon < \varepsilon_0$. By performing a scaling of the space and step size we later use the result for cubes of size comparable to ε in Theorem 5.7; see Remark 5.3.

Lemma 5.2. *Given $0 < \varepsilon_0 \le 1$, there exists $\gamma = \gamma(\varepsilon_0) > 0$ such that if $\varepsilon_0/2 \le \varepsilon < \varepsilon_0$, $x \in Q_1$, $A \subset Q_1$, we have*

$$\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) \ge \gamma |A|.$$

Proof. We define $N_0 = \lceil \frac{2\sqrt{N}}{\varepsilon_0} \rceil + 1$ and consider the event E of the first N_0 movements to be uniformly distributed. That is $E = \{c_1 = \cdots = c_{N_0} = 1\}$. We have

$$\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) \ge \mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}} | E)\mathbb{P}(E).$$

Observe that $\mathbb{P}(E) = \beta^{N_0}$. If a uniform random step occurs, then the step size is at most ε. Hence, after N_0 uniform random steps the token is at a distance of at most

$$N_0\varepsilon \le N_0\varepsilon_0 < \Big(\frac{2\sqrt{N}}{\varepsilon_0} + 2\Big)\varepsilon_0 \le 2\sqrt{N} + 2 \le 4\sqrt{N}$$

from the starting point. Therefore, we have that all the steps until time N_0 are inside $Q_{10\sqrt{N}}$ and we have $\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}} | E) \ge \mathbb{P}^x(X_{N_0} \in A | E)$. We have proved that

$$\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) \ge \mathbb{P}^x(X_{N_0} \in A | E)\beta^{N_0}.$$

We consider U_i a sequence of independent random variables uniformly distributed in B_1. And we define $Y = \sum_{i=1}^{N_0} U_i$. Let f denote the density of Y; it is a radial decreasing function strictly positive in the ball of radius N_0. Given $x_0 = x \in Q_1$ and $y \in Q_1$ we can bound

$$f_{X_{N_0}|E}(y) = \frac{1}{\varepsilon^n} f((y-x)/\varepsilon) \ge \frac{1}{\varepsilon_0^n} f(\sqrt{N}/\varepsilon) \ge \frac{1}{\varepsilon_0^n} f(2\sqrt{N}/\varepsilon_0).$$

Since $\frac{2\sqrt{N}}{\varepsilon_0} < N_0$ we have $f(2\sqrt{N}/\varepsilon_0) > 0$.

Finally we obtain

$$\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) \ge \mathbb{P}^x(X_{N_0} \in A | E)\beta^{N_0} \ge \frac{1}{\varepsilon_0^n} f(2\sqrt{N}/\varepsilon_0)|A|\beta^{N_0}.$$

Therefore the result holds for $\gamma = \frac{1}{\varepsilon_0^n} f(2\sqrt{N}/\varepsilon_0)\beta^{N_0}$. ∎

Remark 5.3. Given a cube Q there exists an affine transformation $h(x) = ax + b$ such that $h(Q) = Q_1$. Given the process X_k we can consider the process $h(X_k)$. Observe that this new process is of the type that we are considering for $\tilde{\varepsilon} = a\varepsilon$ and the pushforward measure $\tilde{\nu}$ given by $\tilde{\nu}_x(A) = \nu_{h^{-1}(x)}(A)$. Then results established for Q_1 such as Lemma 5.2 can be applied to cubes of any size. Moreover, if $\varepsilon_0/2 \le \tilde{\varepsilon} < \varepsilon_0$, then the constant γ only depends on ε_0.

Now we state our version of the Calderón–Zygmund lemma. In the discrete setting, the ε sets a natural limit for the scale. To control the error when stopping the decomposition at the level ε, we introduce an additional condition. When applying the decomposition, a careful choice of the parameters allows us to guarantee the two opposite goals: there are enough cubes in the decomposition and the share of A in measure is still large enough.

First we introduce some notation. We denote by $\mathcal{D}_\ell$ the family of dyadic open subcubes of Q_1 of generation $\ell \in \mathbb{N}$. That is, $\mathcal{D}_0 = \{Q_1\}$, $\mathcal{D}_1$ is the family of 2^N dyadic open cubes obtained by dividing Q_1, and so on. Given $\ell \in \mathbb{N}$ and $Q \in \mathcal{D}_\ell$ we define $\operatorname{pre}(Q) \in \mathcal{D}_{\ell-1}$ as the unique dyadic cube in $\mathcal{D}_{\ell-1}$ containing Q.

Let $0 < \tilde{\delta} < \delta < 1$ and $A \subset Q_1$ a measurable set, such that

$$|A| \leq \delta.$$

Next we will construct a collection of (open) cubes $\mathcal{Q}_B$, containing subcubes from generations $\mathcal{D}_0, \mathcal{D}_1, \ldots, \mathcal{D}_L$, and a set

$$B := \bigcup_{Q \in \mathcal{Q}_B} Q.$$

By the assumption we first observe that

$$|Q_1 \cap A| \leq \delta |Q_1|.$$

Then we split Q_1 into 2^N dyadic cubes $\mathcal{D}_1$. For those dyadic cubes $Q \in \mathcal{D}_1$ that satisfy

$$|A \cap Q| > \delta |Q|, \tag{5.2}$$

we select $\operatorname{pre}(Q)$ into $\mathcal{Q}_B$.

For other dyadic cubes that do not satisfy (5.2) and are not contained in any cube already included in $\mathcal{Q}_B$, we keep splitting, and again repeat the selection according to (5.2). We repeat splitting $L \in \mathbb{N}$ times. At level L, in addition to the previous process, we also select those cubes $Q \in \mathcal{D}_L$ (not the predecessors) into $\mathcal{Q}_B$ for which

$$\delta |Q| \geq |A \cap Q| > \tilde{\delta} |Q| \tag{5.3}$$

and are not contained in any cube already included in $\mathcal{Q}_B$. Now the following lemma holds.

Lemma 5.4 (Calderón–Zygmund). *Let $A \subset Q_1$, $0 < \tilde{\delta} < \delta < 1$, $L \in \mathbb{N}$ and B be as above. It holds that*

$$|A| \leq \delta |B| + \tilde{\delta}.$$

Proof. Observe that for $\operatorname{pre}(Q)$ selected according to (5.2) into $\mathcal{Q}_B$, it holds that

$$|A \cap \operatorname{pre}(Q)| \leq \delta |\operatorname{pre}(Q)|$$

since otherwise we would have stopped splitting already at the earlier round. Also for cubes Q selected according to (5.3) into $\mathcal{Q}_B$, it holds that $|A \cap Q| \leq \delta |Q|$. Summing up, for all the cubes $Q \in \mathcal{Q}_B$, it holds that

$$|A \cap Q| \leq \delta |Q|. \tag{5.4}$$

Moreover, by construction, the cubes in $\mathcal{Q}_B$ are disjoint.

We define $\mathcal{G}_L$ as a family of cubes of $\mathcal{D}_L$ that covers $Q_1 \setminus B$ a.e. It immediately holds a.e. that

$$A \subset Q_1 = \bigcup_{Q \in \mathcal{Q}_B} Q \cup \bigcup_{Q \in \mathcal{G}_L} Q.$$

By this, using (5.4), as well as observing that $|A \cap Q| \leq \tilde{\delta} |Q|$ by (5.3) for every $Q \in \mathcal{G}_L$, we get

$$\begin{aligned} |A| &= \sum_{Q \in \mathcal{Q}_B} |A \cap Q| + \sum_{Q \in \mathcal{G}_L} |A \cap Q| \\ &\leq \sum_{Q \in \mathcal{Q}_B} \delta |Q| + \sum_{Q \in \mathcal{G}_L} \tilde{\delta} |Q| \\ &\leq \delta |B| + \tilde{\delta}. \end{aligned}$$

■

Before proceeding to the main result, we need to show that if the stochastic process starts in a certain cube, it will reach any subcube in the next level of the dyadic decomposition with positive probability. We also need to show that for any starting point in Q_1, the process reaches $Q_{1/2}$ with positive probability. We obtain these results as a corollary of the following lemma.

Lemma 5.5. *Given $0 < R_1 < R_2 < R_3$, there exist $\varepsilon_2 = \varepsilon_2(R_1, R_2, R_3) > 0$ and $p = p(R_1, R_2, R_3) > 0$ such that for $x \in B_{R_2}$ we have*

$$\mathbb{P}^x(T_{B_{R_1}} < \tau_{B_{R_3}}) \geq p$$

for $\varepsilon < \varepsilon_2$.

Proof. For $c > 0$ we consider the radial increasing function $\varphi(x) = -|x|^{-c}$. For $x \in B_{R_3} \setminus B_{R_1}$ we have

$$\begin{aligned} &\beta \fint_{B_\varepsilon(x)} \varphi(y)\, dy + \alpha \int \frac{\varphi(x + \varepsilon z) + \varphi(x - \varepsilon z)}{2}\, d\nu_x(z) \\ &\qquad \leq \varphi(x) + \frac{\beta \varepsilon^2}{2(N+2)} \Delta \varphi(x) + \alpha \varepsilon^2 \Lambda^2 \sup_{z : |z| = 1} \langle D^2 \varphi(x) z, z \rangle + o(\varepsilon^2), \end{aligned}$$

where we have used the second-order Taylor expansion for φ in $B_{R_3 + \Lambda \varepsilon_2} \setminus B_{R_1 - \Lambda \varepsilon_2}$. Observe that φ is smooth in that region for ε_2 small enough.

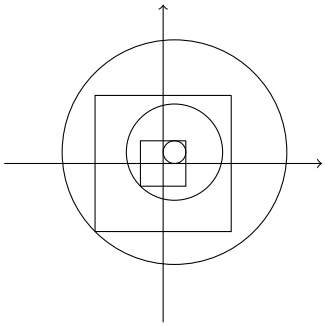


Figure 1. The inclusions of balls and cubes defined in Corollary 5.6 in the case that $P = Q_1 \cap \{x : x_i > 0 \text{ for } i = 1, \dots, N\}$.

We consider $s = |x|$ and $\phi(s) = \varphi(x)$. Recall that for a radial function the eigenvalues of $D^2\varphi$ are $\phi''(s)$ with multiplicity 1 and $\phi'(s)/s$ with multiplicity $N-1$. We obtain

$$\begin{aligned}\beta \fint_{B_\varepsilon(x)} \varphi(y)\,dy + \alpha \int \frac{\varphi(x+\varepsilon z)+\varphi(x-\varepsilon z)}{2}\,d\nu_x(z)\\ \leq \phi(s) + \varepsilon^2\Big(\Big(\frac{\beta(N-1)}{2(N+2)} + \alpha\Lambda^2\Big)\frac{\phi'(s)}{s} + \frac{\beta}{2(N+2)}\phi''(s)\Big) + o(\varepsilon^2). \qquad (5.5)\end{aligned}$$

We have $\phi'(s) = cs^{-c-1}$ and $\phi''(s) = -c(c+1)s^{-c-2}$. Therefore, the right-hand side in (5.5) is smaller than $\varphi(x)$ for every $x \in B_{R_3} \setminus B_{R_1}$ for c large enough and ε_2 small enough. Hence, $\varphi(x_n)$ is a supermartingale.

If $q = \mathbb{P}^x(T_{R_1} < \tau_{R_3})$, we obtain

$$\varphi(R_3)(1-q) + \varphi(R_1 - \Lambda\varepsilon_2)q \leq \mathbb{E}^x[\varphi(x_{\tau_{B_{R_3}\setminus B_{R_1}}})] \leq \varphi(x) \leq \varphi(R_2).$$

Hence $q \geq \frac{\varphi(R_3)-\varphi(R_2)}{\varphi(R_3)-\varphi(R_1-\Lambda\varepsilon_2)} > 0$. ∎

Corollary 5.6. *There exist $\varepsilon_2 > 0$ and $p > 0$ such that if $\varepsilon < \varepsilon_2$, $x \in Q_1$ and $P \subset Q_1$ is a cube of side length $1/2$, we have*

$$\mathbb{P}^x(T_P < \tau_{Q_{10\sqrt{N}}}) \geq p.$$

Proof. For y the center of P, $R_1 = 1/4$, $R_2 = 3\sqrt{N}/2$ and $R_3 = \sqrt{N}(5\sqrt{N} + 1/4)$, we have

$$B_{R_1}(y) \subset P \subset Q_1 \subset B_{R_2}(y) \subset Q_{10\sqrt{N}} \subset B_{R_3}(y),$$

as shown in Figure 1.

Since $B_{R_1}(y) \subset P$ and $Q_{10\sqrt{N}} \subset B_{R_3}(y)$ we have

$$\mathbb{P}^x(T_P < \tau_{Q_{10\sqrt{N}}}) \geq \mathbb{P}^x(T_{B_{R_1}(y)} < \tau_{B_{R_3}(y)}).$$

As $x \in Q_1 \subset B_{R_2}(y)$, the result follows from Lemma 5.5. ∎

We are ready to prove our main result, that a set of positive measure is reached by the process with positive probability. The idea is the following: given a suitable set A we construct B using the Calderón–Zygmund lemma such that B is larger than A in measure. Using this, we prove that the process reaches B (estimate (5.7) below) and then A by considering two alternatives (estimates (5.9) and (5.10) below).

Theorem 5.7. *There exist $\varepsilon_0 > 0$ and a nondecreasing function $\varphi\colon (0,1) \to (0,1)$ such that for every $\varepsilon < \varepsilon_0$, $A \subset Q_1$, $|A| > 0$ and $x \in Q_1$, we have*

$$\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) > \varphi(|A|).$$

Proof. We define $\varepsilon_0 = \min\{\varepsilon_1, \varepsilon_2, 1\}$ where ε_1 and ε_2 are given by Lemma 5.1 and Corollary 5.6, respectively. We also define

$$\varphi(\xi) = \inf\big\{\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) : \nu, \varepsilon < \varepsilon_0,\ A \subset Q_1,\ |A| \geq \xi \text{ and } x \in Q_1\big\}$$

for $\xi \in (0,1)$. Observe that $\varphi(\xi)$ is nondecreasing since for a larger ξ the set where the infimum is taken is smaller. We set

$$q_0 := \inf\big\{\xi \in (0,1) : \varphi(\xi) > 0\big\}. \tag{5.6}$$

Since we want to prove that $q_0 = 0$, we suppose, aiming for a contradiction, that $q_0 > 0$.

First, observe that

$$\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) \geq \mathbb{P}^x(T_{Q_{\frac{1}{2}}} < \tau_{Q_{10\sqrt{N}}}) \inf_{y \in Q_{1/2}} \mathbb{P}^y(T_A < \tau_{Q_1}).$$

By Corollary 5.6 we have that $\mathbb{P}^x(T_{Q_{\frac{1}{2}}} < \tau_{Q_{10\sqrt{N}}})$ is positive and by Lemma 5.1 we have a positive lower bound for $\mathbb{P}^y(T_A < \tau_{Q_1})$ for $y \in Q_{1/2}$ whenever $|Q_1 \setminus A|$ is small enough. Therefore the probability $\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}})$ is uniformly bounded from below for A such that $|Q_1 \setminus A|$ is small enough. We get $1 > q_0$.

By the previous observation, we may choose $q > q_0$ such that $(q + q^2)/2 < q_0$. Thus, for $\eta := (q - q^2)/2$ we have

$$q - \eta = \frac{q + q^2}{2} < q_0 < q.$$

Given $A \subset Q_1$ with $q \geq |A| > q - \eta$, we consider the union of cubes B constructed in Lemma 5.4 for $\delta = q$, $\tilde{\delta} = \eta$ and $L \in \mathbb{N}$ such that $2^L \varepsilon < \varepsilon_0 \leq 2^{L+1}\varepsilon$. Observe that L depends on ε, that is, the depth of the Calderón–Zygmund decomposition depends on ε. This is what allows us to have the smaller cubes in the decomposition of side length comparable to ε. All the other constants are independent of ε. With these choices, by the Calderón–Zygmund lemma, Lemma 5.4, we have $|A| \leq q|B| + \eta$, that is

$$|B| \geq \frac{|A| - \eta}{q} \geq \frac{q - 2\eta}{q} = \frac{q - (q - q^2)}{q} = q.$$

Hence, by the definition of φ and (5.6), we have

$$\mathbb{P}^x(T_B < \tau_{Q_{10\sqrt{N}}}) \geq \varphi(q) > 0, \tag{5.7}$$

since by the choice of q we had $q > q_0$. We can estimate

$$\mathbb{P}^x(T_A < \tau_{Q_{10\sqrt{N}}}) \geq \mathbb{P}^x(T_B < \tau_{Q_{10\sqrt{N}}}) \inf_{y \in B} \mathbb{P}^y(T_A < \tau_{Q_{10\sqrt{N}}}). \tag{5.8}$$

Now we estimate $\mathbb{P}^y(T_A < \tau_{Q_{10\sqrt{N}}})$, separating into two cases depending on $y \in B$. Because of the construction of B we know that one of the following must hold:

- There exists a dyadic cube Q with side length equal to $1/2^L$ such that $y \in Q \subset B$ and $|A \cap Q| > \tilde{\delta}|Q| = \eta|Q|$, or
- there exists a dyadic cube Q with side length larger than or equal to $1/2^L$ such that $y \in \operatorname{pre}(Q) \subset B$ and $|A \cap Q| > \delta|Q| = q|Q|$.

In the first case, by scaling the cube Q to Q_1 (see Remark 5.3) we obtain a process for $\tilde{\varepsilon}$ with $\varepsilon_0/2 \leq \tilde{\varepsilon} = 2^L\varepsilon \leq \varepsilon_0$. By applying Lemma 5.2 we obtain for $y \in Q \subset B$,

$$\mathbb{P}^y(T_A < \tau_{Q_{10\sqrt{N}}}) \geq \eta\gamma(\varepsilon_0). \tag{5.9}$$

Observe that γ depends on ε_0 but not on ε.

In the second case we scale $\operatorname{pre}(Q)$ to Q_1 and obtain a version of the process for $\tilde{\varepsilon} \leq \varepsilon_0$ and some $\tilde{\nu}$. We may assume that the scaled version of Q is $P = Q_1 \cap \{x : x_i > 0 \text{ for } i = 1, \dots, N\}$. We can bound the probability of reaching P by Corollary 5.6 and then the probability of reaching A using that $|A \cap Q| > q|Q|$. By the choice of q, we obtain for $y \in \operatorname{pre}(Q) \subset B$,

$$\mathbb{P}^y(T_A < \tau_{Q_{10\sqrt{N}}}) \geq p\varphi(q) > 0. \tag{5.10}$$

Using (5.9), (5.10) and (5.7) in (5.8), we conclude that

$$\mathbb{P}^y(T_A < \tau_{Q_{10\sqrt{N}}}) \geq \varphi(q)\min\{\eta\gamma, p\varphi(q)\} > 0.$$

Hence, $\varphi(\xi) > 0$ for every $\xi > q - \eta$, which is a contradiction. ∎

Now we state a version of the classical De Giorgi oscillation lemma for the subsolutions of the DPP.

Lemma 5.8 (De Giorgi oscillation lemma). *There exist $k > 1$ and $C, \varepsilon_0 > 0$ such that for every $R > 0$ and $\varepsilon < \varepsilon_0 R$, if u is a subsolution to the DPP in B_{kR} with $u \leq M$ in B_{kR} and*

$$|B_R \cap \{u \leq m\}| \geq \theta|B_R|,$$

for some $\theta > 0$ and $m, M \in \mathbb{R}$, then there exists $\eta = \eta(\theta) > 0$ such that

$$\sup_{B_R} u \leq (1 - \eta)M + \eta m + CR^2\|f\|_\infty.$$

Proof. Recall that u is a subsolution, that is,

$$u(x) \le \alpha \int u(x+\varepsilon z)\, dv_x(z) + \beta \fint_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 f(x).$$

We define $\tilde{u}(x) = u(2Rx)$; we have

$$\tilde{u}\Big(\frac{x}{2R}\Big) \le \alpha \int \tilde{u}\Big(\frac{x+\varepsilon z}{2R}\Big)\, dv_x(z) + \beta \fint_{B_\varepsilon(x)} \tilde{u}\Big(\frac{y}{2R}\Big)\, dy + \varepsilon^2 f(x).$$

We consider $\tilde{x} = \frac{x}{2R}$ and define $\tilde{\varepsilon} = \frac{\varepsilon}{2R}$, $\tilde{\nu}$ such that $\nu_x = \tilde{\nu}_{\tilde{x}}$ and $\tilde{f}$ such that $\varepsilon^2 f(x) = \tilde{\varepsilon}^2 \tilde{f}(\tilde{x})$. We get

$$\tilde{u}(\tilde{x}) \le \alpha \int \tilde{u}(\tilde{x}+\tilde{\varepsilon} z)\, d\tilde{\nu}_{\tilde{x}}(z) + \beta \fint_{B_{\tilde{\varepsilon}}(\tilde{x})} \tilde{u}(y)\, dy + \tilde{\varepsilon}^2 \tilde{f}(\tilde{x}).$$

That is, $\tilde{u}$ is a subsolution to the DPP in $B_{k/2}$ for $\tilde{\varepsilon}$, $\tilde{\nu}$ and $\tilde{f}$ as defined above. We consider the value of ε_0 given by Theorem 5.7. Observe that $\tilde{\varepsilon} = \frac{\varepsilon}{2R} < \varepsilon_0$. Also observe that $\tilde{u} \le M$ in $B_{k/2}$ and

$$|B_{1/2} \cap \{\tilde{u} \le m\}| \ge \theta |B_{1/2}|.$$

We have $B_{1/2} \subset Q_1$. We take $k = 2(5N + \Lambda\varepsilon_0)$, such that $X_{\tau_{Q_{10\sqrt{N}}}} \in B_{k/2}$. We define $A = B_{1/2} \cap \{\tilde{u} \le m\}$ and consider the stopping time

$$T = \min\{T_A, \tau_{Q_{10\sqrt{N}}}\}.$$

For every $\tilde{x} \in Q_1$, we have

$$\begin{aligned}
\tilde{u}(\tilde{x}) &\le \mathbb{E}^{\tilde{x}}\Big[\tilde{u}(X_T) + \tilde{\varepsilon}^2 \sum_{i=0}^{T-1} \tilde{f}(X_i)\Big] \\
&\le \mathbb{E}^{\tilde{x}}[\tilde{u}(X_T) | T_A < \tau_{Q_{10\sqrt{N}}}] \mathbb{P}^{\tilde{x}}(T_A < \tau_{Q_{10\sqrt{N}}}) \\
&\quad + \mathbb{E}^{\tilde{x}}[\tilde{u}(X_T) | T_A > \tau_{Q_{10\sqrt{N}}}](1 - \mathbb{P}^{\tilde{x}}(T_A < \tau_{Q_{10\sqrt{N}}})) + \|\tilde{f}\|_\infty \mathbb{E}^{\tilde{x}}[\tilde{\varepsilon}^2 T] \\
&\le M \mathbb{P}^{\tilde{x}}(T_A < \tau_{Q_{10\sqrt{N}}}) + m(1 - \mathbb{P}^{\tilde{x}}(T_A < \tau_{Q_{10\sqrt{N}}})) + C\|\tilde{f}\|_\infty,
\end{aligned}$$

where the first inequality holds since $\tilde{u}$ is a subsolution to the DPP and we have bounded $\mathbb{E}^{\tilde{x}}[\tilde{\varepsilon}^2 T]$ by Lemma 2.9.

Observe that $\inf_{\tilde{x} \in B_{1/2}} \mathbb{P}^{\tilde{x}}(T_A < \tau_{Q_{10\sqrt{N}}})$ is positive as stated in Theorem 5.7. Also observe that $\|\tilde{f}\|_\infty = (2R)^2 \|f\|_\infty$. Therefore, we have proved the result since bounding $\tilde{u}(\tilde{x})$ for every $\tilde{x} \in B_{1/2} \subset Q_1$ is equivalent to bounding $u(x)$ for every $x \in B_R$. ∎

Observe that the values of k and η do not depend on ε nor R. And an analogous statement holds for supersolutions.

6. Proof of the Hölder estimate

The Hölder estimate follows from the De Giorgi oscillation lemma, Lemma 5.8, after a finite iteration. We include the details here as we have to take special care of the role of ε in the arguments.

Given a function u we define

$$M(R) = \sup_{B_R} u, \quad m(R) = \inf_{B_R} u \quad \text{and} \quad \mathcal{O}(R) = M(R) - m(R).$$

We also define

$$\operatorname{osc}(A) = \sup_{A} u - \inf_{A} u.$$

Observe that $\operatorname{osc}(B_R) = \mathcal{O}(R)$.

Lemma 6.1. *There exist $\lambda < 1$ and $k > 1$ such that for every solution u to the DPP defined in B_{kR} we have*

$$\mathcal{O}(R) \leq \lambda \mathcal{O}(kR) + CR^2 \|f\|_\infty$$

for every R and $\varepsilon < R\varepsilon_0$.

Proof. We can assume that $\mathcal{O}(kR) \neq 0$ where k is given by Lemma 5.8. We consider $l = (M(kR) + m(kR))/2$. Either

$$|\{u \geq l\} \cap B_R| \geq |B_R|/2$$

or

$$|\{u \leq l\} \cap B_R| \geq |B_R|/2.$$

Suppose that the first holds (the proof is completely analogous in the other case); then since $u \geq m(kR)$ and $l = m(kR) + \frac{\mathcal{O}(kR)}{2}$, Lemma 5.8 implies that

$$m(R) \geq m(kR) + \eta \frac{\mathcal{O}(kR)}{2} - CR^2 \|f\|_\infty$$

for some $\eta = \eta(\frac{1}{2}) > 0$. Then, since $M(R) \leq M(kR)$, we have

$$\begin{aligned} \mathcal{O}(R) &= M(R) - m(R) \\ &\leq M(kR) - m(kR) - \frac{\eta}{2}\mathcal{O}(kR) + CR^2 \|f\|_\infty \\ &= \left(1 - \frac{\eta}{2}\right)\mathcal{O}(kR) + CR^2 \|f\|_\infty. \end{aligned}$$

Thus, the statement holds for $\lambda = 1 - \eta/2$. ∎

By iterating the oscillation estimate from Lemma 6.1 we can obtain the Hölder regularity. To that end we prove the following lemma.

Lemma 6.2. *If $\mathcal{O}(s) \geq 0$ is a nondecreasing function and $\mathcal{O}(s) \leq \lambda \mathcal{O}(ks) + Cs^2 \|f\|_\infty$ for every $s > \xi$ for some $\lambda \in (0,1)$, $k > 1$ and $\xi > 0$ such that $\lambda k^2 > 1$, then*

$$\mathcal{O}(\rho) \leq \frac{1}{\lambda}\Big(\frac{\rho}{R}\Big)^\gamma (\mathcal{O}(R) + CR^2\|f\|_\infty)$$

for every $R \geq \rho > \xi$ where $\gamma = \frac{\log \frac{1}{\lambda}}{\log k}$.

Proof. Since $k > 1$ there exists a unique $m \in \mathbb{N}_0$ such that

$$k^m \leq \frac{R}{\rho} < k^{m+1}.$$

By repeatedly using that $\mathcal{O}(s) \leq \lambda \mathcal{O}(ks) + Cs^2\|f\|_\infty$, for $s = \frac{R}{k^m}, \frac{R}{k^{m-1}}, \dots, \frac{R}{k}$ we obtain

$$\begin{aligned}
\mathcal{O}(\rho) &\leq \mathcal{O}\Big(\frac{R}{k^m}\Big) \\
&\leq \lambda \mathcal{O}\Big(\frac{R}{k^{m-1}}\Big) + C\Big(\frac{R}{k^m}\Big)^2\|f\|_\infty \\
&\leq \lambda^2 \mathcal{O}\Big(\frac{R}{k^{m-2}}\Big) + \lambda C\Big(\frac{R}{k^{m-1}}\Big)^2\|f\|_\infty + C\Big(\frac{R}{k^m}\Big)^2\|f\|_\infty \\
&\leq \lambda^m \mathcal{O}(R) + CR^2\|f\|_\infty\Big(\frac{\lambda^{m-1}}{k^2} + \cdots + \frac{\lambda}{k^{2(m-1)}} + \frac{1}{k^{2m}}\Big) \\
&= \lambda^m \mathcal{O}(R) + CR^2\|f\|_\infty \lambda^m\Big(\frac{1}{\lambda k^2} + \cdots + \frac{1}{(\lambda k^2)^{m-1}} + \frac{1}{(\lambda k^2)^m}\Big) \\
&\leq \lambda^m \mathcal{O}(R) + CR^2\|f\|_\infty \lambda^m \frac{1}{\lambda k^2 - 1}.
\end{aligned}$$

Observe that we have used the hypothesis for values larger than ξ, in fact

$$\frac{R}{k} \geq \frac{R}{k^2} \geq \cdots \geq \frac{R}{k^m} \geq \rho > \xi.$$

We have

$$\begin{aligned}
&\log \frac{R}{\rho} < (m+1)\log k, \\
&\frac{\log \frac{R}{\rho}}{\log k} < (m+1), \\
&\lambda^{\frac{\log \frac{R}{\rho}}{\log k}} > \lambda^{m+1}.
\end{aligned}$$

Thus, the inequality follows since

$$\lambda^m = \frac{1}{\lambda}\lambda^{m+1} < \frac{1}{\lambda}\lambda^{\frac{\log \frac{R}{\rho}}{\log k}} = \frac{1}{\lambda}\Big(\frac{\rho}{R}\Big)^\gamma.$$

■

Observe that Lemmas 6.1 and 6.2 prove that, given u a solution to the DPP defined in $B_R(x)$,

$$\operatorname{osc}(B_\rho(x)) \leq \frac{1}{\lambda}\Big(\frac{\rho}{R}\Big)^\gamma \big(\operatorname{osc}(B_R(x)) + CR^2\|f\|_\infty\big)$$

for $\varepsilon < \rho\varepsilon_0$. We are ready to prove the Hölder estimate.

Proof of Theorem 1.1. Given $x, z \in B_R$ we consider $\rho = |x - z|$.

If $\rho \geq R$, we have

$$|u(x) - u(z)| \leq 2 \sup_{B_{2R}} |u| \frac{|x - z|^\gamma}{R^\gamma}.$$

If $\rho < R$ and $\varepsilon < \rho\varepsilon_0$, we employ the previous lemma and obtain

$$\begin{aligned}
|u(x) - u(z)| &\leq \operatorname{osc}(\overline{B_\rho(x)}) \\
&\leq \frac{1}{\lambda}\Big(\frac{\rho}{R}\Big)^\gamma \big(\operatorname{osc}(\overline{B_R(x)}) + CR^2\|f\|_\infty\big) \\
&\leq \frac{1}{\lambda}\Big(\frac{\rho}{R}\Big)^\gamma \big(\operatorname{osc}(B_{2R}) + CR^2\|f\|_\infty\big) \\
&\leq \frac{2\sup_{B_{2R}} |u| + CR^2\|f\|_\infty}{R^\gamma \lambda} |x - z|^\gamma.
\end{aligned}$$

In the case $\varepsilon \geq \rho\varepsilon_0$ then we can estimate $\operatorname{osc}(\overline{B_\rho(x)})$ by $\operatorname{osc}(\overline{B_{\rho'}(x)})$ for $\rho' = \frac{\varepsilon}{\varepsilon_0}$. We get

$$\begin{aligned}
|u(x) - u(z)| &\leq \operatorname{osc}(\overline{B_\rho(x)}) \\
&\leq \operatorname{osc}(\overline{B_{\rho'}(x)}) \\
&\leq \frac{1}{\lambda}\Big(\frac{\rho'}{R}\Big)^\gamma \big(\operatorname{osc}(\overline{B_R(x)}) + CR^2\|f\|_\infty\big) \\
&\leq \frac{1}{\lambda}\Big(\frac{\rho'}{R}\Big)^\gamma \big(\operatorname{osc}(B_{2R}) + CR^2\|f\|_\infty\big) \\
&\leq \frac{1}{\lambda}\Big(\frac{\varepsilon}{\varepsilon_0 R}\Big)^\gamma 2\Big(\sup_{B_{2R}} |u| + CR^2\|f\|_\infty\Big).
\end{aligned}$$

Observe that we have $\rho' \leq R$ since $\varepsilon < \varepsilon_0 R$. ∎

7. Generalization to Pucci-type operators and inequalities

Here we explain how to modify our arguments to include solutions to Pucci-type operators and inequalities. Our method is robust and essentially the same arguments remain valid.

We start by defining the operator and then a stochastic process associated to it.

Definition 7.1. Let $u\colon \mathbb{R}^N \to \mathbb{R}$ be a bounded Borel measurable function. We define the maximal Pucci-type operator

$$\begin{aligned}\mathcal{L}_\varepsilon^+ u(x) &:= \frac{1}{2\varepsilon^2}\left(\alpha \sup_{\nu\in\mathcal{M}(B_\Lambda)} \int \delta u(x,\varepsilon z)\,d\nu(z) + \beta \fint_{B_1} \delta u(x,\varepsilon y)\,dy\right)\\ &= \frac{1}{2\varepsilon^2}\left(\alpha \sup_{z\in B_\Lambda} \delta u(x,\varepsilon z) + \beta \fint_{B_1} \delta u(x,\varepsilon y)\,dy\right),\end{aligned}$$

where $\delta u(x,\varepsilon y) = u(x+\varepsilon y) + u(x-\varepsilon y) - 2u(x)$ for every $y \in B_\Lambda$; $\mathcal{L}_\varepsilon^-$ is defined analogously, just replacing sup by inf.

For related operators, see [6].

For each $\varepsilon > 0$ we consider a stochastic process starting at $x_0 \in \mathbb{R}^N$. The process is driven by a controller. Given the value of x_k, the next position of the process x_{k+1} is determined as follows. A biased coin is tossed. If we get heads (probability α), the controller chooses $z \in B_\Lambda$ and we have $x_{k+1} = x_k \pm \varepsilon z$, each with probability $1/2$. If we get tails (probability β), x_{k+1} is distributed uniformly in the ball $B_\varepsilon(x_k)$.

To be more precise, a strategy S for the controller is a measurable function defined on the partial histories, that is,

$$S(x_0, x_1, \dots, x_k) = z \in B_\Lambda.$$

Then the process is moved according to this choice. That is, given $A \in \mathcal{B}$ and $c = 0$ or 1, we have the following transition probabilities:

$$\pi_S(x_0,(c_1,x_1),\dots,(c_k,x_k),\{c\}\times A) = \begin{cases} \alpha \dfrac{\delta_z + \delta_{-z}}{2}\left(\dfrac{A - x_k}{\varepsilon}\right) & \text{if } c = 0,\\[2ex] \beta \dfrac{|A \cap B_\varepsilon(x_k)|}{|B_\varepsilon|} & \text{if } c = 1,\end{cases}$$

where $z = S(x_0, x_1, \dots, x_k)$.

For a fixed strategy we have a process as before. The only difference is that now the measure ν may depend not only on x but S (this does not introduce any difference into our arguments). For a fixed S we can consider $\mathbb{E}_S^{x_0}$, the corresponding expectation. All the estimates obtained for $\mathbb{E}^{x_0}$ hold for $\mathbb{E}_S^{x_0}$ and are independent of S.

We consider a game where the controller is paid $g(x_\tau)$ at the end and therefore it is their goal to maximize that value. The expectation for earnings is given by

$$u(x_0) = \sup_S \mathbb{E}_S^{x_0}[g(x_\tau)],$$

where $\mathbb{E}_S^{x_0}$ stands for the expectation with respect to the process and S is the strategy adopted by the controller. The function $u\colon \mathbb{R}^N \to \mathbb{R}$ satisfies the DPP given by

$$u(x) = \alpha \sup_{z\in B_\Lambda} \frac{u(x+\varepsilon z) + u(x-\varepsilon z)}{2} + \beta \fint_{B_\varepsilon(x)} u(y)\,dy$$

for $x \in \Omega$, and $u(x) = g(x)$ for $x \notin \Omega$.

We can consider a version of the game where whenever the token leaves a point x_i, the controller is paid $\varepsilon^2 f(x_i)$. In this case the expectation for earnings is given by

$$u(x) = \sup_S \mathbb{E}_S^x \left[\varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i) + g(X_\tau) \right]. \tag{7.1}$$

It turns out, as will be shown below, that u is the unique bounded Borel measurable function that satisfies

$$u(x) = \alpha \sup_{z \in B_\Lambda} \frac{u(x+\varepsilon z) + u(x-\varepsilon z)}{2} + \beta \fint_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 f(x) \tag{7.2}$$

for $x \in \Omega$ and $u(x) = g(x)$ for $x \notin \Omega$. Or equivalently $\mathcal{L}_\varepsilon^+ u + f = 0$.

The existence of a solution to equation (7.2) can be seen as before. Next we prove the equivalent to Theorem 3.4.

Theorem 7.2. *The function u given by (7.1) is the unique bounded solution to equation (7.2) with boundary values g.*

Proof. Let v be a solution to equation (7.2). Given a strategy S_0 we have

$$\begin{aligned}
&\mathbb{E}_{S_0}^{x_0}\left[v(X_{k+1}) + \varepsilon^2 \sum_{i=0}^{k} f(X_i) \,\middle|\, \mathcal{F}_k \right](\omega) \\
&\quad \le \alpha \sup_{z \in B_\Lambda} \frac{v(x_k+\varepsilon z) + v(x_k-\varepsilon z)}{2} + \beta \fint_{B_\varepsilon(x_k)} v(y)\, dy + \varepsilon^2 \sum_{i=0}^{k} f(x_i) \\
&\quad = v(x_k) + \varepsilon^2 \sum_{i=0}^{k-1} f(x_i).
\end{aligned}$$

Thus $\{v(X_k) + \varepsilon^2 \sum_{i=0}^{k-1} f(X_i)\}_k$ is a supermartingale. Then, by Doob's stopping time theorem (recall that v and f are bounded), we have

$$v(x) \ge \mathbb{E}_{S_0}^x \left[v(X_\tau) + \varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i) \right].$$

Since this holds for every strategy and $v(x_\tau) = u(x_\tau)$, we get

$$v(x) \ge \sup_S \mathbb{E}_S^x \left[v(X_\tau) + \varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i) \right] = u(x).$$

On the other hand, given $\eta > 0$ we consider a strategy S_0 that almost maximizes the right-hand side of (7.2), that is, $S_0(x_0, \dots, x_k) = \tilde{z} \in B_\Lambda$ such that

$$\frac{v(x_k+\varepsilon\tilde{z}) + v(x_k-\varepsilon\tilde{z})}{2} \ge \sup_{z \in B_\Lambda} \frac{v(x_k+\varepsilon z) + v(x_k-\varepsilon z)}{2} - \eta 2^{-(k+1)}.$$

The strategy can be taken measurable similarly to [25, Lemma 3.1].

We have

$$\begin{aligned}
&\mathbb{E}_{S_0}^{x_0}\left[v(X_{k+1}) + \varepsilon^2 \sum_{i=0}^{k} f(X_i) - \eta 2^{-(k+1)} \,\middle|\, \mathcal{F}_k\right](\omega) \\
&\quad \geq \alpha \sup_{z \in B_\Lambda} \frac{v(x_k + \varepsilon z) + v(x_k - \varepsilon z)}{2} + \beta \fint_{B_\varepsilon(x_k)} v(y)\, dy \\
&\qquad + \varepsilon^2 \sum_{i=0}^{k} f(x_i) - \eta 2^{-(k+1)} - \eta 2^{-(k+1)} \\
&\quad = v(x_k) + \varepsilon^2 \sum_{i=0}^{k-1} f(x_i) - \eta 2^{-k}.
\end{aligned}$$

Thus $\{v(X_k) + \varepsilon^2 \sum_{i=0}^{k-1} f(X_i) - \eta 2^{-k}\}_k$ is a submartingale. Then

$$\begin{aligned}
v(x) - \eta &\leq \mathbb{E}_{S_0}^{x}\left[v(X_\tau) + \varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i) - \eta 2^{-\tau}\right] \\
&\leq \sup_S \mathbb{E}_S^x\left[v(X_\tau) + \varepsilon^2 \sum_{i=0}^{\tau-1} f(X_i)\right] \\
&= u(x).
\end{aligned}$$

Since this holds for every $\eta > 0$ we conclude that $v \leq u$. Thus $v = u$, and we have proved that u is a solution to (7.2) and that every solution coincides with it, so there is a unique solution. ∎

Finally, we state Theorem 1.2 again, which is our main result of this section: one only needs Pucci-type inequalities in order to obtain the regularity result.

Theorem 7.3. *Let f be a bounded Borel function. There exists $\varepsilon_0 > 0$ such that if u satisfies*

$$\mathcal{L}_\varepsilon^+ u \geq -|f|, \quad \mathcal{L}_\varepsilon^- u \leq |f| \tag{7.3}$$

in B_{2R} where $\varepsilon < \varepsilon_0 R$, there exist $\gamma > 0$ and $C > 0$ such that

$$|u(x) - u(z)| \leq \frac{C}{R^\gamma}\left(\sup_{B_{2R}} |u| + R^2 \|f\|_\infty\right)(|x - z|^\gamma + \varepsilon^\gamma)$$

for every $x, z \in B_R$.

Remark 7.4. Observe that the ε-ABP estimate (Theorem 4.1), as well as all the results from Section 4, is valid if we consider the maximal Pucci-type operator $\mathcal{L}_\varepsilon^+$ instead of $\mathcal{L}_\varepsilon$. This is due to the fact that (similarly to equations (4.4) and (4.5)), if u is a bounded Borel measurable function satisfying $\mathcal{L}_\varepsilon^+ u + f \geq 0$ in Ω, then $\mathcal{L}_\varepsilon^+ \Gamma + f \geq 0$ in K_u, where Γ

is the concave envelope of u and K_u is the set of contact points defined in (4.1). Hence, using this together with the fact that the second differences satisfy $\delta\Gamma(x_0, y) \le 0$ for each $x_0 \in K_u$, we can estimate the α-term in $\mathcal{L}_\varepsilon^+\Gamma(x_0)$, so

$$\mathcal{L}_\varepsilon^+\Gamma(x_0) \le \frac{\beta}{2\varepsilon^2} \fint_{B_\varepsilon} \delta\Gamma(x_0, y)\, dy.$$

This is analogous to inequality (4.12) in the proof of Lemma 4.4, and it is indeed the cornerstone in all the estimates from Section 4.

With the analogous results of Sections 2 and 4 for $\mathcal{L}_\varepsilon^+$ in hand, those of Section 5 follow. However, there is a key modification needed in the analogous version of Lemma 5.8 where, after establishing some estimates related to the stochastic process, solutions to the DPP are considered. Here we adapt our argument to functions satisfying (7.3).

Lemma 7.5. *There exist $k > 1$ and $C, \varepsilon_0 > 0$ such that for every $R > 0$ and $\varepsilon < \varepsilon_0 R$, if u satisfies $\mathcal{L}_\varepsilon^+ u \ge -|f|$ in B_{kR} with $u \le M$ in B_{kR} and*

$$|B_R \cap \{u \le m\}| \ge \theta |B_R|$$

for some $\theta > 0$ and $m, M \in \mathbb{R}$, then there exists $\eta = \eta(\theta) > 0$ such that

$$\sup_{B_R} u \le (1-\eta)M + \eta m + CR^2 \|f\|_\infty.$$

Proof. The function u satisfies $\mathcal{L}_\varepsilon^+ u \ge -|f|$, that is,

$$u(x) \le \alpha \sup_{z \in B_\Lambda} \frac{u(x+\varepsilon z) + u(x-\varepsilon z)}{2} + \beta \fint_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 |f(x)|.$$

We define $\tilde{u}(x) = u(2Rx)$. We have

$$\tilde{u}\Big(\frac{x}{2R}\Big) \le \alpha \sup_{z \in B_\Lambda} \frac{\tilde{u}(\frac{x+\varepsilon z}{2R}) + \tilde{u}(\frac{x-\varepsilon z}{2R})}{2} + \beta \fint_{B_\varepsilon(x)} \tilde{u}\Big(\frac{y}{2R}\Big)\, dy + \varepsilon^2 |f(x)|.$$

We consider $\tilde{x} = \frac{x}{2R}$ and define $\tilde{\varepsilon} = \frac{\varepsilon}{2R}$ and $\tilde{f}$ such that $\varepsilon^2 f(x) = \tilde{\varepsilon}^2 \tilde{f}(\tilde{x})$. We get

$$\tilde{u}(\tilde{x}) \le \alpha \sup_{z \in B_\Lambda} \frac{\tilde{u}(\tilde{x}+\tilde{\varepsilon} z) + \tilde{u}(\tilde{x}-\tilde{\varepsilon} z)}{2} + \beta \fint_{B_{\tilde{\varepsilon}}(\tilde{x})} \tilde{u}(y)\, dy + \tilde{\varepsilon}^2 |\tilde{f}(\tilde{x})|, \tag{7.4}$$

where $\tilde{u}$ is defined in $B_{k/2}$. We consider the value of ε_0 given by Theorem 5.7. Observe that $\tilde{\varepsilon} = \frac{\varepsilon}{2R} < \varepsilon_0$. Also observe that $\tilde{u} \le M$ in $B_{k/2}$ and

$$|B_{1/2} \cap \{\tilde{u} \le m\}| \ge \theta |B_{1/2}|.$$

Given $\eta > 0$ we consider the strategy S_0 the almost maximizes the right-hand side of (7.4), that is, $S_0(x_0, \dots, x_k) = \tilde{z} \in B_\Lambda$ such that

$$\frac{u(x+\varepsilon\tilde{z}) + u(x-\varepsilon\tilde{z})}{2} \ge \sup_{z \in B_\Lambda} \frac{u(x+\varepsilon z) + u(x-\varepsilon z)}{2} - \eta 2^{-(k+1)}.$$

As in the proof of Theorem 7.2 we get that $\{v(X_k) + \varepsilon^2 \sum_{i=0}^{k-1} f(X_i) - \eta 2^{-k}\}_k$ is a submartingale for $\mathbb{E}^{\tilde{x}}_{S_0}$.

We have $B_{1/2} \subset Q_1$. We take $k = 2(5N + \Lambda\varepsilon_0)$, such that $X_{\tau_{Q_{10\sqrt{N}}}} \in B_{k/2}$. We define $A = B_{1/2} \cap \{\tilde{u} \le m\}$ and consider the stopping time

$$T = \min\{T_A, \tau_{Q_{10\sqrt{N}}}\}.$$

For every $\tilde{x} \in Q_1$ (and in particular for those in $B_{1/2}$) we have

$$\begin{aligned}
\tilde{u}(\tilde{x}) - \eta &\le \mathbb{E}^{\tilde{x}}_{S_0}\left[\tilde{u}(X_T) + \tilde{\varepsilon}^2 \sum_{i=0}^{T-1} |\tilde{f}(X_i)| - \eta 2^{-T}\right] \\
&\le \mathbb{E}^{\tilde{x}}_{S_0}[\tilde{u}(X_T) | T_A < \tau_{Q_{10\sqrt{N}}}] \mathbb{P}^{\tilde{x}}_{S_0}(T_A < \tau_{Q_{10\sqrt{N}}}) \\
&\quad + \mathbb{E}^{\tilde{x}}_{S_0}[\tilde{u}(X_T) | T_A > \tau_{Q_{10\sqrt{N}}}](1 - \mathbb{P}^{\tilde{x}}(T_A < \tau_{Q_{10\sqrt{N}}})) + \|\tilde{f}\|_\infty \mathbb{E}^{\tilde{x}}_{S_0}[\tilde{\varepsilon}^2 T] \\
&\le M \mathbb{P}^{\tilde{x}}_{S_0}(T_A < \tau_{Q_{10\sqrt{N}}}) + m(1 - \mathbb{P}^{\tilde{x}}_{S_0}(T_A < \tau_{Q_{10\sqrt{N}}})) + C\|\tilde{f}\|_\infty,
\end{aligned}$$

where we have bounded $\mathbb{E}^{\tilde{x}}_{S_0}[\tilde{\varepsilon}^2 T]$ by Lemma 2.9.

Observe that $\inf_{\tilde{x} \in B_{1/2}} \mathbb{P}^{\tilde{x}}(T_A < \tau_{Q_{10\sqrt{N}}})$ is positive as stated in Theorem 5.7. Also observe that $\|\tilde{f}\|_\infty = (2R)^2 \|f\|_\infty$. Therefore, we have proved the result since bounding $\tilde{u}(\tilde{x})$ for every $\tilde{x} \in B_{1/2}$ is equivalent to bounding $u(x)$ for every $x \in B_R$. Finally since the inequality holds for every $\eta > 0$ it holds without it. ∎

Remark 7.6. Given nonempty subsets $\mathcal{M}_x \subset \mathcal{M}(B_\Lambda)$ for each $x \in \mathbb{R}^N$ with suitable measurability requirements, we can consider solutions to the equation

$$u(x) = \alpha \sup_{\nu \in \mathcal{M}_x} \int u(x + \varepsilon z)\, d\nu(z) + \beta \fint_{B_\varepsilon(x)} u(y)\, dy + \varepsilon^2 f(x).$$

Observe that such functions would satisfy (7.3) and therefore would be in the hypothesis of Theorem 7.3.

Our results also cover tug-of-war games with noise. Indeed, the value functions satisfy

$$\frac{1}{2\varepsilon^2}\left(\alpha\Big(\sup_{B_\varepsilon(x)} u + \inf_{B_\varepsilon(x)} u - 2u(x)\Big) + \beta \fint_{B_1} \delta u(x, \varepsilon y)\, dy\right) + f(x) = 0. \tag{7.5}$$

Since

$$\sup_{B_\varepsilon(x)} u + \inf_{B_\varepsilon(x)} u \le \sup_{z \in B_1} (u(x + \varepsilon z) + u(x - \varepsilon z))$$

we have $0 \le f + \mathcal{L}^+_\varepsilon u$ and similarly $0 \ge f + \mathcal{L}^-_\varepsilon u$. Therefore, solutions to (7.5) satisfy (7.3). Moreover, we can consider solutions to the DPP associated to the normalized $p(x)$-Laplacian given by

$$u(x) = \frac{\alpha(x)}{2}\Big(\sup_{B_\varepsilon(x)} u + \inf_{B_\varepsilon(x)} u\Big) + \beta(x) \fint_{B_\varepsilon(x)} u(z)\, dz + \varepsilon^2 f(x). \tag{7.6}$$

Let $\beta^- := \inf_{x\in\Omega} \beta(x) > 0$. Then we observe that

$$\begin{aligned} 0 &\le f(x) + \frac{1}{2\varepsilon^2}\left(\alpha(x) \sup_{z\in B_\Lambda} \delta u(x,\varepsilon z) + \beta(x) \fint_{B_1} \delta u(x,\varepsilon y)\,dy\right) \\ &\le f(x) + \frac{1}{2\varepsilon^2}\left((1-\beta^-) \sup_{z\in B_\Lambda} \delta u(x,\varepsilon z) + \beta^- \fint_{B_1} \delta u(x,\varepsilon y)\,dy\right). \end{aligned}$$

Similarly,

$$\begin{aligned} 0 &\ge f(x) + \frac{1}{2\varepsilon^2}\left(\alpha(x) \inf_{z\in B_\Lambda} \delta u(x,\varepsilon z) + \beta(x) \fint_{B_1} \delta u(x,\varepsilon y)\,dy\right) \\ &\ge f(x) + \frac{1}{2\varepsilon^2}\left((1-\beta^-) \inf_{z\in B_\Lambda} \delta u(x,\varepsilon z) + \beta^- \fint_{B_1} \delta u(x,\varepsilon y)\,dy\right). \end{aligned}$$

Thus solutions to (7.6) satisfy the hypotheses of Theorem 7.3.

In a similar way, there is a large family of discrete operators associated to different PDEs that are in the hypothesis of our main result.

Funding. Á. A. is partially supported by a UniGe starting grant "Curiosity driven" and grants MTM2017-85666-P, 2017 SGR 395. B. P. and M. P. are partially supported by the Academy of Finland project #298641.

References

[1] Á. Arroyo, J. Heino, and M. Parviainen, Tug-of-war games with varying probabilities and the normalized $p(x)$-Laplacian. *Commun. Pure Appl. Anal.* **16** (2017), no. 3, 915–944 Zbl 1359.35062 MR 3623556

[2] Á. Arroyo and M. Parviainen, Asymptotic Hölder regularity for the ellipsoid process. *ESAIM Control Optim. Calc. Var.* **26** (2020), Paper No. 112 Zbl 1459.35061 MR 4185066

[3] R. F. Bass, *Diffusions and elliptic operators*. Probab. Appl. (New York), Springer, New York, 1998 Zbl 0914.60009 MR 1483890

[4] P. Blanc, C. Esteve, and J. D. Rossi, The evolution problem associated with eigenvalues of the Hessian. *J. Lond. Math. Soc. (2)* **102** (2020), no. 3, 1293–1317 Zbl 1462.35136 MR 4186128

[5] P. Blanc, J. P. Pinasco, and J. D. Rossi, Maximal operators for the p-Laplacian family. *Pacific J. Math.* **287** (2017), no. 2, 257–295 Zbl 1362.35143 MR 3632889

[6] K. K. Brustad, P. Lindqvist, and J. J. Manfredi, A discrete stochastic interpretation of the dominative p-Laplacian. *Differential Integral Equations* **33** (2020), no. 9-10, 465–488 Zbl 1474.35292 MR 4149517

[7] L. Caffarelli and L. Silvestre, Regularity theory for fully nonlinear integro-differential equations. *Comm. Pure Appl. Math.* **62** (2009), no. 5, 597–638 Zbl 1170.45006 MR 2494809

[8] L. Caffarelli, R. Teymurazyan, and J. M. Urbano, Fully nonlinear integro-differential equations with deforming kernels. *Comm. Partial Differential Equations* **45** (2020), no. 8, 847–871 Zbl 1452.35055 MR 4126323

[9] L. A. Caffarelli and X. Cabré, *Fully nonlinear elliptic equations*. Amer. Math. Soc. Colloq. Publ. 43, American Mathematical Society, Providence, RI, 1995 MR 1351007

[10] L. A. Caffarelli, R. Leitão, and J. M. Urbano, Regularity for anisotropic fully nonlinear integro-differential equations. *Math. Ann.* **360** (2014), no. 3-4, 681–714 Zbl 1304.35730 MR 3273642

[11] H. Chang Lara and G. Dávila, Regularity for solutions of nonlocal, nonsymmetric equations. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **29** (2012), no. 6, 833–859 Zbl 1317.35278 MR 2995098

[12] D. Gilbarg and N. S. Trudinger, *Elliptic partial differential equations of second order*. Classics in Mathematics, Springer, Berlin, 2001 Zbl 1042.35002 MR 1814364

[13] N. Guillen and R. W. Schwab, Aleksandrov-Bakelman-Pucci type estimates for integro-differential equations. *Arch. Ration. Mech. Anal.* **206** (2012), no. 1, 111–157 Zbl 1256.35190 MR 2968592

[14] H. Hartikainen, A dynamic programming principle with continuous solutions related to the p-Laplacian, $1 < p < \infty$. *Differential Integral Equations* **29** (2016), no. 5-6, 583–600 Zbl 1374.35020 MR 3471974

[15] N. V. Krylov and M. V. Safonov, An estimate for the probability of a diffusion process hitting a set of positive measure. *Dokl. Akad. Nauk SSSR* **245** (1979), no. 1, 18–20 MR 525227

[16] N. V. Krylov and M. V. Safonov, A property of the solutions of parabolic equations with measurable coefficients. *Izv. Akad. Nauk SSSR Ser. Mat.* **44** (1980), no. 1, 161–175, 239 MR 563790

[17] H. J. Kuo and N. S. Trudinger, Linear elliptic difference inequalities with random coefficients. *Math. Comp.* **55** (1990), no. 191, 37–53 Zbl 0716.39005 MR 1023049

[18] S. Kusuoka, Hölder continuity and bounds for fundamental solutions to nondivergence form parabolic equations. *Anal. PDE* **8** (2015), no. 1, 1–32 Zbl 1316.35064 MR 3336920

[19] G. F. Lawler, Estimates for differences and Harnack inequality for difference operators coming from random walks with symmetric, spatially inhomogeneous, increments. *Proc. London Math. Soc. (3)* **63** (1991), no. 3, 552–568 Zbl 0701.39002 MR 1127149

[20] M. Lewicka, *A course on tug-of-war games with random noise*. Universitext, Springer, Cham, 2020 Zbl 1452.91002 MR 4174394

[21] T. Lindvall and L. C. G. Rogers, Coupling of multidimensional diffusions by reflection. *Ann. Probab.* **14** (1986), no. 3, 860–872 Zbl 0593.60076 MR 841588

[22] H. Luiro and M. Parviainen, Gradient walks and p-harmonic functions. *Proc. Amer. Math. Soc.* **145** (2017), no. 10, 4313–4324 Zbl 1376.35082 MR 3690615

[23] H. Luiro and M. Parviainen, Regularity for nonlinear stochastic games. *Ann. Inst. H. Poincaré C Anal. Non Linéaire* **35** (2018), no. 6, 1435–1456 Zbl 1398.91058 MR 3846232

[24] H. Luiro, M. Parviainen, and E. Saksman, Harnack's inequality for p-harmonic functions via stochastic games. *Comm. Partial Differential Equations* **38** (2013), no. 11, 1985–2003 Zbl 1287.35044 MR 3169768

[25] H. Luiro, M. Parviainen, and E. Saksman, On the existence and uniqueness of p-harmonious functions. *Differential Integral Equations* **27** (2014), no. 3-4, 201–216 Zbl 1324.49029 MR 3161602

[26] J. J. Manfredi, M. Parviainen, and J. D. Rossi, On the definition and properties of p-harmonious functions. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **11** (2012), no. 2, 215–241 Zbl 1252.91014 MR 3011990

[27] Y. Peres, O. Schramm, S. Sheffield, and D. B. Wilson, Tug-of-war and the infinity Laplacian. *J. Amer. Math. Soc.* **22** (2009), no. 1, 167–210 Zbl 1206.91002 MR 2449057

[28] Y. Peres and S. Sheffield, Tug-of-war with noise: a game-theoretic view of the p-Laplacian. *Duke Math. J.* **145** (2008), no. 1, 91–120 Zbl 1206.35112 MR 2451291

[29] C. Pucci and G. Talenti, Elliptic (second-order) partial differential equations with measurable coefficients and approximating integral equations. *Advances in Math.* **19** (1976), no. 1, 48–105 MR 419989

[30] E. Ruosteenoja, Local regularity results for value functions of tug-of-war with noise and running payoff. *Adv. Calc. Var.* **9** (2016), no. 1, 1–17 Zbl 1331.35069 MR 3441079

[31] R. W. Schwab and L. Silvestre, Regularity for parabolic integro-differential equations with very irregular kernels. *Anal. PDE* **9** (2016), no. 3, 727–772 Zbl 1349.47079 MR 3518535

Received 1 December 2020; revised 4 October 2021; accepted 11 November 2021.

Ángel Arroyo
MOMAT Research Group, Interdisciplinary Mathematics Institute, Department of Applied Mathematics and Mathematical Analysis, Universidad Complutense de Madrid, 28040 Madrid, Spain; ar.arroyo@ucm.es

Pablo Blanc
Department of Mathematics and Statistics, University of Jyväskylä, PO Box 35, 40014 Jyväskylä, Finland; pblanc@dm.uba.ar

Mikko Parviainen
Department of Mathematics and Statistics, University of Jyväskylä, PO Box 35, 40014 Jyväskylä, Finland; mikko.j.parviainen@jyu.fi